Early detection of Breast Cancer in microscopic scales by following frequencies changes in mammography images Using Cellular Automata

Fatemeh Esmaeili¹, Saeed Setayeshi², Mohammad Esmaeil Akbari³

¹ Faculty of Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran
² Dept. of Medical Radiation, Engineering, Amirkabir University of Technology, Tehran, Iran
³ Cancer Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran

ABSTRACT
Breast cancer is one of the most frequent fatal cancers among women. It is one of the cancers in which the early diagnosis at early stages increases the chance for a successful treatment and patients’ complete recovery. Although studies have shown that early diagnosis through mammography increases patients’ five-year survival, it has 9% false negative. Along with the advances of computer technology, radiologists can improve their diagnosis using the ability of computers to improve the quality of images, select the relevant area, detect suspected areas, and so forth. Recently, Many diagnostic methods using computers, known as CAD (computer-aided design), have been suggested for detection of breast lesions and consequently diagnosis of the cancer, and each method has its own advantages and disadvantages. This study was conducted to follow the changes in breast tissue. Homogeneity, similarity, and presence of repeated patterns in frequencies of mammographic images, taken in different years from patients who are eventually diagnosed with breast cancer, were compared to those in a normal tissue using image processing techniques, such as Fast Fourier Transform and Autocorrelation mathematical concept. Afterwards, we used cellular automata to classify results to normal and cancerous tissues. Finally, considerable changes were observed in autocorrelation of frequencies of mammographic images of cancer patients. Whereas, autocorrelation of frequencies of mammographic images of normal people were rather identical. In this respect, changes in frequencies of mammographic images can be followed to predict the incidence of cancer at early stages. The data of the study included mammographic images taken with minimum one year intervals from 14 patients who are eventually diagnosed with breast cancer. The images were adopted from the Center for Cancer Research, and the accuracy of results was obtained as 98%.

Keywords: Early detection, Breast Cancer, Fourier Transform, Autocorrelation, Cellular Automata

1. INTRODUCTION
Although important progresses have occurred in early diagnosis and treatment of breast cancer and reduction of mortality in women in the last two decades, breast cancer is still one of the most frequent cancers among Iranian women. This type of cancer occurs earlier than other cancers and other important causes of death, such as the cardiovascular diseases and is consequently the major cause of women’s life lost and the main threat to their health [1]. As the main cause of breast cancer is unknown, it is impossible to prevent it. The early diagnosis is the most appropriate way to treat it [2, 3].

One of the most reliable methods for early diagnosis of breast cancer is mammography [2, 3]. A 1987 study showed that Five-year survival rate has been 82% for women whose tumor is diagnosed early through mammograms and 60% in other women for a control group. However, this technique may have 9% false negative mainly because of poor quality of mammograms. Also, Mammography is generally difficult for young women whose breast tend to be dense, firm and cluttered with natural fibrous masses and spotting tumors. Recently, the American Cancer Society has recommended that all women between the ages of 40 and 49 should get their mammograms every 2 years and every year after ages of fifty. If this recommendation is followed, each radiologist must consider about 75 mammograms per day that may have only a few revealing abnormalities. However, when there are a lot of mammographic images, radiologists face with difficulty in reaching correct accuracy and estimations. Previous studies have shown that 10-30% of breast lesions are not visible in mammography due to human visual limitations [4].
Having used the progresses in digital image processing, pattern recognition and artificial intelligence, radiologists could improve their diagnosis through computer systems. On average, their sensitivity has increased by 10% with the assistance of computer-aided detection (CADe) and computer-aided diagnosis (CADx) systems [3]. Recent studies have presented many methods of CADe and CADx with different strategies and techniques for diagnosis of lesions and microcalcifications [5-22]. The most widely used methods are introduced in related works section.

The method proposed in this study was a CAD system examining mammographic images in order to detect a tissue that would be eventually infected with cancer against a normal tissue through following changes in autocorrelation of frequencies. Given that the frequency of the area where a lesion begins to develop changes, the fast Fourier transform (FFT) of image segments was computed. FFT is a very useful and practical function in mathematics and used for calculation of digital image frequencies. In case of observing any changes in the values of autocorrelation of FFT matrices of mammographic images taken at different times, it can be concluded that a lesion is getting developed. For implementation of this algorithm, we use image processing techniques consist of a few typical steps, namely: Segmentation, Feature extraction and selection and classification. Cellular Automata (CA) is used for classification.

This paper begins with a description of several related works. Then we present the methods and materials which we used to develop our research, describing the stages from segmentation to classification. Finally, the experimentation phase is described and the assessment of performance of the classifiers is presented in the results.

2. Related Works

An important first step CAD system is the segmentation of the image to diagnose breast lesions. Dominguez and Nandi [23] specified regions by binarized images to segment at multiple threshold levels. Their method’s sensitivity obtained 80% and 2.3 false-positives per image. Singh and Bovis [24] presented a method based on the concept of using Gaussian mixture models, for estimating the probability of each pixel in an image as being part of a lesion.

After the segmentation step found suspicious regions of interest (ROI) containing lesions, feature extraction step is performed. Gray Level Co-occurrence Matrix (GLCM) is one of the most used techniques in feature extraction. In [25] authors used five Co-occurrence matrices were acquired from four spatial orientations to diagnose masses in mammograms, with \( \theta \in (0, \pi/4, \pi/2, 3\pi/4) \) and pixel distance \( (d=1) \). Since, shape and margins of cancerous tumors are not uniform, it is very difficult to extract features. So, the range of pixel distances and the number of spatial orientations should be increased. Yuan et al. [26] proposed three categories of features in their study. The first category comprised of features characterizing speculation, margin, shape and contrast of the lesion. The second category consisted tissue features and the third category included a distance feature computed as a Euclidean distance from the nipple to the center of the lesion. They presented a linear stepwise feature selection method for the classification with a Wilks Lambda criterion to select a subset features.

The next step after feature extraction is selecting subset of features that leads to the highest discriminant power to lesion classification. Sometimes, using more features may decrease the performances of the algorithm and increase the complexity. Therefore, an optimized subset of features from a large number of available features must be extracted. Li et al. [27] used general criterion for feature extraction and selection of significant features: independence, reliability, discrimination and optimality. They divided features into three categories: geometric features, texture features and intensity features. Cheng et al. [28] presented some detailed list of features in each category. They proposed two major methods for feature selection: genetic algorithm (GA) and stepwise feature selection. By GA, population of solutions were created through the chromosomes and obtained the solutions by applying genetic operators such as mutation and crossover to find best solutions considering to the predefined fitness function. Zyout et al. [6] used a method called PSO-kNN through Particle Swarm Optimization (PSO) that extract parameters in a heuristic way by using a k-nearest neighbor (kNN) Method for classification of lesions.

The final step is classification. Some of the most popular techniques are artificial neural networks (ANN), linear discriminant analysis and support vector machines (SVM). Lopez et al. [21] used two different Artificial Neural Networks model: generalized Regression and Feedforward Backpropagation and were tested on 100 images from MiniMammographic Image Analysis Society (MiniMIAS) database [29] and acquired a precision of 94% and 80% of true positives, respectively. Campos et al. [30] used multilayer neural networks and independent component analysis. The best performance was earned with probabilistic neural networks. The method provides a success rate of 97.3%, 96% of sensitivity and 100% of specificity. Nandi et al. [31] presented genetic programming and acquired accuracies above 99.5% for training and
above 98% for testing. In [7] authors compare the results of artificial neural networks, linear discriminant analysis and support vector machines to classify suspicious regions. The results show that the SVM presented the highest overall accuracy and specificity microcalcification clusters, while the NN was more successful for mass-classification in the same parameters. A study on the performance of several classifier is presented in [8], including ANN, SVM, Bayesian and kNN techniques. Another extensive survey of related literature can be found in [9].

3. Method and Materials

The proposed method consists of several stages for automatic differentiation of normal tissues from those that will eventually turn cancerous. In this respect, the selected patients included patients who were eventually inflicted with cancer, underwent a surgical operation, and had at least two mammographic images taken before their operation in different years. Once the data were collected, the method was performed through segmentation of images and feature selection and extraction and finally classification.

After collecting the data, the suspected area was selected in segmentation step and divided into small areas for simplicity in computing. Then, in Feature extraction and selection step, in order to acquire the value of tissue uniformity, Fast Fourier Transform (FFT) function is applied on each area and FFT matrixes is computed for each area. The result of this function is a matrix containing important information and determining the values of frequencies and consequently the uniformity of the tissue. Next, the value of Autocorrelation for each FFT matrix was calculated. The obtained number, determine the value of tissue uniformity. The less number is, the more uniformity tissue has. Then, this process is applied on mammogram images of normal tissue. In classification step, all values were acquired from images of tissues with cancer eventually and normal tissue, classified by Cellular Automata (CA). Comparison of the results shows that frequencies in images of tissues with cancer eventually are less similar to one another and have many changes. Whereas, frequencies in images of normal tissues are more similar to one another and have few changes. As a result, if Autocorrelation value of some images from the people whose tissues are eventually diagnosed with cancer, was closed to Autocorrelation values of normal tissues images. It means that person had normal tissue on the time that mammography image was taken, and if after some times the Autocorrelation value from an image had significant difference with it’s before image, means breast tissue was in the stage of cancer. Consequently, when we receive some images from suspicious person, the images will be given to the algorithm. If the obtained value from the algorithm was in the range of normal tissue, means that person have not any cancer. But, if the obtained value was in the range of cancerous tissue it means that person has cancer or the tissue will be cancerous.

3.1. Data set

The first step was making an image data base. To this end, the researcher contacted 231 patients with cancer who underwent surgery from June 2011 to February 2014. The medical records of those patients were available in the Center for Cancer Research of Shahid Beheshti University of Medical Sciences. Among the patients, 14 had at least two mammographic images taken before their operation in different years with at least one year interval. The mammographic images of these patients were collected, scanned and resized to a specific size and resolution of 2656*4000. Eleven patients had tumors and three patients had microcalcifications. Totally, 47 images were collected from 14 patients, as one patient had 7 images, one patient had 5 images, two patients had 4 images, seven patients had 3 images, and three patients had 2 images. The mammographic images had been taken at CC and MLO views. In this study, only images taken at CC view were examined for simplicity. Also, six mammographic images of normal tissues were collected from internet databases and resized to a specific size and 2656*4000 resolution.

3.2. Segmentation, Feature Extraction and Selection

As the volume of mammographic image processing computations is very high, and there is no need to include the image background with labels, marks or black points, the studied area should be separated from the image using segmentation methods. Therefore, the background has to be removed. For this reason, all images were resized to a specific size and resolution of 2656*4000. Then, the image width is divided into five areas and the image length is divided into four areas. The Region of Interest (ROI) is depicted in figure 1(a). Because, the first and fifth areas of the image width and the fourth area of the image length did not contain any significant information, they were excluded from computation. Remained segments are shown in Figure 1(b). Although excluded segments might develop tumors in future, the major part of the tissue that contained important information was selected because this study was conducted to examine changes in tissue, and the area that might develop tumors in future was not known at baseline.
Because, the algorithm has long-running time, the selected image is divided into nine areas, three areas in the width and three areas in the length. These areas numbered from left to right and top to bottom. It is described in the figure 1(b)

As mentioned before, the objective of this study was to predict breast cancer at early stages when physicians could not diagnose the lesions in mammographic images and assumed that the tissue was normal. In this regard, the most suitable criterion for determining the heterogeneous or non-heterogeneous tissues was the computation of image frequencies and detection of their similarity or differences. Then, autocorrelation function is calculated from matrix of frequencies to achieve the similarity or differences between frequencies. If the number of autocorrelation was great, it means that the frequencies of images are differ largely from each other and not similar, as a result lesion will begin to develop. However, if the number of autocorrelation was small, it means that the frequencies of images are similar to each other and their values are close, as a result the tissue is normal.

Autocorrelation is the equality and similarity between observations as a function of the time lag between a signal with itself. It means that how much similarity is between a signal in a specific time and the same signal at different times in the few steps in back or farther. It is a mathematical tool for finding similarity or repeating patterns, such as the presence of a periodic signal vague by noise. In fact, it is the cross-correlation of a signal with itself. It is often used in signal processing for analyzing functions or series of values.

\[
R(s, t) = \frac{E[(X_t - \mu_t)(X_s - \mu_s)]}{\sigma_t \sigma_s}
\]

Different fields of study describe autocorrelation differently, and not all of these definitions are the same. In some fields, the term “autocovariance” is used interchangeably that is given by:

\[
C_{xx}(s, t) = E[(X_t - \mu_t)(X_s - \mu_s)] = E[X_tX_s] - \mu_t\mu_s
\]

The autocovariance is the covariance of the variable compared with a time-shifted version of itself [32].

3.3. Classification

Since the cellular automata, despite their simplicity and high accuracy, are powerful tools for computational to optimize and analyze of complex systems in artificial environments and they are used widely in mathematical modeling of biological systems. Therefore, cellular automata are used widely in data modeling and classification.

3.3.1. Cellular Automata

Cellular automata are discrete, dynamical system specified by local rules, composed of very simple, uniformly interconnected cells. Each cell takes its new state considering its old state and neighbors’ state through the rules, at each stage of time. One cell is called neighbor of another cell, if it could effect on another cell in one step based on the rule. The cells could have only one state from a set of finite and discrete states and initial values of cells is based on initial conditions that are defined only by local rules. Also, other properties of CA are definitive rules and uniformity cells. It means that all of cells have same states and changing between states is followed from a set of same rules.

The main difference between cellular automata and other common automata is simplicity, local relation between cells, being same connections pattern for all cells, update synchronized cells by same rule simultaneously and produce interesting and complex behavior.

In Cellular Automata, cells are connected according to a range of feature values and create an instance space based on CA’s function. Initial core is formed by training data and finally CA will be convergence. The state of each cell displays a class is assigned to that cell. The goal of classification is assigning the cells into the areas regarding to same classes.

To classify cells into their corresponding classes, voting rule is used that locally reduces entropy. A voting rule looks at a cell’s neighbors and sets the cell regarding the maximum number of
neighbors that are set to a given class, also voting rules are insensitive to the location of these neighbors [33]. In this study, von Neumann neighborhood is used because it is linear in the number of dimensions of the instance space and it is simpler. Also, a rule called N4. V1_nonstable is used in which a cell change its class, if the class of majority of its four neighbors is changed. In this paper, it is supposed that the cells have to be assigned into two classes and CA has two dimensions. Specifically, the rule is defined as: N4_v1_nonstable=

\[
\begin{align*}
0 & : \text{class 1 neighbors + class 2 neighbors} = 0 \\
1 & : \text{class 1 neighbors > class 2 neighbors} \\
2 & : \text{class 1 neighbors < class 2 neighbors} \\
\text{rand}([1,2]) & : \text{class 1 neighbors = class 2 neighbors}
\end{align*}
\]

The term \text{rand}([1,2]) selects randomly from the components with equal probability [33]. It has to mention that at first all cells have zero value and the cells only were assigned to training data have 1 or 2 value that means they are assigned into class 1 or class 2.

3.3.2. Method of Feature Extraction and Classification

After segmentation step, the segmented image is converted to black and white image for simply computation by the use of MATLAB function: rgb2gray. Then, as it is mentioned before, the image is divided to nine areas. Afterwards, the following steps are done for each nine areas:

1. Fast Fourier Transform of that area is computed by fft MATLAB function. The result is a 2D matrix with elements of complex numbers in correspondence with that area.
2. Conjugate of Fast Fourier transform matrix is computed by conj MATLAB function. The result is a same size matrix with elements that are complex numbers but they are conjugate of elements in Fast Fourier Transform Matrix.
3. Covariance between Fast Fourier Transform and its conjugate is calculated by cov MATLAB function. The result is a 2×2 matrix with four elements. Because, each matrix has 2 dimensions and covariance is calculated between each dimension and others.
4. The mean of Fast Fourier Transform Matrix is calculated. The result is a number but since it has to sum with covariance matrix in further, a 2×2 matrix is built by repetition of this number.
5. The mean of conjugate of Fast Fourier Transform Matrix is calculated. The result is a number but since it has to sum with covariance matrix in further, a 2×2 matrix is built by repetition of this number.
6. The value of Autocorrelation is computed from this equation:

\[
\text{Autocorrelation} = E[(X_e - \mu_e)(X_s - \mu_s)] + \mu_e\mu_s \tag{3}
\]

In fact, above equation is equation 2. But, it is extracted from another definition of autocorrelation. \(E[(X_e - \mu_e)(X_s - \mu_s)]\) is the covariance of Fast Fourier Transform matrix and its conjugate. It is earned from step 3. \(\mu_e\) is the mean of Fast Fourier Transform matrix and \(\mu_s\) is the mean of conjugate of Fast Fourier Transform matrix. They are earned from steps 4 and 5. Equation 3 is applied on the covariance matrix and the means matrixes. Then, the result is a 2×2 matrix and the mean of this matrix is our target number and a criterion for determining homogeneity or non- homogeneity of the breast tissue. If the number was big, it means that covariance and changes of tissue are very large. Therefore, breast tissue has more non-homogeneity. But, if the number was small, it means that covariance and changes of tissue are little. Therefore, breast tissue has homogeneity.

According to achieved nine features from Feature extraction and selection step that they were the numbers of autocorrelation related to nine areas, nine dimensions cellular automata has to be used. But, unfortunately in this case the classifier calculation complexity increases and showing nine dimensions CA is impossible. Therefore, nine features have to be reduced to two features for possibility of showing results in two dimensions and decreasing calculation complexity.

After reducing features that will be discussed in section 4, the values of features should be shown as the cells of cellular automaton. As it is mentioned before, features is obtained from 53 images that 47 images belonged to the people who are eventually diagnosed with cancer and 6 images belonged to normal tissue they are downloaded from internet. To show feature values and using them in CA, they are converted to integer numbers from 1 to 100. A 100×100 2D matrix was defined for exhibition of 2D cellular automata. Also, feature values could be converted to integer numbers from 1 to 1000 and a 1000×1000 2D matrix was assigned to them. But in this case, computational complexity will be increased whereas it will have little effect on the accuracy of the results. The next steps of algorithm are as follows:

1. At first, all of elements in 100×100 2D matrix that show 2D cellular automata are zero.
2. For every 53 data, the value of first feature is searched in matrix rows and the value of second feature is searched in matrix columns. After reaching the desired cell, if that data belonged to the people who are eventually diagnosed with cancer, number 2 is set on that cell. It means that data belongs to class 2 and if that data belonged to
normal tissue, number 1 is set on that cell it means that data belongs to class 1.

3. Considering to voting rule that is defined in previous section, if the cell has zero value and its more neighbors belonged to class 1, number 1 is set on that cell and if its more neighbors belonged to class 2, number 2 is set on that cell and if the number of its neighbors belonged to class 1 is equal to the number of its neighbors belonged to class 2, number 1 or number 2 is set on that cell, randomly.

4. Step 3 is repeated until there were not any cells with zero value.

4. Experiments and Results

The proposed method is tested by Intel Atom Dual Core CPU with 1.66 GHZ and 1.67 GHZ frequencies, 2GB RAM and Windows Starter SP1 operating system version 2009 and MATLAB R2009a software is used for implementation of proposed algorithm.

After Autocorrelation values were calculated for every nine areas by proposed algorithm in feature extraction and selection section. The result was a 2D matrix with 53 rows and 9 columns. The nine columns contained values of autocorrelation of nine areas. The first 47 rows belonged to the people who are eventually diagnosed with cancer. Also the next 6 rows belonged to normal tissue and were downloaded from the internet databases. The maximum values of nine attributes of normal tissues and the nine attributes of tissues are eventually diagnosed with cancer, were compared in 47 diagrams. Some of these diagrams are shown in figures 2 to 15.

As can be seen, the differences between the values of images belonged to the people who are eventually diagnosed with cancer and the values of images belonged to normal tissue are very obviously.

Figure 2. The Results related to the first person

Figure 3. The Results related to the second person

Figure 4. The Results related to the third person

Figure 5. The Results related to the fourth person
Figure 6. The Results related to the fifth person

Figure 7. The Results related to the sixth person

Figure 8. The Results related to the seventh person

Figure 9. The Results related to the eight person

Figure 10. The Results related to the ninth person

Figure 11. The Results related to the tenth person
It may be questionable in the Figure 9, the values of 4 and 7 areas belonged to person who is eventually diagnosed with cancer are less than the values of 4 and 7 areas belonged to the normal tissues. It could be realized by comparing related mammography images. The image belonged to the eight person who is eventually diagnosed with cancer is shown in figure 16(a) and the image belonged to one of the normal tissues is shown in figure 16(b). As can be seen, 4 and 7 areas in two images are very similar and close to each other.

According to obtained results, it can be realized that the values of 1, 3, 7 and 9 areas are very close in some cases or even overlap each other and make noises. To solve this problem, the values of 1, 3, 7 and 9 areas are eliminated in calculation because of having noise and interfere. Finally, only 2, 4, 5, 6 and 8 areas that have important information are considered in calculations. But, since these five features have to be reduced to two features, two features are computed by
sum the values of row and column areas. It means that the sum of 4, 5 and 6 areas (row areas) is assumed as first feature and the sum of 2, 5 and 8 areas (column areas) is assumed as second feature. Sum operation is used because when the sum of values belonged to the tissue is eventually diagnosed with cancer and the sum of values belonged to the normal tissue are compared, shows a large difference and classification task will be easier.

The result values of these two features are real numbers between 0 and 10. In figure 17 the data with blue color are related to 47 images of people whose tissues are eventually diagnosed with cancer and the data with green color are related to 6 images of normal tissues. The same result is obtained by using cellular automata and proposed algorithm in classification section. The results are shown in figure 18. Gray area is related to data belonged to normal tissues and white area is related to data of images belonged to people whose tissues are eventually diagnosed with cancer. As it mentioned in classification section, for showing these numbers in 2D cellular automata, these real numbers have to be converted to integer numbers. So, they are multiplied by 10 and rounded. Thereby, the numbers between 1 and 100 was obtained and they could be shown in 100×100 2D matrix as data structure of a 2D cellular automaton with the values of zero (class 1) and one (class 2).

In fact, the concept of figure 17 is the same concept of figure 18. The difference is that the first row and the first column of corresponding matrix with cellular automata is in the left bottom corner in figure 17 but they are in the left top corner with similar concept in figure 18.

**Figure 17.** The result of classifying 53 data with different colors

**Figure 18.** The result of classifying 35 data with cellular automata

5. **Discussion**

When a microscopic lesion begins to develop in an area, the frequency of that area begins to change. In other words, considering the changes in frequency, physicians can guess whether a lesion is developing or not. This study examined the frequencies of mammographic images using image processing techniques and their tool in MATLAB software. Then, Autocorrelation equation was computed from frequencies matrix for achieving homogeneity, similarity and uniformity between frequencies. If autocorrelation number is small, similarity and uniformity between frequencies is more and probability of lesion creation is less and if autocorrelation number is large, similarity and uniformity between frequencies is less and probability of lesion creation is more. This means that the normal tissue does not change a lot, and values of frequencies are very close and similar to one another. However, when a lesion begins to develop in an area, the frequency of that area will change. Therefore, the development of tumors and microcalcifications can be detected some years before being observed by physicians.

Further work will include research towards the enhancement of the classification process, by gathering more data. One of the difficulties facing the researcher in this study was the shortage of data. The data were collected with difficulty because few patients had mammographic images from previous years before their operation. Using more data, the proposed method can be improved, higher accuracy
can be achieved, and more features can be even extracted from images when seeking higher accuracy. Moreover, the researcher did not have the data related to patients with lesions other than cancer, such as cysts. The data of such patients can improve the proposed method.

6. Conclusion

This study was conducted to introduce a CAD model predicting breast cancer at its early stages when lesions could not be seen in mammographic images, and physicians could not diagnose any lesions in mammographic images. In this study, the image processing techniques were used in order to diagnose the incidence of breast cancer at microscopic scales some years before it could appear as a tumor or microcalcification.

In this study, the data were collected from The Cancer Research Center of Shahid Beheshti University of Medical Sciences. The data included 47 mammography images of 14 patients who are eventually diagnosed with breast cancer and had two, three, four, five or seven mammographic images before their operation with at least one year interval. These images are scanned and resized to a specific resolution of 2656*4000. Also six mammographic images of a normal tissue are collected from internet databases and resized to a specific resolution of 2656*4000. Region of interest (ROI) was selected in segmentation step. Then, it was divided to nine areas for simplicity in computing. Then, in feature extraction and selection step, frequencies matrix of every nine areas were computed using a function in MATLAB software and the mathematic equation of Autocorrelation was calculated from frequencies matrix for obtaining similarity and uniformity of frequencies. If the autocorrelation number was small, it means that elements of frequencies matrix are very similar to each other and the breast tissue has more homogeneity. It can be concluded that the breast tissue is normal. However, if the autocorrelation number was great, it means that elements of frequencies matrix have less similarity and the tissue has non-homogeneity. It can be concluded that there is a lesion. In other words, when a lesion begins to develop in an area, frequencies of that area begin to change, and the changes can be detected using the method proposed in this study some years before the changes are observed and diagnosed as a tumor or microcalcification by physicians. The number of Autocorrelation for every nine areas in all images was computed and these nine features were reduced to two features for simply computation. In classification step, cellular automata were used. Because, they are very powerful tools in optimization and analysis of complex systems in artificial environments and widely used in mathematical modeling of biological systems and finally accuracy of the proposed method was obtained 98%.

7. References


29. http://peipa.essex.ac.uk/info/mias.html


28th IEEE EMBS Annual International Conference, New York City, USA, August 30-September, 3.
