A New Approach for Potentially Breast Cancer Detection Using Extracted Features and Artificial Neural Networks

Malik Braik[1], Alaa Sheta[2]
[1]Department of Information Technology
Al-Balqa' Applied University
Al-Salt, Jordan
[2]Department of Computers and Systems
Electronics Research Institute (ERI)
Cairo, Egypt
m_fjo@yahoo.com, asheta66@gmail.com

ABSTRACT*: Breast cancer (B-cancer) detection is still complex and challenging problem, and in that case, we propose and evaluate a four-step approach to segmentation and detection of B-cancer disease. Studies show that relying on pure nakedeye observation of experts to detect such diseases can be prohibitively slow and inaccurate in some cases. Providing automatic, fast, and accurate image-processing-and artificial intelligence-based solutions for that task can be of great realistic significance. As a testbed we use a set of mammogram images taken from the Medical Hussein City in Jordan. This study utilizes morphological operations as a segmentation approach with Artificial Neural Network (ANN) as a classifier tool. The presented approach itself scans the whole mammogram and performs filtering, segmentation, features extraction, and detection in a succession mode. The feasibility of the proposed approach was explored on 32 commonly virulent images, and the recognition rate achieved in the detection step is 100%; moreover, the overall accuracy is convinced and satisfied in all cases. Further, the approach is able to give reliable results on distorted medical images, since the approach is subjected to a rectification step. Finally, this study is very effectual in decreasing mortality and increasing the quality of treatment of early onset of B-cancer.*

Keywords: Extracted Features, Artificial Neural Networks, Morphological Operations, Mammogram Images

Received: 21 February 2011, Revised 28 March 2011, Accepted 3 April 2011

1. Introduction

Breast cancer (B-cancer) is the most frequent tumefaction in women and is the leading cause of cancer deaths among women. Most people turn away to think of B-cancer as a woman's disease. But men get B-cancer too. Research has shown that women with a family history of B-cancer have a higher hazard for evolving the disease. That is true whether the family history is on the mother's side or the father's. In spite of the increasing number of cancers being diagnosed recently, the death rate has been reduced obviously due to the advanced screening programs [1]. Premature detection of B-cancer increases the potentiality of survival rate whereas postponed diagnosis considerably encounters the patient to a critical stage and occasionally results in death [2]. Proper screening programs and diagnostic techniques dramatically increase the survival rate of diagnosed women. A number of methods are currently being developed to detect B-cancer in early stages, including: Ultrasound,Magnetic Resonance Imaging (MRI), Thermal Imaging (TI), and Positron Emission Tomography (PET). It is not yet evident if these methods will play an effective role in Bcancer screening [3]. Ultrasound and MRI are the major tools that could be used. Breast ultrasounds are too expensive as a screening option for the general public. MRI tests are more sensitive, and can be too expensive as a general

screening test. Ultrasound uses sound waves to make images of tissue,and it is commonly used as a follow-up test to mammography and clinical breast exam. On the one hand, ultrasound is not used alone for B-cancer screening, since it does not produce an accurate image of the entire breast. On the other hand, MRI uses nuclear resonance to examine the breast, and there is promising evidence that it could ultimately become a good tool for B-cancer screening [4], [5]. Mammography is a technique that uses X-rays to provide an image of the breast. These images, called mammograms, are used to find potential signs of B-cancer: like tumors and abnormal changes in the skin. Contemporarily, screening mammography of the breast are the most effective tools for premature detection of B-cancer. Hence it is considered the most dependable technique for premature detection, and it is not the optimal method of screening for B-cancer. Moreover, mammography based screening programs are carried out in Jordan and their effectiveness has had a great impact on diagnosis. Overall, mammography can find cancers at an early stage, when they are small and most responsive to treatment [4], [6]. As a result, the issue of adverse consequences of screening for women who do not have B-cancer, as well as women who have an early stage of B-cancer that will not in progress, has become one of the core issues in recent debates about mammography [3]. Still, studies have proved that all Bcancers that are retrospectively detected on the mammograms are not exactly detected by radiologists [7], [8]. Due to the subtle and complex nature of the radiographic findings concerned with B-cancer, human factors such as distraction by image features and simple oversight can be responsible for the errors in radiological diagnosis [9], [10]. Authors in [11]–[13] were used morphological features to provide automated cancer diagnosis tools using statistical feature set. Another study used Artificial Neural Networks (ANNs) and the k-nearest neighborhood algorithm to distinguish the healthy and cancerous tissues of breast [14]–[16]. Also, some authors have been widely adopted NNs in the field of B-cancer segmentation [17], [18]. Brzakovic [19] used fuzzy pyramid linking to identify homogeneous regions in mammograms, and then used a statistical model to classify regions as non-tumor, benign tumor, or malignant tumor. The digital mammograms have already been classified as normal, cancerous or benign. Figure 1 shows an example of a typical mammogram.

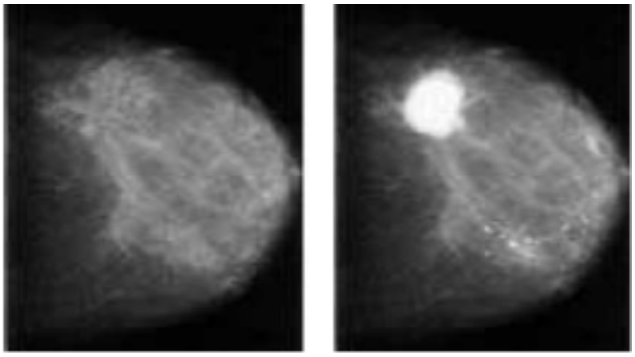

Figure 1. Typical mammogram; Left: Normal breast, Right: Cancerous breast

This paper proposes a new algorithm addressed to perform prescreening of digital mammograms for the presence of virulent regions in breast tumors. Our algorithm partitions a mammogram image into homogeneous regions for possible locations of virulent regions in breast based on some morphological operations. Then, ANN is involved as a diagnostic technique depending upon some textural features composed of statistical gray level and Haralick features of the virulent regions. In conclusion, the aim of this work is threefold: 1) identifying region of interests in breast tumors; 2) analyzing the underlying features extracted from images of breast; 3) distinguishing between ”cancer” and ”non-cancer” images of B-cancer.

The performance achieved with the proposed scheme was 94% sensitivity and 94% specificity on test set; also, results show 100% classification accuracy of both training set and testing set. Furthermore, promising results of both segmentation and detection phases were achieved.

2. Artificial Neural Networks

Artificial neural networks are basically parallel working systems that are similar to biological neural networks; hence they are also called neural networks. These are simply consisting of large number of processing units interconnected in a network which are linked by weighted connections. These units are called neurons. The neuron receives input signals, applies some weights from other processing units, aggregates these signals, feeds the result to an activation function, and generates the output signal [20]. An illustration of neuron is shown in Figure 2 [20].

In Figure 2, the inputs to the processing unit can be external inputs (φ_1, φ_2, φ_3, and φ_4) or outputs of proceeding units. Bias inputs such as w_{i0} can also be used to represent a displacement or a constant input to the unit. The activation function f_i can be any kind of singular valued function and can take any form linear or nonlinear. h_i and y_i represent the hidden and output signals, respectively [20], [21].

Nodes are the basic units in ANN, which are connected to each other by links known as synapses, associated with each synapse there is a weight factor, also called connection strength. Usually, neural networks have user-defined inputs and produce, as nearly as possible, a specific set of target outputs, which reflect the information stored in connections during training [22]. The combination of neurons and connections help the network to exhibit complex behavior, that is, neural networks perform a highly adaptive nonlinear decision function from training examples, and can be used to solve problems that are not amenable to conventional statistical and mathematical methods. In general, a neural network consists of at least three layers: input, hidden, and output layers [23].

Recently, NNs have met with applications in medical imaging [17], [24], and constitute a very large research field, also there is a vast amount of literature on neural networks. The most common types of neural networks are: feedforward networks, feedback networks, and feedforward backpropagation networks.

2.1 Feedforward Networks

Feedforward networks are also called ”static networks”. These networks are the simplest type of ANNs. Here, the information moves in only one direction, forward, from the input nodes, through the hidden nodes, and to the output nodes. There are no cycles in this type of network. There are layers and nodes at each layer. Each node at the input and the inner layers receives input values, processes, and passes it to the next layer. This process is conducted by weights. Weight is the connection strength between two nodes. The numbers of neurons in the input and output layers are determined by the numbers of input and output parameters, respectively [25].

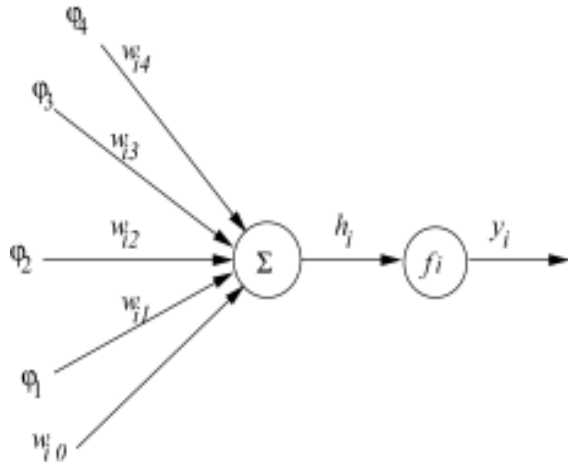

Figure 2. A neuron

2.2 Feedback Networks

Feedback networks are also called ”dynamic networks”. In these types of networks, the information flow is bidirectional (forward and backward). These networks have cycles that feedback information to the hidden layers and from the output layer to the input layer or to the hidden layer. Both types of networks are trained either using supervised learning, in which the network is provided with the desired outputs and then trained to match those outputs, or unsupervised learning, in which the network is trained without providing the desired outputs [26].

2.3 Feedforward Backpropagation Network

The feedforward backpropagation network is a feedforward Multi-Layer Perceptron (MLP) with the backpropagation algorithm used for training. The backpropagation architecture is the most popular, effective, multi-layered network. The typical network has an input layer, at least one hidden layer, and an output layer. Each layer in the network is fully connected to the succeeding layer [27]. The general illustration of the feedforward neural network with four layers is shown in Figure 3.

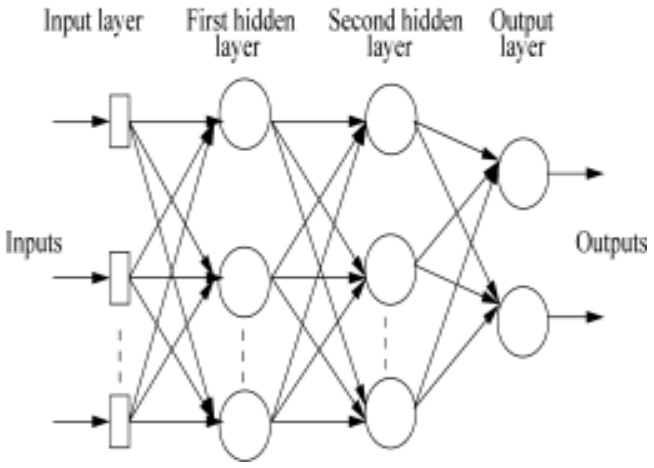

Figure 3. Feedforward neural network

The backpropagation training algorithm network involves two steps: the forward and backward steps. In the forward step, the inputs are presented to each neuron in the hidden layer through a weighted matrix multiplier. Next, the summation of inputs in the hidden layer is performed and the bias is added. The output of each neuron in the hidden layer is processed by an activation function at that layer and the result is propagated to all the output layer neurons. Then, all the weights are adjusted according to the calculated error using steepest descent method. In the backward step, the procedure is repeated many times through the hidden layers to the input layer until the sum of the squared error reaches an acceptable level. The neural networks look attentively to have a good generalization capability [28].

To properly train the neural network, the model is fed with a variety of real life examples, called training sets. The data sets normally contain input and output data. The neural network creates connections and learns patterns based on the input and output data sets. The data set is divided into two sets: training and testing sets. Neural networks are trained by using a training data set, and their generalization capacity is examined by using testing sets. The training data must never be used in the test data.

3. Motivation And B-Cancer Challenges

B-cancer is a tremendous public health problem. The utterance of cancer can be psychologically excruciating. Actuality, B-cancer takes years to develop, and eventually the malignant tumors extend from the breast to the other parts of the body if the cancer has not been treated. According to the Jordan Cancer Society (JCS), B-cancer is the most common cancer among women, with lung cancer being the second most common. B-cancer is one of the major reasons of deaths occurring in women well under the age of 60, with lung cancer being the second leading cause of cancer death in women. Pursuant to JCS, B-cancer tops the list of the cancers that Jordan women suffer from. The other cancers women suffer from are Lung Cancer, Colorectal Cancer, Endometrial Cancer, and Non Hodgkin Lymphoma Cancer. Every year, new cases of B-cancer in women are being detected in Jordan. Right now, many women who have B-cancers have been treated for B-cancer. Furthermore, many women in Jordan will be found to have invasive B-cancer last year, and many of them died from the disease this year. The chance of a woman having invasive B-cancer some time during her life is about one in ten, and the chance of dying from B-cancer is about one in twenty. Fortunately, B-cancer death rates are going down. This decline is often attributed to earlier detection and improved treatment tools according to JCS in 2009. Testing for B-cancer can be a sensitive fact. But there still some questions for proper testing, and based on who you ask, the right answers should be. When should we test? Who should we test? How should we test? What tests should we use?

The important issue in specifying the strength of testing is whether the tests can adequately differentiate between malignant and benign tumors, so that, patients with benign tumors are not subjected to the risks of surgery, radiation or chemotherapy. B-cancer detection and treatment achieve many benefits in estimating B-cancer causes which have led to the increasing interest in periodicity of testing of breast tissues, thus, the Jordan community-based testing should be recommended. Moreover, the

importance of testing for women who have an early stage of B-cancer or who do not have B-cancer, has become one of the core facts concerning B-cancer detection and treatment. Further, the JCS suggests that establishing the testing of breast tissues is an important factor of early detection and can be only indirectly treated. There are many tests that could be used to examine the breast tissue for B-cancer:

1) Self Breast Exams (SBEs) are inexpensive and can be done anytime, but there are varying debates on the effectiveness of these exams [29].

2) Clinical Breast Exams (CBEs) are generally done,but it could be hard to catch early stage B-cancer depending on these exams [29].

Thus, a more sensitive test should be needed. Therefore,B-cancer is one of the greatest challenges of the Ministry of Health in Jordan. Hence, it is an extremely important to catch B-cancer at an early stage. Here, are some of the main indications of occurrence of B-cancer:

- Change in the size or shape of the breast.
- Dimpling of the breast skin.
- The nipple becoming inverted.
- Swelling in the armpit.

Jordan B-Cancer Program (JBCP) was developed in 2007 as new comprehensive initiative patriotism, and has many activities regarding B-cancer prevention and protection. The aims of JBCP are:

- Increasing the rates of cure.
- Increasing the rates of survival.
- Increasing the quality of treatment through increasing the services of early diagnostic.
- Increasing the educational healthy level.

One of the main activities released by JBCP is the enlightenment movement ”every year and you are right”. TheCurrent Enlightenment Group (CEG) plays an important role in reducing the rates of cancer infection. JBCP has many profitable acts to Jordan society, according to the JCS, almost 600 - 650 women were diagnosed with B-cancer every year. On a positive response, the survival rate for the women suffering from B-cancer has increased because of development of improved detection techniques and treatment methods. Furthermore, the risk of developing B-cancer can be reduced by: having children before 30, breast feeding, maintaining a healthy weight, and exercising regularly. Finally, The JCS estimates that, each year, about 650 women are diagnosed with B-cancer, and this number could be increased if no peremptory remedy has been considered. The average old of diagnosis of B-cancer is 51 years, whereas half B-cancer incidences infect the women with an age between 40 and 59 years. B-cancer is the leading cause of death in women under the age of 40. Also, in Jordan, B-cancer accounts for 65% of diagnosis at later stages, and the chance of cure reaches 95% if the cancer could be discovered early. Hence, doctors recommend that women begin regular X-ray mammography screening at the age of 40. Thus, this study concluded that regular B-cancer screenings and testing are advisable despite the odds of misdiagnosis.

4. Proposed Detection Algorithm

Detection abnormalities in mammograms is a difficult task, especially due to the high number of normal cases. Also, a closer inspection of mammograms reveals several difficulties for possible cancer detection, for example: the variations in the recording procedure as well as the global appearance may be both inefficient and difficult in detecting abnormalities in mammograms. A common approach for detecting abnormalities is to use a series of heuristics approaches, like: filtering and thresholding. These methods may include texture analysis which may automatically detect abnormalities [30]. However, those methods possibly suffer from a lack of robustness when the number of images to be classified is large [31]. More that, these approaches act only as a second reader and the final decision is made by the radiologist. By using Computer Aided Detection (CAD) software the number of errors might decrease, both false negatives (malignant cases that were not recalled) and false positives (cases that are recalled unnecessarily). Therefore, our work can contingently provide clues to treat B-cancerin its early stages. Precisely, interpreting mammograms that are used for diagnosing process involves preprocessing and post-processing stages. In this work preprocessing stage deals with image enhancement and image segmentation sub-stages. The post-processing stage consists of another two sub-stages; features extraction and neural network classifier.

An overview of our system is presented in Figure 4. Our system calculates seven textural features from breast tissue images, including statistical gray level features and Haralick features [32].

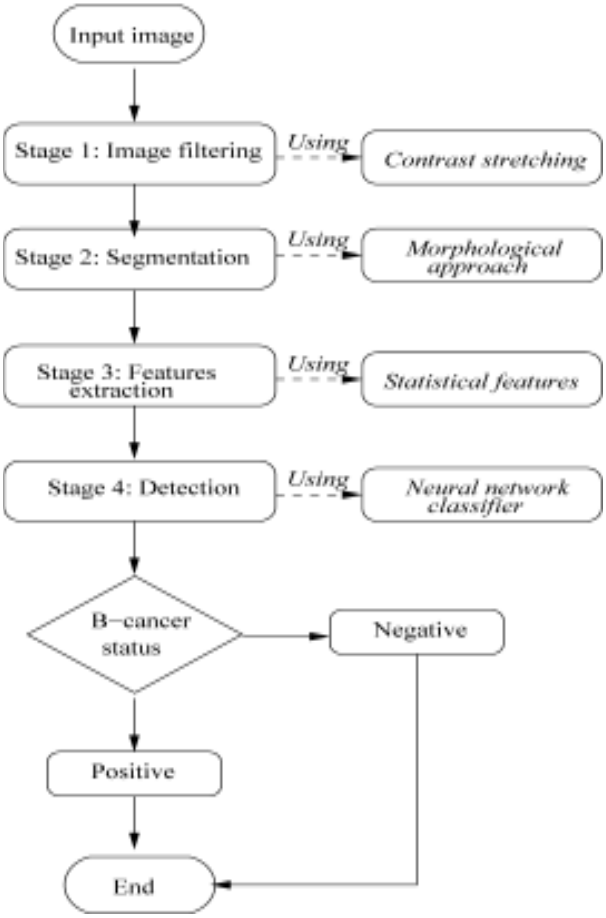

Figure 4. Proposed B-cancer detection system

Figure 4 presents the main steps of the implemented work in this paper. The first step is the image filtering step which is very essential for premature detection of cancers. The first step in turns should be scanned for extraction the region of interests using the second step. The second step is then subjected to the third step in which some texture and statistical features are extracted. Finally, the extracted features are passed through pre-trained neural networks for possible B-cancer detection. We illustrate the general outline and details of the presented work in the following sections.

5. Image Filtering

Medical images suffer from a wide variety of distortions which may cause some difficulties in identifying the abnormalities in breast tissues. As a result, perceptual quality of the images is deteriorated. High quality images and mammographic interpretation are mandatory for the detection of premature and delicate symptoms of B-cancer. Therefore, image filtering is inquired to suppress unwanted noise and both improve the quality and visual appearance of the images.

The essential improvement that is required in mammography is accomplished by contrast stretching method, which particularly beneficent for mammograms with low contrast.

5.1 Contrast Stretching

Contrast stretching is one of many techniques that used to enhance mammograms. It increases the contrast of the image, and adjusts the histogram of the image such that, there is a greater separation between foreground and background gray level distributions [33]. The following designing function f can be used to rescale the output grayscale values.

$$f = \begin{cases} \alpha x & 0 \le x < a \\ \beta(x - a) & a \le x < b \\ \varphi(x - b) & b \le x < L \end{cases} \quad (1)$$

where x is the input image gray level, α, β, and φ are chosen greater than unity in the region of stretch. The parameters a and b are obtained by examining the histogram of the original mammogram image. L is the maximum gray level of the original mammogram image. f shows a typical contrast stretching transformation of the gray-level distribution in the mammogram image. Thus, an is processed with contrast stretching method to obtain an enhanced image. The contrast filter is a spatial filter that has been developed to approximately match contrast variations of typical mammogram. Then, it provides an output with more contrast and detail. Consequently, at the locations of microcalcifications, the pixel values in the contrasted image are increased relative to the values of pixel of normal tissue [34]. Actually, contrast filter improves the readability of the low contrast areas in the image, and destroys the areas in the image where the intensity of the pixels is outside the range of intensities that are being enhanced [35], [36].

6. Image Segmentation

We need a significant technique for image segmentation that attempts to recognize and extract the malignant tumors with the aid of the distribution of gray levels in the image. So, one extremely task in discrimination the malignant tumors from benign tumors depends originally on choosing a proper segmentation process. Malignant tumors can have speculated and/or fuzzy borders, while benign tumors can be lucent at the center and can have well defined borders. However, this step is developed because it can help radiologists to identify the signs of cancer, and mark the suspicious areas in the mammograms, which may indicate the presence of cancer. There is a possibility for the difference in malignant tissues and normal tissues to exist in the mammograms but beyond the threshold of human perception. Therefore, we think deeply not to use thresholds in our proposed segmentation algorithm. The developed segmentation algorithm is based primarily on morphology theorem. Morphology algorithm is frequently employed in various areas to segment digital mammogram and is utilized in the construction of the regions. Once the tumor objects are extracted, these can be categorized as benign or malignant. Briefly, this step provides the following goals:

1) obtains the true locations of suspicious areas which may assist radiologists during the diagnosis.

2) classifies the abnormalities of the mammograms into benign or malignant.

3) spots the salient regions in mammograms, these salient regions correspond to distinctive areas which may include the breast boundary, the pectoral muscle, candidate masses, and some other dense tissue regions.

6.1 Morphological Operations

Morphological operations are a way of extracting image elements like regions shape, image boundaries and so on. Morphological theory is a powerful tool to analyze and describe numerous image processing problems [33]. Dilation and erosion are the two basic morphological operations. *Dilation* is usually used to smooth boundaries of the regions or bridge very small gaps between neighboring regions. *Erosion* has an opposite effect of dilation, as it shrinks the objects uniformly.

With A and S as sets in Z^2, the *dilation* of A by S is defined as given in Equation 2 [33].

$$A \ominus S = \{z \mid (\hat{S})_z \cap A \neq \Phi\} \quad (3)$$

With A and S as sets in Z^2, the *dilation* of A by S is defined as given in Equation 3 [33].

$$A \ominus S = \{z \mid (S)_z \subseteq A\} \quad (2)$$

In this work, A is the image and S is a suitable structuring element. Other morphological operations like closing and opening were also introduced to assist removing some tiny objects from the image. Opening tends to eliminate some of the foreground (i.e. bright pixels) from the edges of object region. Opening is erosion followed by a dilation using the same structuring element for both operations. Closing enlarges the boundaries of foreground regions in an image. Closing is a dilation followed by an erosion using the same structuring element for both operations [33].

The opening of set A by structuring element S, denoted by $A \quad S$ is defined as given in Equation 4 [33], [37].

$$A \circ S = (A \ominus S) \oplus S \quad (4)$$

The closing of set *A* by structuring element *S*, denoted by $A \bullet S$ is defined as given in Equation 5 [33].

$$A \bullet S = (A \oplus S) \ominus S \quad (5)$$

(*S*) is a matrix that defines a neighborhood shape and size for morphological operations consisting of only 0's and 1's. The pixels with values of 1 define the neighborhood. Experimentally, in this work, (*S*) has a disk shape with radius equal to 5.

Morphological operations are used as a post-processing stage after the filtering stage to segment the mammogram image. In this work, the proposed methodology of segmenting Bcancer based on morphological operations can be concisely summarized into a group of steps as follows:

1) Step 1: The Region Of Interests (ROIs) is selected for the purpose of segmentation. ROIs is then identified and labeled. ROIs could be tumor or any other relevant matter.

2) Step 2: After specifying the ROIs, all small undesirable objects are reduced or completely removed.

3) Step 3:After eliminating all tiny objects, the image is smoothed using multidimensional filter.

4) Step 4: Dilation and erosion are applied using flat linear structuring element in each specified neighborhood for the selected ROIs in Step 3.

5) Step 5: Suppressing light structures that are connected to image border to reduce the overall intensity of the ROIs.

6) Step 6: Reconstructing the final segmented image from the previous two steps.

7. Feature Extraction

From each segmented image, I, we extract a set of statistical- and texture-based features to capture the discriminating characteristics of the tissue patterns. A feature vector (*f*) is created for I where each element of (*f*) is a distinct feature value. The feature vectors in our study were extracted from the selected regions of the breast. These values are calculated as described below.

7.1 Textural Features

The proliferation of nuclei in cancerous tissue suggests that textural characteristics can help discriminate between different grades of B-cancers [38]. The following features are extracted from the 2D grayscale mammogram image:

7.1.1 Grey Level Features: We calculate two gray level features from I as described in [39]. These features are:

• **Average (mean)**: mean (*m*) measures the average intensity. Mean is defined as given in Equation 6 [40].

$$m = \frac{1}{n} \sum_{i=1}^{n} x_i \quad (6)$$

Where: x_i is the value of the pixel *i*, and *n* is the total number of pixels.

• **Standard Deviation**: standard deviation (σ) measures the pixel deviation from the mean. (σ) is defined as given in Equation 7 [40].

$$\sigma = \left[\frac{1}{n-1} \sum_{i=1}^{n} (x_i - m)^2 \right]^{\frac{1}{2}} \quad (7)$$

7.1.2 Haralick Features*:* The second-order co-occurrence texture features are described by sixteen Haralick features presented in [32]. In this study we calculate a co-occurrence matrix $Z \in R$ for image *I*, then, we used the matrix to generate five Haralick texture features. Concisely, the method followed for extracting the feature set is called the Color Co-occurrence Method or CCM method in short. CCM is a method, in which both the color and texture of an image are taken into account, to arrive at unique features, which represent that image.

7.2 Co-occurrence Methodology for Texture Analysis

The image analysis technique selected for this study was the CCM method. The CCM methodology was applied to each pixel in the segmented image, in which each pixel map is used to generate a color co-occurrence matrix, resulting in one CCM matrix. The color co-occurrence texture analysis method was developed through the use of Spatial Graylevel Dependence Matrices or

in short SGDM's [32]. The gray level co-occurrence methodology is a statistical way to describe shape by statistically sampling the way certain greylevels occur in relation to other grey-levels. These matrices measure the probability that a pixel at one particular gray level will occur at a distinct distance and orientation from any pixel given that pixel has a second particular gray level. For a position operator *p*, we can define a matrix P_{ij} that counts the number of times a pixel with grey-level *i* occurs at position *p* from a pixel with grey-level *j*. The SGDMs are represented by the function *P(i, j, d,θ)* where i represents the gray level of the location *(x, y)* in the image *I(x, y)*, and *j* represents the gray level of the pixel at a distance *d* from location *(x, y)* at an orientation angle of . All the neighbors from 1 to 8 are numbered in a clockwise direction. Neighbors 1 and 5 are located on the same plane at a distance of 1 and an orientation of 0 degrees. In this research, a one pixel offset distance and a zero degree orientation angle was used. We used GLCM function provided by Mathworks to create gray-level co-occurrence matrix for the images produced from the segmentation phase. The images are in a gray-level representation, with the number of gray levels is set to 8, the symmetric value is set to "true", and the offset value is "0". Consequently, the statistics of the two-dimensional (2D) co-occurrence matrix of an image form the basis of our proposed technique for detection.

7.3 Texture Features Identification

Haralick [32] proposed sixteen features extraction approach from an image focusing on the spatial distribution of gray values. We prefer five features to complete this work, Those features are:

• **Contrast**: contrast (*C*) measures the local variations in the gray level of the image. Contrast is defined as given in Equation 8 [40], [41].

$$C = \sum_{i=1}^{n} \sum_{j=1}^{n} |i-j|^2 \times P(i,j) \tag{8}$$

Where: $P(i, j)$ is the pixel at (i, j) coordinates.

• **Energy**: energy (*E*) measures the image uniformity. Uniformity measures the maximum gray level when they are equal. *E* is defined as given in Equation 9 [40], [41].

$$E = \sum_{i=1}^{n} \sum_{j=1}^{n} P(i,j)^2 \tag{9}$$

• **Dissimilarity**: dissimilarity (*D*) measures the images local variation along a certain orientation and displacement. *D* is defined as given in Equation 10 [40], [41].

$$D = P(i,j) * |i-j| \tag{10}$$

• **Homogeneity**: homogeneity (*H*) measures the relative's smoothness of the intensity in the region. *H* is defined as given in Equation 11 [40], [41].

$$H = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{P(i,j)}{1+|i-j|} \tag{11}$$

• **Entropy**: the entropy (*e*) measures the randomness in the image. *e* is defined as given in Equation 12 [40],[41].

$$e = -\sum_{i} \sum_{j} P(i,j) \log_2 P(i,j) \tag{12}$$

Finally, all the seven extracted features of the segmented objects are passed to the final step of proposed B-cancer detection algorithm.

8. Cancer Detection Using ANN

The detection step has been accomplished based on backpropagation ANN, really, there are many classifiers that can be used in image processing applications, and therefore, careful consideration was given to this particular stage. However, a neural network is chosen in this paper due to its well known technique as a successful classifier for many medical applications [14]. The inputs of the ANN are features vectors consisting of the seven extracted features. Those extracted features fall into three categories that are related to the statistical approach in which they are depending on the pixel gray values, neighborhood of the pixel, and pixel location. The features vectors serve as a classification method. The specified ROIs objects are split into two parts; training and testing sets. Then, once the feature extraction step was completed, two files were obtained. They were: (i)

Training texture feature data, and (ii) Test texture feature data. The training set are used to train the NN model, whilst the testing set are used to verify the accuracy of the trained NN model.

Consequently, The training set file consists of 50% of the total ROIs segmented images with the testing file consists of the remaining 50%. Thus, the extracted features of 16 images were used as the training set, and the extracted features of 16 images were used as the testing set. Furthermore, each dataset has two classified images, where examples of the two classes, i.e. cancer and no cancer, must be provided to train the classification model.

In short the files had 17 columns and 8 rows each; the former 16 columns represent 16 samples of B-cancer, and the latter column represents a benign sample. Each of the 16 columns had 7 values representing the 7 texture features extracted for a particular segmented image. The last column represents the texture features of the benign sample. The first row represents the class of the particular column of data. Thus, the first 16 values are "1" which indicate the malignant samples, and the last value which is '0" represents a nonmalignant sample. Finally, ANN is trained several times for the possible detection of B-cancer cells, after that, ANN can classify the specified ROIs if it has B-cancer or not, and hence the B-cancer object is reconstructed. After that, we formulated a Matlab program aiming to specify if the input image belonging to class 1 or class 0, i.e. contains B-cancer or not.

9. Experimental Results

The experimental results of the implemented detection approach are displayed in the next sub-sections in a succession mode. We proffer the results using some malignant samples.

9.1 Filtering Using contrast Stretching

Filtering is a crucial step for handling the distortions that may exist in the mammograms, such that, the suspicious region could be displayed with higher contrast, which in turn accomplishes two tasks: the former task facilitates the function of the segmentation step for manipulating the breast cells describing the ROIs, the latter task augments the clarity of the object features which in turn paves the way for NN to classify the B-cancer easily. The results of applying contrast stretching to an image containing B-cancer is shown in Figure 5.

It can be seen from Figure 5 that the malignant regions appear almost visibly. In addition, the overall intensity of the processed image is also increased.

9.2 Segmentation Using Morphological Approach

The segmentation step is used to select and identify the ROIs which contains the virulent cells, The procedure given in section VI is applied for the whole set of tested images.Figure 6 shows an example of an original image containing B-caner, and the results of the segmentation procedure.

In most cases, mammograms involve several dark regions and noises, occasionally, those dark regions could not be removed completely in the filtering step, which may trick the segmentation process, besides this, some small objects may be produced as a result of the first step in the segmentation procedure. These small undesirable objects are shown in the upper right image in Figure 6. Step 2 and step 3 in the segmentation approach try to remove all the tiny protuberances appeared on the upper right image as shown in the lower right image in Figure 6. The B-cancer in the lower right image is not fully identified, therefore, the last three steps in the segmentation procedure are inescapable to strengthen the segmentation result as shown in the lower right image in Figure 6. The last three steps in the proposed morphological algorithm eliminate the remnant dark regions and tiny protrusions from the image using the proposed filters, namely: multidimensional filter, erosion filter, and the dilation operator. The dilation operator is applied to bridge the small gaps using the proposed flat structuring element.

It can be seen from Figure 6 that the morphological operations perform very well in extracting the image elements, and it can express the image details in more specific way.

Consequently, the morphological operations produce the best segmentation results of the mammogram images including only the upnormal areas, as a prior task to posterior classification. Hence, the suspicious regions will be accurately segmented. Finally, it can be evaluated that the proposed method is able to segment the mammogram images.

As an additional step, the suspicious object in the input image is identified in the mammogram image by tracing the boundary of the suspicious regions. Through the proposed segmentation method; the binary image contains only 0's that represents the

background and 1's that represents the object itself.

In edge based techniques, segmenting an object is achieved by locating its boundary using image gradient, which has high values at the edges of the objects to locate the boundary of pixels, such that, the edges in the image can occur on the boundary between two pixels when the respective brightness values between two pixels are significantly different [9], [42]. Hence, they are looking for edges between regions with different characteristics such that the edges are placed in the image with strong intensity contrast. The aim is to find object's boundaries and segment regions enclosed by the boundaries. The edge operator used in this work is Sobel edge detector [33]. Tracing the ROIs in the image containing B-cancer with the binary image and the edges of the cancer object are shown in Figure 7.

As a conclusion, the presented segmentation approach gives reliable results; because of that, morphological operations are extensively used in the analysis of medical images.

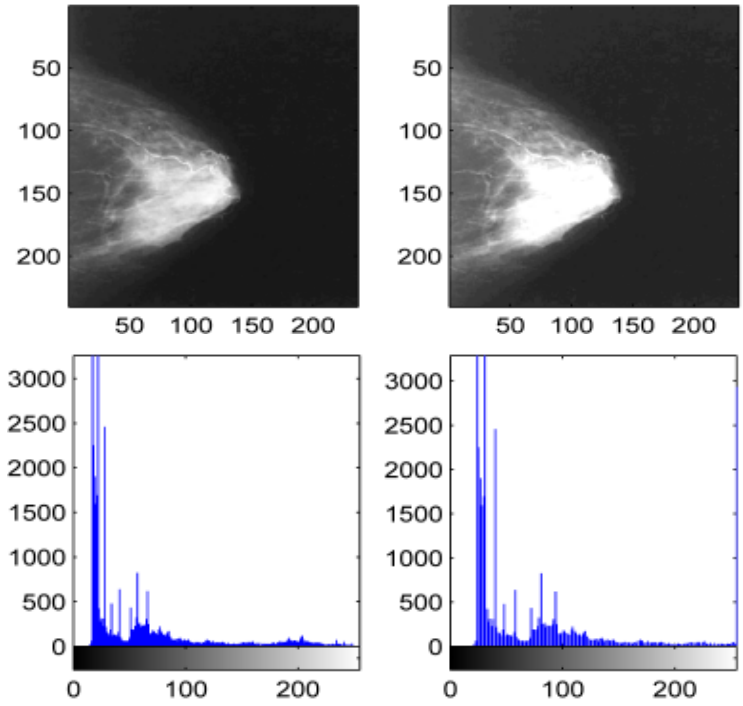

Figure 5. Enhancement results: Upper left: Original image, Upper right:Enhanced image using contrast stretching, Lower left: Histogram of the original image, Lower right: Histogram of the stretched image

9.3 Extracted Features Using Co-occurrence Methodology

We computed some of textural features with the aid of SGDM's using o-occurrence methodology, these features were used as the basis of our detection algorithm. So, they are considered for both train and test files of neural network. The extracted features of the 16 training images are shown in Table 1.

9.4 Detection Using ANN

A software routine was formulated in MATLAB that would take .mat files representing the training and test data; train the classifier using the 'train files', and then use the 'test file' to perform the classification task on the test data. Consequently, the routine would load all the training data files and make modifications to the data according to the proposed model chosen. Anyhow, the detection step is performed based on neural networks, then, the files containing the extracted features are fed into the neural network which serves as classification method. For the purpose of classing, two images had been fed into the NN model, one class is classified as a malignant sample, and the other class is classified as nonmalignant sample. On right positive, ANN is trained several times for the possible detection of B-cancer cells, after that, ANN can identify the specified ROIs if it belongs to class 1 class 0. The architecture of the ANN model commonly includes type of network, number of input and output neurons, transfer function, number of hidden layers as well as number of hidden neurons. Generally, the input neurons and

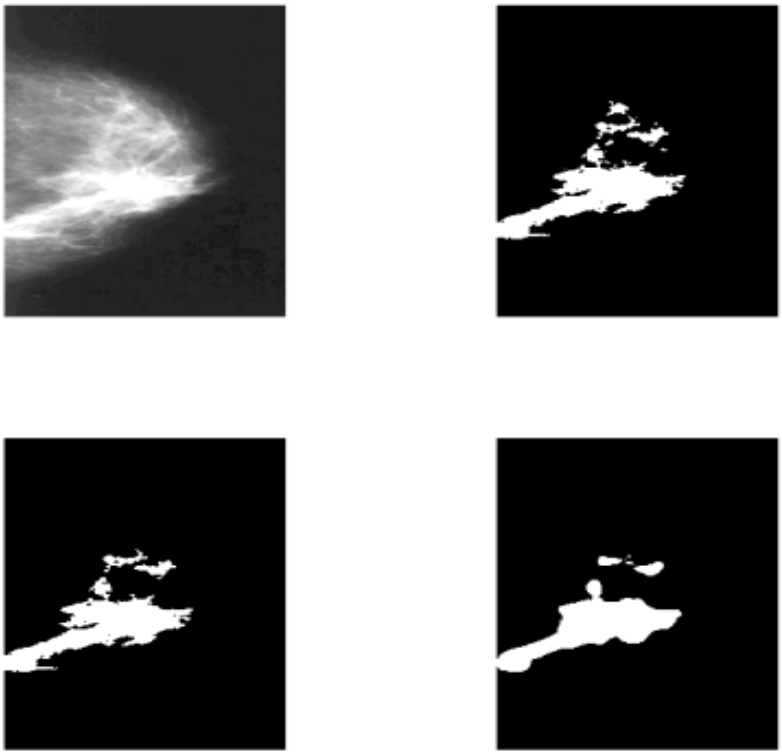

Figure 6. Results of the proposed morphological approach, Upper left: The enhanced image, Upper right: the suspicious ROIs, Lower left: Eliminating all small objects from ROIs, Lower right: the final segmented image

output neurons are problem specific. In this study, ANN is utilized three layer feed-forward network with back-propagation algorithm. Therefore, The architecture of the network used in this study was as follows:

1) Number of hidden layers: 1 with 10 neurons in the hidden layer.

2) Number of inputs to the input layer, 'n' (representing the number of texture features selected) depended on the model used.

3) Number of output layers: 2 (= number of classes)

4) The parameters of the network were as follows:

- Network: Feed forward back propagation
- Performance function: Mean Square Error (MSE)
- Number of iterations: 1000
- Foal error = 10e-5.

The detection method based ANN exhibits a high accuracy of correct categorization. All the trained and tested mammogram images were correctly classified using ANN for possible B-cancer detection. The network was trained 1000 times based on the extracted features of the segmented objects using the feed-forward back-propagation network to minimize the MSE function. The improved results are attained by classifying the disease in the context of its features. Examples of class 1 containing suspicious regions classified by NNs as B-cancer are shown in Figures 8 and 9.

Eventually, it can be seen from Figures 8 and 9 that ANNs has the ability to recognize the mammogram images if the breast is infected with B-cancer or not.

The convergence rate of classifying the B-cancer based ANN is recognized using the performance function. The MSE value of 0.0001 was used to measure the prediction accuracy;hence, the convergence rate achieved a high accuracy with 15 iterations.

Image number	Mean	Standard Deviation	Contrast	Energy	Dissimilarity	Homogeneity	Entropy
1	5.2709	10.2929	0.0753	0.2300	0.0751	0.9624	1.9569
2	5.9365	11.8716	0.0776	0.1950	0.0774	0.9613	1.9858
3	6.7574	15.6565	0.0785	0.1572	0.0783	0.9609	2.1596
4	5.6114	10.2621	0.0601	0.2069	0.0600	0.9670	1.8651
5	5.9365	6.2504	0.0347	0.4875	0.0336	0.9833	1.2266
6	3.7516	5.7628	0.0624	0.4080	0.0597	0.9704	1.4567
7	6.7542	14.4621	0.0767	0.1783	0.0767	0.9617	2.0554
8	5.1980	11.1229	0.0763	0.2410	0.0761	0.9794	2.0001
9	6.0935	12.0016	0.0798	0.1980	0.0784	0.9721	1.9877
10	5.7124	13.0015	0.0801	0.1762	0.0791	0.9801	1.9996
11	4.9204	9.1761	0.0541	0.1269	0.0599	0.8710	1.7611
12	5.8950	10.1204	0.0747	0.2875	0.0436	0.8833	1.3260
13	3.7015	5.6128	0.0612	0.4071	0.0588	0.9604	1.4123
14	6.8742	14.7621	0.0798	0.1800	0.0766	0.9707	2.0654
15	6.6940	15.6666	0.0784	0.1569	0.0783	0.9439	2.1686
16	5.6345	10.5421	0.0698	0.2159	0.0612	0.9800	1.8771

Table 1.Values Of The Extracted Features For The Training Images

The convergence curve of the detection process is shown in Figure 10.

Figure 10 shows that the MSE value is converged to an optimum value, therefore, ANNs has the ability to recognize the B-cancer in the mammogram images. This rate of detection indicates that the proposed method may be used as the basis for an effective prompting tool to assist radiologist in diagnosis of B-cancer. The experimental results based the convergence curve indicate that the concept of automatic detection is very promising. That is, ANN is a powerful tool to model complex medical data which may capable of high classing accuracies.

10. Evaluation Criteria

10.1 Recognition Rate

The classification accuracy as an evaluation metric is adequate provided that the class distribution of B-caner status is relatively balanced. The overall accuracy or recognition rate (Re_rate) is defined as the ratio between the total number of correctly classified instances and the test set size. Re_rate is defined as given in Equation 13.

$$Re\_rate = \frac{N_r}{N} \times 100\% \quad (13)$$

Where: N_r is the number of correctly classified samples during the test run. N is the complete number of test samples. Our test was performed for a total of 17 test samples, 16 samples are malignant and 1 sample is nonmalignant, and the same test was also performed for the trained samples, the rate of recognition for training and testing samples is shown in Table 2.

In Table 2 the recognition rate was computed using Equation 13, where $N_r = N_{r1} + N_{r2}$. It can be observed from Table 2 that the recognition rate is 100%, hence, all training and testing mammogram images were classified correctly as class 1 or class 0. In the long run, the experimental results show the fact that this new approach can automatically identify the cancer cells. Thus, the screening experiments have shown that the concept of automatic detection was very convincing.

10.2 Sensitivity versus Specificity

The importance of B-cancer segmentation and detection should not be depreciated. Increasing survival rates have been

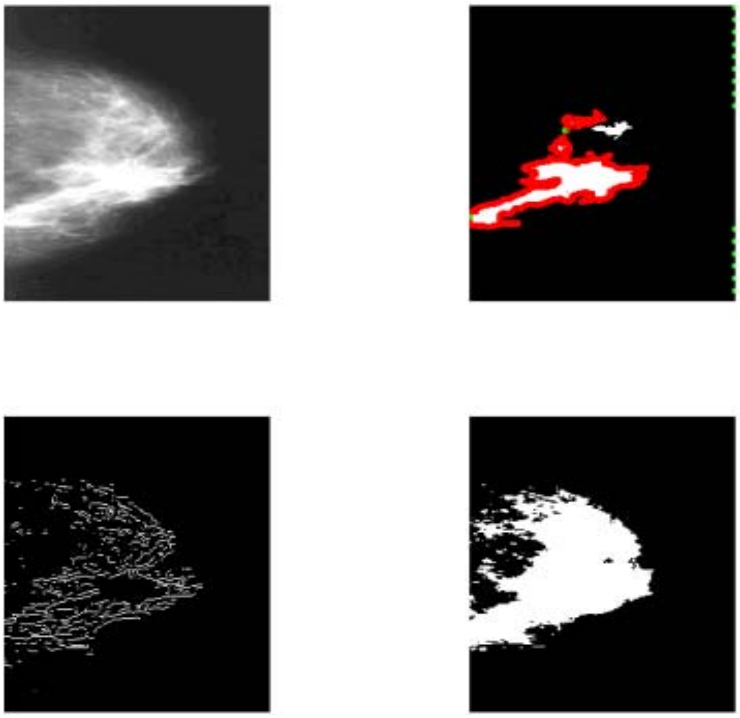

Figure 7. Edge image segmentation: Upper left: the original mammogram image, Upper right: tracing the boundary of the ROIs, Lower left: edge detection based Sobel operator, Lower right: The binary segmented image

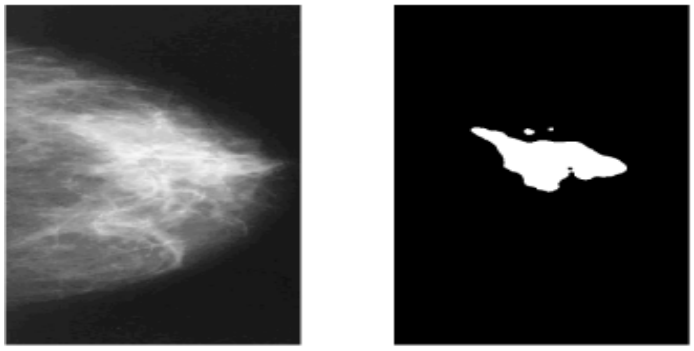

Figure 8. Training case: A mammogram image containing B-cancer classified by ANN

attributed to healthcare awareness campaigns and developing medical technologies. Regular examinations are keys to early B-cancer detection; however, as with any examination, there exists the possibility of misdiagnosis due to error. There are two classes of scientific error; systematic error and statistical error. A systematic error is the difference between what has been estimated of something and its actual state. For instance, human error is an example of systematic error which can be a factor in misdiagnosis. Also, a statistical error is the difference between what has been computed of something and its actual state. Though, statistical errors are due to fluctuations in the measurement apparatus that are not predictable. Statistical errors are subdivided into two types of errors: Type 1 errors and Type 2 errors. A type I error is also known as a False Positive (FP). A FP occurs when a hypothesis states something is true when it is actually not true. For example, a patient receives a false alarm and is told she has cancer when she in fact does not. Type II errors, also referred to as False Negative (FN). A FN occurs when the status of something is reported as false when it is actually true. An example of this is when a malignant tumor is diagnosed as benign and the patient is told that she is clear of cancer when she actually is not. The following four relative metrics are calculated when comparing the classifier output of the constructed systems that were determined by biopsy:

• False Positive: tumors which were not marked as tumor, but were classified as tumor.

• False Negative: tumors which were marked as tumor, but which were not classified as tumor.

• True Positive (TP): tumors which were not marked as malign, and that were also classified as malign.

• True Negative TN): tumors which were not marked as align, and that were also not classified as malign.

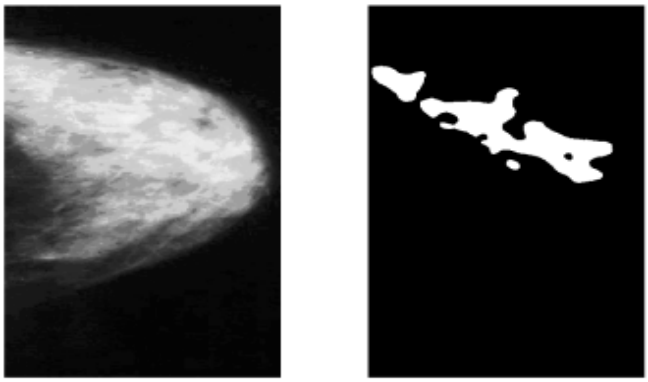

Figure 9. Testing case: A mammogram image containing B-cancer classified by ANN

Images set	Cancer (N_{r1})	Non-cancer(N_{r2})	Tot No (N)	Recognition rate
Training set	16	1	17	100%
Testing set	16	1	17	100%

Table 2. Recognition Rate Based On Classification Results

	Tumor marked malign	Tumor marked as not malign
Tumor classified as malign	True positives	False positives
Tumor classified as not malign	False negatives	True negatives

Table 3. The Confusion Matrix. A "Perfect" System Detects With 100% Sensitivity And 100% Specificity

				Status
		Negative	Positive	
Histopathology	Negative	17	1	SP: 94%
Margin status	Positive	2	18	SE: 90%
		NPV: 89%	PPV: 94%	

Table 4. Confusion Matrix For Margin Analysis On Clinical Specimens

The performance of such a diagnostic system is best described in term of its sensitivity and specificity quantifying their performance related to false positive and false negative instances. Sensitivity (SE) is the ratio of tumors which were marked and classified as tumor, to all marked tumors. Thus, sensitivity is expressed as the ratio of number of true positives to the sum of true positives and false negatives; sensitivity is defined as given in Equation 14.

$$SE = \frac{TP}{TP + FN} \tag{14}$$

Specificity (SP) is the ratio of tumors which were not marked and also not classified as tumor, to all unmarked tumors. Specificity is thus defined as the ratio between the number of true negative predictions to the sum of true negatives and false positives as given in Equation 15.

$$SP = \frac{TN}{TN + FP} \quad (15)$$

In reality, no inspection method can evaluate with perfect accuracy. Typically an instrument is considered to have an acceptable performance if its specificity and sensitivity are above 0.90. Table III displays a confusion matrix, which illustrates the relationship between the four possible values of an evaluation.

The results of a test performed with perfect sensitivity and specificity will all be either TP or TN' and never FP or FN. We wrote a small program to evaluate our metric criteria on 8 images, then, we test our identified results with the actual results provided by radiologist who have a high experience in diagnosis B-cancer. The data and the actual results are borrowed from Medical Hussein City with high efforts. Table 4 shows the confusion matrix for evaluation margin status in breast specimens, with sensitivity, specificity, Negative Predictive Value (NPV), and Positive Predictive Value (PPV).

The computed sensitivity (true positive (TP) detection rate) was 90.0%. The corresponding computed specificity was reduced to 94.4%

9. Conclusion And Future Works

In this paper an image-processing-based approach is proposed for Breast cancer (B-cancer) segmentation and detection from digital mammogram images. The approach has been carried out with the assistance of morphological operations and Artificial Neural Networks (ANNs). This type of approach will help the diagnostics of B-cancer as a useful view. The proposed technique is composed of four main steps; in the first two steps the images at hand are enhanced and segmented using contrast stretching and a morphological approach, respectively, in the third step some texture features were computed, finally, the extracted features were passed through a pre-trained neural networks. Our experimental results indicate that the proposed approach can significantly support accurate and automatic detection of Bcancer disease. Based on our experiments, the developed neural network classifier that is based on statistical texture classification perform well and can successfully detect and classify the malignant breasts with high precision. In future works we plan to use support vector machine technique into the classification process for possible hiking up the efficiency and accuracy of the implemented tool.

References

[1] Peto, R., Boreham, J., Clarke, M., Davies, C., Beral, V. (2000). Uk and usa breast cancer deaths down 25% in year 2000 at ages 20-69 years, *International Journal of Innovative Computing and Applications*, 355 (2).

[2] Ranadhir, G., Ghosh, M.,Yearwood, M, J. (2004). A modular framework for multi category feature selection in digital mammography, *In: Proceedings of the12th European Symposium On Artificial Neural Networks ESANN2004, Bruges(Belgium)*, p.175–180.

[3] Smith, R., Saslow, G., Sawyer, K., Burke, W.,Costanza, M., Evans, W., Foster, J.R.,Hendrick, E., Eyre, R. F. H., Sener, A., (2003). American cancer society guidelines for breast cancer screening, *A Cancer Journal for Clinicians*, 53 (3)141–169.

[4] Orel, S.G. (2000). Mr imaging of the breast, *Radiol Clin North Am*, No. 38, p. 899–913.

[5] Hiremath, P. S., Iranna,Y. H. (2005). Automated detection of esophageal cancer (necrosis) in endoscopic images using color image segmentation, *In*: *GESTS International Transactions on Computer Science and Engineering*, 16, p. 107–118.

[6] Delbeke, D., Martin, W, H. (2001).Positron emission tomography imaging in oncology, *Radiol Clin North Am*, No. 38, p. 883–917.

[7] Thurfjell, E. L., Lernevall, K. A., Taube, A. A. (1994). Benefit of independent double reading in a population-based mammography screening program, *Radiology*, 191, p. 241244.

[8] Elmore, J. G., Wells, C. K., Lee, C. H., Howard, D. H., Feinstein, A. H. (1994). Variability in radiologists interpretations of mammograms, *The New England journal of Medicine*, 331 (22) 14931499.

[9] Murugavalli, S., Rajamani, V. (2006). A high speed parallel fuzzy c-mean algorithm for brain tumor segmentation, *ICGST International Journal on Bioinformatics and Medical Engineering, BIME*, 06, 29–34.

[10] Alata, M., Molhim, M., Ramini, A. (2008). Optimizing of fuzzy c-means clustering algorithm using ga, *In*: *Proceedings of World Academy Of Science, Engineering, and Technology*, 29, p. 224– 229.

[11] Jain, R., Abraham, A. (2004). Comparative study of fuzzy classification methods on breast cancer data, *Australiasian Physical and Eng. Sciences in Medicine*.

[12] Chen, H. H., Duffy, S. W., Tabar, L. (1997). A mover-stayer mixture of Markov chain models for the assessment of dedifferentiation and tumour rogression in breast cancer, *j-J-APPL-STAT*, 24.

[13] Yau, K. K. W., Ng, S. K., Cheung, S. T. and Tung, M. C. (1999). Estimation of surgeon effects in the analysis of post-operative colorectal cancer patients data, *j-J-APPL-STAT*, 26.

[14] Zhou, Z., Jiang, Y., Yang, Y. B., Chen, S. F. (2002). Lung cancer cell identification based on artificial neural network ensembles, *Artificial Intelligence in Medicine*, 24 (1) 25–36.

[15] Tasoulis, D. K., Spyridonos, P., Pavlidis, N. G., Cavouras, D., Ravazoula, P., Nikiforidis, G., Vrahatis, M. N. (2003). Urinary bladder tumor grade diagnosis using on-line trained neural networks, *In*: *Proc. Knowledge Based Intelligent Information Eng. Systems Conf*, p. 199– 206.

[16] F. S. C. and E. G. T. Ng Y.-K. (2003). Developing case-based reasoning for discovery of breast cancer, *Journal of Mechanics in Medicine and Biology*.

[17] Sajda, P. and Spence, C. (2002). Learning contextual relationships in mammograms using a hierarchical pyramid neural network, *In*: *IEEE Transactions on Medicaln Imaging*, 2.

[18] Armatur, S. C., Piraino, D., Takefuji, Y. (1992). Optimization neural networks for the segmentation of magnetic resonance images, *In*: *IEEE Transactions on Medical Imaging*, 11, p. 215–220.

[19] Brzakovic, D., XM, L., Brzakovic, P. (1990). An approach to automated detectionof tumors in mammograms, *IEEE Transactions on Medicaln Imaging*, 9. p. 233–241.

[20] Busch, C. (2002). *Whole-Arm Grasping with Hyper-Redundant Planer Maniplulators Using Neural Networks*. Vorarlberg University of Applied Sciences, Diplomathesis.

[21] Al-Hiary, H. (2005). Identification and model predictive controller design of nonlinear systems using artificial neural networks, Master's thesis, Al-Balqa Applied University, Al-Salt, Jordan.

[22] Rebecca, C. W. (1997). Neural network models: Foundations and applications to an audit decision problem, *Department of Accounting, College of Management, National Taiwan University, Taipei, Taiwan 106, Republic of China Annals of Operations Research*, 75, 291– 301.

[23] ASCE, (2000). Artificial neural networks in hydrology. *In*: Preliminary concepts, *Journal of Hydrologic Engineering*, 2 (5) 115–123.

[24] Lin, D. and Yan, C. (2002). Lung nodules identification rules extraction with neural fuzzy network, *In*: *IEEE, Neural Information Processing*, V. 4.

[25] Chiras, N., Evans, C., Rees, D. (2002). Non-linear gas turbine modeling using feedforward neural networks, *In:Proceedings of ASME TURBO EXPO June 3-6, Amsterdam, The Netherlands GT-30035, University of Glamorgan, publisher of Electronics, Pontypridd, CF37 1DL, Wales, UK*.

[26] Prechelt, L. (1995). Adaptive parameter pruning in neural networks, *International Computer Science Institute I 1947 Center St.- Suite 600- Berkeley, California 94704-1198 - (510) 643-9153 - FAX (510) 643- 7684*.

[27] McAvoy, T. J., Wang, N. S., Naidu, S., Bhat, N., Gunter, J., Simmons, N. (1989). Interpreting biosensor data via backpropagation, *Proc. Int. Joint Conf. on Neural Networks, Washington D.C.*, V. 1, p. 227– 233.

[28] Leonard, J., Kramer, M. A. (1990). Improvement of the backpropagation algorithm for training neural networks,*Computer Chemical Engineering*, 14, p. 337–343.

[29] Barron, K. A. (2006). Modeling and uncertainty in breast cancer decision making, Master's thesis, The Pennsylvania State University, the Graduate School, Industrial Engineering.

[30] Vyborny CJ, G. M. (1994). Computer vision and arti cial intelligence in mammography, *Am J Roentgenol*, V. 162, p. 699–708.

[31] Kegelmeyer, W. J. (1992). Computer detection of stellate lesions in mammograms, *In*: Proceedings of the SPIE Conference on Biomedical Image Processing and Three-Dimensional Microscopy, February 10-13, San Jose, California, p. 446 –454.

[32] Haralick, R. M., Shanmugam, K. and Dinstein, I. (1973). Textural features for image classification, *In*: *IEEE Transactions on Systems, Man and Cybernetics*, vol. 3, p. 610621.

[33] Gonzalez, R. and Woods, R. (2008). *Digital Image Processing*. Prentice Hall, 3rd Edition.

[34] Gulsrud, T. O. (2000). *Texture analysis of Digital Mammograms*. PhD thesis.

[35] Pallikarakis, N. and Badea, C. (2008). *Medical Image Processing,Biomedical Technology Group, Dept. of Medical Physics*. current URL is http://bme.med.upatras.gr/improc/enhancement point processing. See Topic: Contrast Stretching, Last accessed on May 23.

[36] Braik, M., Sheta, A. and Ayesh, A. (2007). Particle swarm optimisation enhancement approach for improving image quality, *Int. J. Innovative Computing and Applications*, vol. 1, no. 2, pp. 138–145, 2007.

[37] Dhandra, B. V., Nagabhushan, P., Hangarge, M., Hegadi, R. and Malemath, V. S. (2006). Script identification based on morphological reconstruction in document images, *In*: *Proceeddings of the 18th International Conference on Pattern Recognition*, vol. 2, pp. 950–953.

[38] B. W. et al. (1998). Automated breast tumor diagnosis and grading based on wavelet chromatin texture description, *Cytometry*, vol. 33, no. 1, pp. 32–40.

[39] Doyle, S. A. and Madabhushi. (2007). Automated grading of prostate cancer using architectural and textural image features, *In*: *ISBI 2007*, p. 12841287.

[40] Gonzalez, R., Woods, R. (2002). *Digital Image Processing*. Prentice Hall, second ed..

[41] Pydipati, R., (2004). Evaluation of classifiers for automatic disease detection in citrus leaves using machine vision, Master's thesis, University Of Florida.

[42] Scarpa, V., Haindl, M. (2006). Unsupervised texture segmentation by spectral-spatial-independent clustering, *In*: *Proceedings of the 18th International Conference on Pattern Recognition*, pp. 151–154.