Classification of Breast Masses in Digital Mammograms Using Support Vector Machines

Nasseer M. Basheer,
Computer Engineering Department
Technical College/ Mosul, Iraq

Mustafa H. Mohammed
M. Tech Student, Computer Engineering Department,
Technical College/ Mosul, Iraq

Abstract— Breast cancer is the most common form of cancer diseases among women and the second leading cause of cancer deaths worldwide. Early detection is a key factor to a successful cancer treatment. Currently, the most common method for breast cancer screening is the X-ray mammography due to its simplicity, portability and cost effectiveness. The aim of developing Computer Aided Detection (CADe) and Computer Aided Diagnosis (CADx) systems is to aid radiologists to improve breast cancer screening and diagnosis. These systems act as a second opinion to the radiologist for better reading and understanding of mammography images. In this paper, a Computer Aided Diagnosis (CADx) system is implemented under the MATLAB (version 7.14.0.739− R2012a) environment for classifying abnormal masses in digital mammograms using Support Vector Machines (SVM). The proposed system successfully achieved 93% classification accuracy, which is considered as a good result when compared with similar works in the same research field.

Keywords— Breast cancer, X-ray mammography, Computer Aided Diagnosis, feature extraction, SVM classifier

I.   Introduction and Review

The high incidence of breast cancer has increased significantly in the recent years. It is a leading cause of fatality in women, with approximately 1 in 12 women affected by the disease during their lifetime [1]. In the last two decades, breast cancer has been the second leading cause of cancer deaths among women in the United States of America, following lung cancer [2]. According to [3], an estimated 232,340 new cases of invasive breast cancer are expected to occur among women in the United States of America during 2013; about 2,240 new cases are expected in men. Out of the 232,340 cases, 39,620 breast cancer deaths are expected. The most effective way to reduce breast cancer deaths is to treat it earlier. However, earlier treatment requires the ability to detect breast cancer in early stages [4]. The survival rate and the disease prognosis depend greatly on the cancer stage. And the treatment is more efficient when the disease is detected earlier, as the evolution into a more severe stage is avoided [5].

Currently, the mammogram is the most efficient system to detect clinically occult illness, being the only image-based method recommended for breast cancer screening. Mammography can greatly reduce the breast cancer mortality in a well-organized screening program over the population [5]. It is currently the best technique for reliable detection of early, non-palpable, potentially curable breast cancer [6]. Also, the process is easy and has a few side effects. However, because of the limitations of the human visual system, it is difficult for radiologists to provide both accurate and uniform evaluation for the enormous number of mammograms generated in widespread screening [2]. A major class of problems in medical sciences involves the diagnosis of a disease, based upon various tests performed upon the patient. When several tests are involved, the ultimate diagnosis may be difficult to obtain, even for a medical expert. Specifically, where breast cancer is concerned, the treating physician is interested in ascertaining whether the patient under examination exhibits the symptoms of a benign case, or whether the case is a malignant one [6].

With the advances of computer technology, radiologists have an opportunity to improve their image interpretation using computer capabilities that can enhance the image quality of mammograms. Over the past two decades, many attempts have been made to assist the radiologists in the detection and diagnosis of masses by developing computer-aided tools for mammography interpretation. Image processing and intelligent systems are two mainstreams of computer technologies that have been constantly explored in the development of computer-aided mammography systems [7]. Some of these studies in the field of mass diagnosis are listed below:

Cascio D. *et al.* 2006 [9] presented an algorithm for detecting massive lesions in mammographic images. A reduction of the surface under investigation was achieved, without loss of meaningful information, through segmentation of the whole image, by means of a ROI Hunter algorithm. Once the features were computed for each ROI, they were used as inputs to a supervised neural network with momentum. The output neuron provides the probability that the ROI is pathological or not. Results are provided in terms of Receiver Operating Characteristics (ROC) curve, the area under the ROC curve was found to be $A_Z = 85.6$ %.

Si X. and Jing L. 2009 [10] proposed an approach using Twin Support Vector Machine (TW-SVM) for automated detection of masses in digital mammograms. The proposed scheme was evaluated by a data set of 100 clinical mammograms from the Digital Database for Screening Mammography (DDSM) database. The results showed that the

sensitivity could reach 89.7% with 0.31 false positive per image. Further analysis showed that the proposed CAD system achieved 94% sensitivity for malignant masses in the test sets but the detection accuracy for benign masses was much lower, only 78%.

Rejani Y. and Selvi S. 2009 [13] presented a tumor detection algorithm from digital mammograms. The mammogram enhancement procedure included filtering, top hat operation, DWT and contrast stretching which is used to increase the contrast of the image. For the segmentation stage thresholding is used. The features were extracted from the segmented breast area. Next stage classifies the regions using the SVM classifier. The method was tested on 75 mammographic images, from the mini-MIAS database. The methodology achieved a sensitivity of 88.75%.

Islam M. *et al.* 2010 [8] presented an efficient computer aided mass classification method in digitized mammograms using Artificial Neural Network (ANN), which performed benign-malignant classification on a region of interest (ROI) that contained a mass. Three layers artificial neural network (ANN) with seven features was proposed for classifying the marked regions into benign and malignant. The authors achieved 90.91% sensitivity and 83.87% specificity which was claimed to be very much promising as compared to the radiologist's sensitivity, claimed to be 75%.

Eddaoudi F. *et al.* 2011 [11] proposed a masses detection algorithm using SVM classification and texture analysis. The segmented regions of interest were classified to normal and abnormal tissue using Haralick features calculated from the co-occurrence matrix. The results obtained with original mammograms, showed that 65 malignant out of 76 were classified true positive while 13 mammograms out of 19 were classified true negatives which correspond to a classification rate of 77% in average.

Singh B. 2011 [12] tested content based retrieval techniques for tissue classification and analysis of breast images. Analysis was carried out on 56 images collected from open source mini-MIAS database. Euclidean distance was used to compare the features of query image with stored images in database. Results showed that the suggested features can be used for both classification and retrieval of mammographic images. The retrieval accuracy was obtained to be 85.7%.

In this paper, a mass classification algorithm is developed based on the principle of Support Vector Machines (SVM) for normal-abnormal and benign-malignant classification of breast masses in digital mammograms. This paper is organized as follows, section 2 discusses the extraction of various features from the image, demonstrating the feature types used in this work, section 3 gives a brief introduction to Support Vector Machines with the associated equations, section 4 offers an overview on the parameters used to evaluate the proposed classification algorithm, section 5 shows the implementation steps of the proposed algorithm with the acquired results. Finally, section 6 concludes the article.

II.   The Feature Extraction Stage

An *image feature* is a distinguishing primitive characteristic or attribute of an image. Image features are of major importance in the isolation of regions with common property in an image (image segmentation) and subsequent identification or labeling of such regions (image classification) [14]. Also, feature extraction is a key phase in most pattern recognition systems. The features should have similar values for the patterns within the same class and have significant difference in different classes [10].

The development of algorithms for gathering a set of representative and concise features extracted from medical images that can truly represent an image remains a challenge for the researchers in the image processing and information retrieval fields. Mathematical and statistical methods aimed at extracting specific information from the images have been adapted in this regard. These features are usually placed in a *feature vector*, also referred to as *signature* [15]. This feature vector consists of various classes of features, some of the important classes include wavelet features, statistical intensity-based features and shape features, which will be described as follows:

A. Wavelet features:

The extracted wavelet features were calculated from the 2-D discrete wavelet transform that is implemented using the instruction (dwt2), which receives two arguments. The first argument is the input image, corresponding to the ROI the features being extracted from, and the second one specifies the particular wavelet that is applied. The Cohen-Daubechies-Feauveau (CDF) wavelet is used, since this wavelet class is widely applied in image processing researches [16,17,18].

The instruction(dwt2) returns four arguments corresponding to the four subbands (LL, LH, HL, and HH) that are produced from one level of DWT. Four wavelet features were extracted from this step, corresponding to the variance factor of each subband, respectively [12].

B. Statistical intensity-based features:

Statistical features based on the pixels making up the segmented region have been used extensively by many researchers in the field of digital mammography [10]. These features can be derived directly from the pixels values of the segmented ROI according to a specific formula for each feature, or can be computed indirectly in terms of histogram of the segmented ROI. The following measures have been formulated as quantitative shape descriptions of the image histogram [14,8]:

1) Mean value: The mean μ of the pixels values in the segmented ROI, estimates the value in the ROI in which central clustering occurs. In other words, it represents the average of all the pixels in the segmented ROI, as shown in the following equation:

$$\mu = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} p(i,j)$$

(2-1)

Where p(i,j) is the pixel value at point (i,j), in an ROI of size M×N. The instruction (mean2) can be used to calculate the mean value of the foreground ROI only after neglecting the black background pixels from the calculation.

2) Standard Deviation: The standard deviation σ is the estimate of the mean square deviation of grey pixel value p(i,j) from its mean value μ. Standard deviation describes the dispersion within the ROI. That is why the standard deviation feature is sometimes called the *image dispersion*. The standard deviation is defined as:

$$\sigma = \sqrt{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(p(i,j)-\mu)^2}$$

(2-2)

The standard deviation is calculated directly using the instruction (std2), which is equivalent to applying (eq.2-2) on the segmented ROI.

3) Smoothness: Relative smoothness R is a measure of grey level contrast that can be used to establish descriptors of relative smoothness in the segmented ROI. Smoothness can be calculated as:

$$R = 1 - \frac{1}{1+\sigma^2}$$

(2-3)

Where σ is the standard deviation of the segmented ROI.

4) Entropy: The entropy h can be described as a measure of the maximal amount of potential information given by the segmented ROI. Overall entropy of the image can be calculated as:

$$h = -\sum_{k=0}^{L-1} Pr_k \left(\log_2 Pr_k\right)$$

(2-4)

Where Pr_k is the probability of the k-th grey level, which can be calculated as $Z_k/M\times N$, Z_k is the total number of pixels with the k-th grey level and L is the total number of the available grey levels in an ROI of size M×N.

5) Skewness: The skewness *S* characterizes the degree of asymmetry of a pixel distribution in the specified ROI around its mean. Skewness is a pure number that characterizes only the shape of the distribution, and is given by:

$$S = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left[\frac{p(i,j)-\mu}{\sigma}\right]^3$$

(2-5)

Skewness is calculated directly using the instruction (skewness), however, since this instruction handles 1-D arrays only, it is applied twice on 2-D images.

6) Kurtosis: The kurtosis *K* measures the peakness or flatness of a distribution relative to a normal distribution. The conventional formula of kurtosis is:

$$K = \left\{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left[\frac{p(i,j)-\mu}{\sigma}\right]^4\right\} - 3$$

(2-6)

The additional (-3) term makes the value of *K* zero for a normal distribution. As with skewness, kurtosis can be measured directly using the instruction (kurtosis), which is applied twice for the same reasons given in the skewness feature to deal with the 2-D ROI.

7) Uniformity: The uniformity *U* is a texture measure based on the histogram of the segmented ROI, and is defined as:

$$U = \sum_{k=0}^{L-1} Pr_k^{\,2}$$

(2-7)

Where Pr_k is the probability of the k-th grey level. Because the Pr_k has values in the range of (0 to 1) and their sum equals 1, *U* is maximum when the numbers of pixels in all grey levels are equal, resulting in all the gray levels to be equal probable and their distribution to be uniform, and decreases otherwise.

C. Shape features:

Since the shape of the segmented ROI is important for proper diagnosis, shape features are extracted from each ROI prior to classification. In particular, the excessive lengthening is often a symptom of the absence of abnormality while the loss of chaotic dynamics in the lesion structures can mark a tumor [9]. Four shape features were extracted from the segmented ROI, which are listed as follows [5]:

1) Compactness: The compactness *C* is a measure of shape, which indicates how much the ROI is compact, and it is determined as:

$$C = \frac{P^2}{4\pi A}$$

(2-8)

Where *P* is the perimeter, *A* is the area of the segmented ROI in pixels. The 4π factor is added to the denominator such that the compactness of a complete circle is 1.

2) Aspect Ratio: The aspect ratio *AR* corresponds to the aspect ratio of the smallest window fully enclosing the ROI in both directions (as shown in figure 1), namely:
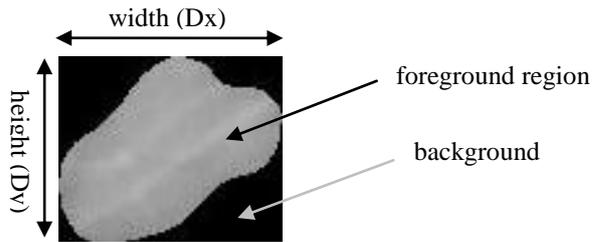
$$AR = \frac{D_y}{D_x} \tag{2-9}$$

Fig. 1 An ROI window example from which features will be extracted

Where D_y, D_x are the height and width of the previously mentioned window. When *AR* is bigger than 1, the value of *AR* is reciprocated, giving the value of 1/*AR* instead for consistency.

3) Area to background percentage: The percentage *perc* is specified by dividing the area of the segmented ROI in pixels by the area of the same window given in figure 1, this can be written as:

$$perc = \frac{Area_{ROI}\ (in\ pixels)}{Area_{window}\ (in\ pixels)} \tag{2-10}$$

The value of this feature will range from (0 to 1). It takes small values for ROI with appendices and branches emitted from it, and larger values for more compacted and rounded objects.

4) Perimeter ratio: The perimeter ratio *r_prim* presents the ratio between the perimeter of the segmented ROI to the perimeter of the same rectangular window of figure 1, which is written as:

$$r\_prim = \frac{primete\ r_{ROI}\ (in\ pixels)}{primeter_{window}\ (in\ pixels)} \tag{2-11}$$

The perimeter of the ROI is calculated by the instruction (bwperim), and then divided by the window perimeter, given from its height and width.

III.   Introduction to Support Vector Machines

One classification technique that is widely used for the diagnosis of breast tumors is the Support Vector Machine (SVM). SVM is one of the shining peaks among the many learning algorithms deeply inspired by statistical learning theory and appeared in the machine learning community in the last decades.

The theoretical advantage of SVMs is that by choosing a specific hyperplane among many that can separate the data in the feature space, the problem of overfitting the training data is reduced. They are often able to characterize a large training set with a small subset of the training points. Also, SVMs can work on features with arbitrary distributions, without the need to make any independence assumptions [24].
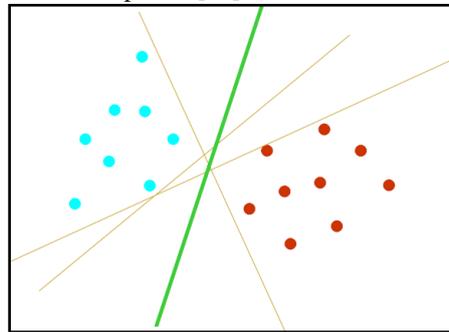

Fig.2 The optimal separating hyperplane

Consider the example in figure 2, here there are many possible linear classifiers that can separate the data, but there is only one that maximizes the margin (maximizes the distance between the classifier and the nearest data point of each class). This linear classifier is termed the *optimal separating hyperplane*. Intuitively, this boundary is expected to generalize well as opposed to the other possible boundaries [21].

Mathematically, the purpose of SVM is to find the optimal values for the hyperplane parameters *w* (e.g. w_0) and *b* (e.g. b_0). After finding the optimal separating hyperplane, such as $w_0 \cdot x + b_0 = 0$, an unseen pattern, x_t, can be classified by the decision rule [22]:

$$f(x) = sign\ (w_0 \cdot x + b_0) \tag{3-1}$$

Where *x* is a vector of the dataset mapped to a high dimensional space. Each x_i, belonging as it does to one of two classes, has a corresponding value y_i, where $y_i \in \{1, -1\}$, while *w* and *b* are parameters of the hyperplane that the SVM will estimate. The nearest data points to the maximum margin hyperplane lie on the planes:

$(w \cdot x) + b = +1$ for $y = +1$

$(w \cdot x) + b = -1$ for $y = -1$ (3-2)

By rescaling w and b, with no loss of generality, and grouping the above constraints in a single formula [35]:

     $\forall i$, $y_i f(x_i) \geq 1$ (3-3)

Where $y = +1$ for class w_1 and $y = -1$ for class w_2. The optimal separating hyperplane is enforced to separate the two classes of examples with the largest margin because, intuitively, a classifier with a larger margin is more noise-resistant. SVMs identify the datapoints near the optimal separating hyperplane which are called support vectors. The distance between the separating hyperplane and the nearest of the positive and negative datapoints is called the margin of the SVM classifier [23].

IV. Evaluation Methods

In CAD research, it is often necessary to assess the diagnostic performance of CAD systems and evaluate the differences between them [24]. Also, it is most important to test the performance of CAD systems against practically proven results, obtained by experienced radiologists [5]. The performance of a diagnostic system is usually measured by its sensitivity and specificity.

Sensitivity is the measure of how reliable a system is at making positive identifications, or, in other words, correctly identifying that which is inspected as being specifically that which is sought. A highly sensitive system will recognize what it is looking for most of the time, and rarely produces a false negative [19]. On the other hand, *Specificity* is a measure of how well a system can make a negative identification, or indicate when something inspected is not what is being sought, but something else. A classification system with high specificity will rarely make the mistake of identifying what is being inspected as what is being sought. There are other criteria that include both Sensitivity and Specificity, that is the *Accuracy* which measures the global performance of the algorithm about the correct decisions [19].

Using sensitivity and specificity, the results are usually defined in terms of Receiver Operating Characteristic (ROC) curve, Figure 2, which corresponds to the trade-off between the true-positive rate and the false-positive rate inherent in selecting specific thresholds on which predictions might be based [5]. The receiver operating characteristic analysis is a statistical method for analyzing, visualizing and comparing the performance of binary classification tasks. To evaluate the performance of a CADx system, where the task is to classify a suspicious region into benign or malignant, ROC analysis is typically used [24]. The Receiver Operating Characteristic curve expresses sensitivity as a function of $1-$ specificity, and utilizes the Az metric, where Az is the area under the ROC curve [15].
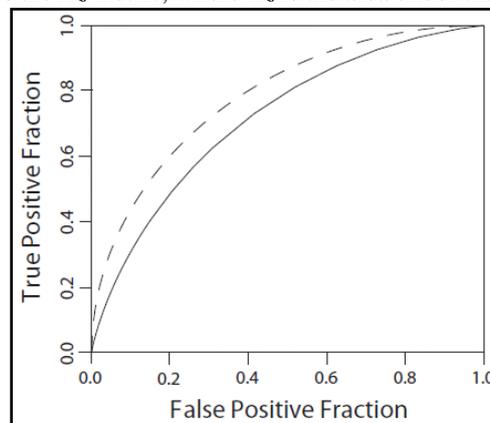

Fig.2 Two examples of a receiver operating characteristic (ROC) curve. The system with the dashed ROC curve performs better than the system with the solid ROC curve [23]

The overall performance is evaluated in terms of the area under the ROC curve [5]. The performance requirements of CADx systems are targeted to increase specificity without scarifying sensitivity [15].

V. Implementation and Results

In this work, a set of 105 mammography images are acquired from two databases (61 images from the online available MIAS database and 44 images from the local Al-joumhory hospital (in Mosul/Iraq) database). The images were segmented in a previous work by means of adaptive median filtering and texture analysis [20]. The features extracted from the segmented regions are fed into the classifier algorithm. In this work, the concept of Support Vector Machine (SVM) is used for classification, which is applied for two times. The first time involves the classification of the segmented ROI into normal or abnormal ROI, the second time involves the classification of abnormal ROI into benign and malignant masses. In each classification stage, distinctive train and test sets are created. For each set, the $1 \times M$ feature vectors of the extracted ROIs are vertically put together, resulting in an $N \times M$ feature matrix, where M corresponds to the number of features applied, N to the number of the extracted ROIs in each set apart, respectively. Therefore, two feature matrices are given for the two sets. On the other hand, two arrays of characters representing ground truth values (actual diagnosis) for each mass in train and test images are given. These characters can take one of two possible options, which are 'N' for normal and 'A' for abnormal mass in the first classification stage and 'B' for benign and 'M' for malignant mass in the second one. The instruction (svmtrain) is used for training the classifier in both stages, which accepts at least two input arguments. These arguments are the feature matrix calculated for the training set, and the characters array formed

from the training images. The output argument of (svmtrain) is a structure representing the knowledge base acquired from the training phase. The testing phase continues from the point where the training phase stops. The instruction (svmclassify) is used to classify the ROIs of the test set using the knowledge previously acquired from training. This instruction also accepts at least two arguments, which are the structure representing the knowledge base produced from the training instruction and the feature matrix calculated for the testing set.

The output argument of (svmclassify) is an array of characters representing the classifier expectation about the feature matrix calculated for the test set. This output is compared with the ground truth (actual diagnosis) of the test set in both classification steps (the masses classified as 'abnormal' in the first classification step are passed to the next classification step to be classified as benign or malignant masses) in order to evaluate the performance of the classifier in terms of sensitivity, specificity and accuracy.

In the first classification step, the 105 images representing the overall number of images are divided into 54 images for training and 51 images for testing. Normal-abnormal classification results of the test set can be presented as follows:

total number of test images: 51, number of ROI in the test images: 101
classification results (in regions):

True Positive (TP)= 45,True Negative (TN)=49, False Positive (FP)=6,False Negative (FN)=1

Sensitivity, Specificity and Accuracy [19] can be calculated as:

$\text{Sensitivity} = \frac{TP}{TP+FN} \times 100 = \frac{45}{45+1} \times 100 = 97.826 \%$

$\text{Specificity} = \frac{TN}{TN+FP} \times 100 = \frac{49}{49+6} \times 100 = 89.091 \%$

$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \times 100 = \frac{45+49}{45+49+6+1} \times 100 = 93.069 \%$

$\text{False Positive per Image (FPI)} = \frac{FP}{N} = \frac{6}{51} = 0.1176 \text{ FP/ image}$

Next, the Receiver Operating Characteristics (ROC) curve of the implemented classifier can be plotted by providing several values of sensitivity and specificity based on multiple classification attempts. The calculated area under the ROC for the proposed classifier is $A_z=0.9$, which is considered a promising result when compared with other classifiers in the field of breast mass diagnosis [9,25]. The ROC of the proposed classifier is shown in figure 3.

In the second classification step, only the ROIs classified as abnormal in the first step are accepted here. Again, the overall available abnormal masses are divided for training and testing, where 44 regions are allocated for training and 45 images are taken for testing. After applying the suggested classification procedure, the classifier produced classification accuracy of (92.3%). The reason behind the drop in the accuracy of the classification is due to the propagation of misclassified regions from the first step on the second step of the classification
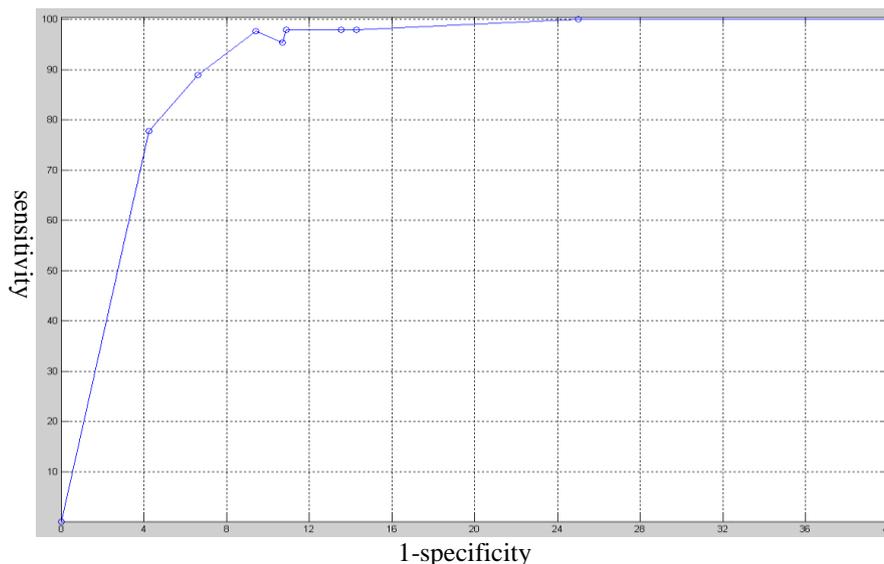

Fig. 3 The Receiver operating characteristic curve of the proposed classifier

VI. Conclusion:

In this paper, a breast mass classification system is implemented under the MATLAB environment for normal-abnormal and benign-malignant classification of a previously segmented mass candidates. From the suggested classification procedure, it is realized that intensity and shape driven features have the greatest effect on the later classification stage. Also, the inclusion of a normal-abnormal classification stage before benign-malignant classification, has a positive effect on the results of classification, greatly reducing the amount of false positive regions (from 55 FP regions in the test set before normal-abnormal classification into only 6 afterwards), thus, increasing the specificity of classification.

Furthermore, it is noticed that the classification accuracy is slightly reduced after performing benign-malignant classification (from 93.069 % to 92.3%), the reason behind this drop in classification accuracy is due to the propagation of classification error from the normal-abnormal to the benign-malignant classification stage. Finally, it can be noticed that the sensitivity measure is usually inversely proportional with the specificity measure. This inverse relationship is, however, obvious since increasing the sensitivity will yield in a system that identifies more regions as abnormal in order to achieve higher sensitivity. Since many of these regions are actually negative, the system will identify them as positive. This will increase the number of (FP), reducing the specificity of the system.

References
[1] Dehghani S. and Dezfooli M.; "Breast Cancer Diagnosis System Based on Contourlet Analysis and Support Vector Machine", World Applied Sciences Journal, Vol. 13, No. 5, pp. 1067-1076, 2011.

[2] Moayedi F. et al.; "Subclass Fuzzy-Svm Classifier As An Efficient Method To Enhance The Mass Detection In Mammograms", Iranian Journal of Fuzzy Systems, Vol. 7, No. 1, pp. 15-31, 2010.

[3] Siegel R., Naishadham D., Jemal A.; "Cancer Statistics, 2013", CA: A Cancer Journal for Clinicians, Vol. 63, No. 1, pp. 1-30, 2013.

[4] Paulin F. and Santhakumaran A.; "Classification of Breast cancer by comparing Back propagation training algorithms", International Journal on Computer Science and Engineering (IJCSE), Vol. 3 No. 1 Jan 2011.

[5] Cruz C.; "Automatic Analysis of Mammography Images:Enhancement and Segmentation Techniques", Engineering Faculty-Porto University, Porto, July 2011.

[6] Reyes C. and Sipper M.;"A fuzzy-genetic approach to breast cancer diagnosis", Artificial Intelligence in Medicine, Vol. 17, pp. 131-155, 1999.

[7] Martins L. et al; "Detection of Masses in Digital Mammograms using K-means and Support Vector Machine", Electronic Letters on Computer Vision and Image Analysis, Vol. 8, No. 2, pp. 39-50, 2009.

[8] Islam M., Ahmadi M. and Sid-Ahmed M., "An Efficient Automatic Mass Classification Method In Digitized Mammograms Using Artificial Neural Network", International Journal of Artificial Intelligence & Applications (IJAIA), Vol.1, No.3, July 2010.

[9] Cascio D. et al.; "Mammogram Segmentation by Contour Searching and Massive Lesion Classification with Neural Network", Institute of Electrical and Electronic Engineering (IEEE), 2006.

[10] Xiong S. and Jing L.; "Mass Detection in Digital Mammograms Using Twin Support Vector Machine-based CAD System", International Conference on Information Engineering, 2009.

[11] Eddaoudi F., Regragui F. and Lamouri N.;"Masses Detection Using SVM Classifier Based on Textures Analysis", Applied Mathematical Sciences, Vol. 5, No. 8, 367 – 379, 2011.

[12] Singh B.; "Mammographic Image Enhancement, Classification and Retrieval using Color, Statistical and Spectral Analysis", International Journal of Computer Applications, Vol. 27, No.1, pp. 18-23, August 2011.

[13] Rejani Y. and Selvi S.;"Early Detection of Breast Cancer Using SVM Classifier Technique", International Journal on Computer Science and Engineering, Vol. 1, No. 3, pp. 127-130, 2009.

[14] Pratt W.; "Digital Image Processing", Fourth Edition, Wiley, 2007.

[15] Deserno T.; "Biomedical Image Processing", Springer, 2011.

[16] Katsigiannis S.; "Acceleration of the Contourlet Transform", The university of Economy and Business, Athens July 2011.

[17] Beladgham M. et al.; "Improving Quality of Medical Image Compression Using Biorthogonal CDF Wavelet Based on Lifting Scheme and SPIHT Coding", Serbian Journal Of Electrical Engineering, Vol. 8, No. 2, pp. 163-179, May 2011.

[18] Toufik B. and Mokhtar N.; "The Wavelet Transform for Image Processing Applications", Advances in Wavelet Theory and Their Applications in Engineering, Physics and Technology, No.17, pp. 395-422.

[19] Lee N.; "Computer-Aided Diagnostic Systems For Digital Mammograms", Louisiana State University, December, 2006.

[20] Basheer N. and Mohammed M.;"Segmentation of Breast Masses in Digital Mammograms Using Adaptive Median Filtering and Texture Analysis", International Journal of Recent Technology and Engineering (IJRTE), Vol. 2, No. 1, pp. 39-43, 2013.

[21] Gunn S.; "Support Vector Machines for Classification and Regression", School of Electronics and Computer Science, University of Southampton, 1998.

[22] Shih F.; "Image Processing And Pattern Recognition Fundamentals and Techniques", Wiley, 2010.

[23] Bordes A.; "New Algorithms for Large-Scale Support Vector Machines", university of Paris, 2010.

[24] Samulski M.; "Computer Aided Detection as a Decision Aid in Medical Screening", university of Nijmegen, December 2011.

[25] Chan H., Wei J. and A. Helvie; "Computer-aided Detection System for Breast Masses on Digital Tomosynthesis Mammograms: Preliminary Experience", Radiology, Vol. 237, No. 3, December 2005.