



Computer-Aided Detection System for Breast Cancer Based on GMM and SVM

N. El-Sokary⁽¹⁾, A.A. Arafa⁽¹⁾, A.H. Asad⁽²⁾ and H.A. Hefny⁽²⁾

⁽¹⁾ Radiation Engineering Department, National Center for Radiation Research and Technology, Atomic Energy Authority, Cairo, Egypt

⁽²⁾ Computer Science Department, Institute of Statistical Studies and Researches, Cairo University, Cairo, Egypt

Received 16th Jan. 2019
Accepted 25th Mar. 2019

Region-of-interest (ROI) segmentation is an important critical step and challenging task in the evolution of computer-aided detection (CAD) system for breast cancer. The discovery of breast cancer in early stages can save many women lives. However, most of the early detection systems are costly in terms of complexity, price and processing time; that make it unsuited for developing countries. The digital mammography is proven to be one of the most important diagnostic techniques for breast cancer tumors. Therefore, this work proposes a CAD system for breast cancer detection from digital mammography based on Gaussian Mixture Model (GMM) followed by Support Vector Machine (SVM). The best contribution of our proposed system is the usage of GMM for the first time in the literature for mammogram images segmentation into ROI areas. Besides, the discrimination between the three classes of tissues as normal, benign or malignant, is used without previous knowledge of mammogram images' type. Moreover, the proposed system is fully automated in all of its stages with reduced computation compared with recent used methods. Hence, it offers a suitable early detection system to our country regarding moneywise, timewise, and reduced complexity. A non-linear multi-class SVM is used for classifying the ROI into three classes: normal, benign or malignant tissue. The experiments show overall average classification accuracy of 90% for detecting normal, malignant or benign randomly chosen 90 cases from the benchmark mini-MIAS dataset. On the other hand, the proposed method achieves 92.5% accuracy when classifying the benign from malignant cases. The confusion matrix was used to assess the proposed CAD system overall performance as it is considered to be a good evaluation metric method.

Keywords: Breast cancer, Diagnosis, CAD system, GMM, EM algorithm, Mammogram, Mini-MIAS, SVM classifier

Introduction

Nowadays, breast cancer is one of the most widespread cancers among a large number of women in the world. Every year, a large number of women lose their lives because of breast cancer. Every year, more than a million people are diagnosed with breast cancer in the world with about 400 thousand deaths. From previous researches, it is shown that, early detection could save lives and increase treatment opportunities. Mammography, ultrasound and Magnetic

Resonance Imaging (MRI) are the most commonly used diagnostic techniques for the detection of breast cancer. Mammography is one of the most commonly used and efficient techniques for the early detection of breast cancer, that increases the chances of treatment of patients and save their lives. Mammography is a specific type of the imaging that uses a low dose X-ray based on the difference in the X-rays absorption rates among the various breast tissues, such as fatty tissue, tumor

tissue, glandular tissue and calcifications. Besides, the mammograms are usually of high resolution images that facilitate the assessment of the breast cancer cases[1].

In recent years, many researches focused on development of the CAD systems to help radiologists in examining mammogram images. The CAD is an alternative method for double checking, that assists in the interpretation of mammograms and increase the detection accuracy of breast cancer from mammography. Besides, CAD systems are used to discover various suspicious areas within the breast tissue that the radiologist should detect. These systems usually ensure the detection of candidate lesions: masses or calcification. Therefore, the main purposes of CAD systems are to help in evaluating the risk of cancer and discriminating the abnormalities type based on the analysis of detected lesions. These systems usually involve describing the suspected lesions features and building a classifier based on these features to distinguish between each type of lesions.

Hence, the breast cancer's CAD system is considered a successful tool to detect and confront breast cancers in early stages. It increases the detection rate particularly in younger women, in which the suspicious areas interfere with dense breast tissue [2]. Among all the abnormalities that are encountered in the mammograms, detection and diagnosis of tumor as benign or malignant are one of the hard tasks, because of the variation in tumor in size, shape and margin, and the interfere with dense breast tissue. However many researches have relied mainly on many features of shape, margin, and texture to classify the masses [3].

In this present, a proposed system for the detection of breast cancer is developed, based on GMM clustering algorithm for the ROI segmentation and non-linear SVM for the ROI classification on randomly selected 90 mammograms images from the benchmark mini-MIAS dataset. This proposed system consists of five steps. First; the pre-processing step for removing the background and label from mammogram images. Second; the pectoral muscle removal. Third; the segmentation of the ROI from the breast region using GMM. The fourth step; the extraction of the features set from the segmented ROI. Finally; classifying the

ROI into normal, benign or malignant using multi-class SVM.

This paper is arranged into six sections, namely:

- The literature survey which presents a review of previous works involves the related to segmentation and classification for breast cancer.
- The proposed CAD system section, which present the full description of the proposed CAD system and the materials used.
- The experiment section discusses the conducted experiment.
- The result and discussion" section which presents the results obtained and discusses their pros and cons.
- The conclusion section concludes the paper contribution and finally.
- Future work section presents some ideas for further investigations.

Literature survey

Many researches focused on improving the mass detection and segmentation in order to accurately detect the abnormalities regions from mammogram images. Soulami, K.B., et al. [4] used the Particle Swarm Optimization algorithm (PSO) to extract the ROI and then used Fourier transform and Gray Level Co-Occurrence Matrix (GLCM) to extract shape and texture features from the abnormalities. The classification step of the ROI into normal or abnormal has been performed by the use of the SVM. Their algorithm was tested on mini-MIAS dataset and achieved 83.87% accuracy. There are some weaknesses in this system, such as drawbacks in removing the pectoral muscle area from the dense mammogram images. Moreover, the level of entropy thresholding is chosen manually to solve the over segmentation problem of the breast region.

Elmoufidi, A., et al.[5] used the dynamic K-means clustering algorithm with local binary pattern for the detection of the ROI in mammograms. In this system the contrast enhancement step is necessary; therefore a two-dimensional median filtering was used for enhancement of the contrast in mammogram images. Their algorithm was tested on mammograms from the MIAS dataset and the results showed that the algorithm was efficient for detection of the region of interest in mammography with an average of 85% and 2.84% False Positive Per Image (FPPI). One disadvantage

of this work was the dependence on the enhancement step to increase the contrast of images and improve the image for preparation for the followed steps in the system.

Guo, M., et al. [1] proposed a new model for the ROI segmentation to avoid the classical Vector Field Convolution (VFC) problem in image segmentation. Combined with Laplacian operator morphological filter, a new addition force by the user, defines the vector field kernel and was calculated by convolving the edge map. They used mini-MIAS database for testing their method and compared it with Gradient Vector Flow (GVF) snake and classical VFC snake methods. The advantage of their proposed method is shown in achieving the best performance and the lowest computational cost. The disadvantage in this work was that, it depends on finding more accurate boundaries of the tumor regions, especially for the abnormal cases. Also, the evaluation for their method was based on only comparing contour iteration time and computational cost with other methods.

Ismahan, H., F. Amel, and B. Abdelhafid [6] presented an effective approach for ROI segmentation based on the morphological method for detection of masses in digital mammograms. Multiple morphological watersheds were carried out for detection and localization of several types of masses in mammograms to avoid the problem of over-segmentation resulted from single morphological watershed. A set of 38 images from the Mini-MIAS data set was used for testing the efficiency of the algorithm and the results were compared with the manually marked masses by the radiologist. The obtained results showed promising performances of their proposed algorithm. The weaknesses of this work involve its high dependence on the enhancement process to improve images in order to increase the contrast of mammogram before segmentation step, and its massive computation results from applying two different segmentation techniques to improve ROI segmentation results.

Jothilakshmi, G.R., and Arun Raaza [7] proposed an effective method for the detection of masses in digital mammogram images and classification using multi- SVM classifier. They used a region-based segmentation method for extracting the ROI from mammogram images. Afterwards, they extract the texture features from the ROI using

GLCMs. The classification of the ROI into malignant and benign utilizes the texture features as an input to a multi-SVM. The algorithm has been evaluated on a set of 50 mammogram images from MIAS dataset and achieved accuracy up to 94%. However, this work focuses only on the differentiation between benign and malignant tumor. Algorithm efficiency was tested on mammogram images with the previous knowledge of its abnormal type of tissue.

Makandar, Aziz, and Bhagirathi Halalli [8] used multiple combined techniques for segmentation to extract the suspicious region from the mammogram images. The segmentation step consisted of watershed and morphological operations hybrid with the active contour technique. The aim from their segmentation technique was proposed to help minimizing the over segmentation of traditional watershed segmentation technique. Algorithm performance was tested on mammogram images from mini-MIAS database. The efficiency of the algorithm was measured by the use of the efficiency equation for calculating the number of true identification images from total number of abnormal images, and error equation for calculate number of false identification images from total number of abnormal images. The algorithm achieved an accuracy of 95.83% and the error is 4.17%. The weakness in this work was the segmentation step that was implemented using more than one technique and it has been implemented on more than one stage aiming to improve the ROI segmentation accuracy.

Proposed CAD system

The proposed CAD system for breast cancer detection consists of five consecutive steps without any previous knowledge of the breast cancer case or any pre-assistance from radiologist, as shown in Fig.(1).

A. Image preprocessing

A preprocessing step is needed for the mammogram images to eliminate the noise. The breast region in mammogram represents only 30% of the whole size of image. In addition, scanning mammography generate artifacts as in the mini-MIAS database the mammograms images contain high and low intensity labels, tape artifacts, and scanning artifacts, as shown in Fig.(2). These artifacts are one of the problems that are facing the

ROI segmentation process and causing many false positives. Therefore, in this step, only the area containing the breast region is selected automatically and removing these artifacts [16].

B. Pectoral removal

In some cases, pectoral muscles are easy to segment from the breast tissue. But in most cases, the entire pectoral muscle or parts of pectoral muscle are not visible and not easy to be removed[9]. In this section, a fully automatic pectoral muscle removal is proposed for successful identification of pectoral region.

The proposed removal technique of pectoral muscle is based on some characteristics concerning the nature of the mammogram images as follows: first, the pectoral muscles appear in the upper right corner or upper left corner of the mammogram. Second, the pectoral muscle appears a triangular shaped region. Afterwards, the pectoral muscle area is decreasing in size from top to bottom. Finally, it is necessary to remove the pectoral muscle because of its similarity to the abnormalities. In other words, the pectoral muscle is considered to be the brighter region in the breast tissue and it represents the very dense region in the mammograms. These properties provide features of the pectoral muscle that are similar to the abnormalities' features and hence influence the detection of the abnormalities region.

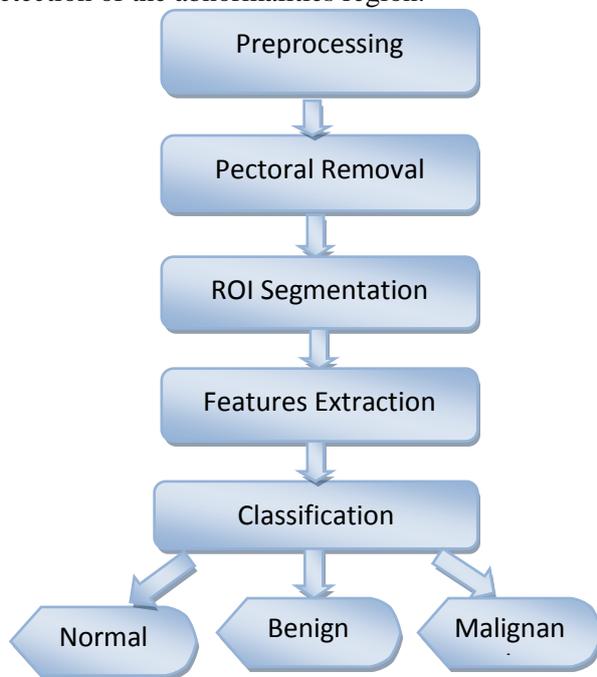


Figure (1): Hierarchical structure of the proposed CAD system

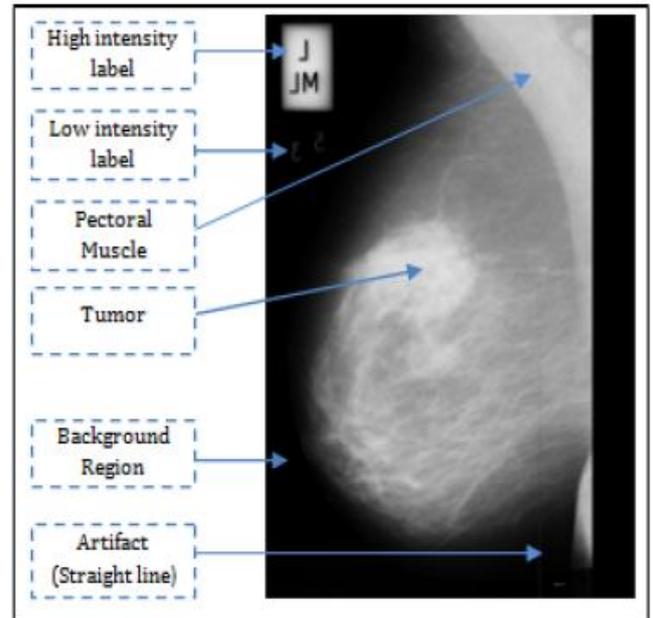


Fig.(2): The elements that constitute a mammogram image [5]

Therefore, a new simple method is proposed for removing the pectoral muscle region automatically. This method is summarized in the following steps:

- 1- Change image orientation: the orientation of all images is checked and changed (if necessary) to assure that the location of the pectoral muscle in the image at the upper left corner.
- 2- Check the image rows from top and left: in this step, the rows are searched to find the valid values that assure the start of pectoral region. Afterwards, the image is shifted to the top to discard the existence of bad rows. The same process is done for the left column of the image. Hence, the first point of the pectoral region is asserted to be at (0, 0). That point represents the right-angled vertex of the pectoral muscle triangle region.
- 3- Calculate the second point in the triangle by finding first point exceeding the mean value of the whole row.
- 4- Calculate the third point in the triangle by finding first point exceeding the mean value of the whole column.
- 5- Apply mask for pectoral muscle region by using the three points which have been selected in previous steps.

- 6- Subtract the pectoral muscle area from the original mammogram image.
- 7- As shown in Fig.(3), the pectoral muscle area is accurately extracted from the original mammogram image after changing the image orientation.

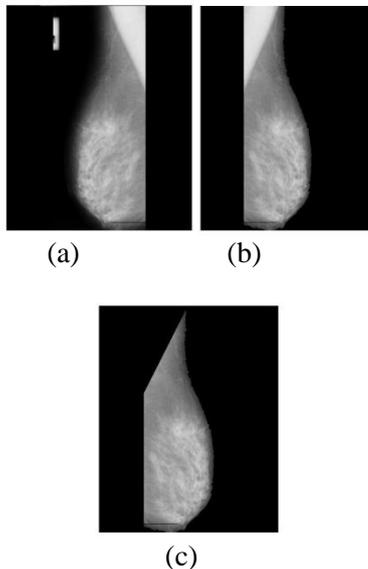


Figure (3): (a) The original image 'mdb081'. (b) After preprocessing and image orientation change. (c) After pectoral muscle region removal.

C. Segmentation

The segmentation step is a critical and very important process because all subsequent steps (features extraction and classification) depend on its result. The GMM is suggested to extract the ROI from the breast tissue. GMM uses the Expectation-Maximization (EM) algorithm for estimating the GMM parameters [10]. The segmentation of the ROI region is carried out by maximizing the ROI component likelihood [11].

In this work, GMM is selected to perform this step for the following reasons: (1) It is the statistical inference techniques studied. (2) It gives more flexibility in selecting component distribution. (3) It gives the density estimate for each group. (4) GMM can construct soft aggregate boundaries. This means that points in space may belong to any category with a given probability (i.e. rather than K-means as an example). The GMM is formulated as follow:

$$\sum_{i=1}^k p_i N(x_i | \mu_i, \sigma_i) \quad (1)$$

Where $N(x_i | \mu_i, \sigma_i)$ is a Gaussian distribution given by

$$N(x_i | \mu_i, \sigma_i) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp \left[-\frac{1}{2\sigma_i^2} (x_s - \mu_i)^2 \right] \quad (2)$$

The parameters μ_i, σ_i^2, p_i are respectively the mean, variance and the chance that a voxel x_s belongs to the class c_i . There are priori known as number of classes which can be used. This means that $p_i = (x_s \in c_i), 0 < p_i < 1, \sum_i^k p_i = 1, p_i = 1$. The GMM then is defined as following:

$$\theta = (p_1, \dots, p_k, \mu_1, \dots, \mu_k, \sigma_1^2, \sigma_2^2) \quad (3)$$

These parameters must be estimated. The EM algorithm is considered one of the most common methods that are used for estimation of the GMM parameters [12]. This algorithm is an iterative process to calculate the maximum probability when the observed data can be considered incomplete. Each frequency in the EM algorithm consists of two processes: the expectation (i.e. E-step) and the maximization (i.e. M-step). In the E-step, the observed data and the current estimates of the model parameters are used for estimating the missing data. These model parameters are the conditioned expectation that clarify the choice of terminology. In the M-step, the probability function is maximized under the pre-assumption that the missing data are known. The estimate of the missing data from the E-step is used in apposition of the real missing data. The algorithm ensures the maximization of the likelihood in each iteration, therefore convergence is guaranteed [12, 13].

In our proposed GMM based segmentation, the numbers of clusters k is decided automatically using thresholding to be suitable for each image according to its density which is calculated. After computation of the GMM parameters using the EM algorithm, the mammogram image is segmented into k cluster regions where each pixel belongs to a cluster. Figure (4) shows the ROIs segmented by GMM from three classes of normal, benign and malignant tissues.

D. Features extraction

The features extraction step is applied on the set of ROIs which are obtained from the segmentation of the mammogram image into regions to differentiate among their characteristics. Generally, the patterns present in a ROI are vary in contrast, high in contrast or low or tend to be coarse. Therefore, the analysis of intensity distribution of a ROI is necessary for successful and effective classification. [14]. The pattern existent in the ROI exhibits significant texture

features. Thus, in this paper fourteen features based on texture, shape and statistics are extracted from segmented ROIs different sizes from all images. GLCM is utilized for the texture features extraction from the ROIs[15]. For shape-based features, there are four features calculated from the ROI; circularity, brightness, compactness, and volume. Statistics-based features are the mean, standard deviation, correlation, skewness, smoothness, kurtosis, entropy energy and histogram of the ROI [16].

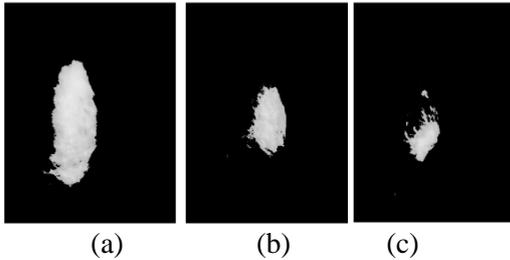


Figure (4): (a) Normal ROI segmented by GMM from original mammogram image 'mdb003'. (b) Benign ROI segmented by GMM from original image 'mdb212'. (c) Malignant ROI segmented by GMM from original image 'mdb072'

E. Classification

The classification of the ROI as either normal, benign tumor or a malignant one is achieved using the techniques of machine learning. This classification is based on the features that are extracted from the segmented ROI in previous step. The more related and indicative the features are of the ROI in the image, the more accurate the classifier model can perform. Many studies have been conducted on the use of variable machine learning models to classify breast cancer images. A Support Vector Machine (SVM) is a supervised learning classifier that distinguishes between different classes by finding a hyperplane that disconnect between them.

For a labeled training set of vectors in the form $(x_i; y_i)$; $i = 1, 2, \dots, k$, where $x_i \in R^n$ are the feature values, $y_i \in \{1, -1\}$ is their categories label, n is the number of features (i.e. dimension of the classification problem) and k is the number of sample[16, 17]. The equation of a hyperplane is given by

$$f(x) = x' \beta + b = 0 \quad (4)$$

where $\beta \in R^d$ and b is a real number. Finding β and b that minimize $\|\beta\|$ such that formulating the best separating hyperplane (i.e., the decision boundary). for all test feature points $(x_j; y_j)$, $y_j f(x_j) \geq 1$. On the other hand, the

vectors (x_s) that are on the boundary, and satisfy the condition $y_s f(x_s) = 1$ are called the support vectors[16].

In this work, a nonlinear multi-class SVM classifier is suggested to be used in the classification of the segmented ROI into normal, benign and malignant cases of mammogram images based on extracted features as inputs to the classifier.

Experiment

For experimental evaluation of the proposed system, a publicly available database of the Mammographic Image Analysis Society (Mini-MIAS) dataset was used. This is a publicly available digitized dataset with a subset of 322 mammograms. Every image is 1024×1024 pixels gray level image. The tumors have different shapes and sizes. The tumors have been labeled by a black rectangle. In general, the contrast of masses with the background is poor. The masses are often different in shapes and sizes and the boundary are irregular and ill-defined. Therefore, it is hard to distinguish between normal and abnormal breast tissue.[18]. In the proposed work, a collection of ninety mammograms images are randomly chosen from mini-MIAS database. Taking into consideration that those samples include images from three different types of breast density categories which are fatty, glandular and dense; for each category of images in Mini-MIAS database; normal, benign and malignant. The images are segmented using GMM with EM algorithm for estimating the parameters of the model. Fourteen features, based on texture, shape and statistics are extracted for each segmented ROI.

A nonlinear multi-class SVM classifier is used to classify segmented ROIs in mammograms images into the three classes. The chosen kernel for SVM is radial basis function (rbf) kernel which is the popular kernel among researches for its high accuracy among other kernels. From the total of 90 mammogram images that are randomly chosen from mini-MIAS dataset, 30 are normal cases, 30 are benign cases and the remaining 30 are malignant cases. The set of 30 cases in each class was randomly divided into two subsets; one set of 10 samples representing 33% was chosen as a training data and the other set of 20 samples representing 67% was used as testing data.

Table (1): Comparison with recent Malignant-Benign classification methods

Method	Segmentation	Features	classification method	Number of image	Radiologist intervention	2-class based Accuracy	3-class based Accuracy
Jothilakshmi, et. al. [7]	split and merge technique	Texture features	Multi-Support Vector Machine	50	YES	94%	--
Ismahan, et. al. [6]	watershed transform, morphological gradient		Visual Comparison between region detected and region selected by radiologist	38	YES	90%	--
Proposed Method	GMM	Texture, Statistical and shape features	Nonlinear m SVM	90	NO	92.5%	90%

Result and Discussion

Figure (5) presents the results from the first 3 steps of the proposed algorithm as discussed in section III. Each row in that figure shows an example of each studied case (i.e. normal glandular tissue, benign dense tissue and malignant fatty tissue respectively from top to down). On the other hand, each column of that figures illustrates the output of the corresponding step. Column (a) figures out the mammogram image after the preprocessing step in which the unwanted parts are removed and orientation is fixed. Column (b) illustrates the output of pectoral removal step. Finally, column (c) shows the segmented ROI using GMM.

To validate the proposed CAD system ninety mammogram cases are considered. First, the extracted fourteen features for all thirty training cases are used to train the classifier. Then, the rest of sixty cases are tested by the classifier. The proposed algorithm detects accurately nineteen cases as true normal with an accuracy of 95%, seventeen samples as true benign with an accuracy of 85% and fourteen samples as true malignant with an accuracy of 90%. Hence, the overall average accuracy is approximately 90%.

A confusion matrix is used to describe the performance of the classification model on the tested set of mammograms images for which the true values are recognized. Figure (6) shows the resultant confusion matrix for the model's performance when tested with the test set reserved for this purpose. The diagonal values of the matrix

indicate the true positive values of each class. The other values indicate the false negatives. As shown in Figure (6), the number of the falsely predicted classes (by the classifier) as abnormal are one. Furthermore, the number of malignant tumors that are falsely predicted classes as benign is 2, while the number of malignant tumors that are falsely predicted as normal is 0. Moreover, the number of benign tumors that are falsely predicted as malignant is 2, while the number of benign tumors that are falsely predicted as normal is 1. The precision is then defined as $TP/TP+FP$; where TP is the abnormal cases that are classified as abnormal cases and FP is the normal cases that are classified as abnormal cases. The sensitivity, is computed as $TP/TP+FN$; where FN denotes the abnormal cases that are classified as normal cases. The sensitivity is shown in the last column of the confusion matrix, while the accuracy is shown in the last row of the confusion matrix and illustrated in Fig. (6).

Makandar and Halalli have used a complicated segmentation step, in which they utilize three techniques with more computational time and cost. They detected the abnormal tissues with accuracy 95.86%. On the other hand, the proposed CAD system used only one technique for the segmentation step with less computational complexity and reduced time. The system achieved 95% accuracy for abnormality detection.

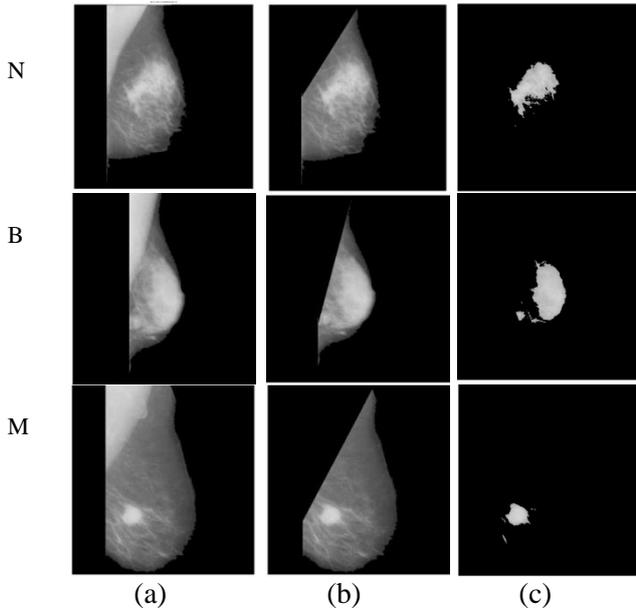


Figure (5): Result of each stage for three classes. The row (N) is for Normal image 'mdb209', row (B) is for Benign image 'mdb290', and row (M) is for Malignant image 'mdb028'. The column (a) shows the mammogram image after preprocessing step, column (b) shows the image after pectoral removal step, and column (c) shows the image after ROI segmentation using GMM

Breast Cancer Detection Confusion Matrix

Predicted Class	Normal	19 31.7%	1 1.7%	0 0.0%	95.0% 5.0%
	Benign	0 0.0%	17 28.3%	2 3.3%	89.5% 10.5%
	Malignant	1 1.7%	2 3.3%	18 30.0%	85.7% 14.3%
		95.0% 5.0%	85.0% 15.0%	90.0% 10.0%	90.0% 10.0%
	Normal	Benign	Malignant		Actual Class

Figure (6): The confusion matrix of the classification result

To the best of our knowledge, most of the previous works were interested only in the classification rate of malignant and benign tumors, in which there is a previous knowledge of abnormality cases. This classification accuracy is deduced for the proposed CAD system from the confusion matrix in Fig.(6) as 87.5%. Table (1) shows the comparison of tumors detection in previously reported data [6, 7]. Jothilakshmi and Raaza [7] selected 50

mammogram images that were previously known as abnormal and detected the tumor type at a rate of 94%. Ismahan et al., [6] used two techniques for segmentation and test type of abnormality at the rate of 90%. On the other hand, running the proposed CAD system for detecting the type of abnormality achieve an accuracy of 92.5% as shown in Table (1).

The proposed CAD system has the advantages over the methods in [6, 7] and other, that it is fully automated in all its stages and uses less complex and fast computation procedure without any radiologist contribution.

Conclusion

In our proposed system, for the mammogram input image, the segmentation of breast tissue process is implemented in one step just by using GMM for the first time in the literature. The nonlinear multi-class SVM is used for classification of the ROI into three categories; normal, benign and malignant tissue. Although the proposed system achieved an overall average accuracy of approximately 83%. This could be a promising result taking into consideration that our proposed system was evaluated on just only 90 images which are randomly chosen from mini-MIAS dataset. Besides, the proposed CAD system offers a suitable early detection system in Egypt, at a low cost, less time losing and of less complexity.

Future work

For future work, the authors intend to operate and test the present proposed system on the whole Mini-MIAS dataset. Another useful approach would be replacing the currently used step of pectoral muscle region removal step with GMM for segmenting also the pectoral muscle region from breast boundary, which in turn will give more accurate diagnosis and better performance for the whole system. Applied GMM for breast tissue segmentation and classification of breast density in mammogram images into three categories Glandular, density and fatty.

References

1-M. Guo, M. Dong, Z. Wang, Y. Ma, and Y. n. Guo, "A new method for mammographic mass segmentation based on parametric active contour model," in Wavelet Analysis and Pattern Recognition (ICWAPR), 2015 International Conference on, 2015, pp. 27-33.

- 2-R. Ghongade and D. Wakde, "Detection and classification of breast cancer from digital mammograms using RF and RF-ELM algorithm," in *Electronics, Materials Engineering and Nano-Technology (IEMENTech)*, 2017 1st International Conference on, 2017, pp. 1-6.
- 3-A. Midya and J. Chakraborty, "Classification of benign and malignant masses in mammograms using multi-resolution analysis of oriented patterns," in *Biomedical Imaging (ISBI)*, 2015 IEEE 12th International Symposium on, 2015, pp. 411-414.
- 4-K. B. Soulam, M. N. Saidi, and A. Tamtaoui, "A CAD System for the Detection of Abnormalities in the Mammograms Using the Metaheuristic Algorithm Particle Swarm Optimization (PSO)," in *Advances in Ubiquitous Networking 2*, ed: Springer, 2017, pp. 505-517.
- 5-A. Elmoufidi, K. El Fahssi, S. Jai-Andaloussi, and A. Sekkaki, "Detection of regions of interest in mammograms by using local binary pattern and dynamic K-means algorithm," *International Journal of Image and Video Processing: Theory and Application*, vol. 1, pp. 2336-0992, 2014.
- 6-J.H. Ismahan, F. Amel, and B. Abdelhafid, "Mass segmentation in mammograms for computer-aided diagnosis of breast cancer," in *Control, Engineering & Information Technology (CEIT)*, 2015 3rd International Conference on, 2015, pp. 1-5.
- 7-G. Jothilakshmi and A. Raaza, "Effective detection of mass abnormalities and its classification using multi-SVM classifier with digital mammogram images," in *Computer, Communication and Signal Processing (ICCCSP)*, 2017 International Conference on, 2017, pp. 1-6.
- 8-A. Makandar and B. Halalli, "Combined segmentation technique for suspicious mass detection in mammography," in *Trends in Automation, Communications and Computing Technology (I-TACT-15)*, 2015 International Conference on, 2015, pp. 1-5.
- 9-R. Nithya and B. Santhi, "Computer-aided diagnosis system for mammogram density measure and classification," *Biomedical Research*, vol. 28, 2017.
- 10-S. Ragothaman, S. Narasimhan, M. G. Basavaraj, and R. Dewar, "Unsupervised segmentation of cervical cell images using gaussian mixture model," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2016, pp. 70-75.
- 11-M. Balafar, "Gaussian mixture model based segmentation methods for brain MRI images," *Artificial Intelligence Review*, vol. 41, pp. 429-439, 2014.
- 12-A. M. Santos, A. O. de Carvalho Filho, A. C. Silva, A. C. de Paiva, R. A. Nunes, and M. Gattass, "Automatic detection of small lung nodules in 3D CT data using Gaussian mixture models, Tsallis entropy and SVM," *Engineering applications of artificial intelligence*, vol. 36, pp. 27-39, 2014.
- 13-M. A. Mahjoub, "Image segmentation by adaptive distance based on EM algorithm," *arXiv preprint arXiv:1204.1629*, 2012.
- 14-S. J. S. Gardezi, I. Faye, F. Adjed, N. Kamel, and M. Hussain, "Mammogram Classification Using Chi-Square Distribution on Local Binary Pattern Features," *Journal of Medical Imaging and Health Informatics*, vol. 7, pp. 30-34, 2017.
- 15-JA. M. Anter and A. E. Hassenian, "Computer aided diagnosis system for mammogram abnormality," in *Medical Imaging in Clinical Applications*, ed: Springer, 2016, pp. 175-191.
- 16-JM. Abdel-Nasser, H. A. Rashwan, D. Puig, and A. Moreno, "Analysis of tissue abnormality and breast density in mammographic images using a uniform local directional pattern," *Expert Systems with Applications*, vol. 42, pp. 9499-9511, 2015.
- 17-T. Subashini, V. Ramalingam, and S. Palanivel, "Automated assessment of breast tissue density in digital mammograms," *Computer Vision and Image Understanding*, vol. 114, pp. 33-43, 2010.
- 18-J. Suckling, J. Parker, D. Dance, S. Astley, I. Hutt, C. Boggis, et al., "The mammographic image analysis society digital mammogram database," in *Exerpta Medica. International Congress Series*, 1994, pp. 375-378.