Automatic Breast Cancer Detection Using Digital Thermal Images

Ola O. Soliman
Dept. of Mathematics,
Faculty of Science,
Cairo University
Giza, Egypt
ola.omar@ckes.cu.edu.eg

Nasser H. Sweilam
Dept. of Mathematics,
Faculty of Science,
Cairo University
Giza, Egypt
nsweilam@sci.cu.edu.eg

Doaa M. Shawky
Dept. of Engineering Mathematics,
Faculty of Engineering,
Cairo University
Giza, Egypt
doaashawky@staff.cu.edu.eg

***Abstract*—Breast cancer is one of the most common and main reason of death for women all over the world. About one in eight of women is subject to breast cancer over the course of her lifetime. There is no effective method to prevent or know the reasons of growing these cancerous cells, however the number of deaths can be reduced by early detection. Breast cancer detection and classification is one of the most important fields that the researchers are working on. Thermal breast images are considered as an efficient type of screening strategies. The aim of this study is to develop an efficient system to detect breast cancer by using image processing techniques. The proposed system extracts the characteristic features of the breast from the region of interest that is segmented using a novel approach from the thermal input image. Then the image is classified based on these features to normal or abnormal using a neural network classifier. The system is evaluated on a benchmark dataset and a success rate of 96.51% is obtained.**

Keywords—Cancer detection, thermal images, image segmentation, neural network.

I. Introduction

Cancer is a given name to a collection of abnormal cells that begin to divide and grow out of control spreading into the surrounding tissues forming tumors. Those tumors may be growing among any organ at the human body and anybody can get breast cancer [1]. Breast cancer is the most common cancer in women worldwide. In 2018, an estimated 330,080 new cases of breast cancer are expected to be diagnosed in women in the U. S., and about 40,920 women in the U. S. are expected to die from breast cancer [2]. Therefore, early detection and treatment of cancer minimize the risk of death and increase survival rates [2]. There are several imaging techniques to detect breast cancer such as mammography, ultrasound, Magnetic Resonance Imaging (MRI) and thermography. Mammogram is less effective and has some drawbacks such as it is painful, invasive and not useful in the cases of women who have dense breast. MRI also, as mammography, cannot identify the difference between a cancerous lump and a benign cyst. Breast thermal images, on the other hand, are considered as an effective screening tool to detect breast cancer. Since thermography is a non-invasive (it has no contact with the body of any kind and no radiation), painless, low cost screening test and women of all ages can be examined by it, particularly women with dense breast. In addition, breast thermography allows us to detect tumors 10 years earlier than mammography [3], [4].

Depending on the expertise; mistakes can be made by medical professionals while identifying a disease. With the help of technology such as machine learning, diagnosis can be more accurate when related to a diagnosis made by an experienced doctor [5]. Any Computer-Aided Detection (CAD) system for breast cancer detection can be divide mainly into three stages: pre-processing and region of interest segmentation (ROI), feature extraction, classification and performance analysis. ROI segmentation aims to separate breast region from the rest of thermal images. After such segmentation, some features are extracted. In order to classify the breast into normal or abnormal, some tools of artificial classifications algorithms are usually applied. Usually, obtained classification results of breast thermal images depend mainly on the accuracy of segmentation result. Accurate segmentation of region of interest from medical thermal images still remains as an open problem. The absence of clear edges, low contrast nature and low signal to noise ratio are among the inherent limitations of thermal image. Hence, complex pre-segmentation steps remain a problem due to the more intricate intensity field of breast thermogram.

In this work, an automatic breast cancer detection technique is presented. The main contribution of the presented work is the proposed novel segmentation technique that utilizes the concave and convex shape of the frontal view of the breast. In order to evaluate the accuracy of the presented approach, different selective features were extracted from the segmentation regions and then a neural network (NN) classifier was used to detect the breast abnormalities.

II. Background

In this section, we briefly discuss the related work for the early detection of breast cancer.

978-1-5386-8154-1/18/$31.00 ©2018 IEEE

Authorized licensed use limited to: NATIONAL INSTITUTE OF TECHNOLOGY CALICUT. Downloaded on October 11,2023 at 05:38:24 UTC from IEEE Xplore. Restrictions apply.

One of the popular ways for separating the normal from abnormal breast is the application of asymmetric analysis in which the left and right breast features are compared together. If asymmetry between the left and right breasts in each patient is more than a specified threshold, then this will increase the probability of a cancer case in that patient. In these algorithms, the comparison of left and right breasts is performed based on the extracted features from the breast image or by using pixel temperature as a feature. For instance, in [6], a breast cancer detection algorithm based on asymmetric analysis as primitive decision and decision-level fusion by using Hidden Markov Model (HMM) is proposed. The authors defined a novel texture feature based on Markov Random Field (MRF) model that is called MRF-based probable texture feature. To determine the cancerous cases from normal cases, a classification contemporary with decision level fusion by means of Hidden Markov Model (HMM) is employed to fuse different extracted features from an image. To improve the classification accuracy, a decision level fusion is used so that the results of all features can be used for an image simultaneously. To evaluate the proposed system, 65 breast thermography images available online have been used and false negative rate of 8.3% and false positive rate of 5% are obtained. Also in [7], an algorithm has been introduced to extract the breast characteristic features which are extracted from the thermal images based on image analysis. An NN is used as a classifier to classify the breast to normal or abnormal. The RGB color model is used to display the color breast thermal image and HIS color model is used in all processing stages. This model contains 6 stages all applied to the gray level image, the first one is background elimination (optimum threshold is applied to make the pixels belong to the background region or to the body); then the region of interest (ROI) is identified. At this step, the images are segmented and separated into left and right breast region depending on finding 3 points (first minimum pixel, maximum pixel and second minimum pixel). The top, bottom, left and right boundaries are generated based on these points as well as the middle line of the body to separate the left and right breast regions. Then breast mass contour detection algorithm based on Hough transformation is applied. A gray level co-occurrence matrix is constructed to extract the texture features, then the feature vector is extracted (4 statistical features for each color space (4*3=12) and 5 texture features in the four directions; horizontal, vertical, and two diagonal directions of the gray level co-occurrence matrices (5*4=20)), a vector of features is passed to the classifier which is the final stage. A benchmark database was used to evaluate the proposed system. The reached success rate was 96.12%.

Moreover in [8], a system for breast cancer detection has been developed using backpropagation neural network. The sensitivity, specificity and accuracy have been calculated then the authors compare their results with radial basis function network. The authors concluded that the backpropagation neural network is the best technique to detect breast cancer. In [9], four different neural networks have been implemented: Back Propagation Algorithm (BPA), Learning Vector Quantization (LVQ), Radial Basis Function Network (RBFN) and Competitive Learning Network (CL). LVQ has been found to record the best result as a classifier to detect breast cancer. Also, in [10], a hybrid system for breast cancer tumor recognition using three modules is proposed. The models include feature extraction with use of fuzzy feature, a hybrid bee algorithm (BA), and multi-layer perceptron (MLP) neural network is used. In [11], statistical features were extracted from thermal images using wavelet transformation. The authors used NN as a supervised classifier. A feed-forward multilayer perceptron network (MLP) that consists of a four layers has been developed which has 15 neurons in the input layer, 30 neurons in the first hidden layer, 15 neurons in the second hidden layer and one neuron in the output layer. Total 156 samples of 52 individuals are used to train the network and remaining 150 samples of 50 individuals are used to test and validate the system. The obtained accuracy was 90.48%. In addition, high-order spectral features have been extracted by taking random transform of pseudo-color breast thermal image in [12]. A feed forward artificial neural network and support vector machine (SVM) are used to classify the input data into normal and abnormal. NN presented better sensitivity, specificity and accuracy values of 92%, 88% and 90%, respectively, in comparison with SVM.

III. METHODOLOGY

In this section, we briefly discuss the main steps of the proposed approach to detect breast cancer.

A. *Patient prepation and image data acquisition*

Since we are working with a type of tests that depends on the human body temperature, the patient must be examined under specific conditions. As mentioned in [7], the design and environmental conditions of the room should conform to the thermodynamic attributes required in thermal image acquisition as part of image quality control. The room itself should be of adequate size to maintain a homogenous temperature. The patient should be instructed to refine from sun exposure, breasts treatment, cosmetics, antiperspirants, lotions and bathing before the examination. Lastly, the patient must undergo 15 min of waist-up nude acclimation. The last 5 min of this acclimation period the patient placed with hands on top of their head in order to facilitate an improved anatomic presentation of the breasts for imaging. These images were captured using a special type of camera which is used to convert the infrared radiation emitted from the skin into electronical impulses that are visualized in color on a monitor [13]. The captured images are a 640×480×8 bits digital color mapped infrared thermal image which are converted to 24-bit RGB image.

B. *pre-processing and region of interest segmentation*

Noises are always undesirable. So removing noise with preserving edges of the image plays a vital role in image processing. Median filter is one of the simplest and most popular approaches for removing noise like salt and pepper.

Authorized licensed use limited to: NATIONAL INSTITUTE OF TECHNOLOGY CALICUT. Downloaded on October 11,2023 at 05:38:24 UTC from IEEE Xplore. Restrictions apply.

In this work, the original thermal images are converted into gray-scale by the following equation [7]:

$$Gray\text{-}scale= 0.21R + 0.71G + 0.07B \qquad (1)$$

Where R, G, and B are the red, green and blue bands of color image, respectively. Figure 1 shows a sample of RGB and gray-scale breast thermal image.

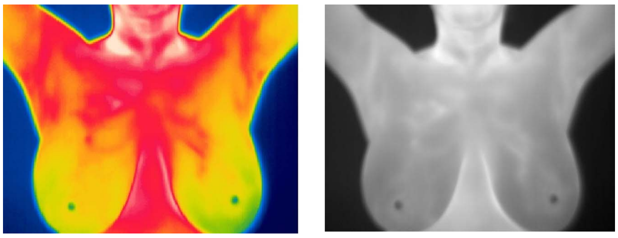

Fig. 1. RGB and Gray-Scale Breast thermal image.

Segmentation of the region of interest (ROI) is an important stage to detect the cancer from the thermal breast image, the aim is to separate the regions of the left and right breast from the human body. The proposed algorithm is based on three predefined parameter: the distance is 1 meter between the patient and the camera, the image contains only of the upper part of the human body and the breast has a specific position at the image as shown in Fig. 1.

The input breast thermal image is smoothed using Gaussian filter with σ = 1.4, then Canny edge detector is applied to detect the outer boundary of the breast, in order to make this detector suitable for further processing to find the boundaries, a morphological dilation operation has been used. Detecting lower border is performed by identifying infra-mammary line of the breast using horizontal projection profile (HPP) approach. HPP is a histogram of one dimensional array with a number of entries equal to the number of white pixels in each row; it is applied on the image from the bottom to the top. Due to fold of the breast the value of HPP is high in the infra-mammary line. The corresponding row number to the first high HPP is taken as the lower limit of the breast thermal image segmentation. Depending on the concave and convex shape of the frontal view of the breast, right, left and upper borders of the breast are defined. Region of the breast is obtained based on detecting left, right, lower and upper borders. By finding the bifurcation point which is the intersection of the tow infra-mammary curves, we can separate the left and right breast. It is too difficult to find complete edges of the infra-mammary curves because of the unclear nature of thermogram images. Thus, a Gaussian smoothing filter is applied followed by adaptive thresholding Canny filter, to extract the curves of the segmented image of the breast. A morphological closing operation is done to join the broken edges on the detected infra-mammary curves (The two longer curves found in breast thermogram at the lower half projecting towards the image center). The pixel locations of these curves are found by performing a boundary tracing. The detected curves are in ascending order based on its length. Hence, we select the first two longer curves then apply a polynomial curve fitting and extended upwards to find the bifurcation point. Figure 2 shows the segmentation result.

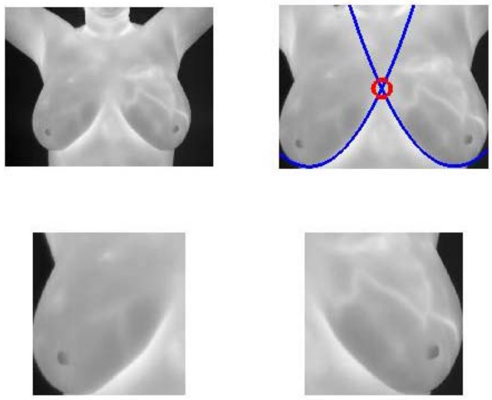

Fig. 2. Segmentation process with polynomial curves.

C. Feature extraction and classification

Feature extraction is performed on the segmented thermogram. Extracting effective features is a necessary step for asymmetry analysis. Therefore, first- and second-order statistical features are extracted by computing histogram and Gray-Level Co-occurrence Matrices (GLCM) of the segmented image. There are four color moments that are used as features. These are computed from the histogram of the image which are the mean, variance, skewness and kurtosis. GLCM is a statistical method of examining texture that considers the spatial relationship of pixels [14]. It is created by calculating how often various combinations of gray levels co-occur in an image. In our work we compute GLCM in four directions, vertical, horizontal, left and right diagonal at a distance of 1 pixel. Mean, variance, energy, entropy, contrast, homogeneity and correlation features are computed from the four GLCMs. Since the growth of the cancer in the breast is chaotic, we calculate the average value of features obtained from GLCM at the four directions. For measuring the asymmetry between right and left breast, absolute difference of each feature from the region of interest is computed and normalized to be a feature vector which will be passed to the classifier.

A neural network (NN) classifier is used for training and testing the significant features. NNs have been successfully applied across an extraordinary range of problem domains in different areas as engineering, physics and medicine (e.g., [15-20]. For this work, we use a popular algorithm, Backpropagation Neural Network Model (BPNN). BPNN is multi-layer feed-forward NN with no backward loop but it back-propagates the error so that it is called BPNN. It has four main steps: Feed-Forward computation, back propagation to the output layer, Back propagation to the hidden layer and weight updates.

Authorized licensed use limited to: NATIONAL INSTITUTE OF TECHNOLOGY CALICUT. Downloaded on October 11,2023 at 05:38:24 UTC from IEEE Xplore. Restrictions apply.

IV. Results and Discussion

A. Breast cancer Dataset

A benchmark database [21] has been used to evaluate the proposed approach. This public database is constructed by collecting the IR images from UFF University's Hospital. In the proposed approach, 63 IR single breast images (640 × 480 pixels) from this database were used (29 healthy and 34 malignant).

B. Experimental Results

The MATLAB (R2017a) is used for the implementation. A total of 63 thermography of the breast were segmented and analyzed as discussed in the Methodology section. A feature vector of thirty two features representing the breast characteristic is used, (Twelve statistical features and twenty texture features). Table [1] shows the Statistical analysis of extracted features.

TABLE 1 STATISTICAL ANALYSIS OF EXTRACTED FEATURES

Features	Normal breast		Abnormal breast	
	Mean	***SD***	***Mean***	***SD***
Histogram				
Mean	0.0993	0.0803	0.2651	.02133
Variance	0.2204	0.1590	0.3020	0.2538
Shewness	0.1362	0.1358	0.3049	0.2874
Kurtosis	0.0955	0.0755	0.3493	0.2458
GLCM				
Mean	0.1897	0.0972	0.2505	.1815
Variance	0.2204	0.1428	0.5320	0.3056
Entropy	0.2566	0.1066	0.3177	0.2250
Contrast	0.0942	0.0583	0.1295	.1524
Energy	0.1341	0.0875	0.3876	0.2627
Correlation	0.1936	0.1500	0.2880	0.2524
Homogeneity	0.1095	0.0661	0.1498	0.1515

A BPNN with 8 neurons in hidden layer has been trained as a classifier. The obtained accuracy, sensitivity, and specificity are 96.51%, 79.7%, and 98.25%, respectively.

V. Conclusions

In this paper, an approach for the early detection of breast cancer using thermal images is proposed. Using thermography technique and a developed segmentation algorithm, a novel set of useful features has been extracted and classified using backpropagation neural network. A promising classification result has been obtained, which shows the effectiveness of the proposed approach.

References

[1] http://www.srspharma.com/breast-cancer.htm, 4th Jan, 2013.

[2] http://www.cancer.org/cancer/breastcancer/detailedguide/breast-cancer-key-statistics, 9th Jan, 2018.

[3] B. F. Jones, "A reappraisal of the use of infrared thermal image analysis in medicine," IEEE Trans. Med. Imaging, 1998 Dec;17(6):1019-27.

[4] D. Kennedy, T. Lee, D. Seely, "A comparative review of thermography as a breast cancer screening technique," Integr Cancer Ther.;8(1):9-16, 2009 Mar.

[5] R. W. Brause, "Medical Analysis and Diagnosis by Neural Networks," in Proceeding ISMDA '01 Proceedings of the Second International Symposium on Medical Data Analysis, pp. 1–13, 2001.

[6] R. Rastghalam, H. Pourghassem, "Breast cancer detection using MRF-based probable texture feature and decision-level fusion-based classification using HMM on thermography images" Pattern Recognition51, 176–186, 2016.

[7] N. A. Mohamed, "Breast cancer risk detection using digital infrared thermal images," international journal of bioinformatics and biomedical engineering, 2015.

[8] S. Singh, P. R. Gupta, "Breast cancer detection and classification using neural network," International Jornal Of Advanced Engineering Sciences And Technologies, Vol No. 6, Issue No. 1, pp. 004-009, 2011.

[9] R. Elhamdi, M. Njah, M. Chtourou, "Breast Cancer Diagnosis Using a Hybrid Evolutionary Neural Network Classifier," 18th Mediterranean Conference on Control & Automation, pp. 1308-1315, Jun. 2010 .

[10] A. Khosravi, J. Addeh, J. Ganjipour, "Breast Cancer Detection Using BA-BP Based Neural Networks and Efficient Features," &th iranian Conference on Machin Vision amd Image Processing, Iran, 2011.

[11] S. Pramanik, D. Bhattacharjee, M. Nasipuri, "Wavelet Based Thermogram Analysis for Breast Cancer Detection," International Symposium on Advanced Computing and Communication (ISACC), India, pp. 205-212, IEEE, 2015.

[12] U. R. Acharya, E.Y. K. NG, S. V. Sree, C. K. Chua, S. Chattopadhyay, "Higher order spectra analysis of breast thermograms for the automated identification of breast cancer," Expert Syst. 31(1), pp. 37-47, 2014.

[13] M. Gautherie, "Atlas of breast thermography with specific guidelines for examination and interpretation, " Milan, Italy, PAPUSA. 1989.

[14] U. R. Acharya, E. Y. K. NG, J. H. Tan, S. V. Sree, "Thermography based breast cancer detection using texture features and support vector machine,". J. Med. Syst. 36(3), 1503–1510 , 2012.

[15] M Shawky, Doaa, and Ahmed F Seddik. "On the Temporal Effects of Features on the Prediction of Breast Cancer Survivability." *Current Bioinformatics* 12.4 (2017): 378-384.

[16] Seddik, Ahmed F., and Doaa M. Shawky. "Logistic regression model for breast cancer automatic diagnosis." *SAI Intelligent Systems Conference (IntelliSys), 2015*. IEEE, 2015.

[17] Seddik, Ahmed F., and Doaa M. Shawky. "A low-cost screening method for the detection of the carotid artery diseases." Knowledge-Based Systems 52 (2013): 236-245.

[18] Abdel-Rahman, A. S., Ahmed Farag Seddik, and Doaa M. Shawky. "An affordable approach for detecting drivers' drowsiness using EEG signal analysis." Advances in Computing, Communications and Informatics (ICACCI), 2015 International Conference on. IEEE, 2015.

[19] Shawky, Doaa M. "Employing Spectral Domain Features for Efficient Collaborative Filtering." arXiv preprint arXiv:1703.01093 (2017).

[20] Mohamed, Zainab, et al. "Characterizing Focused Attention and Working Memory Using EEG." Sensors 18.11 (2018): 3743.

[21] L. Silva, D. Saade, G. Sequeiros, A. Silva, A. Paiva, R. Bravo, and A. Conci, "A new database for breast research with infrared image," Journal of Medical Imaging and Health Informatics, vol. 4, no. 1, pp. 92–100,2014.

Authorized licensed use limited to: NATIONAL INSTITUTE OF TECHNOLOGY CALICUT. Downloaded on October 11,2023 at 05:38:24 UTC from IEEE Xplore. Restrictions apply.