IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727
PP 39-45
www.iosrjournals.org

Automatic Detection of Breast Cancer in Mammograms using Optimized Cascaded Network

Shamy Shabeer[1], J. Dheeba[2]
[1](Lecturer, Information Technology, Al Khobar Female College, Khobar)
[2](Associate Professor, CSE, College of Engineering Perumon, Kollam)

Abstract: *This paper presents a novel method for accurate early detection of breast cancer to reduce the mortality rate. Masses and micro calcification clusters are an important early signs of breast cancer. However, it is often difficult to distinguish abnormalities from normal breast tissues because of their subtle appearance and ambiguous margins. Computer aided diagnosis (CAD) helps the radiologist in detecting the abnormalities in an efficient way. This paper investigates a new classification approach for detection of breast abnormalities in digital mammograms using Particle Swarm Optimized cascaded coreelation Neural Network (SOCCN). The proposed abnormality detection algorithm is based on extracting Laws Texture Energy Measures from the mammograms and classifying the suspicious regions by applying a pattern classifier. The method is applied to real clinical database of 216 mammograms collected from mammogram screening centres. The detection performance of the CAD system is analysed using Receiver Operating Characteristic (ROC) curve.*

Keywords: *Breast cancer, cascaded correlation, laws, mammograms, particle swarm optimization*

I. Introduction

Breast cancer (BC) is the second leading cause of cancer death in women. Worldwide, breast cancer comprises 22.9% of all cancers in women in developed countries [1]. Breast tissues can belong to normal, benign or malignant. Benign stages are not considered cancerous, their cells are close to normal in appearance, they grow slowly and they do not invade nearby tissues or spread to other parts of the body. The benign stage conditions include fibrocystic changes, cysts, fibro adenomas, infections and trauma. Malignant stages are cancerous if left unchecked, these malignant cells eventually spread beyond the original cancer to other parts of the body. One proven way of reducing mortality from breast cancer is the screening of asymptotic women by mammography.

Mammography plays a major role in early detection of breast cancers, detecting about 75% of cancers at least a year before they can be felt [2]. It is estimated that 48 million mammograms are performed each year in US. Mammography is a special type of x-ray imaging used to create detailed images of the breast. Breast abnormalities that can indicate breast cancer are masses, calcifications, architectural distortion. A mass is defined as a space occupying lesion seen in at least two different projections.

Overall, screening mammograms miss up to 20 percent of breast cancers that are present at the time of screening [3,12]. False-negative results occur more often among younger women than among older women because younger women are more likely to have dense breasts. An automated system can overcome these problems by reducing the number of false positive and false negative readings from radiologists and increase the chance of detecting abnormalities early. This is a favorable prognosis for patients, as incorrect or late detections often result in mortality. As with many labor-intensive occupations, radiologists use computer-aided detection (CAD) systems that can identify potential cancers on mammography images. Recent studies have also shown that CAD systems, when used as an aid, have improved radiologists' accuracy of detection of breast cancer.

The studies indicate the importance in analyzing the problem and efforts done to improve the performance of the cancer detection in digital mammograms. Researchers are responsible to conceive new and improved analytical tools to solve a problem. The application of classifiers in medical diagnosis is increasing gradually. There is no doubt that evaluation of data taken from patients and decisions of medical experts are the most important factors in diagnosis. Classification systems can help in minimizing possible errors and also can provide instant examination of medical data in shorter time and in a more detailed manner.

This paper concentrates on developing a CAD system as artificial second radiologists. Texture helps to understand image content based on textural properties in images. Texture is the most important visual cue in identifying different types of homogeneous regions and gives information about the surface property, depth and orientation. This texture information helps to extract specific characteristics from a data. Mammographic images possess textural information that could bear discriminant features. Designing optimal neural network architecture for classification of abnormal and normal tissues requires a tedious trial and error process.

Especially automatic determination of artificial neural network parameters is the most critical task. Particle swarm optimization technique is used for optimizing the initial network parameters like hidden layer neurons, learning rate and momentum constant. This paper focuses mainly on designing a CAD system based on the optimized cascaded correlation neural network (CCNN) [4] evaluated using particle swarm optimization [] to improve the classification accuracy in breast cancer detection thereby reducing the misclassification rate.

II. Related works

A lot of researches in the area of CAD systems for breast cancer and developing intelligent techniques for improving classification accuracy have been conducted in last few decades. Different studies have demonstrated that Computer Aided Detection (CAD) of breast cancer can improve the detection rate from 4.7% to 19.5% compared to radiologists [13]. Detection and classification of microcalcification clusters and masses from mammograms play important roles in early diagnosis of breast cancer. The difficulty for the detection of abnormally is due to its small size, varying shapes, low contrast and unclear boundary from surrounding normal tissue, etc. Regarding classification of abnormalities in mammogram, a number of techniques have been presented using machine learning approaches to classify samples as normal and abnormal.

Anna Karahaliou et al [5] investigates multi-scale texture properties of the tissue surrounding microcalcifications (MCs) for breast cancer diagnosis using probabilistic neural network. Texture features are extracted from the original and multiscale representation of the tissue surrounding the MCs, aiming at capturing coarse and fine tissue alterations that may be associated with a malignant or a benign biological process. Gray-level first order statistics, gray-level co-occurrence matrices features, and Laws texture energy measures are extracted from original image. Wavelet coefficient first-order statistics and wavelet coefficient co-occurrence matrices features are extracted from Region of interest. The classifying power of these features is analyzed using probabilistic neural network.

Kupinski and Giger et al [6] presents a radial gradient index based algorithm and a probabilistic algorithm for detecting lesions in digital mammograms. These techniques are seeded segmentation algorithms; they begin with a point, called the seed point, which is defined to be within the suspect lesion. A series of partitions containing the seed point is created by thresholding, and rules determine which partition best segments the suspect lesion. Shape constraints are used to regularize the partitions analyzed, and simplifying the partition selection process by using utility functions based either on a single feature or probabilities.

Berkman Sahiner et al [7] used a Convolution Neural Network (CNN) classifier to classify the masses and the normal breast tissue. First, the Region of Interest (ROI) of the image is taken and it was subjected to averaging and subsampling. Second, gray level difference statistics (GLDS) and spatial gray level dependence (SGLD) features were computed from different subregions. The computed features were given as input to the CNN classifier.

Huai Li et al [8] presents a statistical model supported approach for enhanced segmentation and extraction of suspicious mass areas from mammographic images. A morphological operation is done to clean out the unwanted background region and model-based image segmentation is performed to localize the suspected mass areas using stochastic relaxation labeling scheme. Stochastic model-based image segmentation is a technique for partitioning an image into distinctive meaningful regions based on the statistical properties of both gray level and context images.

Naga Mudigonda et al [9] proposed a segmentation method for finding the suspected mass regions in mammograms. Prior to detection the mammograms are enhanced using Gaussian smoothing and subsampling operations. Oriented flow-like textural information in the adaptive ribbons of the pixels across the margins of the masses is taken as features. The adaptive ribbons of pixels thus extracted are used to classify the regions detected as masses or false positives at first and further to discriminate between benign masses and malignant tumors at a later stage.

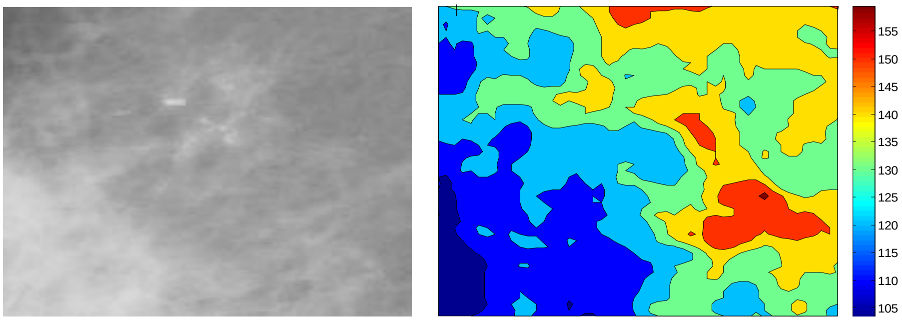
Nevine Eltonsy et al [10] presents a method based on the presence of concentric layers surrounding a focal area with suspicious morphological characteristics and low relative incidence in the breast region. The suspicious focal area is localized using the morphological features and based on the minimum distance criterion the redundant focal areas are eliminated. The presence of concentric layers around the suspected focal regions is analyzed using multiple concentric layer criterions to detect the suspicious regions.

Among existing CAD techniques, the main problem of developing an acceptable CAD system is inconsistent and low classification accuracy [9]. In order to improve the training process and accuracy, this research work investigates novel intelligent classifiers that use texture information as input to classify the normal and abnormal tissues in mammograms. Moreover, the intelligent machine learning classifiers are optimized using heuristic algorithms for finding appropriate hidden neurons, learning rate and momentum constant during the training process.

III. Database Description

A real time clinical study was taken to analyze screening mammograms of breast cancer patients. The goal of this study is to reduce the number of false positive rates which help to avoid unnecessary biopsies and emotional stress to many women. Women after the age of 40 are advised to take mammograms every year and hence the total number mammograms evaluated worldwide in one year may be in the order of millions. All clinical mammograms that were collected from screening clinics were positive for presence of cancer. Mammograms were collected from 54 patients and all these patients have agreed to have their mammograms to be used in research studies. For each patient 4 mammograms were taken in two different views, one is the Craniocaudal (CC) and the other is the Mediolateral Oblique (MLO) view. The two projections of each breast (right and left) were taken for every case. A total of 216 mammograms were taken, all the mammograms were digitized to a resolution of 290 x 290 Dots per Inch (DPI). The real clinical mammograms were digitized with a CADPRO *advantage* digitizer. Each digitized mammograms was incorporated into a 2020 x 2708 pixel image (5.47 Mpixels).

The abnormality detection in mammograms is difficult as breast contains a variable amount of connective, glandular and fatty tissue. In Fig. 1 the breast region is presented as an isoline, for visualizing the shape of the abnormality boundaries. Isolines are contour plots which forms closed loops and these loops help to identify clusters with high or low pixel value.



(a) (b)

Fig. 1 Representation of Abnormality in Mammogram using Isoline (a) Breast Region with Abnormalities (b) Isolines Showing Clusters of Breast Tissues

IV. Feature Extraction

The texture energy measures developed by Kenneth Ivan Laws at the University of Southern California have been used for many diverse applications. These texture features are used to extract Laws texture energy measures from the ROI containing abnormality and normal tissue patterns [10]. These measures are computed by first applying small convolution kernels to the ROI and then performing a windowing operation.

A set of nine 5 x 5 convolution masks is used to compute texture energy, which is then represented by a vector of nine numbers for each pixel of the image being analyzed. The 2-D convolution kernels for texture discrimination are generated from the following set of 1-D convolution kernels of length five. The texture descriptions used are level, edge, spot, wave and ripple.

$$
\begin{aligned}
L5 &= [\,1 \quad 4 \quad 6 \quad 4 \quad 1\,] \\
E5 &= [-1 \quad -2 \quad 0 \quad 2 \quad 1\,] \\
S5 &= [-1 \quad 0 \quad 2 \quad 0 \quad -1\,] \\
W5 &= [-1 \quad 2 \quad 0 \quad -2 \quad 1\,] \\
R5 &= [\,1 \quad -4 \quad 6 \quad -4 \quad 1\,]
\end{aligned}
$$

From this above 1-D convolution kernels 25 different two dimensional convolution kernels are generated by convoluting a vertical 1-D kernel with a horizontal 1-D kernel. Example for generating a 2-D mask from a 1-D is given below.

$$
\begin{bmatrix} -1 \\ -2 \\ 0 \\ 2 \\ 1 \end{bmatrix} \times \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \end{bmatrix} = \begin{bmatrix} -1 & -4 & -6 & -4 & -1 \\ -2 & -8 & -12 & -8 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 8 & 12 & 8 & 2 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}
$$

Similarly, 25 different two dimensional masks can be formed.

L5L5	E5L5	S5L5	W5L5	R5L5
L5E5	E5E5	S5E5	W5E5	R5E5
L5S5	E5S5	S5S5	W5S5	R5S5
L5W5	E5W5	S5W5	W5W5	R5W5
L5R5	E5R5	S5R5	W5R5	R5R5

The following steps will describe how texture energy measures are identified for each pixel in the ROI of a mammogram image.

Step 1: Apply the two dimensional mask to the preprocessed image i.e the ROI to get $F(i, j)$, where $F(i, j)$ is a set of 25 $N \times M$ features.

Step 2: To generate the LTEM at the pixel, a non-linear filter is applied to $F(i, j)$. The local neighbourhood of each pixel is taken and the absolute values of the neighbourhood pixels are summed together. A 15×15 square matrix is taken for doing this operation to smooth over the gaps between the texture edges and other micro-features. The non linear filter applied is,

$$E(x, y) = \sum_{j=-7}^{7} \sum_{i=-7}^{7} |F(x+i, y+j)|$$

By applying the above equation 25 energy features per pixel are obtained. The TEM images are represented as,

L5L5T	E5L5T	S5L5T	W5L5T	R5L5T
L5E5T	E5E5T	S5E5T	W5E5T	R5E5T
L5S5T	E5S5T	S5S5T	W5S5T	R5S5T
L5W5T	E5W5T	S5W5T	W5W5T	R5W5T
L5R5T	E5R5T	S5R5T	W5R5T	R5R5T

Step 3: The texture features obtained from step 2 is normalized for zero-mean.

V. PSO optimized CCNN

Cascaded Correlation Neural Network (CCNN) was developed by Fahlman and Libiere in the year 1990. Cascade correlation neural networks are "self organizing" networks and are similar to traditional networks in that the neuron is the most basic unit. Training the neurons however is novel. The CCNN is a supervised learning architecture that builds a near-minimal multilayer network topology in the course of training. Initially the network contains only inputs, output units, and the connections between them. This single layer of connections is trained using the Quickprop algorithm.

In an attempt to improve the classification accuracy of the WNN classifier PSO is used to tune the initial network parameters. Particle Swarm Optimization (PSO) was introduced by Kennedy and Eberhart [14] as a population based stochastic search and optimization process. PSO simulates the behaviour of bird flocking or fish schooling and used it to solve the optimization problems. In the basic PSO algorithm the system is initialized with a population of random solutions and searches for optima by updating positions and velocity. The potential solutions called particles fly through the problem space by following the current optimum particles. All of the particles have fitness values which are evaluated by the fitness function to be optimized, and have velocities which direct the flying of the particles. Each particle is updated after every iteration using two values $pbest$ and $gbest$. $pbest$ is the personal best value, which indicates the best solution achieved so far (i.e lowest fitness value) and the global best solution achieved so far by any particle in the population. In a *n*-dimensional search space, $\vec{X}_i = (X_{i1}, X_{i2}, ..., X_{in})$ and $\vec{V}_i = (V_{i1}, V_{i2}, ..., V_{in})$ are the positions and velocities respectively and they are updated for the d^{th} dimension of the i^{th} particle and is given by,

$$V_{id}(t+1) = V_{id}(t) + c_1.rand_1.(pbest_{id} - X_{id}(t)) + c_2.rand_2.(gbest_d - X_{id}(t)) \quad (1)$$

$$X_{id}(t+1) = X_{id}(t) + V_{id}(t+1) \quad (2)$$

c_1 and c_2 are the acceleration constants, $rand_1$ and $rand_2$ are the random numbers, $pbest_{id}$ is the individual's personal best i.e the local best solution found so far. $gbest_d$ is the neighbourhood's best solution found in the entire global community or in some neighbourhood of the current particle. $V_{id}(t)$ Velocity of individual at iteration t, $X_{id}(t)$:Position of individual at iteration t, $pbest_{id}$:Best position of individual until iteration t, $gbest_d$: Best position of the group until iteration t. The fitness function sought for optimal training is the Mean Square Error (MSE) formulated as,

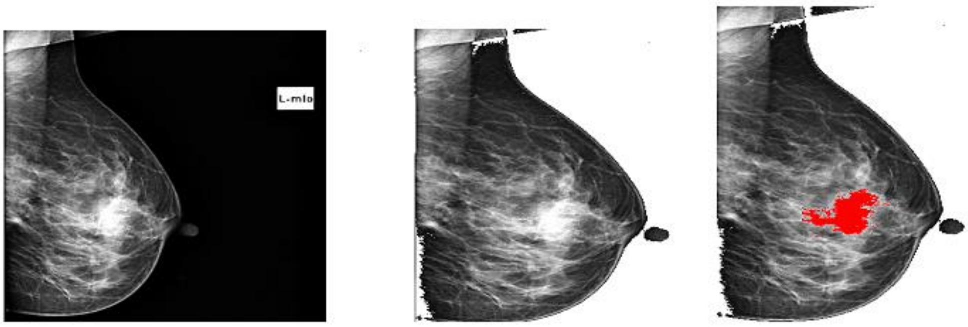
$$MSE = \sum_{p \in T} \sum_{k=1}^{N_o} (t_k^P - y_k^{p,o})^2 \tag{3}$$

where t_k^p is the target (desired) output, $y_k^{p,o}$ is the actual output from the k^{th} neuron in the output layer o, for the pattern p in the training set. With the framed fitness function the SOCCN algorithm automatically evolve a best solution.

VI. Classification Based On SOCCN

Real time clinical mammograms were collected from 54 patients and all these patients have agreed to have their mammograms to be used in research studies. For each patient 4 mammograms were taken in two different views, one is the CranioCaudal (CC) and the other is the Mediolateral Oblique (MLO) view. The two projections of each breast (right and left) were taken for every case. For this study a total of 216 mammograms were taken, all the mammograms were digitized. In this study, we focus on development of a CAD system for the detection of cancers with the help of different views (MLO and CC views) using mammographic images collected from mammogram screening centers.

Training the patterns in the real clinical database is done using SOCCN. A total of 1064 patterns were collected from the ROI to train the classifier. The network was designed with 25 input neurons and one output neuron. The hidden layer neurons is optimally selected and it uses a tansigmoid activation function. The output layer neuron uses linear activation function. The learning rate is set to 0.01 and momentum constant to 0.9. The training was stopped when the root mean square error per training was less than 0.1 or when it reaches 500 epochs. After training the patterns, the network was evaluated with the test cases. Figure 2 shows the detection results of the SOCCN classifier. Figure 2 (a) shows the original mammogram from real clinical database and the ROI image of a digital mammogram in Figure 2 (b). The detection results are shown in Figure 2 (c) in which abnormalities are marked with red circles. The SOCCN algorithm was trained with the same parameters used for analyzing the MIAS database. An optimized SOCCN is achieved with Nh =116, Lr = 0.00114 and Mc = 0.9238. Testing is done for all the 216 real time clinical images.



(a) **(b)** **(c)**

Fig. 2. Classification results of abnormalities using SOCCN approach. From top to bottom are the cases of circumscribed masses, Spiculated masses and microcalcification (a) Original Image (b) ROI image (c) Classified abnormalities

A comparative performance analysis based on ROC curve of various classifiers is shown in Fig. 3. The result emphasizes the potential of the SOCCN learning algorithm to be used as a breast cancer classifier. The classifier focuses the experimentation on trying to improve the classification rate by focussing on initial neural-network settings. Consequently, by using optimized parameter settings for wavelet neural networks the classification accuracy is improved drastically. A very high classification rate was achieved for the optimally tuned wavelet neural networks.

Table 1 Performance Measures for Optimally Tuned Classifier models

Classifiers/ Metrics	Performance Measures for Real Clinical Database	
	CCNN	**SOCCN**
Sensitivity (%)	90.756	90.843
Specificity (%)	79.487	86.732
Accuracy (%)	87.97	89.973
AUC	0.8966	0.9238
Misclassification Rate	0.1371	0.0824

The performance evaluation demonstrates that the result of the optimized wavelet neural network classifier is generally better than that of other well known classifiers. This is consistent with the fact that optimization is useful in initial parameter setting of the network. The result shows that the best area under the ROC curve is found to be 0.9138 for real clinical Database. The SOCCN approach generates a sensitivity of 90.843 % with a specificity of 86.732 %. The misclassification rate is found to be 0.0824 which is less when compared with the classification done by other classifiers. When optimized learning is introduced, there is an improvement in classification accuracy which is 89.973% . Therefore, SOCCN have a great potential to be applied in automatic detection of abnormalities in mammograms by reducing the misclassification rate.

Summarizing the results for real clinical mammogram the classification accuracy of SOCNN is higher than that of the other well known classifier models. This is because of the fact that the SOCNN incorporates the wavelet neural network and optimally designs the neural network using swarm intelligence algorithm. This superior performance makes SOCNN suitable for efficiently detecting abnormalities in mammograms.

VII. conclusion

The novel approach presented in this paper demonstrated that the SOCCN classifier produces an improvement in classification accuracy to the problem of computer-aided analysis of digital mammograms for breast cancer detection. The algorithm developed here classifies mammograms into normal & abnormal. First, the ROI of the image is chosen then laws texture features are extracted and classified using SOCCN. The optimized CCNN based classifiers using the properties of both wavelet and neural network provides good classification accuracy by reducing the false positives and false negatives. The SOCCN classifier designed using PSO algorithm applied to CCNN are investigated for detecting breast cancer in mammograms. The results give better classification accuracy than the traditional classifiers. The optimized wavelet neural network accelerates the convergence of the error back propagation algorithm and also it avoids major disruptions in the direction of learning.

References

[1] Freer, T.W. and Ulissey, M.J. "Screening Mammography with Computer-aided Detection: Prospective Study of 12,860 Patients in a Community Breast Center", *Radiology, Vol. 220, pp. 781-786, 2001.*

[2] Giger, M.L., Karssemeijer, N. and Armato, S.G. "Computer aided diagnosis in Medical Imaging'', *IEEE Transactions on Med. Imaging Vol. 20, No. 12, pp. 1205-1208, 2001.*

[3] Thurfjell, E.L., Lernevall, K.A. and Taube, A.A. "Benefit of Independent Double Reading in a Population based Mammography Screening Program", *Radiology, Vol. 191, No.1, pp. 241-244, 1994.*

[4] Dheeba, J. and Tamil Selvi, S. "Cascaded Correlation Neural Network based Microcalcification detection in Mammographic Images", *Advances in Intelligent and Soft Computing, Springer, Vol.132, pp.153-160, 2012.*

[5] Anna N. Karahaliou, Ioannis S. Boniatis, Spyros G. Skiadopoulos, Filippos N. Sakellaropoulos, Nikolaos S. Arikidis, Eleni A. Likaki, George S. Panayiotakis and Lena I. Costaridou, "Breast Cancer Diagnosis: Analyzing Texture of Tissue Surrounding Microcalcifications*", IEEE Transactions on Information Technology in Biomedicine, Vol. 12, No. 6, pp. 731-738, 2008.*

[6] Kupinski, M.A. and Giger, M.L. "Automated Seeded Lesion Segmentation on Digital Mammograms", *IEEE Transactions on Medical Imaging, Vol. 17, No. 4, pp. 510-517, 1998.*

[7] Berkman Sahiner., Heang-Ping Chan., Nicholas Pretrick., Datong Wei., Mark, A., Dorit, D. and Mitchell, M. "Classification of Mass and Normal Breast Tissue- A convolution Neural Network Classifier with Spatial Domain and Texture Images", *IEEE Transactions on Medical Imaging, Vol. 15, No. 5, pp. 598-609, 1996.*

[8] Huai Li, Yue Wang, K. J. Ray Liu, Shih-Chung B. Lo and Matthew T. Freedman, "Computerized Radiographic Mass Detection Part I: Lesion Site Selection by Morphological Enhancement and Contextual Segmentation", *IEEE Transactions on Medical Imaging, Vol. 20, No. 4, pp. 289-301, 2001.*

[9] Nico Karssemeijer, Johannes D. M. Otten, Andre L. M. Verbeek, Johanna H. Groenewoud, Harry J. de Koning, Jan H. C. L. Hendriks, Roland Holland, "Computer-aided Detection versus Independent Double Reading of Masses on Mammograms", *Radiology, pp. 192 – 200, 2003.*

[10] Christodoulos I. Christodoulou, Silas C. Michaelides and Constantinos S. Pattichis, "Multifeature Texture Analysis for the Classification of Clouds in Satellite Imagery", *IEEE Transactions on Geoscience and Remote Sensing, Vol. 41, No. 11, pp. 2662-2668, 2003.*

[11] Alain Tiedeu, Christian Daul, Aude Kentsop,PierreGraebling , Didier Wolf, "Texture-based analysis of clustered microcalcifications detected on mammograms", *Digital Signal Processing, vol. 22, 124–132, 2012.*

[12] American cancer society, cancer facts and fgures, American Cancer Society,Atlanta, Ga, 2005-2015.

[13] Chunxiao Chen, Gao Liu, Jing Wang, and Gail Sudlow, Shape-based automatic detection of pectoral muscle boundary in mammograms. *Journal of medical and biological engineering, 35, pp. 315-322, 2015.*

[14] Zengxi Huang, Yiguang Liu, Chunguang Li, Menglong Yang, Liping Chen A robust face and ear based multimodal biometric system using sparse representation *Pattern Recogn., vol. 46, pp. 2156–2168, 2013.*