Chapter 2

A Decision Support System (DSS) for Breast Cancer Detection Based on Invariant Feature Extraction, Classification, and Retrieval of Masses of Mammographic Images

Mahmudur Rahman and Nuh Alpaslan

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.81119

Abstract

This paper presents an integrated system for the breast cancer detection from mammograms based on automated mass detection, classification, and retrieval with a goal to support decision-making by retrieving and displaying the relevant past cases as well as predicting the images as benign or malignant. It is hypothesized that the proposed diagnostic aid would refresh the radiologist's mental memory to guide them to a precise diagnosis with concrete visualizations instead of only suggesting a second diagnosis like many other CAD systems. Towards achieving this goal, a Graph-Based Visual Saliency (GBVS) method is used for automatic mass detection, invariant features are extracted based on using Non-Subsampled Contourlet transform (NSCT) and eigenvalues of the Hessian matrix in a histogram of oriented gradients (HOG), and finally classification and retrieval are performed based on using Support Vector Machines (SVM) and Extreme Learning Machines (ELM), and a linear combination-based similarity fusion approach. The image retrieval and classification performances are evaluated and compared in the benchmark Digital Database for Screening Mammography (DDSM) of 2604 cases by using both the precision-recall and classification accuracies. Experimental results demonstrate the effectiveness of the proposed system and show the viability of a real-time clinical application.

Keywords: breast cancer, mammogram, NSCT, HOG, classification, diagnostic aid, computer-aided diagnosis (CAD), content-based image retrieval

1. Introduction

Breast cancer is considered a major health problem since it is the second most common cause of cancer among women only after lung cancer in both developed and developing countries [1, 2]. Based on the estimates of American cancer society (ACS), there were 231,840 new cases of invasive breast cancer and 40,290 breast cancer deaths among U.S. women in 2015 [3]. Although, the breast cancer mortality has declined among women of all ages during the last two decades due to the result of treatment improvements, earlier detection, and awareness, the incidence rate has increased significantly during this time [2, 3].

The diagnosis and detection of breast tumor in the early stages is the best opportunity to increase the chances of survival. Therefore, women of age 40 or older are recommended to get mammograms regularly. However, such a recommendation results in the generation of a very large number of mammograms that need to be processed. In addition, the interpretation of mammogram images mostly depends on the experience of the radiologists, and the tumors may be overlooked easily while viewing the image in early stages of breast cancer as the clinical indications are varied in appearance [4]. Screening mammograms is also a repetitive task that causes fatigue and eye strain since for every thousand cases analyzed by a radiologist, only 3–4 are cancerous and thus an abnormality may be overlooked [5]. It has been seen that between 60 and 90% of the biopsies of cancer cases predicted by radiologists found benign later and those biopsied women are exposed to needless fear and anxiety [6].

To support radiologists in the process of visually screening mammograms to avoid miss-diagnosis, computer-aided detection and/or diagnosis (CADe and CADx) systems have been proposed for analyzing digital mammogram due to the rapid advancement of digital imaging, computer vision, pattern recognition and machine learning technologies [7–9]. The CADe systems are responsible for highlighting or cueing the locations depicting the suspicious micro calcification clusters and masses and CADx systems deal with classifying classification between malignant and benign masses. Many methods have been proposed in the literature to assist radiologists in accurate interpretation of mammogram for detection of suspicious areas of micro-calcification clusters and breast masses often hidden in dense breast tissues and classification to benign and malignant lesions utilizing a wide variety of algorithms [10–15]. Previous studies have shown that using CAD improves radiologists' efficiency in searching for and detecting micro-calcification clusters as well as helps them detect more cancers associated with malignant micro-calcifications [10]. However, current CAD has no or little impact in helping radiologists detect more subtle cancers associated with mass-like abnormalities due to the relatively low performance in mass detection due to large variation in shape and size and are often indistinguishable from surrounding tissues [15].

Also, the majority of the research efforts in this domain has focused on the problem of the cancer detection, in which the likelihood of malignancy is computed based on some feature extraction and classification schemes [15–19]. These systems are non-interactive in nature and the prediction represents just a cue for the radiologist without the ability to explain the reasoning of the decision-making (the "black-box" type approach), as the final decision regarding the likelihood of the presence of a malignant mass is left exclusively to him/her. Hence, the clinical benefit of using current commercially available CAD systems is still under debate and test.

In the last several years, developing CAD schemes that use content-based image retrieval (CBIR) approach to search for the clinically relevant and visually similar mammograms (or regions) depicting suspicious lesions has also been attracting research interest [20–22]. CBIR-based CAD schemes have potential to provide radiologists with "visual aid" and increase their confidence in accepting CAD-cued results in the decision making. Furthermore, CBIR would be a useful aid in the training of students, residents, and less experienced radiologists since it would allow them to view images of lesions that appear similar, but may have differing pathology and help them see how the pattern in their current case closely resembles a pattern in cases previously proven to be non-cancerous, thereby, improving specificity. Although preliminary studies have suggested that using CBIR might improve radiologists' performance and/or increase their confidence in the decision making, this technology is still in the early development stage with lack of benchmark evaluation in ground-truth datasets. Despite of great research interest and the significant progress made in the last several years, developing CBIR approaches for breast cancer detection that can be routinely accepted and applied in the clinical practice is still not a reality.

To address the limitation of the current CAD in detecting mass-like abnormalities and due to the ongoing success of CBIR to provide clinical decision support for medical images of different modalities, we proposed to develop an integrated and interactive retrieval system. It will be able to respond to image based visual queries of automatically segmented suspicious mass region by displaying mammograms of relevant masses of past cases that are similar to the queried region as well as predicting the image categories (e.g., malignant, benign and normal masses). The performance and reliability of such a CBIR system depends on a number of factors (or issues) including the optimization of lesion segmentation, feature extraction, classification, similarity measures and relationship between the clinical relevance and visual similarity, database quality and sizes.

The most challenging problem in this task is detecting the mass from the background and extracting the discriminative local features of clinical importance. A Graph-Based Visual Saliency (GBVS) method [23] is utilized for the segmentation of the regions of interest (ROIs) as breast masses. To extract invariant features from masses, Non-Subsampled Contourlet transform (NSCT) [24] is utilized due to its powerful capability in image representation compared to wavelets and contourlet transform. Furthermore, in order to distinguish normal and abnormal tissues, eig(Hess) HOG [25] features are extracted based on computation of eigenvalues of the Hessian matrix in a histogram of oriented gradients in addition to several geometric features from the masses in mammographic images. HOG is known as a keypoint descriptor in literature which expresses the local statistics of the gradient orientations around a keypoint [46]. The HOG feature can express object appearance due to the reason that the histogram process gives translational invariance the gradient orientations are strong to lighting changes. It is also useful for the classification and retrieval of textured breast masses with different shapes.

For classification, two-class separation (normal and abnormal) and three-class study (normal, benign, and malignant cases) are carried out on the individual and combined input feature spaces by utilizing Support Vector Machines (SVM) and Extreme Learning Machines (ELM) with 10-fold cross validation. For retrieval, performances are evaluated and compared in different feature spaces in the benchmark DDSM dataset [26] using precision and recall curves

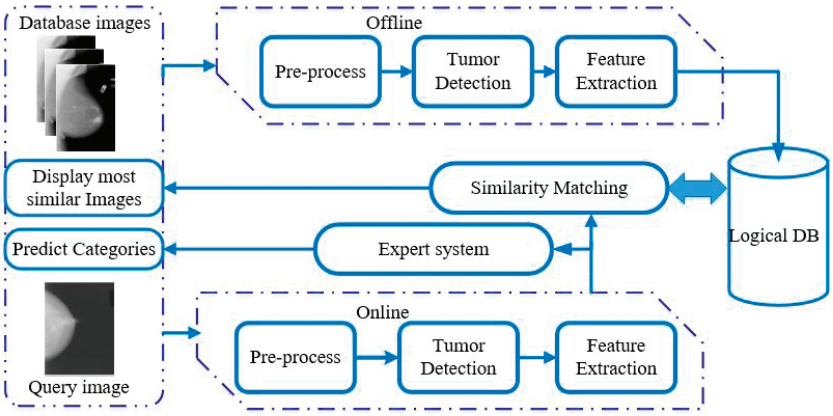

Figure 1. Dataflow diagram of the integrated decision support system (DSS).

obtained from comparison between the query and retrieved images. The system performance is compared with other state-of-the-art algorithms where experimental results indicate that the framework achieved a noticeable increase in recognition rates.

Figure 1 shows the dataflow diagram of the proposed integrated decision support system based on image pre-processing, mass segmentation, feature extraction, classification, and retrieval.

The rest of the paper is organized as follows. Section 2 describes the related works, specially talk about the CAD-based CBIR systems for mammographic mass retrieval. The mass detection based on marker-controlled watershed segmentation and feature extraction are described in Sections 3 and 4 respectively. Our classification and similarity matching methods are described in Section 5, while all discussions on the obtained experimental results are given in Section 6. The last section comprises of conclusions.

2. Background review

There is a clear need to create effective tools and techniques to search, browse and retrieve images from large repositories to aid diagnoses and research due to the phenomenal growth in recent years in the volume of digital mammograms produced in hospitals and clinical centers. Due to the freely available access to datasets of digital mammograms, such as the Digital Database for Screening Mammography (DDSM), interest in developing CAD schemes for mammograms that use CBIR has been attracting continued research interest during the last several years [20–22]. Although mammography-based CAD is one of the mature and widely adopted fields, there have been only a limited number of studies devoted to CBIR-based CAD systems for the detection and retrieval of breast masses in mammograms. Alto et al. [21] proposed the use of the shape, gradient, and texture features for mammography image retrieval

and that was one of the earliest researches on CBIR for mammograms. Linear discriminant analysis, logistic regression, and the Mahalanobis distance were used to evaluate the features for classifying the masses. Kinoshita et al. [22] used the breast density to retrieve images from a mammogram dataset available at the Clinical Hospital of the University of São Paulo at Ribeirão Preto, Brazil. Shape, texture features, moments, Radon transform, and histograms were used to describe breast masses, and the Kohonen self-organizing map (SOM) neural network was used for image retrieval. Wang et al. [27] has utilized histograms for the characterization of breast mass in a set of mammogram database at the Medical Center of Pittsburgh in order to automatically evaluate breast mass. They obtained 71% of correct classification rate with the use of a neural network. Muramatsu et al. [28] proposed a psychophysical similarity measure based on neural networks for evaluation of similar images with mammographic masses. The major drawback is that a large amount of data is required to train an artificial neural network (ANN). Oliveira et al. proposed a CBIR system called MammoSys; the novelty of this study is to present a two-dimensional principal component analysis (2DPCA) method [29] for the description of mass texture and thereby also a dimensionality reduction is performed. Wei et al. [30] proposed an adaptive classification scheme in the context of SVM assisted by content-based image retrieval to improve the classification accuracy in the computer aided diagnosis for breast cancer. A CBIR scheme is proposed in [31] that utilizes SVMs capable of optimally exploiting the distribution of input samples in the feature space on the basis of BI-RADS classifications of masses as carried out by the radiologists. In an article by Zhang [20], a number of CBIR-based CAD schemes for mammograms were compared and their performance were assessed and it was concluded that much research work is needed before the CBIR-based CAD schemes can be accepted in the clinical practice.

3. Breast mass detection

The most challenging aspect in developing any CAD based systems for mammograms is to segment the suspicious masses, which are often hidden in dense breast tissues. Since a cancerous region might typically represented by local-oriented patterns, accurately segmenting it is an important first step for the effective performances of the successive feature extraction, similarity matching and classification steps in developing a CAD system as shown in **Figure 1**. A large number of segmentation methods have been proposed in the literature for the detection of breast masses, such as adaptive region growth [32], multi-layer topographic region growth algorithm [33], active contour (snake) modeling [34], level set algorithm [35], dynamic programming [36] etc. However, due to the limitation of benchmark evaluation and testing datasets to compare the performances, it is difficult to find the most robust and effective method in this domain till now.

3.1. Visual saliency based segmentation

The breast anatomy has a complicated structure because of the presence of pectoral muscles and the different mass density. Although it is easy to analyze breast tissues without getting

confused with pectoral muscles for a radiologist, it is always difficult to distinguish between pectoral muscles and mass for an automatic method in a CAD system. For that reason, pectoral muscles are removed usually before the segmentation, which has a huge limitation as it is done manually in most cases [8, 15, 19]. However, automatic segmentation of pectoral muscle is a troublesome process and also an additional workload in analysis of mammography images in cranio-caudal (CC) view, which are generally without pectoral muscles.

In this study, for example a graph-based visual saliency (GBVS) method [37] is utilized for segmentation by applying thresholding on the saliency map, which does not require the removal of pectoral muscles to detect the breast masses. The reason for choosing GBVS is that it has the ability of generating an output showing concentrated saliency maps in the appropriate image regions where the value of an image pixel location corresponds to the saliency of that pixel with respect to the neighbors. The usefulness of saliency models in cases where some structures are implicit with respect to the image such as pectoral muscles in mammograms is demonstrated in [19]. It is also experimentally shown that the GBVS yields the best results for mass detection from screening mammograms [37].

The GBVS calculates the saliency of a region with respect to its local neighborhood using the directional contrast. In mammography images, it has been monitored that the contrast of mass containing regions is significantly different from the remaining breast tissue. As discussed earlier, the mass encircled by dense tissues is difficult to recognize, whereas, the directional contrast with respect to the local neighborhood helps in identifying such masses along with the masses present in fatty regions. The computation of saliency map consists of following stages: Firstly, to differentiate mass from the neighboring regions in contrast, feature maps are computed from contrast values along four different orientations of 2D Gabor filters (0°, 45°, 90°, and 135°). Then, activation maps are computed as the balanced distribution of a Markov chain which is obtained using the initial feature maps [20]. The balanced distribution denotes higher weights only for the edges present in salient regions. The Ergodic Markov chains are modeled on a fully connected directed graph obtained from feature maps. Weighted connections are used to create the graph. It is created by connecting nodes in a feature map. The directed edge node (i, j) to node (k, l) weight is assigned in the graph.

$$w\big((i,j),(k,l)\big) \overset{\mathrm{A}}{=} D \cdot F(i-k, j-l), \tag{1}$$

where $F(a,b) \overset{\mathrm{A}}{=} \exp\left(-\frac{a^2+b^2}{2\sigma^2}\right)$

$$\mathrm{D} \overset{\mathrm{A}}{=} \left|\log\left(\frac{M(i,j)}{M(p,q)}\right)\right| \tag{2}$$

where $M(i, j)$ denotes a node in the feature map and σ is set to 0.15 times the image width.

Due to the fact that activation maps lack the accumulation of weights, a normalization of activation map is performed to avoid uniform saliency maps. Activation maps are normalized using a similar approach as used in the previous step. Markov chains are computed from the

activation maps. The function D in Eq. (2) maps to the value at location (k, l) in activation map (Eq. (5)). The value of the parameter σ in Eq. (1) is 0.06 times the image width [38].

$$D \overset{A}{=} A(p, q) \tag{3}$$

where $A(p, q)$ represents a node in the activation map. In final stage, normalized activation maps are combined using the sum rule to obtain the saliency map. Once the saliency map

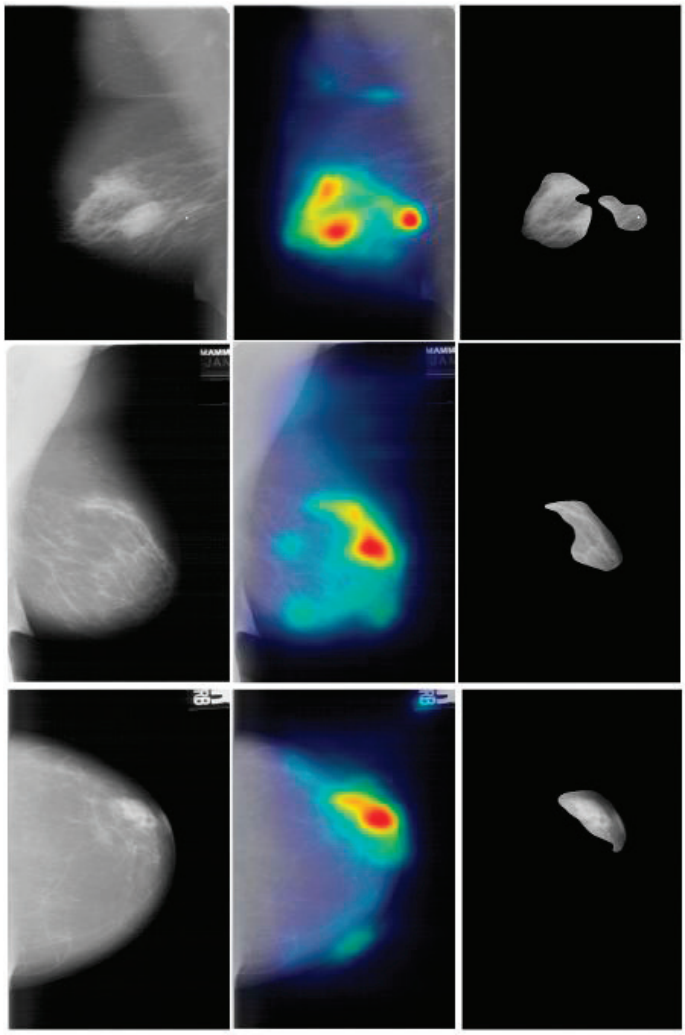

Figure 2. Saliency map generated from enhanced image and suspicious regions obtained after thresholding.

is computed, a threshold is empirically selected to obtain the optimal size ROIs. **Figure 2** illustrates some examples of saliency maps generated from pre-processed images and ROI segmentation from the saliency maps.

4. Feature extraction

Feature extraction plays key role in our breast cancer diagnosis framework. Feature extraction process can only be carried out if the suspicious areas of breast masses are appropriately defined. In selecting effective features from mammogram lesions, great research efforts have been focused on capturing the texture of images and improving correlation to the human visual similarity. Among them, Curvelet transform, Gabor Wavelet, Discrete Wavelet Transform (DWT), and Spherical Wavelet Transform (SWT), Contourlet Transform (CT), local binary pattern (LBP) have been extensively investigated and compared in addition to other popular texture features derived from the co-occurrence matrices and Fourier transformation [39–44]. Since, clinically and visually similar lesions or disease patterns can depict on different locations of the mammograms with different orientations, the selected features should be invariant to the linear shift and rotation of the targeted lesions. To consider these criteria, NSCT and HOG based approaches are used for feature extraction in addition to traditional shape, mass and GLCM based features from region-of-interests (ROIs) with adaptively adjusted size based on the actual mass region segmentation results.

4.1. Non-subsampled contourlet transform

Despite different applications of Wavelet Transform in medical image analysis, it has some limitations in capturing the directional information in images such as smooth contours and the directional edges. For example, orthogonal wavelets consider only horizontally, vertically, and diagonally directed discontinuities. These directions do not express effectively the edges and textures of medical images such as breast mammograms, which have smooth curves that represent benign, malignant masses and micro-calcifications, etc. To express the contour-like smooth edges directly in the discrete domain effectively, the Contourlet Transform was introduced by Do and Vetterli [45], which is an extension of the wavelet transform which uses multi scale transform that is constructed by combining the Laplacian pyramid with directional filter banks (DFB) and has additional characteristics such as directionality and anisotropy in addition to the properties of Wavelet Transform. Although the Contourlet Transform is a more effective method than the Wavelet Transform in image representation, it is not shift-invariant due to down-sampling and up-sampling. Non-Subsampled Contourlet Transform (NSCT) was proposed by Cunha et al. [21] to compensate this limitation and due to its beneficial features, the NSCT has been used in this work for representing the breast masses according to its features.

In NSCT, to keep away from the frequency aliasing of the CT and to obtain the shift-invariance, the non-subsampled Laplacian Pyramids (NSLP) and the non-subsampled directional filter banks (NSDFB) is utilized based on Idealized frequency partitioning obtained with

the structure proposed in [21]. Additionally, the multi-scale and directional decomposition processes are free from each other. The number of decomposition directions is changeable and can be adjusted to any value of 2^{l_j} where the l^j parameter can be represented at scale j, $1 \leqslant j \leqslant J$ and J represents the number of decomposition scales. Different from the classical CT, all subbands of NSCT have the same resolution. That means, the NSCT coefficients of each subband are in one-to-one correspondence with the original surface in the spatial domain. For feature extraction, a combination of k mean, variance, energy, entropy, skewness, and kurtosis parameters from 4-level non-sampled contourlet transform is examined. An example of the NSCT for a mass is shown in **Figure 3**. The image is decomposed into four pyramidal levels, resulting in one, two and eight sub-bands.

4.2. Eig(Hess)-HOG features

The HOG is computed for each key point from a block. The key point denotes the center of the central cell of the block. The adjacent area of each key point is partitioned into cells. One-dimensional histogram of gradient orientations is accumulated for each cell. The histogram of all the cells generates the feature of all key points [22, 46]. A simple 1-D [−1; 0; 1] mask is used

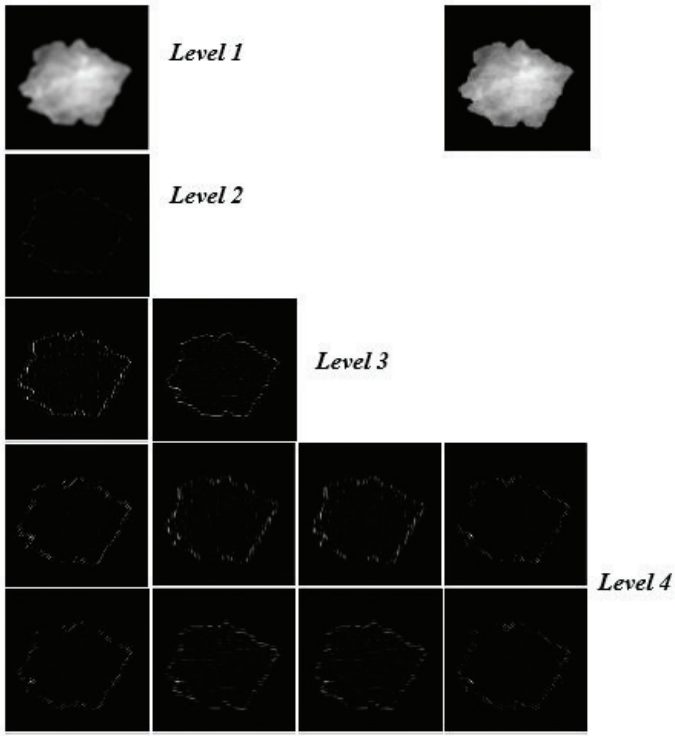

Figure 3. Non-sampled Contourlet transform of an ROI with four-level decomposition.

for the gradient computation. In conventional HOG, firstly the grayscale image is filtered with mask to obtain x and y derivatives of image as in Eq. (4).

$$\begin{aligned} f_x(x,y) &= I(x+1,y) - I(x-1,y) \quad \forall x,y \\ f_y(x,y) &= I(x,y+1) - I(x,y-1) \quad \forall x,y \end{aligned} \tag{4}$$

where f_x and f_y indicates x and y derivatives of image gradient. $I(x,y)$ indicates the intensity at position(x,y). The magnitude and orientation is calculated as in Eq. (5) and (6);

$$m(x,y) = \sqrt{f_x(x,y)^2 + f_y(x,y)^2} \tag{5}$$

$$\theta(x,y) = \tan^{-1}\left(\frac{f_y(x,y)}{f_x(x,y)}\right) \tag{6}$$

The gradient orientations are partitioned into eight bins. For each pixel's orientation bin, the orientation's magnitude $m(x,y)$ is voted to each bin. Then, orientation histogram of every cell and spatial blocks are normalized.

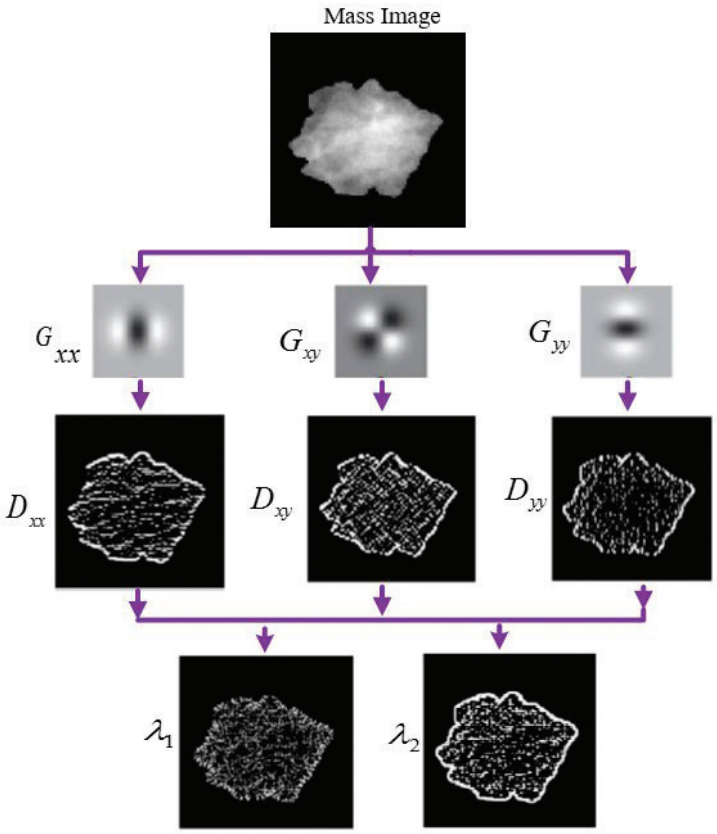

Figure 4. Computation of eigenvalues of the hessian matrix for an ROI as breast mass.

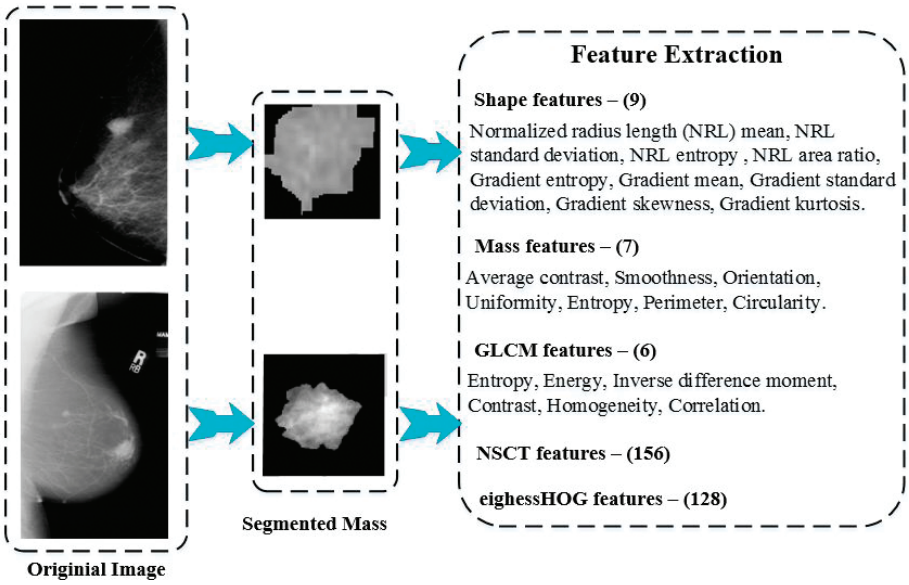

Figure 5. Feature extraction stage with an example mammogram.

Eig(Hess)-HOG uses the Hessian matrix instead of the Gaussian derivative filters to compute the eigenvalues of image surface. The Hessian matrix contains more differential information than the conventional gradient. The Hessian matrix of an image is defined as the second-order partial derivative matrix of the gray scale image. The second order differentials provide more accurate analysis in detail about function curves in breast masses [22].

The number of possible orientation bins is referred to as N_o and the number of cells for direction is referred to as N_c. We set the parameters of the descriptor to $N_c = 4$ cells and $N_o = 8$ bins resulting in a total of $N_o \times N_c^2 = 128$ elements in a HOG feature. **Figure 4** demonstrates the computation of eigenvalues of the Hessian matrix for mass ROI.

In addition as shown in **Figure 5**, a 9-D shape feature and a 7-D mass feature are extracted which representing the mass boundary and the average contrast, smoothness, orientation, uniformity, entropy, perimeter and circularity [17]. Finally, a 6-D texture feature representing the energy, correlation, entropy, inverse difference moment, contrast and homogeneity is obtained from the gray level co-occurrence matrix (GLCM).

5. Classification and similarity matching

For classification of breast masses as either normal and abnormal (two-class separation) or normal, benign, and malignant cases (three-class study), we used SVM and ELM classifiers for excellent generalization performance and little human intervention. The SVM carries out classification between two classes by determining a hyper plane in feature space that is based on the most informative points of the training set [47]. On the other hand, ELM is a single-hidden

layer feed-forward neural network (SLFNs) learning algorithm [48]. It first randomly assigns weights and biases for hidden nodes, and then analytically defines the output weights by using the least square method. Due to the random selection of weights and biases for hidden nodes, the ELM can decrease the learning time considerably and also can achieve superior generalization performance [49].

For similarity matching, it is challenging to find a unique feature representation to compare images accurately for all types of queries. Feature descriptors at different levels of image representation are in diverse forms and may be complementary in nature. The difference between the feature vector of queried mass (or ROI) and the feature vectors of reference images (or ROIs) is calculated to compute the similarity between the query image and the database. Current CAD schemes using CBIR approaches typically use the k-nearest neighbor type searching method which involves searching for the k most similar reference ROIs to the queried ROI. The smaller the difference ("distance"), the higher the computed "similarity" level is between the two compared ROIs. The searching and retrieving result of the CBIR algorithm depends on the effectiveness of the distance metrics to measure the "similarity" level among the selected images.

In this work, a fusion-based linear combination (Eq. (7)) scheme of similarity measure of different features is used with pre-determined weights. The similarity between a query image Iq and target image I_j is described as:

$$\mathrm{Sim}\left(I_q, I_j\right) = \sum_F \alpha^F S^F\left(I_q, I_j\right) = \sum_F \alpha^F S\left(f_q^F, f_j^F\right) \tag{7}$$

where $F \in \{NSCT, HOG, Shape, Mass, and GLCM\}$ and S^F (Iq, Ij) are the Euclidean similarity matching function in individual feature spaces and α^F are weights (determined experimentally) within the different image representation schemes.

6. Result evaluation

To evaluate the effectiveness of the proposed classification and retrieval-based decision support system, the experiments are performed on mammographic digitized images taken from the Digital Database for Screening Mammography (DDSM), a collaboratively maintained public dataset at the University of South Florida [23]. The DDSM database has been widely used as a benchmark for numerous articles on the mammographic area, for being free of charge and having a diverse quantity of cases. The database contains approximately 2500 cases where each case includes two image view anatomy (CC and MLO) of each breast (right and left). The size of the images varies from 1024 × 300 pixels to 1024 × 800 pixels. The DDSM database offers more than 9000 images and from where we selected a total of 5880 images for experiments and result evaluation.

6.1. Experiment design

To experiment with the classification systems, the entire collection of mammograms is divided where 40% of the images are chosen as the training set and the remaining 60% as the test set, and a 10-fold Cross Validation (CV) has been used in the experimental design. The SVM learning approach was examined with the Gaussian radial basis function (GRBF) (r = 2, C = 100). The

overall performance of a classifier is guaranteed by a 10-fold CV in all evaluation indices. The performance of ELM classifier depends on the selection of number of neurons in hidden layer L, which was determined as L = 700 by trials with increments within the range of 100–1000. It was found that both the training and testing errors were decreased when L increased to around 700 and after that the training and testing performance did not improve and kept almost fixed as shown in **Figure 6**. We also tested with different activation functions, such as sigmoid, tangent sigmoid, sin, and radial basis and the tangent-sigmoid found to be the optimal one.

6.2. Performance evaluation

For the performance evaluation of the proposed classification approaches in different feature spaces, we computed the sensitivity (true positive rate) and specificity (true negative rate) for each of the confusion matrices. The accuracy, sensitivity and specificity parameters are employed for the performance evaluation of our classification approaches. The specificity measures the percentage of positive instances that were predicted as positives, while sensitivity measures the percentage of negative instances that were predicted as negatives. The retrieval effectiveness is measured with the precision-recall (PR) graphs that are commonly used in the information retrieval domain. For the experiments, each image in the testing dataset is served as a query image. A retrieved image is considered to be a correct match if it belongs to the same category to which the query image belongs. The performances of the two image categories (e.g., normal and abnormal) and the three image categories (e.g., benign, malignant, normal) are compared, based on the PR graphs.

Finally, for both classification and retrieval evaluation, different combination of concatenated feature vectors are utilized as shown in **Table 1**. For example, the f_1 feature set consists of all five different features: Shape, Mass, GLCM, NSCT, eig(Hess)HOG features, whereas f_6 feature set consists of eig(Hess)HOG feature only.

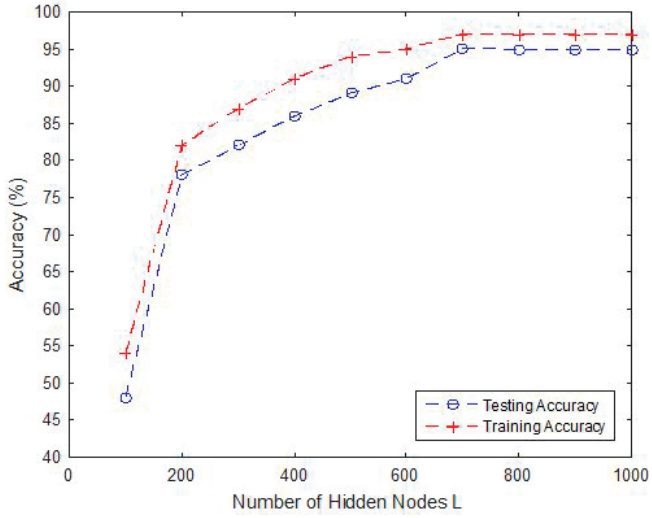

Figure 6. The number of hidden layer nodes L determination with L = 700.

6.3. Classification results

As mentioned, the classification accuracies in different feature sets of **Table 1** are compared and with both SVM and ELM classifier and it was found out that the f_1 feature set with ELM classifier is the most effective feature set in terms of accuracy, sensitivity, specificity both two and three class study. For example, **Tables 2** and **3** demonstrate the results for both 2- and 3- class studies for the f_1 and f_3 feature sets respectively.

From **Tables 2** and **3**, it can be observed that f_1 feature set is more effective than f_3 feature set in terms of accuracy, sensitivity, specificity both two and three class study. In fact, f_1 feature set with ELM as classifier achieved the highest classification accuracy rate in terms of mean accuracy, sensitivity and specificity parameters after 10-fold CV as shown in **Table 4**.

The classification efficiency, training and testing the performances of SVM and ELM were compared independently. As shown in **Table 4**, it can be seen that the proposed methods are effective on the DDSM database and ELM is a highly-effective classification technique

f	Feature set
f_1	Shape, Mass, GLCM, NSCT, eig(Hess)HOG
f_2	Shape, Mass, GLCM, NSCT
f_3	Shape, Mass, GLCM, eig(Hess)HOG
f_4	NSCT
f_5	NSCT, eig(Hess)HOG
f_6	eig(Hess)HOG

Table 1. Utilized different types of feature sets.

2 class				3 class		
Fold #	**Accuracy**	**Sensitivity**	**Specificity**	**Accuracy**	**Sensitivity**	**Specificity**
Fold 1	98.0167	96.5553	95.0939	86.1817	92.8298	92.7950
Fold 2	97.7035	95.8768	94.0501	94.1176	92.5861	93.5955
Fold 3	96.7641	95.5637	94.3633	93.5607	92.7602	92.5513
Fold 4	97.7035	95.9290	94.1545	93.1778	97.7035	98.9039
Fold 5	96.9729	94.6159	94.2589	97.7035	98.4342	98.0689
Fold 6	97.4948	95.9812	94.4676	98.7474	98.5386	98.6952
Fold 7	96.9729	95.8768	94.7808	98.5908	98.2255	91.3271
Fold 8	97.3904	96.2422	95.0939	92.6854	94.3574	95.7158
Fold 9	97.4948	97.0856	94.6764	94.1484	92.4765	92.8945
Fold 10	97.3904	95.6681	93.9457	93.4169	93.9394	93.8349
Ave	96.9729	95.5480	94.1232	92.4155	98.3612	93.4796

Table 2. Classification performance of f_1 feature set for the 2 and 3 class study.

2 class				3 class		
Fold #	**Accuracy**	**Sensitivity**	**Specificity**	**Accuracy**	**Sensitivity**	**Specificity**
Fold 1	95.0417	95.3549	95.0417	91.6464	89.7668	92.3773
Fold 2	94.7807	95.9812	95.1983	91.1939	91.3679	91.1939
Fold 3	95.2505	95.0939	94.9895	91.2287	92.5513	91.6812
Fold 4	95.4592	97.1816	96.2421	92.1685	98.8512	99.0078
Fold 5	95.9290	96.2421	97.7035	98.6945	99.3212	98.8512
Fold 6	96.2421	94.5720	96.7640	98.1723	98.4334	98.0157
Fold 7	98.3298	98.0167	92.9018	98.2768	99.1645	92.0668
Fold 8	94.4676	94.1544	93.3194	93.3194	92.6931	91.2317
Fold 9	94.2588	94.1544	95.9290	91.1274	92.0668	91.3361
Fold 10	93.4238	91.6492	92.9018	94.1545	94.1545	92.2756
Ave	95.2192	96.7223	93.7161	91.5176	98.6789	92.4426

Table 3. Classification performance of f_3 feature set for the 2 and 3 class study.

		Measures	f_1	f_2	f_3	f_4	f_5	f_6
2 class	ELM	Mean accuracy	96.973	95.709	95.219	93.742	92.228	94.665
		Mean sensitivity	95.548	97.223	96.722	96.743	97.421	99.853
		Mean specificity	94.123	94.196	93.716	90.741	87.035	89.478
	SVM	Mean accuracy	94.086	92.015	94.427	91.536	90.341	91.852
		Mean sensitivity	93.824	95.357	93.118	94.124	94.117	96.741
		Mean specificity	91.216	92.628	90.240	88.687	84.896	86.224
3 class	ELM	Mean accuracy	92.415	92.318	91.517	79.547	78.092	85.788
		Mean sensitivity	98.361	98.877	98.678	98.642	97.409	99.357
		Mean specificity	93.479	91.691	92.442	83.016	83.528	93.152
	SVM	Mean accuracy	90.125	89.654	89.884	76.004	74.011	83.256
		Mean sensitivity	95.244	94.519	95.331	94.875	92.546	95.446
		Mean specificity	89.886	87.359	90.214	79.954	80.820	89.744

Table 4. Performance comparisons of SVM and ELM classifiers using different features.

for this task. It is also proved to be highly efficient as lesser computational time was required compared to SVM with same sets but of different training data sizes.

Table 5 shows the comparison between our proposed system and a number of state-of-the-art classification systems. For the DDSM database, our classification system obtained comparable performance in accuracy, sensitivity and specificity. The promising results might be owing to the good segmentation algorithm as well as the effective feature extraction methods.

6.4. Retrieval results

A precision–recall curve based on our similarity fusion approach of different feature sets (f_1–f_6) is shown in **Figure 7**, which represents that f_1 feature set outperform the all other feature sets in terms of precision and recall. **Figure 8** shows that, when f_1 feature set are used, the cancer masses are the most discriminative among three types of masses.

The results verify that the characteristic of breast masses was better represented by the f_1 feature set, which was able to capture the difference between the gray level intensities of the breast densities. Concerning f_1 feature set, for a 4% of recall, a precision of 96% means that from 5880 mammogram images returned by the our proposed CBIR system, 5644 were relevant.

Literature	Database	Classifier	Accuracy (%)	Sensitivity	Specificity
Wang et al. [16]	DDSM	SVM	92.74	—	—
Liu and Tang [17]	DDSM	SVM	93.00	92.00	92.00
Dong et al. [18]	DDSM	GA-SVM	93.24	94.78	91.76
Jen and Yu [19]	DDSM	ADC	—	86.0	84.0
Wei et al. [30]	DDSM	ELM	95.73	—	—
Rouhi et al. [50]	DDSM	MLP	95.01	96.25	93.78
Jiao et al. [51]	DDSM	CNN	96.7	—	—
Proposed scheme	DDSM	ELM	96.973	95.548	94.123

Table 5. Classification performance of different methods.

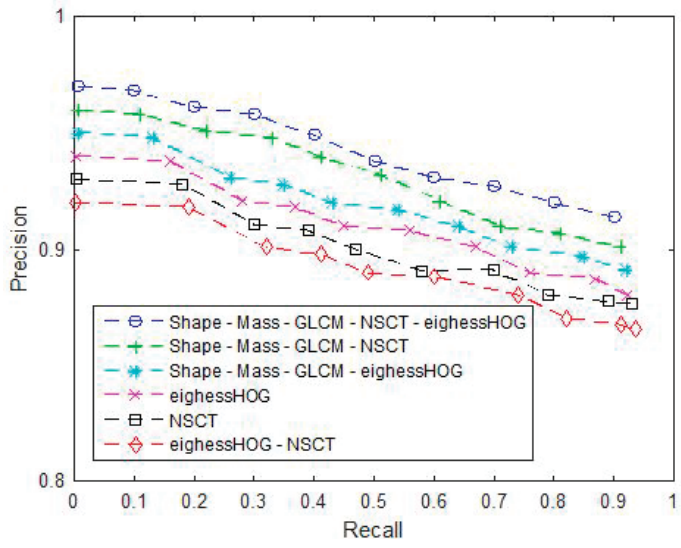

Figure 7. Performance evaluation of six feature sets on similar mass retrieval.

Figure 9 shows an example of a query image from benign category and retrieved images based on the f_1 feature set. The system retrieved all the top eight images from the same category of the query image and from the same direction (right). However, the views of top retrieved image are different where images 2, 3, 4 and 5 (left to right and top to bottom) are from MLO view while the others are from CC view. Another example of retrieval is demonstrated in

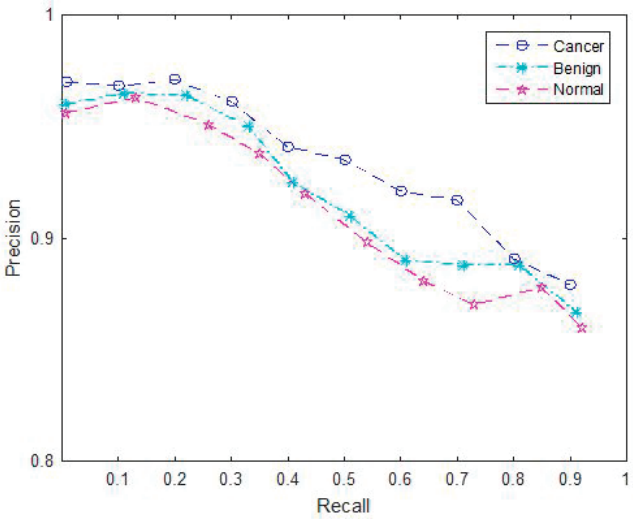

Figure 8. Performance evaluation of f_1 feature set on three different mass types of mass lesions.

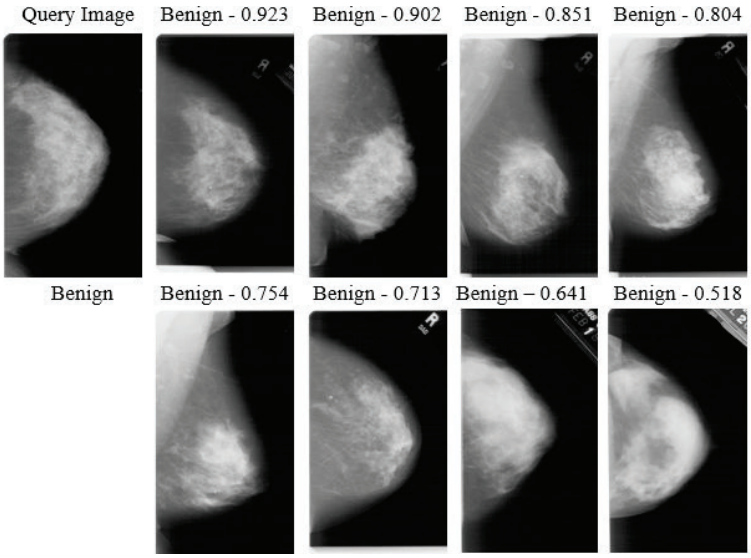

Figure 9. Retrieval example of one benign query of our proposed system using the f_1 feature set.

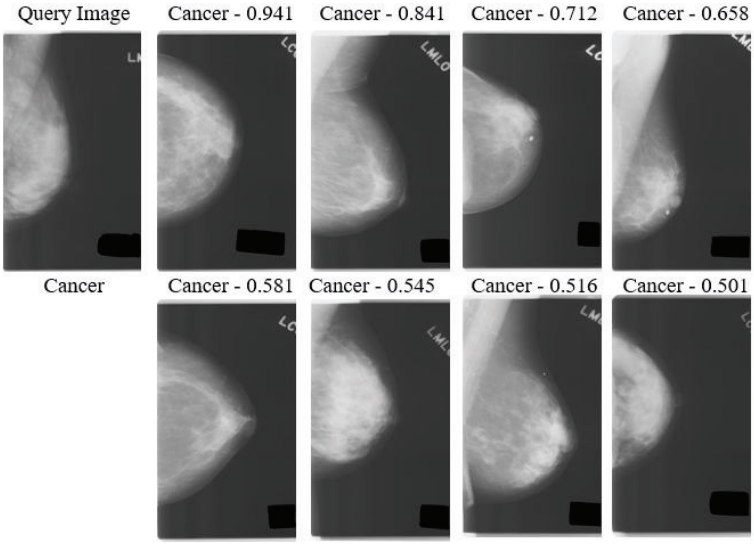

Figure 10. Retrieval example of one cancer query of our proposed system using the f_1 feature set.

Figure 10, using an image of cancer category. All the retrieved images are from the same category of the query image and from the same direction (right); however from both MLO (images, 2, 4, 7) and CC views. This might occur due to the fact that the ROI selected from these MLO images contains a good portion of pectoral muscle that was confused with the white part of the breast density.

The proposed system was implemented using the MATLAB through its image processing toolbox. Feature extraction and image retrieval were performed on an Intel i7 2.9 GHz processor with 8 GB of RAM under Microsoft Windows operating system.

7. Conclusions

In this paper, an integrated decision support system is proposed for the automated mass detection, classification and retrieval of mammograms. The system is evaluated for the retrieval and classification of the mammographic images. The experimental results indicate that the approach is effective to retrieve visually similar lesions from a database and to predict the categories of images for diagnostic correctness. The main objective of this paper is to demonstrate how the image retrieval and classification can be integrated and effectively utilized as a diagnostic support tool to help the radiologist for the mass detection. However, it is recognized that many other advanced image-based features and features from other sources would be necessary for a complete decision support system. In future, we plan to incorporate more advanced features related to the diagnostic relevance into our system and experiment with other classification and combination techniques as well. However, the presence of an

expert radiologist is still considered necessary for the overall visual assessment of the breast mass and the final diagnosis, based on the objective evaluation suggested by the system and contextual information from the patient data.

Author details

Mahmudur Rahman* and Nuh Alpaslan

*Address all correspondence to: md.rahman@morgan.edu

Department of Computer Science, Morgan State University, Baltimore, MD, USA

References

[1] Na. American Cancer Society published second edition of global cancer atlas. Oncology Times. 2015;**37**:34

[2] Al Mousa DS et al. Mammographic density and cancer detection. Academic Radiology. 2014;**21**:1377-1385

[3] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. CA: A Cancer Journal for Clinicians. 2015;**65**:5-29

[4] Moura DC, Guevara López MA. An evaluation of image descriptors combined with clinical data for breast cancer diagnosis. International Journal of Computer Assisted Radiology and Surgery. 2013;**8**(4):561-574

[5] Jotwani A, Gralow J. Early detection of breast cancer. Molecular Diagnosis & Therapy. 2009;**13**(6):349-357

[6] Demircioğlu Ö, Uluer M, Arıbal E. How many of the biopsy decisions taken at inexperienced breast radiology units were correct? The Journal of Breast Health. 2017;**13**(1):23-26. DOI: 10.5152/tjbh.2016.2962

[7] Ganesan K, Acharya UR, Chua CK, Min LC, Abraham KT, Ng K-H. Computer-aided breast cancer detection using mammograms: A review. IEEE Reviews in Biomedical Engineering. 2013;**6**:77-98

[8] Rangayyan RM, Ayres FJ, Leo Desautels JE. A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs. Journal of the Franklin Institute. 2007;**344**:312-348

[9] De Santo M, Molinara M, Tortorella F, Vento M. Automatic classification of clustered microcalcifications by a multiple expert system. Pattern Recognition. 2003;**36**:1467-1477

[10] Zhang Y, Tomuro N, Furst J, Raicu DS. Building an ensemble system for diagnosing masses in mammograms. International Journal of Computer Assisted Radiology and Surgery. 2011;**7**:323-329

[11] Buciu I, Gacsadi A. Directional features for automatic tumor classification of mammogram images. Biomedical Signal Processing and Control. 2011;**6**:370-378

[12] Tahmasbi A, Saki F, Shokouhi SB. Classification of benign and malignant masses based on Zernike moments. Computers in Biology and Medicine. 2011;**41**:726-735

[13] Tan M, Pu J, Zheng B. Reduction of false-positive recalls using a computerized mammographic image feature analysis scheme. Physics in Medicine and Biology. 2014; **59**:4357-4373

[14] Nascimento MZ, Martins AS, Neves LA, Ramos RP, Flores EL, Carrijo GA. Classification of masses in mammographic image using wavelet domain features and polynomial classifier. Expert Systems with Applications. 2013;**40**:6213-6221

[15] Nishikawa Robert M. Current status and future directions of computer-aided diagnosis in mammography. Computerized Medical Imaging and Graphics. 2007;**31**(4-5):224-235

[16] Wang Y, Li J, Gao X. Latent feature mining of spatial and marginal characteristics for mammographic mass classification. Neurocomputing. 2014;**144**:107-118

[17] Liu X, Tang J. Mass classification in mammograms using selected geometry and texture features, and a new SVM-based feature selection method. IEEE Systems Journal. 2014;**8**:910-920

[18] Dong M, Lu X, Ma Y, Guo Y, Ma Y, Wang K. An efficient approach for automated mass segmentation and classification in mammograms. Journal of Digital Imaging. 2015;**28**:613-625

[19] Jen C-C, Yu S-S. Automatic detection of abnormal mammograms in mammographic images. Expert Systems with Applications. 2015;**42**:3048-3055

[20] Zheng B. Computer-aided diagnosis in mammography using content-based imageretrieval approaches: Current status and future perspectives. Algorithms. 2009;**2**(2):828-849

[21] Alto H, Rangayyan RM, Desautels JEL. Content-based retrieval and analysis of mammographic masses. Journal of Electronic Imaging. 2005;**14**(2):023016

[22] Kinoshita SK, de Azevedo-Marques PM, Pereira RR, Rodrigues JAH, Rangayyan RM. Content-based retrieval of mammograms using visual features related to breast density patterns. Journal of Digital Imaging. 2007;**20**(2):172-190

[23] Tie L et al. Learning to detect a salient object. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2011;**33**:353-367

[24] Da Cunha AL, Zhou J, Do MN. The nonsubsampled Contourlet transform: Theory, design, and applications. IEEE Transactions on Image Processing. 2006;**15**:3089-3101

[25] Hanbay K, Alpaslan N, Talu MF, Hanbay D, Karci A, Kocamaz AF. Continuous rotation invariant features for gradient-based texture classification. Computer Vision and Image Understanding. 2015;**132**:87-101

[26] Heath M. et al. Current Status of the Digital Database for Screening Mammography. In: Karssemeijer N, Thijssen M, Hendriks J, van Erning L, editors. Digital Mammography. vol 13. Dordrecht: Computational Imaging and Vision, Springer; 1998

[27] Wang XH et al. Automated assessment of the composition of breast tissue revealed on tissue-thickness-corrected mammography. American Journal of Roentgenology. 2003;**180**:257-262

[28] Muramatsu C et al. Determination of subjective similarity for pairs of masses and pairs of clustered microcalcifications on mammograms: Comparison of similarity ranking scores and absolute similarity ratings. Medical Physics. 2007;**34**:2890

[29] de Oliveira JEE, Machado AMC, Chavez GC, Lopes APB, Deserno TM, Araújo AA. MammoSys: A content-based image retrieval system using breast density patterns. Computer Methods and Programs in Biomedicine. 2010;**99**:289-297

[30] Wei L, Yang Y, Nishikawa RM. Microcalcification classification assisted by content-based image retrieval for breast cancer diagnosis. Pattern Recognition. 2009;**42**:1126-1132

[31] Wei C-H, Li Y, Huang PJ. Mammogram retrieval through machine learning within BI-RADS standards. Journal of Biomedical Informatics. 2011;**44**:607-614

[32] Brake GM, Karssemeijer N. Segmentation of suspicious densities in digital mammograms. Medical Physics. 2001;**28**:259-266

[33] Zheng B, Chang YH, Gur D. Computerized detection of masses in digitized mammograms using single image segmentation and multi-layer topographic feature extraction. Academic Radiology. 1995;**2**:959-966

[34] Lobregt S, Viergever MA. A discrete dynamic contour model. IEEE Transactions on Medical Imaging. 1995;**14**:12-24

[35] Yuan Y, Giger ML, Li H, Suzuki K, Sennett C. A dual-stage method for lesion segmentation on digital mammograms. Medical Physics. **34**:4180-4193

[36] Dominguez AR, Nandi AK. Improved dynamic-programming-based algorithms for segmentation of masses in mammograms. Medical Physics. 2007;**34**:4265-4268

[37] Agrawal P, Vatsa M, Singh R. Saliency based mass detection from screening mammograms. Signal Processing. 2014;**99**:29-47

[38] Avraham T, Lindenbaum M. Esaliency (extended saliency): Meaningful attention using stochastic image modeling. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2010;**32**(4):693-708

[39] Meselhy Eltoukhy M, Faye I, Belhaouari Samir B. A comparison of wavelet and curvelet for breast cancer diagnosis in digital mammogram. Computers in Biology and Medicine. 2010;**40**:384-391

[40] Ergin S, Kilinc O. A new feature extraction framework based on wavelets for breast cancer diagnosis. Computers in Biology and Medicine. 2014;**51**:171-182

[41] Choi JY, Ro YM. Multiresolution local binary pattern texture analysis combined with variable selection for application to false-positive reduction in computer-aided detection of breast masses on mammograms. Physics in Medicine and Biology. 2012;**57**:7029-7052

[42] Lahmiri S, Boukadoum M. Hybrid discrete wavelet transform and Gabor filter banks processing for features extraction from biomedical images. Journal of Medical Engineering. 2013;**2013**:1-13

[43] Reyad YA, Berbar MA, Hussain M. Comparison of statistical, LBP, and multi-resolution analysis features for breast mass classification. Journal of Medical Systems. 2014;**38**(9):100. DOI: 10.1007/s10916-014-0100-7. [Epub 2014 Jul 19]

[44] Shah VP, Younan NH, King RL. An efficient pan-sharpening method via a combined adaptive PCA approach and Contourlets. IEEE Transactions on Geoscience and Remote Sensing. 2008;**46**:1323-1335

[45] Do MN, Vetterli M. The contourlet transform: An efficient directional multiresolution image representation. IEEE Transactions on Image Processing. Dec. 2004;**14**:2091-2106

[46] Dalal N, Triggs B. Histograms of Oriented Gradients for Human Detection. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA; 2005;**1**:886-893. DOI: 10.1109/CVPR.2005.177

[47] Burges JC. A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery. 1998;**2**:121-167

[48] Huang GB, Zhu QY, Siew CK. Extreme learning machine: Theory and applications. Neurocomputing. 2006;**70**:489-501

[49] Xie W, Li Y, Ma Y. Breast mass classification in digital mammography based on extreme learning machine. Neurocomputing. 2016;**173**:930-941

[50] Rouhi R, Jafari M, Kasaei S, Keshavarzian P. Benign and malignant breast tumors classification based on region growing and CNN segmentation. Expert Systems with Applications. 2015;**42**:990-1002

[51] Jiao Z, Gao X, Wang Y, Li J. A deep feature based framework for breast masses classification. Neurocomputing. 2016;**197**:221-231