Applied Soft Computing 44 (2016) 267–280

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Optimized Gabor features for mass classification in mammography

Salabat Khan[a,*], Muhammad Hussain[b], Hatim Aboalsamh[b], Hassan Mathkour[b], George Bebis[c], Mohammed Zakariah[d]

[a] Department of Computer Science, Comsats Institute of Information Technology, Pakistan
[b] Department of Computer Science, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia
[c] Department of Computer Science and Engineering, University of Nevada, Reno, USA
[d] Research Centre, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia

ARTICLE INFO

Article history:
Received 9 November 2015
Received in revised form 18 February 2016
Accepted 11 April 2016
Available online 16 April 2016

Keywords:
Breast cancer detection
Gabor filter banks
Directional features
Particle Swarm Optimization (PSO)
Incremental clustering
Digital mammography

ABSTRACT

Gabor filter bank has been successfully used for false positive reduction problem and the discrimination of benign and malignant masses in breast cancer detection. However, a generic Gabor filter bank is not adapted to multi-orientation and multi-scale texture micro-patterns present in the regions of interest (ROIs) of mammograms. There are two main optimization concerns: how many filters should be in a Gabor filter band and what should be their parameters. Addressing these issues, this work focuses on finding optimizing Gabor filter banks based on an incremental clustering algorithm and Particle Swarm Optimization (PSO). We employ an SVM with Gaussian kernel as a fitness function for PSO. The effect of optimized Gabor filter bank was evaluated on 1024 ROIs extracted from a Digital Database for Screening Mammography (DDSM) using four performance measures (i.e., accuracy, area under ROC curve, sensitivity and specificity) for the above mentioned mass classification problems. The results show that the proposed method enhances the performance and reduces the computational cost. Moreover, the Wilcoxon signed rank test over the significance level of 0.05 reveals that the performance difference between the optimized Gabor filter bank and non-optimized Gabor filter bank is statistically significant.

1. Introduction

Breast cancer is the most common form of cancer which affects women all over the world and stands next to lung cancer in mortality among women [1,2]. A computer-aided diagnosis (CAD) system based on mammograms can assist the radiologists in detecting breast cancer. A CAD system for masses involves two main stages for mass extraction and processing: detection and segmentation, false positive reduction, and discrimination between benign and malignant masses. In the first stage, potential mass regions of interests (ROIs) are detected and segmented from the mammogram image. The detected ROIs represent not only masses but also dense breast parenchyma, which appear as white regions like masses in mammograms and result in false positives. The false positive reduction stage classifies the detected ROIs into mass and normal ROIs. The mass ROIs are further classified as benign and malignant.

Different efforts have been made so far for reducing false positives and increasing benign-malignant classification accuracy. Texture descriptors have been shown to represent masses more accurately [3]. Since texture microstructures appear at different orientations and scales, they can be represented more effectively using Gabor filters. Gabor filters have been used for this purpose (e.g., see Refs. [2,6] and references therein) and give better performance for false positive reduction and benign-malignant discrimination [6]. However, a Gabor filter bank is not adopted to multi-orientation and multi-scale texture microstructures present in mammograms. To address this issue, we propose to tune and optimize the filters in a Gabor filter bank in order to extract the local texture descriptors that characterize texture micro-patterns (e.g., edges, lines, spots and flat areas) more effectively. There are two main optimization concerns; first, how many filters are appropriate to be used in the bank (filter selection problem) and second, what should be the parameter values of each Gabor filter included in the bank (filter design problem). Clearly, both of these problems are application oriented and a general setting of a Gabor filter bank (see e.g., Refs. [8,9]) does not perform well in different application scenarios. In this work, we proposed a systematic approach that unifies the filter selection and design processes using Particle Swarm Optimization (PSO) and an incremental clustering algorithm. In particular, PSO, a global optimization technique, is used to search for optimal

* Corresponding author.
E-mail address: salabat.khan@nu.edu.pk (S. Khan).

http://dx.doi.org/10.1016/j.asoc.2016.04.012
1568-4946/

parameters of Gabor filters to address the filter design problem. In addition, an incremental clustering algorithm removes redundant Gabor filters from the bank by combining similar filters (in parameter space), thus addressing the filter selection problem. Filters in the same cluster are represented with a single filter which is the centroid of the cluster. The strategy helps in two ways; first, the recognition accuracy is improved and second, the computational cost is reduced (discussed in Section 4). This idea was initially presented in Ref. [7].

The key idea is optimizing the parameters of Gabor filters such that they respond stronger to features that best discriminate between normal and abnormal tissue, improving the performance of breast cancer recognition. The main contribution of this paper is a strategy based on PSO and incremental clustering for optimizing a Gabor filter bank that responds stronger to multi-scale and multi-orientation texture micro-patterns in a mammogram and enhances the classification rate. For the evaluation of the effect of the optimized Gabor filter bank on the mass classification problems, it is applied on overlapping blocks of the ROIs to collect moment-based features (i.e., mean, standard deviation, skewness) from the magnitudes of Gabor responses.

The rest of this paper is organized as follows. In the next section, we review related research. In Section 3, we present the proposed methodology in detail. Subsequently, in Section 4, we present our experimental results and discuss the effectiveness of the proposed technique. Finally, Section 5 concludes our work and presents directions for future research.

2. Related work

Mass detection has attracted the attention of many researchers, and many detection techniques have been proposed. A detailed review of these methods can be found in Refs. [10–13]. Next, we give an overview of the most related recent mass classification methods.

Most of the existing methods differ in the types of features that are used for mass representation and the way these features are extracted. Different types of features such as texture, gradient, grey-level, and shape [10] features have been employed for mass representation. Texture descriptors have been very effective in detecting normal and lesion regions in mammograms [14–16]. Wei et al. [17] extracted multiresolution texture features from wavelet coefficients and used them for the discrimination of masses from normal breast tissue in mammograms. They used Linear Discriminant Analysis (LDA) for classifying the ROIs as mass or non-mass. This method was tested on 168 ROIs containing biopsy-proven masses and 504 ROIs containing normal parenchyma, resulting in Az (Area under ROC curve) equal to 0.89 and 0.86 for the training and test data sets.

If texture is described accurately, then texture descriptors can perform better than other descriptors [3]. Lladó et al. [3] used a spatially enhanced Local Binary Pattern (LBP) descriptor, to represent textural properties of masses and to reduce false positives; this method achieved an overall Az equal to 0.94 ± 0.02 on 512 ROIs (256 normal and 256 masses) extracted from mammograms from the DDSM database. The LBP-based method outperforms other CAD methods for mass classification. However, the LBP descriptor builds statistics on local micro-patterns (dark/bright spots, edges, and flat areas etc.) and it is not robust to noise. The scheme proposed by Sampaio et al. [18] used geo-statistic functions for extracting texture features and SVM for classification, yielding Az of 0.87.

Oliveira et al. [19] addressed the problem of the classification of mammogram regions as mass and non-mass and proposed a method using texture features and SVM. The texture features in this method are computed using the taxonomic diversity index and the taxonomic distinctness. This method gives an average accuracy of 98.88% on DDSM database. The method developed by Nguyen et al. [20] for the classification of suspicious ROIs into masses and non-masses employs texture features extracted using Block Variance of Local Coefficients (BVLC) and SVM. This method achieved AUC of 0.93 on MIAS database. Junior et al. [21] presented a method for false positive reduction problem using texture features extracted as several diversity indices from ROIs and SVM. This method is reported to achieve 100% accuracy on DDSM database. Reyad et al. [22] compared first order statistics, LBP and multiresolution analysis features for the classification of mass and non-mass. They used SVM classifier and showed that the combination of first order statistics and LBP gives the best accuracy (98.63%) for mass regions from DDSM. Hussain [23] used multi-scale spatial Weber Law Descriptor (WLD) and SVM to propose a method for false positive reduction in mammograms. The accuracy rate of this method was 98.93% for the classification of mass and non-mass ROIs from DDSM database.

Rouhi et al. [24] used intensity, textural, and shape features, genetic algorithm (GA) for feature selection, and artificial neural network (ANN) to develop a method for the classification of benign and malignant masses. The accuracy of this method on DDSM using 10-fold cross validation is 82.06%. Nanni et al. [25] compared three texture descriptors (Local Ternary Pattern (LTP), local phase quantization (LPQ)) for the discrimination of mammogram tissues as benign and malignant. They used SVM for classification and claimed to achieve an AUC of 0.97, but Li et al. [26] compared their method with this method and showed that this method achieved classification accuracy rates of 64.62 (LTP) and 62.3 (LPQ) on 114 mass regions (52 benign and 62 malignant) from DDSM database using 5-fold cross validation. Li et al. [26] proposed a method for the discrimination of benign and malignant masses using texton analysis with multiple sub-sampling strategies. K-nearest Neighbor (KNN) classifier is used for each sub-sampling strategy and majority-vote is used for final decision from all KNN classifiers. This method achieved an accuracy rate of 85.96% on 114 mass regions (52 benign and 62 malignant) from DDSM database.

Gabor filters have been extensively used for texture description in various image processing and analysis approaches [27,28]. They decompose an image into multiple scales and orientations making the analysis of texture patterns more straightforward. Mammograms contain a lot of texture, and as such, Gabor filters are suitable for texture analysis of mammograms [29,30]. Different texture description techniques using Gabor filters differ in the way that the texture features are extracted [31]. Hussain [2] employed Gabor filters to create 20 Gabor images, which were then used to extract a set of edge histogram descriptors. He used KNN along with fuzzy c-means clustering as a classifier. The method was evaluated on 431 mammograms (159 normal cases and 272 containing masses) from the DDSM database using tenfold cross validation. This method achieved a true positive (TP) rate of 90% at 1.21 false positives per image. The data set used for validation is biased towards abnormal cases, which favors the mass cases. It should be mentioned that this method extracts edge histograms which are holistic descriptors, and do not represent the local textures of masses.

Lahmiri and Boukadoum [32] used Gabor filters along with Discrete Wavelet Transform (DWT) for mass classification. They applied Gabor filters at different frequencies and spatial orientations on the HH high frequency sub-band image obtained using DWT, and extracted statistical features (i.e., mean and standard deviation) from the Gabor images. For classification, they used SVM with polynomial kernel. Their method was tested on 100 mammograms from the DDSM database using tenfold cross validation, achieving an accuracy of 98%. Although the detection accuracy achieved was good, the size of the dataset used for testing was relatively small. Costa et al. [33] explored the use of Gabor filters along with Principal Component Analysis (PCA) for feature extraction,

Independent Component Analysis (ICA) for efficient encoding, and LDA for classification. The success rate of their method, using Gabor filters for feature extraction, was 85.05% on 5090 ROIs extracted from mammograms in the DDSM database. Hussain at el. [6] used a Gabor filter bank in a novel way to extract the most representative and discriminative textural properties of masses present at different orientations and scales for mass classification. SVM with Gaussian kernel was employed for classification. Their method was evaluated over 1024 (512 masses and 512 normal) ROIs extracted from the DDSM database. For classifying ROIs as masses and normal tissues, this method resulted in $Az = 0.96 \pm 0.02$; the accuracy for classifying mass ROIs as benign and malignant was $Az = 0.87 \pm 0.05$.

In Ioan and Gacsadi [34], raw magnitude responses of 2D Gabor filters were investigated as features for proximal SVM. A total of 322 mammogram images from the Mammographic Image Analysis Society (MIAS) database were used for three experimental cases: (a) discrimination between the three classes: normal, benign and malignant using one against all SVM classification, (b) normal vs. tumor (benign and malignant) and (c) benign vs. malignant; 80% data samples were used for training and the rest 20% for testing. The dimension of the feature space in this case was equal to the number of pixels present in the mammogram images times the number of Gabor filters; PCA was used for dimension reduction. The best results (in terms of accuracy) for the three experimental cases were: 75%, 84.37% and 78.26%, respectively.

Geralodo et al. [35] used Moran's index and Geary's coefficients as input features for an SVM classifier and tested their approach over two cases: normal vs. abnormal and benign vs. malignant. They obtained an accuracy of 96.04% and Az of 0.946 with Geary's coefficient and an accuracy of 99.39% and Az of 1 with Moran's index for the classification of normal vs. abnormal cases. For the second case (benign vs. malignant), an accuracy of 88.31% and Az of 0.804 were obtained with Geary's coefficient and an accuracy of 87.80% and Az of 0.89 with Moran's index. The method was tested over 1394 ROI images collected from the DDSM database using tenfold cross validation.

Most of the published research works using 2D Gabor filters for feature extraction are mostly concerned with a generic (non-optimized) setting of filters in a bank [2,33,36,37,38,6]. Based on the Gabor feature extraction strategy, there are three major differences between the earlier works and the method proposed in this paper. First, the size of the Gabor filter bank used for filtering the ROIs is different. Parameters of an optimized Gabor filter in the bank are not similar [2,33] resulting in different resultant responses that are basis for eventual feature values. The second major difference is in the way a filter in the bank is applied to the mammogram images. In previous research works (e.g., Refs. [2,33,35,36] and references therein) related to 2D Gabor filtering in mammography, the entire ROI is convolved with each Gabor filter increasing feature dimensionality significantly. In most of these cases, a feature selection or projection strategy [33] is required to handle high dimensionality. In our case, we partition an ROI into overlapping blocks (Section 3.3), apply an optimized filter to each block and compute statistical measures making the feature dimension more manageable. For example, using the feature extraction strategy presented in Ref. [33], the feature dimension for an ROI resolution of 32×32 pixels with 40 Gabor filters is 40,960; in our case, using a block size of 16×16 pixels, yields 1080 features. This strategy also improves the recognition rate of the method significantly as depicted in Section 4. The third major difference has to do with the way the resultant filtered pixel values are used to form a feature vector. Gabor filtering leads to complex values. We use moment-based magnitude values of pixels in blocks to construct a feature vector (Section 3.1); earlier works [2,33], however, use the magnitude of the complex values directly as feature values, without any further processing.

3. Method

In this section, we describe in detail our method for optimizing Gabor filter bank for false positive reduction and benign-malignant classification problems. First, a brief overview of Gabor filter banks is presented. Then, we discuss the optimization of Gabor filter bank using PSO and clustering. Finally, we review the SVM classifier that is used as a fitness function in PSO.

3.1. Gabor filter bank

Masses in an ROI contain micro-patterns at different scales and orientations. These micro-patterns are helpful in the recognition of cancerous regions in a CAD system. Gabor filters can effectively be used to detect these micro-patterns.

Gabor filters are biologically motivated convolution kernels [40] that have been widely used in a number of applications in the fields of computer vision and image processing (e.g., face recognition [41], facial expression recognition, iris recognition, optical character recognition, vehicle detection [42] etc.). In order to represent multi-scale and multi-orientation textural micro-patterns in ROIs, Gabor filters can be tuned using different orientations and scales. The general form $g(x, y)$ of a 2D Gabor filter family is represented by a Gaussian kernel modulated by an oriented complex sinusoidal wave [42]:

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{\left[-1/2\left((\tilde{x}^2/\sigma_x^2)+(\tilde{y}^2/\sigma_y^2)\right)\right]} e^{(2\pi jW\tilde{x})} \quad (1)$$

$$\tilde{x} = x\cos\theta + y\sin\theta \text{and} \tilde{y} = -x\sin\theta + y\cos\theta. \quad (2)$$

where σ_x and σ_y are the scaling parameters (i.e., they describe the neighborhood of a pixel where the weighted summation takes place), W is the central frequency of the complex sinusoid and $\theta \in [0, \pi)$ is the orientation of the normal to the parallel stripes of the Gabor function.

A particular Gabor filters bank (GFB) contains multiple individual filters adjusted with different parameters (scaling, orientation and central frequency) [9]. In this paper, four different GFB configurations have been considerd: GS2O3 (i.e., containing 6 filters, 2 scales{S} $\times$ 3 orientations{O}), GS3O5, GS4O6 and GS5O8 with an initial max frequency equal to 0.2 and an initial orientation set to 0. The orientations and frequencies for a GFB are calculated using following equations [42]:

$$\text{Orientation}(i) = \frac{(i-1)\pi}{O} \text{where } i = 1, 2, \ldots\ldots,$$
$$O(\sharp \text{ total orientations}). \quad (3)$$

$$\text{Frequency}(i) = \frac{f_{\max=0.2}}{(\sqrt{2})^{i-1}} \text{where } i = 1, 2, \ldots\ldots, S(\sharp\text{total scales}). \quad (4)$$

3.2. Gabor filter bank optimization

A Gabor filter is described with four parameters $\psi = \{\theta, W, \sigma_x, \sigma_y\}$. A bank contains several Gabor filters tuned with different parameter settings ψ, yielding quite different filter responses. In general, parameter settings are chosen in a data-independent way. Different classification problems require an optimal set of features and so optimal Gabor filters. Selecting GFB with optimal parameter settings is an important optimization problem. We employ PSO [43] to select the GFB with optimal responses. PSO is a population based stochastic global optimization technique, which was developed in 1995 by Eberhart and Kennedy [43]. PSO is a suitable strategy that can better cope with the continuous nonlinear optimization problems as compared to Genetic Algorithms

(GAs). Further, for continuous valued problems, usually, the PSO implementation is simpler and converges faster than GAs. PSO carries out both global and local searches simultaneously whereas GA mainly involves global search [44].

The idea behind PSO was inspired by social behavior of flocks of birds and school of fish. It is a multi-agent system where an agent is called 'particle'. PSO starts with an initial population of particles with random positions and velocities. The position of a particle is a d-dimensional vector and represents a candidate solution to an d-dimensional problem. The particles fly through the hyperspace by updating their velocities and share their experiences with each other. Though, a single particle is not intelligent enough, working collectively through sharing the experience about the problem space, however, help the swarm in solving complex optimization problems. In the standard PSO, velocities of the particles are stochastically updated based on two factors: first, the personal best position p_b (a cognitive factor) found so far by the particle itself and second, the global best position p_{gb} (a social factor) found so far by the entire swarm. The quality of a particle position is determined using an application specific fitness function. The velocity and position of i-th particle is updated using the following equations:

$$x^i{}_{t+1} = x^i{}_t + v^i{}_{t+1}.$$

$$v^i_{t+1} = w_i v^i_t + c_1 r_1 (p^i_b - x^i_t) + c_2 r_2 (p_{gb} - x^i_t).$$

where $x^i{}_t$ and $v^i{}_t$ are n-dimensional vectors that represent the position and velocity of the i-th particle in iteration t, respectively; w_i is known as the inertia weight. The $w_i v^i{}_t$ is the weighted current velocity, $c_1 r_1 (p^i_b - x^i_t)$ is the weighted deviation from the self-best position, and $c_2 r_2 (p_{gb} - x^i_t)$ is the weighted deviation from the global best position. If w_i has a large value, the particle will be moving faster due to which the search becomes less refined (global exploration); if it is too low, the search will take longer time (local exploitation). The values of c_1 and c_2 (known as acceleration coefficients) are set to some predefined values at initialization and r_1, r_2 are calculated randomly in the range from 0 to 1.

The four parameters in w_i represent a single Gabor filter to be optimized. In order to optimize n filters (see Fig. 1), the particle will fly in $4n$ dimensional space, where each dimension assumes a continuous value. It is worth mentioning that the encoding scheme is quite flexible and any number of pre-specified filters can easily be encoded by simply increasing/decreasing the dimension of the search space. Each of the parameters in ψ has its own constraints and ranges which could be imposed using some prior knowledge. We have adopted the ranges used in Ref. [7] such that the θ parameter should satisfy the range $[0, \pi)$. Radial frequency of the filters is allowed to be in the range $[0 \geq W \leq 0.5]$ and the σ_x, σ_y (i.e., the sharpness of the filter along major and minor axes, respectively) should be in the range of $[0.796 > \sigma < w/5]$ where w is the width of the mask in the filter. In our experiments, we have used a swarm size of 10 particles that are initialized randomly within the giving ranges of each dimension. The error rate of an SVM classifier is used as fitness value to evaluate the effectiveness of the filters encoded in a particle.

The velocities of the particles quickly attain very large values due to which the particles might take longer jumps in the search space and thus the search becomes less refined or coarse. In order to ensure a finer search in the close proximities of the particles' current positions, velocity clamping is used. If the velocity of a particle in any of its dimension exceeds the value of $\pm V_{max}$, the value at that dimension is clamped to $\pm V_{max}$ where $\pm V_{max}$ is the maximum velocity in both positive and negative directions. In our approach, V_{max} is equal to the maximum range values of the parameters in the respective dimensions. To ensure that the particle does not overfly from the given ranges of parameters, velocity resetting is used. After applying the position update equation, if the position value at a particular dimension of the particle does not remain in the given range, the older position value at that dimension is used and the velocity at that dimension is reset to zero (known as velocity resetting).

The particles keep searching for an optimum solution until the stopping criteria are met. Two stopping criteria are used; first, the search process is terminated if the total number of iterations is reached or the optimum is found. Second, the search is stopped if convergence has been detected. The search process is supposed to have converged when the global best fitness value of the entire swarm doesn't improve over the last few iterations. Convergence, as defined in our approach, does not necessarily mean that the optimal solution has been found. Convergence might occur due to stagnation where the particles move back and forth in the positions visited previously and no more exploration of the search space is possible. There are numerous techniques in the PSO literature for handling stagnation (e.g., particles re-initialization, regrouping of the particles, etc.).

3.3. Incremental clustering

During the execution of PSO for Gabor filters parameter optimization, it is possible that some of the filters encoded in a particle become similar or even identical. These similar or possibly identical filters generate similar/identical responses, thus introduce redundancy in the feature space, which might degrade the generalization ability of SVM and increase processing time. To deal with this situation, an incremental clustering algorithm is used that groups together redundant filters in a particle. The centroid of the cluster is then used as a single filter instead of using all the filters in the cluster. Clustering is performed in the parameter space of ψ. The incremental clustering algorithm used in Ref. [7] was also adopted in our work. Specifically, the first filter in a particle is assigned to the first cluster. Next, the n-th filter $\psi^n = \{\theta^n, W^n, \sigma_x{}^n, \sigma_y{}^n\}$ in the particle is assigned to the i-th cluster, only if the following conditions are satisfied.

$$\theta^i - \mathrm{Thr}_\theta \leq \theta^n \leq \theta^i + \mathrm{Thr}_\theta$$

$$W^i - \mathrm{Thr}_W \leq W^n \leq W^i + \mathrm{Thr}_W$$

$$\sigma_x{}^i - \mathrm{Thr}_\sigma \leq \sigma_x{}^n \leq \sigma_x{}^i + \mathrm{Thr}_\sigma$$

$$\sigma_y{}^i - \mathrm{Thr}_\sigma \leq \sigma_y{}^n \leq \sigma_y{}^i + \mathrm{Thr}_\sigma$$

where $i \in [1N]$ and N represents the total number of clusters; $\{\theta^i, W^i, \sigma_x{}^i, \sigma_y{}^i\}$ is the centroid of the i-th cluster and Thr represents the threshold values used to quantify the similarity between the filter and centroid of the cluster. The thresholds are calculated as follows based on the minimum and maximum range values of the parameters in ψ:

$$\mathrm{Thr}_\theta = \frac{\pi}{K} \times 0.5, \mathrm{Thr}_W = \frac{(W_{max} - W_{min})}{K} \times 0.5 \text{ and}$$

$$\mathrm{Thr}_\sigma = \frac{(\sigma_{max} - \sigma_{min})}{K} \times 0.5$$

Here, K is a user defined parameter which is used to handle the extent of threshold values over the tradeoff between model compactness and accuracy [7] (e.g., large values of K result in more compact models).

The block diagram for the GFB optimization is given in Fig. 2. First, a set of validation ROIs, selected from the mammogram database, are used for PSO learning. The ROIs are resized to the same resolution as per the requirement of GFB. The banks (which are encoded as particles) are initialized randomly and the fitness

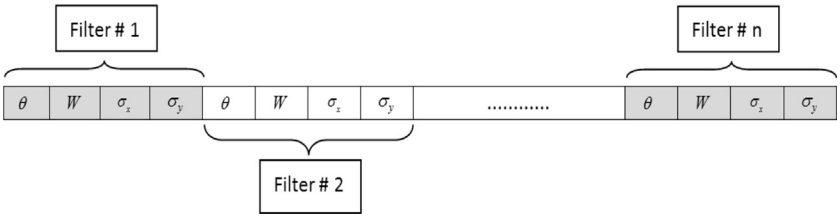

Fig. 1. Particle encoding for a Gabor filter bank containing *n* filters.

of each bank is calculated using the SVM classifier to initialize personal best and global best positions. The *i*-th particle in the swarm is referred to as '*i*' in the block diagram. As PSO learning takes place, it is expected that each new generation would improve the accuracy of the global best particle over the validation ROIs (i.e., measured in terms of the SVM error rate using tenfold cross validation). A generation is completed when all particles update their positions (i.e., when '*i*' exceeds the size of swarm represented as |particles). In a generation, each particle is used to extract the features from the selected validation ROIs. The fitness of *i*-th particle is calculated using the SVM classifier over the features (dataset) extracted with filters represented by the *i*-th particle using tenfold cross validation. This process is repeated until the swarm converges or the allowed number of generations is reached.

3.4. Gabor features extraction

For the optimization process, we use the Gabor feature extraction method proposed in Ref. [6]. Here, we give an overview of this method.

The first step is to divide each ROI into overlapping windows (OW) as shown in Fig. 3. For instance, an ROI of size 512 × 512 pixels is first divided into blocks of equal size of 128 × 128 pixels. In this way, sixteen blocks are created, labeled 1–16 in Fig. 3. Using these blocks, OWs of equal sizes of 256 × 256 pixels are created (e.g., blocks 1, 2, 5 and 6 form the first 256 × 256 pixels OW, blocks 2, 3, 6 and 7 the second, 5, 6, 9 and 10 the fourth, and so on). With this formation, 9 OWs are created. It should be noted that by increasing/decreasing the size of a block, an ROI can be segmented into different sizes and number of OWs

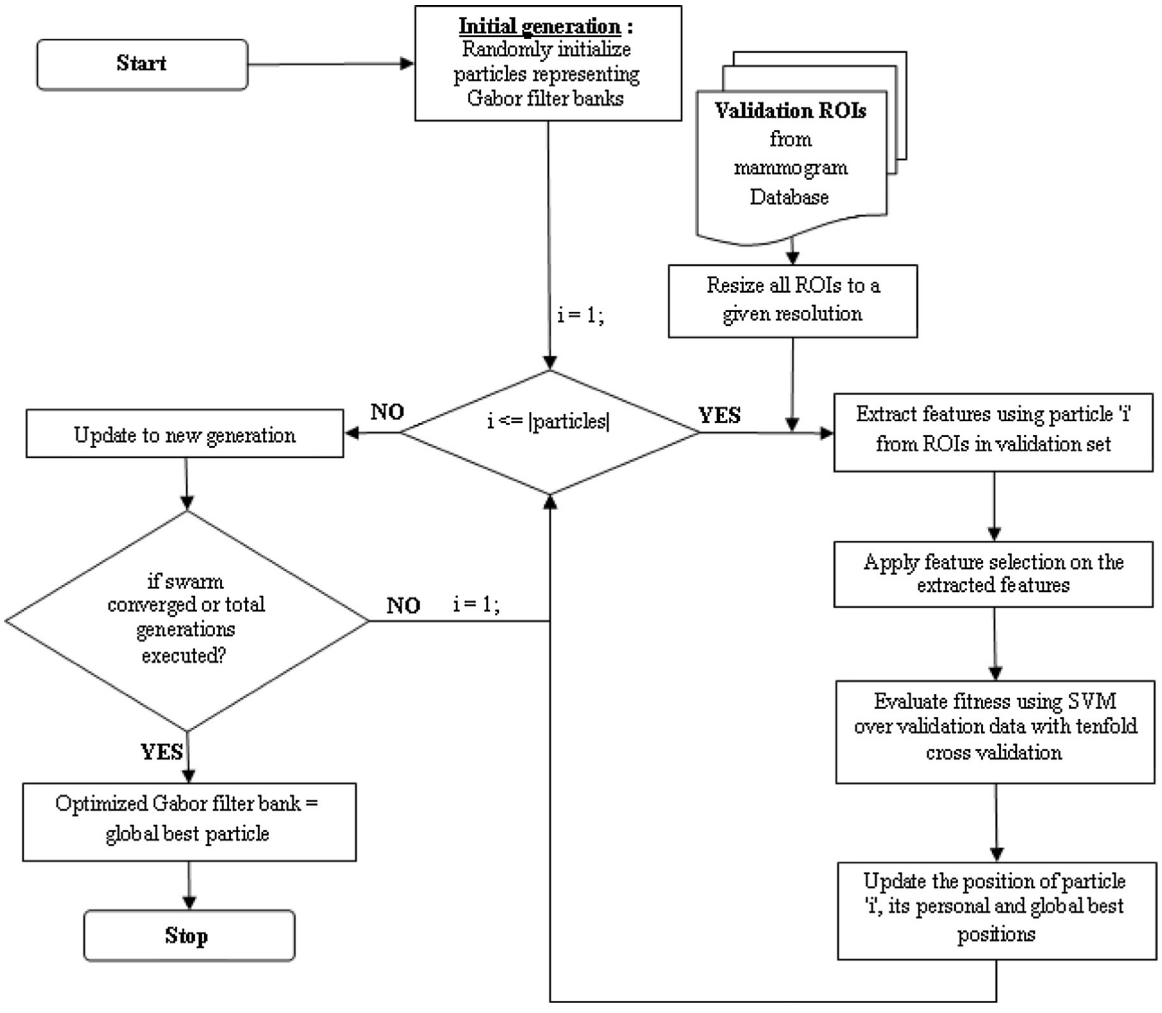

Fig. 2. Block diagram for GFB optimization using PSO.

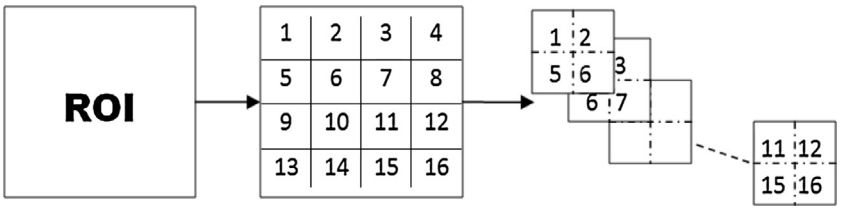

Fig. 3. Segmentation of ROI into overlapping windows.

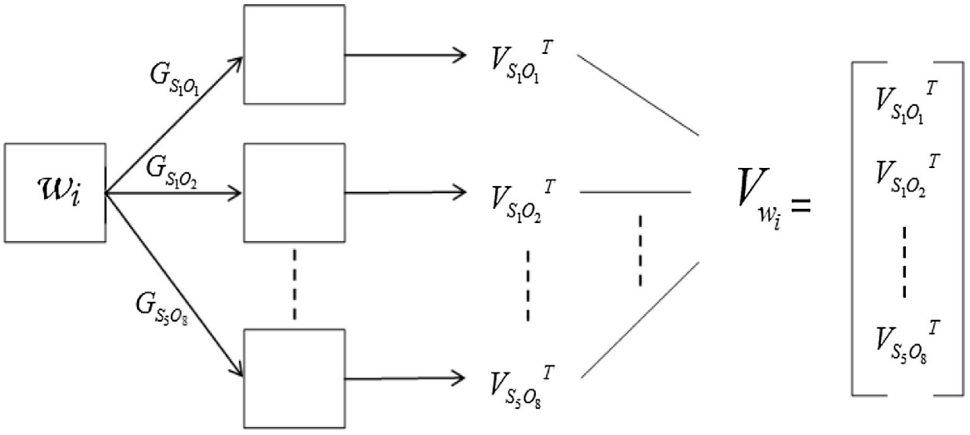

Fig. 4. Extraction of Gabor features from OW W_i using GFB GS5O8, here $V_{S_iO_j} = [\mu_{i,j}, \sigma_{i,j}, k_{i,j}]$.

The second step involves convolving each OW of an ROI with a GFB and computing statistical moments to form the feature vector representing the ROI (see Figs. 4 and 5). In this paper, three statistical moments (SMs) are computed from the magnitude responses of each Gabor filter in an OW. The employed SMs are the mean $\mu_{i,j}$, the standard deviation $\sigma_{i,j}$ and the skewness $k_{i,j}$ (where i corresponds to the i-th filter in the bank and j to the j-th OW of an ROI).

The statistical features V_{wi} computed from all OWs are concatenated to form the feature vector representing the ROI (see Fig. 5). For instance, assuming GS5O8, which consists of 40 filters, and segmenting an ROI into 9 OWs, a feature vector of size 1080 = (9OWs × 40GFs × 3SMs) is obtained which is shown below as a row vector:

$$[\mu_{1,1}, \sigma_{1,1}, k_{1,1}, \mu_{2,1}, \sigma_{2,1}, k_{2,1}, \ldots, \mu_{40,1}, \sigma_{40,1}, k_{40,1}, \mu_{1,2}, \sigma_{1,2}, k_{1,2}, \ldots., \mu_{40,9}, \sigma_{40,9}, k_{40,9}, \text{class}]. \quad (5)$$

3.5. Feature selection

We used SVM as fitness function in PSO and when the size of the OWs is small, the dimension of the feature space becomes too large which makes it difficult to be handled by SVM. This is because many of the features can be irrelevant or redundant, increasing the complexity of the search space and making generalization more difficult (curse of dimensionality). SVMs are very sensitive to noisy training data [45]; they are also prone to over-fitting and poor generalization [46] which can degrade classification performance [47]. In order to cater this shortfall, a fast and robust feature selection strategy is required. Sparse Multinomial Logistic Regression with Bayesian L_1 Regularization (SBMLR_FSS) [48] is a fast and robust feature subset selection (FSS) algorithm, which is employed here. Due to its fast execution, SBMLR_FSS is an ideal algorithm to be used with evolutionary and swarm based optimization methods.

We used the implementation of SBMLR_FSS from Ref. [49]. It returns a list of selected features and provides the weight (based on the information gain) values assigned to all the features in the original feature space. The features selected by SBMLR_FSS are used without any post-processing (e.g., thresholding) in the resultant weight vector.

3.6. Classification and fitness function

For the classification of ROIs, we employ SVM. The error rate of SVM is used as the fitness value to evaluate the effectiveness of the filters encoded by a particle in PSO. Classification of ROIs is a binary classification problem. SVM classifiers [50] are generally designed to solve binary classification problems; thus perfectly suiting our requirements. The idea is to find an optimal hyperplane that can separate the data samples belonging to two different classes with large margins in high dimensional space [51,59,60].

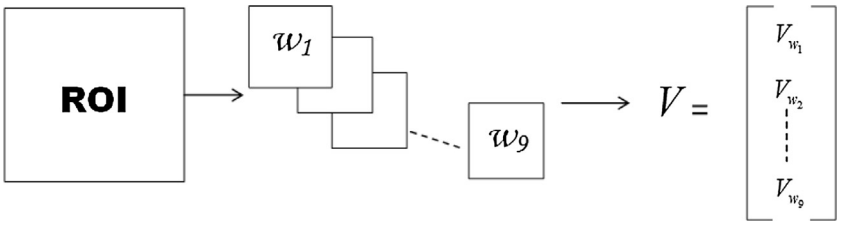

Fig. 5. Extraction of Gabor features from ROI and feature vector V representing the ROI.

The margin is defined as the sum of distances to the decision boundary (hyper-plane) from the nearest samples (support vectors) of the two classes. SVM formulation is based on statistical learning theory and has attractive generalization capabilities in linear as well as non-linear decision problems [50,52]. SVM uses structural risk minimization as opposed to empirical risk minimization [50] by reducing the probability of misclassifying an unseen pattern drawn randomly from a fixed but unknown distribution.

SVM takes classification decisions using the following optimal hyper-plane with maximum margin:

$$g(x) = w^{\mathrm{T}}x + w_0 = 0 \tag{6}$$

where x is the feature descriptor and w and w_0 are unknown parameters, which are computed using training samples $\{(x_i, y_i)|\ 1 \le i \le N\}$, where $y_i \in \{+1, -1\}$ is the class label; the computation involves the solution of an optimization problem based on large margin theory. Once the optimal hyper-plane has been computed, the classification of a test sample x is performed using

$$g(x) = \sum_{i=1}^{N_s} \lambda_i y_i x_i^{\mathrm{T}} x + w_0 \tag{7}$$

where λ_i are Lagrange multipliers and N_s is the number of support vectors i.e., the training samples corresponding to non-zero λ_i's. In case the data samples belonging to two classes are not linearly separable, Kernel trick is used; a kernel function satisfying Mercer's condition [50] is employed that maps the lower dimensional space to a higher dimensional space where the samples become linearly separable. Using a kernel function, the function $g(x)$ is expressed as

$$g(x) = \sum_{i=1}^{N_s} \lambda_i y_i \boldsymbol{K}(\boldsymbol{x_i}, \boldsymbol{x}) + w_0 \tag{8}$$

where $\boldsymbol{K}(\boldsymbol{x}_i, \boldsymbol{x})$ is the kernel function that expresses the inner product of data samples in the higher dimensional space. User defined parameter C is used to control the misclassified penalty or error in the new formulation, the misclassification penalty or error is controlled with a user defined parameter C (regularization parameter, controlling tradeoff between error of SVM and margin maximization), and is tied with the kernel. There are several kernels available to be used e.g., linear, polynomial, sigmoid, radial basis function (RBF), etc. In our experiments, RBF kernel is used as given by:

$$\boldsymbol{K}(\boldsymbol{x}_i, \boldsymbol{x}) = \exp(-\gamma||\boldsymbol{x}_i - \boldsymbol{x}||^2),\ \gamma > 0. \tag{9}$$

The γ is the width of the kernel function. RBF kernel is now tied with two parameters γ and C. Model selection procedure is an attempt to find better hypothesis which is done by tuning these 2 parameters. Loose grid search is the first operation to find the model selection for better regions in the parameter space. After doing the loose grid is followed by finer grid to search the region found by loose grid search. This model selection procedure is recommended in the work of Hsu et al. [53]. SVM parameter groups are used (as found by grid search) only for non-PSO based methods. The best parameters found in these non-optimized methods experiments are saved and used for PSO based experiments in an effort to reduce computational cost of the proposed method. In this way, the parameters of SVM kernel are fixed and there is no need to run grid search (for setting parameters of SVM) when evaluating the fitness of each particle in PSO.

4. Results and discussion

In this section, we present our experimental results to demonstrate the performance of the proposed method. We have conducted experiments for two classification problems: (i) false positive reduction (i.e., to classify candidate ROIs into normal and mass) and, (ii) the classification of mass ROIs into benign and malignant. First, we discuss the mammogram database and the evaluation methodology used in our experiments. Then, we discuss different parameter configurations investigated during our experiments. We statistically compare the optimized and non-optimized GFBs using a non-parametric Wilcoxon signed rank test [54,55] to validate our conclusions. Finally, the proposed method is compared with some recent state-of-the-art methods.

4.1. Database and evaluation methodology

For evaluation, we used the Digital Database for Screening Mammography (DDSM) [56], which consists of more than 2000 cases and is a commonly used benchmark database for breast cancer detection. The database is completely annotated by expert radiologists for each case. The locations of masses in mammograms are available as code chains. We randomly selected 768 cases from the database and using code chains, extracted 256 ROIs, which contain true masses; the sizes of these ROIs vary depending on the size of the mass regions. In addition, we extracted 512 ROIs containing normal but suspicious tissues and 256 benign ROIs. Some sample ROIs are shown in Fig. 6.

For the false positive reduction problem, we used 256 normal and 256 mass (malignant and benign) ROIs. For the discrimination of benign and malignant ROIs, we used 256 benign and 256 malignant mass ROIs. We used 150 ROIs (75 normal, 75 malignant and benign) in the first case and 150 ROIs (75 benign and 75 malignant) in the second case as validation ROIs for the PSO based Gabor filter bank optimization. The GFB was optimized with PSO using tenfold cross validation. The performance of optimized GFB (OGFB) was compared with the non-optimized GFB (NGFB) using an SVM classifier and tenfold cross validation (see Fig. 7).

Using tenfold cross validation, the data set is randomly partitioned into ten non-overlapping and mutually exclusive subsets. For the experiment of fold i, subset i is selected as testing set and the remaining nine subsets are used to train the classifier. In order to have a fair comparison, the same tenfold cross validation subsets were used for different experiments (e.g., with/without feature selection).

We employed commonly used performance evaluation measures for performance evaluation. The SVM classifier assigns a membership value of each class when an unseen pattern is presented to it. The receiver operator characteristics (ROC) curve can be obtained by varying the threshold on this membership value. Then, the area under ROC curve (Az.) can be calculated. The other commonly used evaluation measures in the machine learning community are accuracy or recognition rate (RR) = (tp + tn)/(tp + fp + tn + fn), sensitivity (Sn) = tp/(tp + fn), and specificity (Sp) = tn/(tn + fp), where tn, tp, fp and fn denote, respectively, the number of true negatives, true positives, false positives and false negatives.

4.2. Parameter settings

To thoroughly evaluate the effect of the proposed optimization procedure, we examined its effect on three types of parameters (used for Gabor features): (i) size of ROIs, (ii) size of blocks and OWs and (iii) the number of filters in a GFB. The extracted ROIs are of different sizes; to process them using a Gabor filter bank, they are resized to the same resolution. For resizing the ROIs, Bicubic interpolation method is used where the output pixel value is a weighted average of pixels in the nearest 4×4 neighborhood. We have tested two resolutions: 128×128 and 64×64. In order to extract features at different levels of granularity, the ROIs are partitioned into blocks of three sizes: 8×8, 16×16 and 32×32 pixels. The size of OWs is dependent on the block size and therefore, is not required to be tuned. For the third parameter, we used four config-

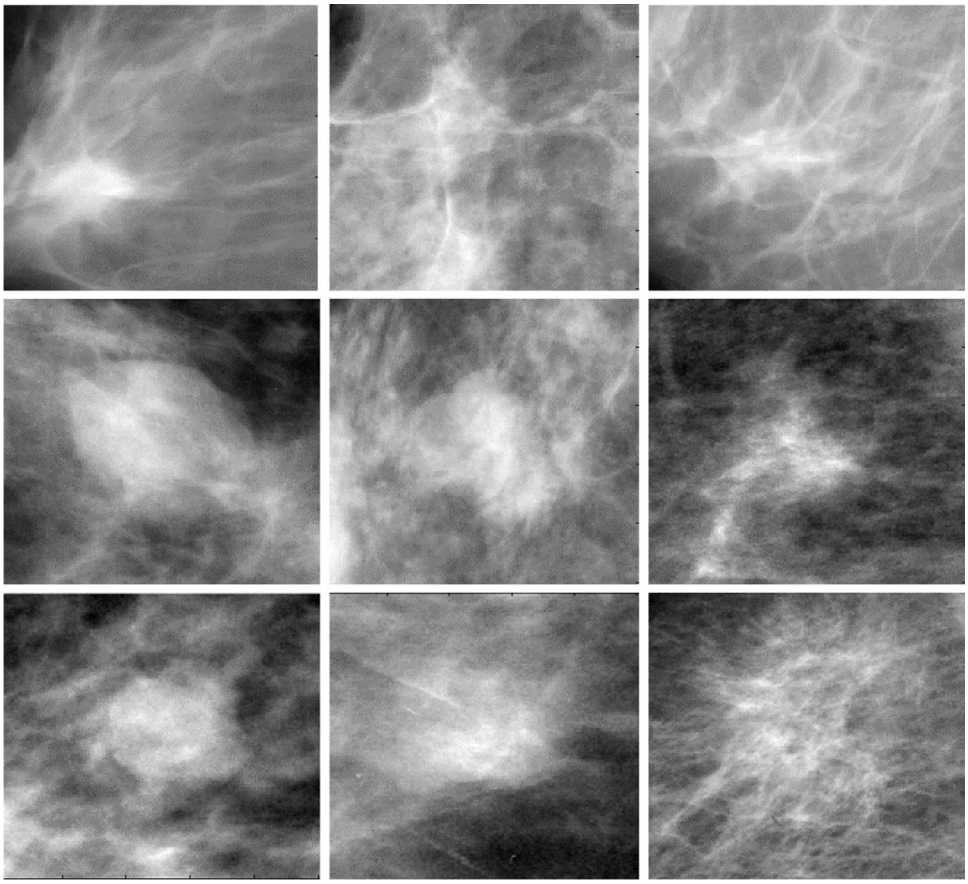

Fig. 6. (Top row) Normal but suspicious ROIs, (middle row) benign mass ROIs, (bottom row) malignant mass ROIs.

urations of Gabor filter banks: GS2O3 (2 scales and 3 orientations), GS3O5, GS4O6 and GS5O8.

For the optimization of GFB, the following PSO parameters were used: swarm size = 10, total generations = 50, iterations_for_converge = 5, c_1 & c_2 = 1.4, k-threshold = 3 and inertia weight = 0.7. These parameter values were chosen because they seems to provide reasonable performance as reported earlier in the literature [43].

In the following discussion, we refer to Gabor feature extraction method using NGFB as **noPSO**. The second feature extraction strategy is referred as **PSO-noFSS** where PSO is used without feature subset selection to find OGFB. We have noticed that PSO results are highly dependent on the performance of SVM classification. The high dimension of features can mislead the learning process by inducing error in the classification phase. We therefore apply SBMLR_FSS before the fitness function evaluation to facilitate the search of particles in the right direction. We refer to this strategy as **PSO-FSS** (i.e., PSO guided by feature subset selection). For a fair comparison, the fourth feature extraction strategy called **noPSO-FSS** is also investigated, where the first feature extraction method **noPSO** is combined with feature subset selection (i.e., SBMLR_FSS).

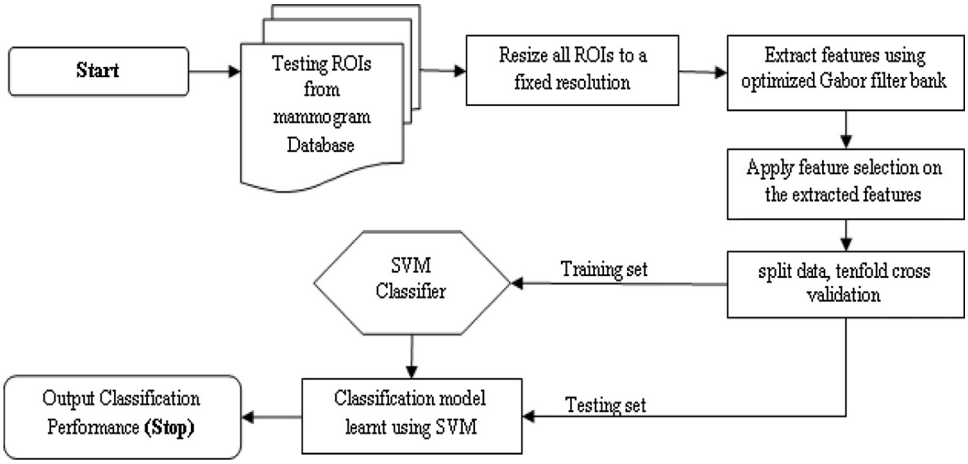

Fig. 7. Performance evaluation using OGFB.

Table 1
Performance in terms of sensitivity (mean ± std) for normal vs. masses (bold highlighted values are best).

ROI size	Block size	noPSO		noPSO-FSS		PSO-noFSS		PSO-FSS	
		GFB	Sn	GFB	Sn	OGFB	Sn	OGFB	Sn
64 × 64	16 × 16	GS3O5	88.04 ± 6.84	GS3O5	84.93 ± 7.93	PSO15 (14)	89.76 ± 4.87	PSO15 (13)	95.81 ± 3.34
		GS4O6	90.99 ± 6.86	GS4O6	91.56 ± 3.10	PSO24 (22)	89.57 ± 5.47	PSO24 (23)	95.81 ± 3.03
		GS5O8	**93.97 ± 4.11**	GS5O8	92.64 ± 3.09	PSO40 (30)	79.16 ± 29.39	PSO40 (31)	95.20 ± 4.54
	8 × 8	GS3O5	91.95 ± 3.15	GS3O5	91.54 ± 6.37	PSO15 (13)	87.90 ± 6.81	PSO15 (15)	96.89 ± 5.42
		GS4O6	91.85 ± 5.42	GS4O6	92.36 ± 4.53	PSO24 (20)	89.60 ± 4.15	PSO24 (19)	97.92 ± 3.08
		GS5O8	92.66 ± 5.56	GS5O8	90.57 ± 5.43	PSO40 (37)	90.17 ± 5.83	PSO40 (30)	96.95 ± 2.25
128 × 128	32 × 32	GS3O5	91.01 ± 4.87	GS3O5	89.16 ± 5.90	PSO15 (14)	83.43 ± 6.46	PSO15 (14)	95.49 ± 4.53
		GS4O6	91.82 ± 6.14	GS4O6	88.08 ± 8.56	PSO24 (21)	83.54 ± 4.17	PSO24 (21)	96.22 ± 2.88
		GS5O8	92.13 ± 4.20	GS5O8	90.77 ± 6.27	PSO40 (33)	89.43 ± 7.26	PSO40 (35)	94.07 ± 4.73
	16 × 16	GS3O5	88.50 ± 4.87	GS3O5	91.34 ± 3.97	PSO15 (14)	88.60 ± 6.05	PSO15 (14)	95.85 ± 5.20
		GS4O6	90.72 ± 5.90	GS4O6	90.92 ± 5.95	PSO24 (20)	90.45 ± 4.74	PSO24 (22)	96.55 ± 2.92
		GS5O8	90.46 ± 4.78	GS5O8	90.03 ± 4.26	PSO40 (35)	**90.51 ± 6.52**	PSO40 (26)	97.70 ± 3.54
	8 × 8	GS3O5	87.26 ± 8.30	GS3O5	88.81 ± 6.53	PSO15 (14)	89.82 ± 6.30	PSO15 (14)	98.19 ± 1.91
		GS4O6	88.47 ± 2.74	GS4O6	89.07 ± 5.59	PSO24 (18)	89.83 ± 6.24	PSO24 (20)	**98.80 ± 2.76**
		GS5O8	91.67 ± 6.97	GS5O8	**94.62 ± 3.10**	PSO40 (38)	89.63 ± 7.96	PSO40 (34)	97.08 ± 3.35

Table 2
Performance in terms of specificity (mean ± std) for normal vs. masses (bold highlighted values are best).

ROI size	Block size	noPSO		noPSO-FSS		PSO-noFSS		PSO-FSS	
		GFB	Sp	GFB	Sp	OGFB	Sp	OGFB	Sp
64 × 64	16 × 16	GS3O5	93.51 ± 5.79	GS3O5	95.63 ± 3.49	PSO15 (14)	92.82 ± 3.56	PSO15 (13)	96.38 ± 3.67
		GS4O6	94.80 ± 5.48	GS4O6	93.85 ± 5.64	PSO24 (22)	91.26 ± 6.00	PSO24 (23)	97.20 ± 1.99
		GS5O8	94.81 ± 5.12	GS5O8	96.01 ± 4.85	PSO40 (30)	90.40 ± 5.93	PSO40 (31)	97.48 ± 3.32
	8 × 8	GS3O5	**96.95 ± 4.31**	GS3O5	93.48 ± 5.38	PSO15 (13)	**95.97 ± 4.96**	PSO15 (15)	98.13 ± 3.21
		GS4O6	96.04 ± 3.24	GS4O6	93.48 ± 6.61	PSO24 (20)	95.08 ± 4.14	PSO24 (19)	98.08 ± 2.03
		GS5O8	95.85 ± 3.70	GS5O8	95.13 ± 3.34	PSO40 (37)	94.67 ± 3.93	PSO40 (30)	98.76 ± 2.03
128 × 128	32 × 32	GS3O5	92.34 ± 3.55	GS3O5	92.35 ± 6.85	PSO15 (14)	87.55 ± 6.64	PSO15 (14)	98.46 ± 1.99
		GS4O6	93.45 ± 4.04	GS4O6	91.94 ± 5.51	PSO24 (21)	92.12 ± 3.47	PSO24 (21)	96.68 ± 2.84
		GS5O8	93.16 ± 5.45	GS5O8	89.63 ± 5.66	PSO40 (33)	93.55 ± 4.93	PSO40 (35)	98.42 ± 2.08
	16 × 16	GS3O5	94.12 ± 1.94	GS3O5	94.70 ± 3.05	PSO15 (14)	95.13 ± 3.70	PSO15 (14)	99.17 ± 1.76
		GS4O6	92.64 ± 6.10	GS4O6	94.59 ± 4.82	PSO24 (20)	94.18 ± 5.99	PSO24 (22)	**99.62 ± 1.22**
		GS5O8	95.27 ± 3.69	GS5O8	97.98 ± 2.15	PSO40 (35)	92.50 ± 3.97	PSO40 (26)	98.03 ± 3.45
	8 × 8	GS3O5	94.47 ± 4.34	GS3O5	96.38 ± 3.39	PSO15 (14)	93.92 ± 5.24	PSO15 (14)	98.49 ± 1.98
		GS4O6	94.21 ± 5.14	GS4O6	94.62 ± 3.65	PSO24 (18)	93.96 ± 3.02	PSO24 (20)	98.90 ± 2.48
		GS5O8	93.56 ± 3.86	GS5O8	**98.34 ± 2.22**	PSO40 (38)	95.06 ± 4.54	PSO40 (34)	98.95 ± 2.44

Table 3
Performance in terms of accuracy (mean ± std) for normal vs. masses (bold highlighted values are best).

ROI size	Block size	noPSO		noPSO-FSS		PSO-noFSS		PSO-FSS	
		GFB	Acc. (%)	GFB	Acc. (%)	OGFB	Acc. (%)	OGFB	Acc. (%)
64 × 64	16 × 16	GS3O5	90.44 ± 4.73	GS3O5	90.05 ± 4.82	PSO15 (14)	91.40 ± 3.50	PSO15 (13)	95.90 ± 2.15
		GS4O6	92.76 ± 3.23	GS4O6	92.59 ± 3.35	PSO24 (22)	90.45 ± 3.65	PSO24 (23)	96.48 ± 1.80
		GS5O8	94.33 ± 3.64	GS5O8	94.52 ± 3.18	PSO40 (30)	84.15 ± 13.10	PSO40 (31)	96.09 ± 2.62
	8 × 8	GS3O5	**94.33 ± 3.00**	GS3O5	92.18 ± 3.35	PSO15 (13)	92.02 ± 4.37	PSO15 (15)	97.46 ± 3.92
		GS4O6	93.94 ± 2.53	GS4O6	92.78 ± 3.57	PSO24 (20)	92.19 ± 2.76	PSO24 (19)	97.85 ± 1.95
		GS5O8	94.13 ± 3.82	GS5O8	92.78 ± 3.18	PSO40 (37)	92.38 ± 2.53	PSO40 (30)	97.86 ± 1.42
128 × 128	32 × 32	GS3O5	91.40 ± 3.25	GS3O5	90.43 ± 4.16	PSO15 (14)	85.77 ± 4.96	PSO15 (14)	96.87 ± 2.48
		GS4O6	92.59 ± 2.85	GS4O6	90.03 ± 4.92	PSO24 (21)	88.10 ± 3.99	PSO24 (21)	96.48 ± 1.55
		GS5O8	92.58 ± 3.02	GS5O8	90.06 ± 3.55	PSO40 (33)	91.62 ± 4.63	PSO40 (35)	96.29 ± 2.15
	16 × 16	GS3O5	91.20 ± 2.83	GS3O5	92.78 ± 2.05	PSO15 (14)	92.01 ± 3.92	PSO15 (14)	97.46 ± 3.07
		GS4O6	91.59 ± 4.26	GS4O6	92.78 ± 3.18	PSO24 (20)	92.20 ± 4.00	PSO24 (22)	98.05 ± 1.59
		GS5O8	92.76 ± 3.60	GS5O8	93.94 ± 2.85	PSO40 (35)	91.60 ± 3.45	PSO40 (26)	97.85 ± 1.95
	8 × 8	GS3O5	90.83 ± 4.77	GS3O5	92.58 ± 3.66	PSO15 (14)	92.01 ± 3.18	PSO15 (14)	98.25 ± 1.43
		GS4O6	91.40 ± 2.81	GS4O6	91.79 ± 3.54	PSO24 (18)	91.99 ± 2.69	PSO24 (20)	**98.82 ± 2.48**
		GS5O8	92.57 ± 4.42	GS5O8	**96.48 ± 2.40**	PSO40 (38)	**92.58 ± 3.30**	PSO40 (34)	98.05 ± 2.25

4.3. False positive reduction problem

In this section, our experimental results for the first classification problem (suspicious normal vs. masses) are presented and discussed. Although most of the benign masses are structurally closer to the malignant masses, there are cases that are structurally closer to the normal masses; this makes the discrimination task challenging. In Table 1, the columns titled OGFB correspond to PSO15, PSO24 and PSO40 while the columns titled NGFB correspond to GS3O5, GS4O6 and GS5O8. The value inside the parentheses on the right side of PSO15 (third row, seventh column of Table 1) is the number of Gabor filters finally selected using the incremental clustering algorithm. In Tables 1–4 , our experimental results are given for the four feature extraction strategies in terms of the four performance measures: sensitivity, specificity, accuracy and Az, respectively. The best results both in terms of average accuracy (**98.82 ± 2.48**) and average Az (**0.99 ± 0.02**) were obtained with PSO24(20), a block size of 8 × 8 pixel and an ROI resolution of 128 × 128 pixels. Considering the results in Tables 1–4, three points are noteworthy. First, the OGFBs yielded a more com-

Table 4
Performance in terms of Az (mean ± std) for normal vs. masses (bold highlighted values are best).

ROI size	Block size	noPSO		noPSO-FSS		PSO-noFSS		PSO-FSS	
		GFB	Az.	GFB	Az.	OGFB	Az.	OGFB	Az.
64 × 64	16 × 16	GS3O5	0.91 ± 0.05	GS3O5	0.90 ± 0.07	PSO15 (14)	0.91 ± 0.03	PSO15 (13)	0.96 ± 0.02
		GS4O6	0.92 ± 0.04	GS4O6	0.93 ± 0.03	PSO24 (22)	0.91 ± 0.04	PSO24 (23)	0.97 ± 0.02
		GS5O8	**0.95 ± 0.03**	GS5O8	0.94 ± 0.03	PSO40 (30)	0.86 ± 0.11	PSO40 (31)	0.97 ± 0.03
	8 × 8	GS3O5	0.94 ± 0.03	GS3O5	0.93 ± 0.04	PSO15 (13)	0.91 ± 0.06	PSO15 (15)	0.98 ± 0.04
		GS4O6	0.95 ± 0.04	GS4O6	0.92 ± 0.04	PSO24 (20)	0.93 ± 0.03	PSO24 (19)	0.97 ± 0.02
		GS5O8	**0.95 ± 0.03**	GS5O8	0.93 ± 0.03	PSO40 (37)	0.93 ± 0.04	PSO40 (30)	0.98 ± 0.02
128 × 128	32 × 32	GS3O5	0.92 ± 0.03	GS3O5	0.92 ± 0.04	PSO15 (14)	0.87 ± 0.05	PSO15 (14)	0.97 ± 0.03
		GS4O6	0.93 ± 0.04	GS4O6	0.90 ± 0.05	PSO24 (21)	0.87 ± 0.04	PSO24 (21)	0.96 ± 0.03
		GS5O8	0.92 ± 0.05	GS5O8	0.89 ± 0.04	PSO40 (33)	0.91 ± 0.04	PSO40 (35)	0.96 ± 0.02
	16 × 16	GS3O5	0.90 ± 0.04	GS3O5	0.92 ± 0.03	PSO15 (14)	0.93 ± 0.05	PSO15 (14)	0.98 ± 0.03
		GS4O6	0.91 ± 0.06	GS4O6	0.93 ± 0.03	PSO24 (20)	0.91 ± 0.05	PSO24 (22)	**0.99 ± 0.02**
		GS5O8	0.93 ± 0.04	GS5O8	0.94 ± 0.03	PSO40 (35)	0.93 ± 0.05	PSO40 (26)	0.98 ± 0.02
	8 × 8	GS3O5	0.91 ± 0.04	GS3O5	0.93 ± 0.03	PSO15 (14)	0.92 ± 0.04	PSO15 (14)	0.98 ± 0.02
		GS4O6	0.89 ± 0.05	GS4O6	0.91 ± 0.04	PSO24 (18)	**0.94 ± 0.04**	PSO24 (20)	**0.99 ± 0.02**
		GS5O8	0.93 ± 0.05	GS5O8	**0.97 ± 0.03**	PSO40 (38)	0.93 ± 0.05	PSO40 (34)	0.98 ± 0.03

pact filter bank (e.g., 20 for the best case PSO24, four redundant filters were removed, which results in a smaller feature space). Second, OGFBs with FSS gave the best performance in almost all cases. Third, GFBs with FSS yielded higher performance than those without FSS.

For a fair comparison between OGFBs and NGFBs, a similar experimental setup was used with the same GFB sizes. For example, using a ROI resolution of 64 × 64 and a block size of 8 × 8 pixels, the performance of PSO15 was compared with that of GS3O5. It is obvious from the statistics given in Tables 3 and 4, that the PSO-FSS method performs significantly better than the corresponding noPSO and noPSO-FSS methods in almost all the cases. It can be observed that even the best results of noPSO are lower than the average results of PSO-FSS. Rather, the worst case results of PSO-FSS are almost comparable to the best results of noPSO and noPSO-FSS methods.

To statistically analyze the difference between NGFB and OGFB with and without FSS, we conducted two tailed paired Wilcoxon signed rank test (SRT) at the 0.05 significance level in terms of all performance measures given in Tables 1–4. This test is a non-parametric alternative to the paired Student's *t*-test that does not make any assumption about the distribution of the measurements being compared. Wilcoxon SRT compares two methods by ranking the pair-wise differences in their performances by absolute value and enumerates the sums for the ranks. The null hypothesis is that the difference between the mean performance values of the two competing algorithms is zero ($H_0 = \mu_1 - \mu_2 = 0$) which is rejected if the observed *p*-value ≤ significance level (S.L). The smaller the *p*-value, the higher the difference between the mean performances of the methods. If the null hypothesis is rejected, a ▲ symbol is placed as an indication that the first method has performed significantly better than the second method (e.g., in the hypothesis 'algo1 vs. algo2, algo1 is the first algorithm); alternatively, a ▼ symbol is placed if the second method has significantly outperformed the first method. The significant differences at the selected significance level are indicated in bold.

According to the statistics of Table 5, OGFBs with FSS (i.e., PSO-FSS method) is significantly better than NGFBs with FSS (i.e., noPSO-FSS method) in terms of all performance measures (in all the cases). The scores and *p*-values for the hypothesis 'PSO-FSS vs. noPSO-FSS' are similar (in all the cases) indicating that the performance of OGFBs is better than that of the corresponding NGFBs. The statistical comparison between PSO-noFSS and noPSO methods reveals that there is no significant difference in terms of average specificity and Az values. However, the noPSO method significantly outperformed the PSO-noFSS in terms of accuracy and sensitivity. It is thus concluded that OGFBs with FSS are significantly better than NGFBs with and without FSS (Table 5).

Table 5
Statistical comparison between OGFBs and NGFBs with and without FSS based on 2-sided paired Wilcoxon signed rank test at 0.05 significance level (normal vs. masses, bold highlighted values indicate significant difference).

Hypothesis	Measure	Score	*p*
PSO-noFSS vs. noPSO	Sensitivity	**▼23/97**	$\mathbf{3.53 \times 10^{-2}}$
	Specificity	30/90	9.46×10^{-2}
	Accuracy	**▼23.5/96.5**	$\mathbf{3.66 \times 10^{-2}}$
	Az. value	17/49	1.74×10^{-1}
PSO-FSS vs. noPSO-FSS	Sensitivity	**▲120/0**	$\mathbf{6.10 \times 10^{-5}}$
	Specificity	**▲120/0**	$\mathbf{6.10 \times 10^{-5}}$
	Accuracy	**▲120/0**	$\mathbf{6.10 \times 10^{-5}}$
	Az. value	**▲120/0**	$\mathbf{6.10 \times 10^{-5}}$

4.4. Discrimination between benign and malignant masses

This section presents the results for a comparatively more difficult classification problem (i.e., the discrimination between benign and malignant masses) in Tables 6–9. This discrimination task is relatively hard due to highly identical microstructures and similar textural patterns of the two classes (benign and malignant). It is therefore expected to have a decline in the recognition rate as depicted in Table 8, when compared to the results presented in Table 3. Tables 8 and 9 indicate that for benign-malignant discrimination, the best average accuracy of (**91.81 ± 3.74**) and best average Az value of (**0.92 ± 0.04**) were obtained with GS5O8, an ROI size of 128 × 128 and a block size of 8 × 8 pixels and the noPSO-FSS method. However, using the same parameter configuration (ROI size, block size and GFB size), almost similar results (i.e., an average accuracy of (**91.61 ± 4.20**) and an average Az value of (**0.92 ± 0.05**)) were obtained with the PSO-FSS method with only 32 Gabor filters (i.e., PSO40 (32)) which contain 8 filters less than GS5O8 (40) that is used by the noPSO-FSS method. This 20% reduction in GFB size is extremely helpful in reducing the computational processing time during filtering the ROIs for extracting features and eventually during learning of the classification model. We can also observe from the results presented in Table 8 that the performance of PSO-FSS method improves for large ROI sizes. We therefore conducted an experiment using ROI resolution of 256 × 256 and block size of 8 × 8 pixels with PSO24(20). In this experiment with the PSO-FSS method, we obtained an average accuracy of (**93.95 ± 3.85**) and an average Az value of (**0.948 ± 0.043**) with just 20 Gabor filters.

As reported previously (for normal vs. masses case), OGFBs are again more compact and recognition rate is up to the mark when

Table 6
Performance in terms of sensitivity (mean ± std) for benign vs. malignant (bold highlighted values are best).

ROI size	Block size	noPSO		noPSO-FSS		PSO-noFSS		PSO-FSS	
		GFB	Sn	GFB	Sn	OGFB	Sn	OGFB	Sn
64 × 64	16 × 16	GS3O5	79.29 ± 4.88	GS3O5	79.19 ± 6.26	PSO15 (13)	79.76 ± 8.63	PSO15 (13)	83.00 ± 5.45
		GS4O6	80.78 ± 6.49	GS4O6	81.04 ± 9.34	PSO24 (21)	80.06 ± 6.18	PSO24 (18)	84.69 ± 7.24
		GS5O8	81.97 ± 10.37	GS5O8	82.76 ± 7.56	PSO40 (34)	79.04 ± 8.24	PSO40 (32)	86.85 ± 8.98
	8 × 8	GS3O5	79.58 ± 9.80	GS3O5	88.08 ± 8.13	PSO15 (12)	80.50 ± 7.95	PSO15 (14)	90.69 ± 6.46
		GS4O6	80.08 ± 8.69	GS4O6	84.68 ± 7.00	PSO24 (19)	80.05 ± 7.56	PSO24 (18)	89.65 ± 3.50
		GS5O8	83.54 ± 6.80	GS5O8	84.01 ± 8.65	PSO40 (29)	80.70 ± 6.76	PSO40 (34)	89.28 ± 7.86
128 × 128	32 × 32	GS3O5	81.11 ± 9.48	GS3O5	79.46 ± 7.02	PSO15 (15)	82.44 ± 6.66	PSO15 (13)	84.45 ± 4.84
		GS4O6	80.49 ± 6.05	GS4O6	81.28 ± 8.79	PSO24 (20)	77.91 ± 7.18	PSO24 (21)	85.44 ± 6.48
		GS5O8	80.03 ± 9.56	GS5O8	83.31 ± 6.83	PSO40 (33)	80.55 ± 6.36	PSO40 (34)	89.94 ± 5.19
	16 × 16	GS3O5	84.40 ± 4.78	GS3O5	80.76 ± 5.45	PSO15 (13)	82.05 ± 3.99	PSO15 (13)	90.20 ± 4.53
		GS4O6	81.20 ± 7.48	GS4O6	83.45 ± 7.02	PSO24 (21)	**84.41 ± 7.52**	PSO24 (22)	90.02 ± 6.04
		GS5O8	81.18 ± 8.06	GS5O8	84.81 ± 10.40	PSO40 (30)	81.82 ± 7.70	PSO40 (34)	88.07 ± 6.65
	8 × 8	GS3O5	78.52 ± 8.23	GS3O5	86.83 ± 5.59	PSO15 (13)	82.21 ± 9.99	PSO15 (15)	89.63 ± 5.59
		GS4O6	79.77 ± 8.40	GS4O6	85.43 ± 5.89	PSO24 (22)	81.17 ± 9.68	PSO24 (20)	**91.79 ± 7.10**
		GS5O8	**85.09 ± 9.42**	GS5O8	**92.83 ± 5.92**	PSO40 (29)	81.30 ± 7.99	PSO40 (32)	89.94 ± 7.04

Table 7
Performance in terms of specificity (mean ± std) for benign vs. malignant (bold highlighted values are best).

ROI size	Block size	noPSO		noPSO-FSS		PSO-noFSS		PSO-FSS	
		GFB	Sp	GFB	Sp	OGFB	Sp	OGFB	Sp
64 × 64	16 × 16	GS3O5	76.94 ± 8.54	GS3O5	80.48 ± 10.45	PSO15 (13)	79.37 ± 8.23	PSO15 (13)	85.49 ± 7.10
		GS4O6	79.29 ± 7.44	GS4O6	79.89 ± 10.09	PSO24 (21)	82.61 ± 8.45	PSO24 (18)	82.59 ± 4.94
		GS5O8	78.08 ± 8.72	GS5O8	82.19 ± 6.19	PSO40 (34)	82.83 ± 10.38	PSO40 (32)	86.71 ± 6.54
	8 × 8	GS3O5	82.86 ± 5.37	GS3O5	70.52 ± 9.66	PSO15 (12)	83.22 ± 10.33	PSO15 (14)	84.70 ± 5.64
		GS4O6	81.74 ± 9.13	GS4O6	82.53 ± 10.01	PSO24 (19)	81.32 ± 6.44	PSO24 (18)	88.57 ± 3.63
		GS5O8	83.97 ± 5.95	GS5O8	83.94 ± 7.53	PSO40 (29)	**85.02 ± 6.40**	PSO40 (34)	87.51 ± 6.33
128 × 128	32 × 32	GS3O5	80.86 ± 6.47	GS3O5	81.42 ± 8.60	PSO15 (15)	83.64 ± 5.13	PSO15 (13)	84.03 ± 9.46
		GS4O6	81.69 ± 10.62	GS4O6	78.55 ± 10.36	PSO24 (20)	81.05 ± 5.60	PSO24 (21)	85.42 ± 5.71
		GS5O8	83.15 ± 5.38	GS5O8	83.16 ± 4.06	PSO40 (33)	81.90 ± 9.18	PSO40 (34)	85.62 ± 7.95
	16 × 16	GS3O5	82.02 ± 8.78	GS3O5	82.66 ± 4.52	PSO15 (13)	84.79 ± 8.03	PSO15 (13)	85.50 ± 9.03
		GS4O6	82.00 ± 5.38	GS4O6	85.81 ± 6.07	PSO24 (21)	84.97 ± 5.06	PSO24 (22)	88.54 ± 8.18
		GS5O8	82.37 ± 8.32	GS5O8	85.67 ± 8.15	PSO40 (30)	83.86 ± 4.77	PSO40 (34)	91.48 ± 7.82
	8 × 8	GS3O5	82.88 ± 7.88	GS3O5	86.83 ± 6.27	PSO15 (13)	80.39 ± 6.34	PSO15 (15)	88.43 ± 11.28
		GS4O6	83.97 ± 6.22	GS4O6	88.40 ± 5.32	PSO24 (22)	79.74 ± 7.88	PSO24 (20)	89.40 ± 5.62
		GS5O8	**85.97 ± 6.58**	GS5O8	**90.92 ± 3.28**	PSO40 (29)	81.22 ± 7.05	PSO40 (32)	**92.99 ± 3.51**

Table 8
Performance in terms of accuracy (mean ± std) for benign vs. malignant (bold highlighted values are best).

ROI size	Block size	noPSO		noPSO-FSS		PSO-noFSS		PSO-FSS	
		GFB	Acc. (%)	GFB	Acc. (%)	OGFB	Acc. (%)	OGFB	Acc. (%)
64 × 64	16 × 16	GS3O5	78.11 ± 5.84	GS3O5	79.86 ± 7.32	PSO15 (13)	79.30 ± 6.04	PSO15 (13)	84.58 ± 4.70
		GS4O6	80.07 ± 5.06	GS4O6	80.44 ± 5.82	PSO24 (21)	81.04 ± 4.84	PSO24 (18)	83.79 ± 2.47
		GS5O8	80.08 ± 7.83	GS5O8	82.41 ± 4.91	PSO40 (34)	80.88 ± 3.97	PSO40 (32)	86.72 ± 5.29
	8 × 8	GS3O5	81.23 ± 6.58	GS3O5	79.11 ± 3.63	PSO15 (12)	81.63 ± 3.28	PSO15 (14)	87.49 ± 2.50
		GS4O6	80.83 ± 6.11	GS4O6	83.21 ± 5.60	PSO24 (19)	80.65 ± 6.84	PSO24 (18)	89.07 ± 2.25
		GS5O8	83.77 ± 4.38	GS5O8	83.79 ± 5.58	PSO40 (29)	82.80 ± 5.63	PSO40 (34)	88.28 ± 5.63
128 × 128	32 × 32	GS3O5	81.04 ± 5.96	GS3O5	80.45 ± 5.64	PSO15 (15)	82.82 ± 3.98	PSO15 (13)	84.36 ± 5.01
		GS4O6	81.05 ± 5.32	GS4O6	80.10 ± 6.59	PSO24 (20)	79.50 ± 4.22	PSO24 (21)	85.34 ± 5.42
		GS5O8	81.63 ± 4.93	GS5O8	83.19 ± 2.99	PSO40 (33)	81.43 ± 4.71	PSO40 (34)	87.71 ± 5.99
	16 × 16	GS3O5	83.21 ± 4.63	GS3O5	81.45 ± 4.20	PSO15 (13)	83.62 ± 4.11	PSO15 (13)	88.48 ± 3.09
		GS4O6	81.64 ± 4.43	GS4O6	84.56 ± 6.45	PSO24 (21)	**84.57 ± 5.01**	PSO24 (22)	89.05 ± 4.99
		GS5O8	81.83 ± 7.19	GS5O8	85.16 ± 4.00	PSO40 (30)	83.02 ± 5.19	PSO40 (34)	89.47 ± 5.64
	8 × 8	GS3O5	80.67 ± 5.37	GS3O5	86.52 ± 3.71	PSO15 (13)	81.04 ± 4.20	PSO15 (15)	88.65 ± 6.48
		GS4O6	81.86 ± 5.76	GS4O6	86.91 ± 2.46	PSO24 (22)	80.08 ± 5.83	PSO24 (20)	90.63 ± 4.00
		GS5O8	**85.53 ± 5.43**	GS5O8	**91.81 ± 3.74**	PSO40 (29)	80.84 ± 5.85	PSO40 (32)	**91.61 ± 4.20**

compared with the corresponding NGFBs with and without FSS. Table 10 shows the results of the Wilcoxon signed rank test for the statistical differences between the performances of OGFB and NGFB with and without FSS. It is obvious that the PSO-FSS method performs significantly better than the noPSO-FSS in terms of all performance measures. PSO-noFSS is only significantly defeated by noPSO when compared on the basis of average sensitivity. This implies that optimization of GFBs can lead to better results in terms of different performance metrics and the difference is statistically significant. Moreover, the generalization ability of SVM is better with PSO-FSS. The results obtained with PSO-FSS using different parameter configurations (e.g., ROI size, block size, initial GFB size) are consistent in the sense that the difference between best and worst cases is very low. This indicates that the performance of PSO-FSS is less dependent on parameters setting and thus more attractive choice.

Table 9
Performance in terms of Az (mean ± std) for benign vs. malignant (bold highlighted values are best).

ROI size	Block size	noPSO		noPSO-FSS		PSO-noFSS		PSO-FSS	
		GFB	Az.	GFB	Az.	OGFB	Az.	OGFB	Az.
64 × 64	16 × 16	GS3O5	0.79 ± 0.06	GS3O5	0.81 ± 0.08	PSO15 (13)	0.81 ± 0.05	PSO15 (13)	0.85 ± 0.05
		GS4O6	0.81 ± 0.05	GS4O6	0.80 ± 0.08	PSO24 (21)	0.82 ± 0.05	PSO24 (18)	0.84 ± 0.04
		GS5O8	0.81 ± 0.07	GS5O8	0.82 ± 0.04	PSO40 (34)	0.82 ± 0.04	PSO40 (32)	0.88 ± 0.04
	8 × 8	GS3O5	0.82 ± 0.07	GS3O5	0.80 ± 0.03	PSO15 (12)	0.81 ± 0.05	PSO15 (14)	0.88 ± 0.04
		GS4O6	0.80 ± 0.07	GS4O6	0.82 ± 0.06	PSO24 (19)	0.80 ± 0.08	PSO24 (18)	0.87 ± 0.05
		GS5O8	0.82 ± 0.06	GS5O8	0.86 ± 0.06	PSO40 (29)	0.82 ± 0.05	PSO40 (34)	0.88 ± 0.07
128 × 128	32 × 32	GS3O5	0.81 ± 0.06	GS3O5	0.80 ± 0.05	PSO15 (15)	0.83 ± 0.05	PSO15 (13)	0.84 ± 0.05
		GS4O6	0.83 ± 0.04	GS4O6	0.81 ± 0.09	PSO24 (20)	0.80 ± 0.04	PSO24 (21)	0.85 ± 0.06
		GS5O8	0.81 ± 0.07	GS5O8	0.83 ± 0.03	PSO40 (33)	0.81 ± 0.07	PSO40 (34)	0.89 ± 0.05
	16 × 16	GS3O5	0.85 ± 0.05	GS3O5	0.81 ± 0.04	PSO15 (13)	0.83 ± 0.05	PSO15 (13)	0.87 ± 0.04
		GS4O6	0.83 ± 0.06	GS4O6	0.85 ± 0.07	PSO24 (21)	**0.86 ± 0.06**	PSO24 (22)	0.91 ± 0.05
		GS5O8	0.82 ± 0.08	GS5O8	0.85 ± 0.04	PSO40 (30)	0.82 ± 0.07	PSO40 (34)	0.90 ± 0.06
	8 × 8	GS3O5	0.81 ± 0.05	GS3O5	0.88 ± 0.05	PSO15 (13)	0.80 ± 0.06	PSO15 (15)	0.90 ± 0.05
		GS4O6	0.79 ± 0.06	GS4O6	0.85 ± 0.03	PSO24 (22)	0.79 ± 0.06	PSO24 (20)	0.92 ± 0.04
		GS5O8	**0.87 ± 0.05**	GS5O8	**0.92 ± 0.04**	PSO40 (29)	0.82 ± 0.05	PSO40 (32)	**0.92 ± 0.05**

Table 10
Statistical comparison between OGFBs and NGFB with and without FSS based on 2-sided paired Wilcoxon signed rank test at 0.05 significance level (benign vs. malign, bold highlighted values indicate significant difference).

Hypothesis	Measure	Score	p
PSO-noFSS vs. noPSO	Sensitivity	**▼13/107**	$\mathbf{5.37 \times 10^{-3}}$
	Specificity	74.5/45.5	4.30×10^{-1}
	Accuracy	71/49	5.52×10^{-1}
	Az. value	25.5/29.5	8.89×10^{-1}
PSO-FSS vs. noPSO-FSS	Sensitivity	**▲117/3**	$\mathbf{3.05 \times 10^{-4}}$
	Specificity	**▲120/0**	$\mathbf{6.10 \times 10^{-5}}$
	Accuracy	**▲119/1**	$\mathbf{1.22 \times 10^{-4}}$
	Az. value	**▲105/0**	$\mathbf{1.22 \times 10^{-4}}$

4.5. Reduction in computation cost

A close look at the results for both classification problems presented in the above sections indicates that the optimization procedure reduces the number of filters in a GB with FSS and improves the classification performance. With less number of filters in a GB will take less time for feature extraction and also, because of reduction in feature space, classification time will also reduce. For example, in case of PSO40 the number of filters is reduced to 32, which in turn decreases the computational cost for Gabor feature extraction by 20%.

4.6. Comparison with existing methods

In the last sections, it has been shown in detail that the optimization technique has significant impact on improving the classification accuracy and reducing the computational cost. In this section, we give a comparison of the mass classification method based on this optimization technique with existing methods. It is to be noted that it is rather difficult to objectively compare different methods for the two classification problems due to many factors. For example, which mammogram database was used for evaluation? Even if the same database was employed, were the same mammograms selected for evaluation? How many samples were used? Which evaluation approach (validation methodology, training and testing set formation with different percentages of ROIs) was used? Were the ratios of ROIs for different classes (e.g., normal, malignant and benign) the same? Even if other methods are implemented and evaluated on the same dataset it might still not be a fair comparison because the tuning of parameters involved in different methods might not necessarily be the same.

In any case, to give a general trend of the performance and give comparison with state-of-the-art methods in terms of accuracy and Az, we have compiled information from various studies as shown in Table 11. The quantities that are not reported in the literature are indicated with a dash symbol. For some methods, standard deviation values are not available. For the two problems, only the best case mean and standard deviation results are reported for all the

Table 11
Comparison with state-of-the-art methods based on average Acc. and Az values (bold highlighted values are best).

Problem	Research work	Database	#ROIs	Avg. Acc. (%) ± std	Avg. Az ± std
Normal vs. masses	Moayedi et al. [57]	MIAS	90	85.9 ± 0.03	–
	Costa et al. [33]	DDSM	5090	90.07	–
	Ioan and Gacsadi [34]	MIAS	322	84.37	0.79
	Geraldo et al. [35]	DDSM	584	99.39(max)	1.00(max)
	Lladó et al. [3]	DDSM	512	–	0.94 ± 0.02
	Nguyen et al. [20]	MIAS	–	–	0.93
	Reyad et al. [22]	DDSM	512	98.63(max)	–
	Oliveira et al. [19]	DDSM	3404	98.88(max)	–
	Hussain et al. [31]	DDSM	512	94.92 ± 2.30	0.96 ± 0.02
	Proposed method	DDSM	512	**98.82 ± 2.48**	**0.99 ± 0.02**
Benign vs. malignant	Geraldo et al. [35]	DDSM	584	88.31	0.89
	Moayedi et al. [57]	MIAS	90	87.00 ± 0.008	–
	Costa et al. [33]	DDSM	3240	84.22	–
	Ioan and Gacsadi [34]	MIAS	114	78.26	0.78
	Li et al. [26]	DDSM	114	85.96	–
	Nanni et al. [25]			0.97	
	Rouhi et al. [24]	DDSM		82.06	–
	Hussain et al. [31]	DDSM	512	85.53 ± 5.43	0.87 ± 0.05
	Proposed method	DDSM	512	**93.95 ± 3.85**	**0.948 ± 0.043**

methods being compared. The results in Table 11 indicate that the mass classification method based on the proposed optimization technique is comparable with the existing.

5. Conclusion

In this paper, a technique has been presented for the optimization of GFBs which are used to extract multi-scale and multi-orientation local texture micro-patterns from mammograms. In particular, PSO has been used for the optimization of GFBs (a design problem), and incremental clustering has been exploited to eliminate redundant Gabor filters from a GFB (a selection problem). The resultant OGFBs are employed to extract Gabor features, which have been evaluated over 768 ROIs using SVM for two classification problems. The results are encouraging as the OGFBs become more compact. The model compactness indirectly implies that the feature space is low dimensional and thus the classification model has better computational efficiency and better generalization.

The effectiveness of the proposed optimization technique was evaluated for the two mass classification problems i.e., false positive reduction problem and discrimination between benign and malignant masses using well-known performance measures. In addition, statistical tests based on different performance metrics were conducted using the Wilcoxon signed rank test. The evaluation confirmed that the overall performance of PSO-FSS method is better than other methods (noPSO, PSO-noFSS, noPSO-FSS) and the difference is statistically significant. For the two classification problems, PSO-FSS achieved encouraging results, reported as 98.82%, with a standard deviation of 0.99, for normal vs. masses and 93.95%, with a standard deviation of 0.95, for benign vs. malignant. The experimental results are comparable with those by state-of-the-art results. It validates that the proposed optimization technique is effective in improving classification rate and reducing the computational cost. Though we examined the effectiveness of the proposed technique with a specific Gabor feature extraction method, it is general and can be employed with other Gabor feature extraction methods.

There are several future avenues in order to enhance the effect of the proposed technique. We employed SBMLR, other feature subset selection techniques can further enhance the effectiveness of the proposed optimization method. It will be interesting to investigate the performance of the proposed method in more complex problem scenarios (e.g., recognition and identification of breast abnormalities like breast structural disorder). Other optimization strategies (e.g., Cuckoo optimization [58]) can also be investigated.

Acknowledgment

This work is supported by the National Plan for Science and Technology, King Saud University, Riyadh, Saudi Arabia under project number 8-INF325-02.

References

[1] S.F. Altekruse, C.L. Kosary, M. Krapcho, et al., SEER Cancer Statistics Review, 1975–2007, National Cancer Institute, Bethesda, MD, 2010.
[2] M. Hussain, False positive reduction using Gabor feature subset selection, 2013 International Conference on Information Science and Applications (ICISA) vol. 0 (2013) 1–5 http://dx.doi.ieeecomputersociety.org/10.1109/ICISA.2013.6579383.
[3] X. Lladó, A. Oliver, J. Freixenet, R. Martí, J. Martí, A textural approach for mass false positive reduction in mammography, Comput. Med. Imaging Graph. 33 (6) (2009) 415–422.
[6] M. Hussain, S. Khan, G. Muhammad, N. Noreen, B. Bebis, Effective extraction of Gabor features for false positive reduction and mass classification in mammography, Appl. Math. Info. Sci. 8 (April (1L)) (2014) 397–412 (An International Journal).
[7] M. Hussain, S. Khan, G. Muhammad, G. Bebis, Mass detection in digital mammograms using optimized Gabor filter bank, Adv. Vis. Comput. 7432 (2012) 82–91 (LNCS Springer, ISVC 2012, July 16–18, 2012, Crete, Greece).
[8] K. Ville, K. Joni-Kristian, Simple Gabor feature space for invariant object recognition, Pattern Recognit. Lett. 25 (3) (2004) 311–318.
[9] B.S. Manjunath, Texture features for browsing and retrieval of image data, IEEE Trans. Pattern Anal. Mach. Intell. 18 (8) (1996) 837–842.
[10] A. Oliver, J. Freixenet, J. Martí, et al., A review of automatic mass detection and segmentation in mammographic images, Med. Image Anal. 14 (2) (2010) 87–110.
[11] A.R. Domínguez, A.K. Nandi, Towards breast cancer diagnosis based on automated segmentation of masses in mammograms, Pattern Recognit. 42 (6) (2009) 1138–1148.
[12] J. Tang, R.M. Rangayyan, J. Xu, et al., Computer-aided detection and diagnosis of breast cancer with mammography: recent advances, IEEE Trans. Info. Technol. Biomed. 13 (2) (2009) 236–251.
[13] M. Elter, A. Horsch, CADx of mammographic masses and clustered micro calcifications: a review, Med. Phys. 36 (6) (2009) 2052–2068.
[14] N. Székely, N. Tóth, B. Pataki, A hybrid system for detecting masses in mammographic images, IEEE Trans. Instrum. Meas. 55 (3) (2006) 944–952.
[15] Y. Wang, X. Gao, J. Li, A Feature Analysis Approach to Mass Detection in Mammography based on RF-SVM, IEEE International Conference on Image Processing, 2007, pp. 9–12.
[16] A.P. Nunes, A.C. Silva, A.C. de Paiva, Detection of masses in mammographic images using geometry, Simpson's Diversity Index and SVM, Int. J. Signal Imaging Syst. Eng. 3 (1) (2010) 43–51.
[17] D. Wei, H. Chan, M. Helvie, B. Sahiner, N. Petrick, D. Adler, M. Goodsitt, Classification of mass and normal breast tissue on digital mammograms: multiresolution texture analysis, Med. Phys. 22 (9) (1995) 1501–1513.
[18] W.B. Sampaio, E.M. Diniz, A.C. Silva, A.C. Paiva, M. Gattass, Detection of masses in mammogram images using CNN, geostatistic functions and SVM, Comput. Biol. Med. 41 (2011) 653–664.
[19] F.S.S. Oliveira, A.O.C. Filho, A.C. Silva, A.C. Paiva, M. Gattass, Classification of breast regions as mass and non-mass based on digital mammograms using taxonomic indexes and SVM, Comput. Biol. Med. 57 (1) (2015) 42–53.
[20] M. Nguyen, Q. Truong, D. Nguyen, T. Nguyen, An alternative approach to reduce massive false positives in mammograms using block variance of local coefficients features and support vector machine, Proc. Comput. Sci. (2013) 399–405.
[21] G.B. Junior, S.V. Rocha, M. Gattass, A.C. Silva, A.C. Paiva, A mass classification using spatial diversity approaches in mammography images for false positive reduction, Expert Syst. Appl. 40 (18) (2013) 7534–7543.
[22] Y.A. Reyad, M.A. Berbar, M. Hussain, Comparison of statistical, LBP, and multi-resolution analysis features for breast mass classification, J. Med. Syst. 38 (2014) 100, http://dx.doi.org/10.1007/s10916-014-0100-7.
[23] M. Hussain, False positive reduction in mammography using multiscale spatial weber law descriptor and support vector machines, Neural Comput. Appl. 25 (July (1)) (2014) 83–93 (Sringer-Verlag, ISSN 0941-0643, ISI Indexed).
[24] R. Rouhi, M. Jafari, S. Kasaei, P. Keshavarzian, Benign and malignant breast tumors classification based on region growing and CNN segmentation, Expert Syst. Appl. 42 (3) (2015) 990–1002.
[25] L. Nanni, S. Brahnam, A. Lumini, A very high performing system to discriminate tissues in mammograms as benign and malignant, Expert Syst. Appl. 39 (2) (2012) 1968–1971.
[26] Y. Li, H. Chen, G.K. Rohde, C. Yao, L. Cheng, Texton analysis for mass classification in mammograms, Pattern Recognit. Lett. 52 (2015) 87–93.
[27] S. Grigorescu, N. Petkov, P. Kruizinga, Comparison of texture features based on Gabor filters, IEEE Trans. Image Process. 11 (10) (2002) 1160–1167.
[28] M.R. Turner, Texture discrimination by Gabor functions, Biol. Cybern. 55 (1986) 71–82.
[29] T. Bhangale, U.B. Desai, U. Sharma, An unsupervised scheme for detection of microcalcifications on mammograms, in: Proc. of IEEE Int. Conf. on Image Proc., Vancouver, BC, Canada, 2000, pp. 184–187.
[30] G.L. Rogova, P.C. Stomper, C. Ke, Microcalcification texture analysis in a hybrid system for computer aided mammography, Proc. SPIE (1999) 1426–1433.
[31] M. Hussain, S. Khan, G. Muhammad, G. Bebis, A comparison of different Gabor features for mass classification in mammography, in: Proc. SITIS 2012, IEEE Computer Society Press, Nov. 25–29, 2012, Naples, Italy, 2012, pp. 142–148.
[32] S. Lahmiri, M. Boukadoum, Hybrid discrete wavelet transform and Gabor filter banks processing for mammogram features extraction, in: Proc. NEWCAS, France, IEEE Computer Society, 2011, pp. 53–56.
[33] D.D. Costa, L.F. Campos, A.K. Barros, Classification of breast tissue in mammograms using efficient coding, Bio-Med. Eng. On-Line 10 (2011) 55 http://www.biomedical-engineering-online.com/content/10/1/55.
[34] B. Ioan, A. Gacsadi, Directional features for automatic tumor classification of mammogram images, Biomed. Signal Process. Control 6 (4) (2011) 370–378.
[35] Braz Geraldo Junior, et al., Classification of breast tissues using Moran's index and Geary's coefficient as texture signatures and SVM, Comput. Biol. Med. 39 (2009) 1063–1072.
[36] R.M. Rangayyan, R.J. Ferrari, J.E.L. Desautels, A.F. Frère, Directional analysis of images with Gabor wavelets, Proc. of XIII Brazilian Symposium on Computer Graphics and Image SIBGRAPI (2000) 170–177.
[37] N. RiyahiAlamet, et al., Computer-aided mass detection on digitized mammograms using a novel hybrid segmentation system, Int. J. Biol. Biomed. Eng. 3 (4) (2009) 51–58.

[38] T. Bhangale, U.B. Desai, U. Sharma, An unsupervised scheme for detection of microcalcifications on mammograms, IEEE International Conference on Image Processing (2000) 184–187.
[40] J.G. Daugman, Two-dimensional spectral analysis of cortical receptive field profiles, Vis. Res. 20 (1980) 847–856.
[41] S. Yu, S. Shiguan, C. Xilin, G. Wen, Hierarchical ensemble of global and local classifiers for face recognition, IEEE Trans. Image Process. 18 (8) (2009) 1885–1896.
[42] S. Zehan, B. George, M. Ronald, Monocular precrash vehicle detection: features and classifiers, IEEE Trans. Image Process. 15 (7) (2006) 2019–2034.
[43] R.C. Eberhart, J. Kennedy, Particle Swarm Optimization, Proc. of IEEE International Conference on Neural Networks (1995) 1942–1948.
[44] R. Rajendra, D. Pratihar, Particle Swarm Optimization algorithm vs genetic algorithm to develop integrated scheme for obtaining optimal mechanical structure and adaptive controller of a robot, Intell. Control Autom. 2 (4) (2011) 430–449.
[45] Z. Gao, G. Lu, M. Liu, M. Cui, A novel risk assessment system for port state control inspection, IEEE International Conference on Intelligence and Security Informatics (2008) 242–244.
[46] L. Bo, L. Wang, L. Jiao, Training hard margin support vector machines using greedy stepwise algorithm, Lect. Notes Comput. Sci. (2005) 632–638.
[47] G.A. Iffat, S.S. Leslie, Feature subset selection in large dimensionality domains, Pattern Recognit. 43 (1) (2010) 5–13.
[48] G.C. Cawley, N.L.C. Talbot, M. Girolami, Sparse multinomial logistic regression via Bayesian L1 regularisation, in: Advances in Neural Information Processing Systems, MIT Press, Cambridge, MA, USA, 2007.
[49] http://featureselection.asu.edu/software.php.
[50] V. Vapnik, Statistical Learning Theory, Springer-Verlag, New York, NY, 1995.
[51] B.E. Boser, I.M. Guyon, V. Vapnik, A training algorithm for optimal margin classifiers, Proc. of the Fifth Annual Workshop on Computational Learning Theory (1992) 144–152.
[52] C. Burges, Tutorial on support vector machines for pattern recognition, Data Min. Knowl. Discov. 2 (2) (1998) 955–974.
[53] C.W. Hsu, C.C. Chang, C.J. Lin, A Practical Guide to Support Vector Classification, Technical Report, Department of Computer Science and Information Engineering, National Taiwan University, 2010.
[54] S. García, F. Herrera, An extension on 'Statistical Comparisons of Classifiers over Multiple Data Sets' for all pairwise coparisons, Mach. Learn. Res. 9 (2008) 2677–2694.
[55] F. Wilcoxon, Individual comparisons by ranking methods, Biometrics 1 (1945) 80–83.
[56] M. Heath, K. Bowyer, D. Kopans, R. Moore, P.J. Kegelmeyer, The digital database for screening mammography, Int. Work. Dig. Mamm. (2000) 212–218.
[57] Fatemeh Moayedi, et al., Contourlet-based mammography mass classification using the SVM family, Comput. Biol. Med. 40 (2010) 373–383.
[58] R. Rajabioun, Cuckoo optimization algorithm, Appl. Soft Comput. 11 (December (8)) (2011) 5508–5518.
[59] M. Hussain, False positive reduction in mammography using multiscale spatial weber law descriptor and support vector machines, Neural Comput. Appl. 25 (July (1)) (2014) 83–93 (Sringer-Verlag, ISSN 0941-0643).
[60] F.S.S. Oliveira, A.O.C. Filho, A.C. Silva, A.C. Paiva, M. Gattass, Classification of breast regions as mass and non-mass based on digital mammograms using taxonomic indexes and SVM, Comput. Biol. Med. 57 (1) (2015) 42–53.