Biomedical Signal Processing and Control 6 (2011) 370–378

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

Directional features for automatic tumor classification of mammogram images☆

Ioan Buciu*, Alexandru Gacsadi

Department of Electronics and Telecommunications, Faculty of Electrical Engineering and Information Technology, University of Oradea, 419987, Universitatii, 1, Romania

ARTICLE INFO

Article history:
Received 13 May 2010
Received in revised form
24 September 2010
Accepted 8 October 2010
Available online 9 November 2010

Keywords:
Gabor features
directional features
mammogram
classification
noise
proximal support vector machines

ABSTRACT

One way for breast cancer diagnosis is provided by taking radiographic (X-ray) images (termed mammograms) for suspect patients, images further used by physicians to identify potential abnormal areas thorough visual inspection. When digital mammograms are available, computer-aided based diagnostic may help the physician in having a more accurate decision. This implies automatic abnormal areas detection using segmentation, followed by tumor classification. This work aims at describing an approach to deal with the classification of digital mammograms. Patches around tumors are manually extracted to segment the abnormal areas from the remaining of the image, considered as background. The mammogram images are filtered using Gabor wavelets and directional features are extracted at different orientation and frequencies. Principal Component Analysis is employed to reduce the dimension of filtered and unfiltered high-dimensional data. Proximal Support Vector Machines are used to final classify the data. Superior mammogram image classification performance is attained when Gabor features are extracted instead of using original mammogram images. The robustness of Gabor features for digital mammogram images distorted by Poisson noise with different intensity levels is also addressed.

1. Introduction

Image manipulation is commonly used in image processing and pattern recognition field. Image manipulation may include quality enhancement, filtering, segmentation, feature selection or extraction and dimensionality reduction, to mention only a few. When it comes to image classification, it is desirable to keep discriminant features and discard non-relevant features that may negatively affects the classification performances.

Mammographic images are X-ray images of breast region displaying points with high intensities density that are suspected of being potential tumors. Early diagnostic and screening is crucial for having a successful medical treatment or cure. Typically, masses and calcium deposits are easily identified by visual inspection. These deposits appear much denser (highly attenuate X-ray) than others types of surrounding soft tissues. Malign tumors are usually associated to unusual smaller and clustered calcification. Other calcification types, including diffuse, regional, segmental or linear, correspond to benign tumors. Such calcification are termed as microcalcification. Automatic tumor classification would require the segmentation of the microcalcification area from the X-ray image, followed by recognition or classification of the segmented area into one of theses three classes: normal tissue (absence of tumor), benign or malign tumor. Automatic tumor detection is extremely challenging as the suspicious calcification or masses appear as free shape and irregular texture, so that no precise patterns can be associated to them. In addition, the presence of more or less prominent blood vessels and muscle fibers may seriously degrade the accuracy of identification or tumor recognition.

Several techniques have been proposed to analyze, detect or to extract features from mammogram images. Strickland and Hahn [1] proposes a two-stage method based on wavelet transforms for detecting and segment calcifications. In the first stage the image is decomposed into four sub-bands ((LL, LH, LH, and HH) without downsampling. Detection is next performed for the HH sub-band and the combination of LH+HL. Four octaves are computed with two inter-octave voices for finer scale resolution. The second stage improves the accuracy of segmentation, where detected pixel sites in HH and LH+HL are dilated and weighted before taking the inverse of the wavelet transform. By so doing, the microcalcifications are greatly enhanced in the resulting image and an appropriate threshold can be used to segment the tumor zone. Haar wavelets along with PCA are proposed by Swiniarski et al. [2] to extract relevant features, and rough sets methods are further employed to classify the resulting features. Recently, the authors extended the work by extracting independent component features, followed by rough sets method for feature selection and data reduction, and, ultimately, a rule-based classifier is employed for a final decision

☆ Expanded Paper from the work presented at the 2nd International Symposium on Applied Sciences in Biomedical and Communication Technologies held 24 September 2009.

* Corresponding author.
E-mail addresses: ibuciu@uoradea.ro (I. Buciu), agacsadi@uoradea.ro (A. Gacsadi).

1746-8094/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bspc.2010.10.003

[3]. Sun et al. [4] developed an approach to decompose regional (local region within an image containing the tumor area) images using 2D Quincunx wavelet transform with the first four even-decomposition levels. For every selected decomposition image, five features, i.e. mean, variance, compactness, fractal dimension and entropy are globally extracted. A decision tree classifier is further utilized to label the images. The wavelet decomposition is carried out on residual image (contrast enhanced abnormalities within a normal structure suppressed background) instead of original raw mammogram images. Zaiane et al. [5] proposed an approach where statistical set of features (including mean, variance, skewness and kurtosis are used as features for classification. To classify a set of features, association-rule based classifiers are applied. Lemaur et. al. [6] developed a method based on wavelets derived from Matzinger polynomial with high Sobolev regularity index. Wavelet decomposition is also used by Ferreira and Borges [7,8] to decompose a mammogram image into different frequency sub-bands. By imposing a certain threshold, some low frequency coefficients are selected followed by a class signature building process. This procedure is applied to the basis image. When a test image (unknown mammogram image) comes, the same steps are performed, i.e. wavelet decomposition, low frequency coefficient selection and class signature assignment. To classify the unknown image, distances between the signature corresponding to the unknown image and each of the signature corresponding to the basis images are computed. Finally, the unknown mammogram is classified based on the lowest distance. Several distance metrics are investigated, including the Euclidean, norm of absolute value and Mahalanobis. To classify mammogram images, Wei et al. [9] employed Support Vector Machines (SVMs), Relevance Vector Machines (RVM) and Kernel Fisher Discriminant (KFD). Sheshadri and Kandaswamy [10] attempted to classify breast tissue using simple image statistics such as its intensity level of histogram (including mean, standard deviation, smoothness, third moment, uniformity and entropy). Amara and Qader [11] developed a system that combines dimensionality reduction module (using principal component analysis), a feature extraction module (using independent component analysis) and a feature subset selection module (using rough set model) to reduce the effect of data inconsistency. Finally, a fuzzy classifier is integrated into the system to label subimages into normal or abnormal regions. K-means algorithm is employed by Martins et al. [12] to segment a mammogram images and co-occurence matrix to describe the texture of segmented structures. However, the goal of that work was rather to discriminate between masses and non-masses structure with the help of support vector machines. Although promising, their approach is heavily dependent on the initialization of the number of clusters used by K-means. Moreover, the computation of co-occurence matrix is quite time consuming limiting the approach for real-time applications.

More related techniques (including both segmentation and classification approaches) can be found on the recent publication [13], which provides an excellent survey on digital mammogram processing techniques.

The aforementioned papers refer to the cases where images are not corrupted by noise. When images are acquired, different noise types may occur in the image acquisition process, affecting both the visual quality and classification performance [14]. Many noise suppression techniques assume that the noise is normally distributed and additive (Gaussian noise) [15]. A specific noise that may occur for the X-ray images satisfies a Poisson distribution, generally caused when the finite number of particles that carry energy (such as electrons in an electronic circuit) is small enough to give rise to detectable statistical fluctuations in a measurement [16]. This noise is typically named quantum noise. Its magnitude varies across the image as it depends on the image intensity. This specific characteristic makes removing such noise very difficult. Aach and Kunz [17] applies several spatial and temporal filters to suppress poisson noise from low-dose X-ray. The authors proposed a novel motion-adaptive temporal filter based on a motion adaptive mean square error criterion. A general framework for studying several noise types (including Poisson) was derived by Bravel et al. [18]. Using local measurements of the mean image intensity, the noise standard deviation and its associated error is approximated using robust statistical estimators. Frosio et al. [19] present a method to estimate the Poisson noise variance based on the maximization of an appropriate likelihood function, which takes into account a sensor model leading to a closed-form analytical expression for the estimate of the sensor gain. A 2D fuzzy Wiener filter is employed by Toprak [20] to suppress the noise. While suppressing the noise, the sharp-edge and the image details are preserved. The novelty resides in using a fuzzy approach to distinguish noise pixels from image ones. Finally, Mencattini et al. [21] used the well known wavelet shrinkage denoising approach to preserve image edges. A proper threshold is picked up to keep or remove image details, thus discarding noise. The noise variance is estimated through a fuzzy logic system under the assumption of white gaussian additive noise.

This paper aims at describing a digital mammogram classification approach, where the manually segmented areas representing abnormal tissues are filtered with the help of Gabor wavelets with several orientations and frequencies. Unlike other wavelets, Gabor functions exhibit specific orientation properties that can be easily tuned by modifying the Gaussian parameters. The resulting database with the filtered images is next decomposed using PCA for dimensionality reduction. Proximal support vector machines are further employed to classify the PCA-Gabor filtered in one of the three classes: normal, benign and malign. PCA is also applied directly to the unfiltered image patches, as baseline. We should note that, Rangayyan et al. [22] also proposed the use of Gabor wavelets. However, there are many differences between the two. The major differences resides in the way we evaluate the Gabor features. Rangayyan et al. used Rose diagram to evaluate the directional properties of Gabor filters using the whole space image. We rather went one step further and analyzed the discriminative characteristics of such directional features with respect to tumor classification, employing advanced classifiers such as proximal support vector machines (PSVM) applied to image patches extracted around image artifacts. While their approach only allows image interpretation with no automatic tumor classification, ours does. Their purpose was to analyze and interpret the symmetry between the left and right mammograms, while our goal was to analyze the discriminant power of Gabor features with respect to various frequency scales and orientations.

The experiments were also conducted when X-ray mammogram images are corrupted by quantum noise with different noise levels. The mammogram images with known labels are assumed to be acquired in ideal conditions, while the mammogram images with unknown classes are corrupted by quantum noise. Therefore, we only altered the test mammogram images. The experiments reveal great robustness for Gabor features against noise. We must stress that no preprocessing step for noise suppression was taken into account. We should also note that, this paper presents an extension of our previous work described in [23], where 30×30 image patches were used for classification. The extension refers to the following issues. The experiments implied higher dimensional patches, i.e. 60×60 to investigate the patch size versus the classification performances. The work in [23] only reported the average recognition rate without taking care of the nature of misclassified patches in terms of false negatives and false positives. We have addressed this issue by computing the confusion matrix. The work described by [23] limited the number of Gabor feature orientations to four. Here, we have extended the orientation range to eight values, for comparison. Another major difference is our investigation

of how the performance of classification is affected by the quantum noise, when both filtered and unfiltered image patch sets are considered.

The remaining of the paper is structured as follows. Section 2 briefly presents the image database used in out experiments followed by the mathematical description of the Gabor wavelets and short PCA description. In Section 3, the experiment setup is described and the experimental outputs are illustrated. Section 4 concludes the work.

2. Material and Methods

2.1. Material: Mammogram Database

The mammogram images used in our experiments were taken from the Mammographic Image Analysis Society (MIAS) [24]. The database contains 322 samples labelled as one of the three categories: normal, benign and malign. There are 208 normal images, 63 benign and 51 malign cases (which are considered abnormal). Each image of 1024×1024 pixels is centered. The abnormal cases are divided into six categories: microcalcification, circumscribed masses, spiculated masses, ill-defined masses, architectural distortion and asymmetry. However, we only considered here the three classes aforementioned when classifying the images. For each abnormal case, the coordinates of center of abnormality are provided along with the approximate radius (in pixels) of a circle enclosing the abnormality. The widest identified abnormality corresponds to a radius of 197 pixels, while the tightest abnormality has a radius of 3 pixels. For some cases, calcification are widely distributed throughout the breast image rather than focused at a single site. Here, the center locations and radii have been omitted. Knowing the location and the approximate size of abnormality allows us to manually extract subimages (patches) with proper dimension representing the tumor zone.

2.2. Methods

2.2.1. Gabor Wavelets

2D Gabor wavelets have been extensively used in computer vision applications for modeling of biological-like vision systems. Studies have shown that Gabor elementary functions are suitable for modeling simple cells in visual cortex [25]. Other nice property is provided by their optimal joint resolution in both space and frequency, suggesting simultaneously analysis in both domains. The Gabor wavelet orientation property makes it very suitable for several applications, including image texture analysis [26,27] or image retrieval [28]. A complex Gabor wavelet is defined as the product of a Gaussian kernel with a complex sinusoid described as:

$$\psi_{\mathbf{k}}(\mathbf{z}) = \frac{\mathbf{k}^T\mathbf{k}}{\sigma^2}\exp(-\frac{\mathbf{k}^T\mathbf{k}}{2\sigma^2}\mathbf{z}^T\mathbf{z})(\exp(ik^T\mathbf{z}) - \exp(-\frac{\sigma^2}{2})), \quad (1)$$

where $\boldsymbol{F}_{PCA}^{kc}$ and $\mathbf{k}$ is the characteristic wave vector:

$$\pi \quad (2)$$

with

$$k_\nu = 2^{-\frac{\nu+2}{2}}\pi, \quad 1ex\varphi_\mu = \mu\frac{\pi}{8}. \quad (3)$$

The parameters ν and μ define the frequency and orientation of the filter. Given an image $\mathcal{I}(\boldsymbol{z})$, a 2D Gabor wavelet transform is defined as the convolution of this image $I(\mathbf{z})$ with a family of Gabor filters with several orientation and frequency values:

$$I_k(\mathbf{z}) = \int\int I(\mathbf{z}')\psi_k(\mathbf{z}-\mathbf{z}')d\mathbf{z}' \quad (4)$$

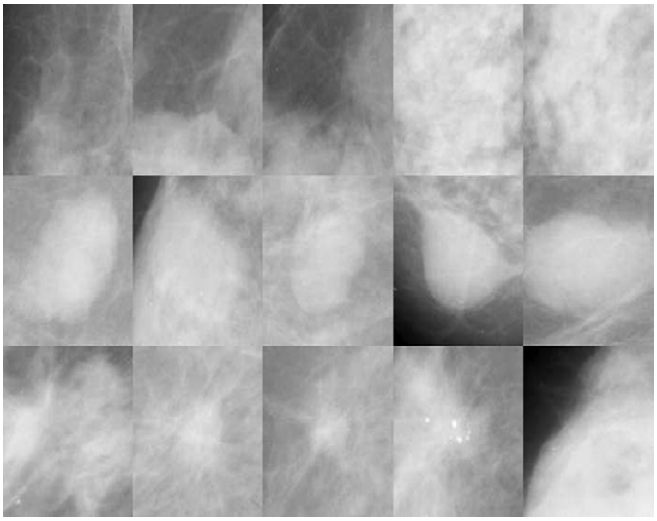

Figure 1. Patches of 140×140 pixels extracted from mammographic images. Top row represents 5 samples for the normal case, middle row represents 5 samples for the benign case, while the bottom row illustrates 5 samples for the malign case.

2.2.2. Principal Component Analysis

Principal Component Analysis (PCA) [29] was picked up as dimensionality reduction for both Gabor features and original raw patch images. The latter case has been chosen as baseline. Given a high dimensional database, PCA is a well-known technique to keep relevant low dimensional information in terms of data covariance matrix, while discarding information associated to low eigenvalues. Particular details on PCA application to our case are described in Subsection 3.3.

3. Results and Discussion

3.1. Experiment Setup

To discard irrelevant (background) information like breast contour, patches of 140×140 pixels surrounding the abnormality region were extracted from the original 1024×1024 pixels images. The patches size assures that, for most abnormal cases not only the abnormality region is captured but also the surrounding area, providing us information about the abnormality shape. For the normal case, the patches were extracted from random position inside the breast area. Figure 1 illustrates 5 samples per class (case).

Each 140×140 image patch is downsampled to either 60×60 or 30×30 pixels, respectively, to reduce the data dimension, as the final Gabor-based feature vector is formed by concatenating the Gabor convolution output for several orientations and frequencies, leading thus, to a very large dimensional feature vector if the initial size is kept. Consequently, the next processing step (PCA) might be computational intractable. Therefore, a gross dimensionality reduction is necessary step to avoid the computational issue. The full mammogram database is split into two disjoint sets to test the generalization ability of the classifier combined with Gabor features.

We ran experiments for three diagnose cases. The first case refers to a three-class classification where classes normal, benign and malign are considered all together. The first set representing 80% samples from the whole database is the set where the three classes are known and the remaining 20% samples are included in the test set with unknown classes (labels) that have to be automatically classified. For the second case, samples corresponding to benign and malign are relabelled into a single class (named here tumor class), while the tumor-free samples are denoted by normal

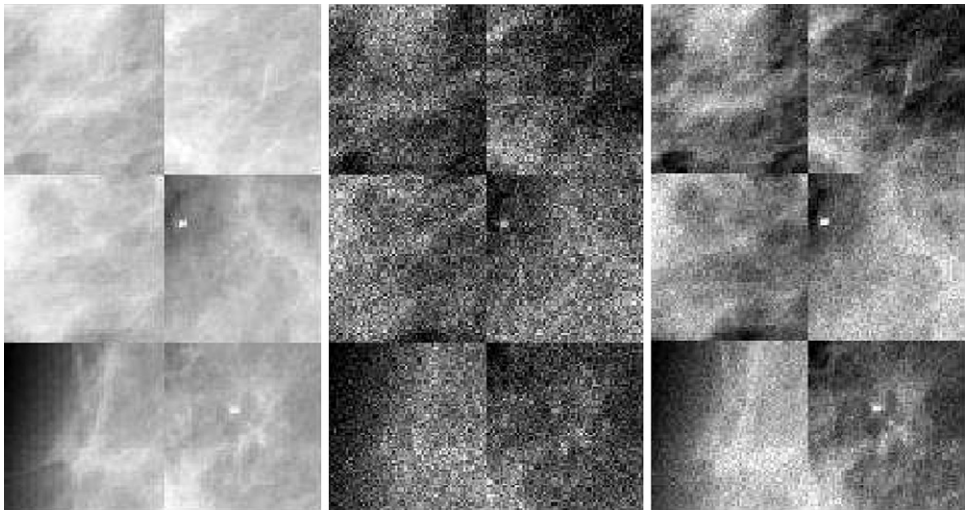

Figure 2. Noise-free and noisy mammogram samples. Each two-column depicts 2 samples for normal, benign and malign cases without noise and with noise corresponding to $\lambda = 0.1$, and λ=1, respectively.

case. This recasts into a normal versus tumor binary classification. The third case is also a binary class problem, where the task was rather to discriminate between benign and malign samples, while the normal samples were discarded from the training and test sets, respectively.

3.2. Quantum noise

MIAS database comprises noise-free image samples. To run experiments with noisy mammogram images we artificially generated quantum noise from each image intensity. Recall that we add noise only to the test mammogram images. The quantum noise obeys a Poisson distribution expressed as:

$$P(p|\rho, T) = (\rho T)^p \cdot \frac{e^{-\rho T}}{p!}, \tag{5}$$

where $\rho T = \lambda$ is the probability of measuring p photons in T seconds, ρ is the number of photons emitted (on average) by a source, and symbol ! denotes the factorial. The mean and standard deviation are given respectively by $\mu = \rho T$ and $\sigma = \sqrt{\rho T}$. Two different λ values are considered, i.e., $\lambda = \{0.1, 1, 4\}$, to generate quantum noise. As the noise is image intensity dependent, the 256 gray level mammogram images are rescaled into the pixel value interval proportional to λ values. Figure 2 depicts two noise-free samples for normal, benign and malign cases along with their noisy samples for those two different λ.

3.3. Experimental Results for the Three-Class Case

Each image was convolved with several Gabor wavelets. We have used eight orientations: 0, $\frac{pi}{8}$, $\frac{\pi}{4}$, $\frac{3\pi}{8}$, $\frac{\pi}{2}$, $\frac{5\pi}{4}$, $\frac{3\pi}{4}$, $\frac{7\pi}{8}$ and two frequency ranges: high frequencies (hfr) with ν = 0,1,2 and low frequencies (lfr) with ν = 2,3,4. In addition, we combined all orientations in one larger feature vector for each image. Gabor filtering leads to complex values. However, we only took the magnitude of the convolution results. When lfr (or hfr) and one orientation is used, 3 output images were formed that have been row-wise scanned and concatenated to yield a final 2700 - dimensional feature vector. For lfr (or hfr) and full orientation range, the convolution is applied to 24 Gabor filters, resulting a 86400 - dimensional vector for each image. Figure 3 depicts the result of the convolution

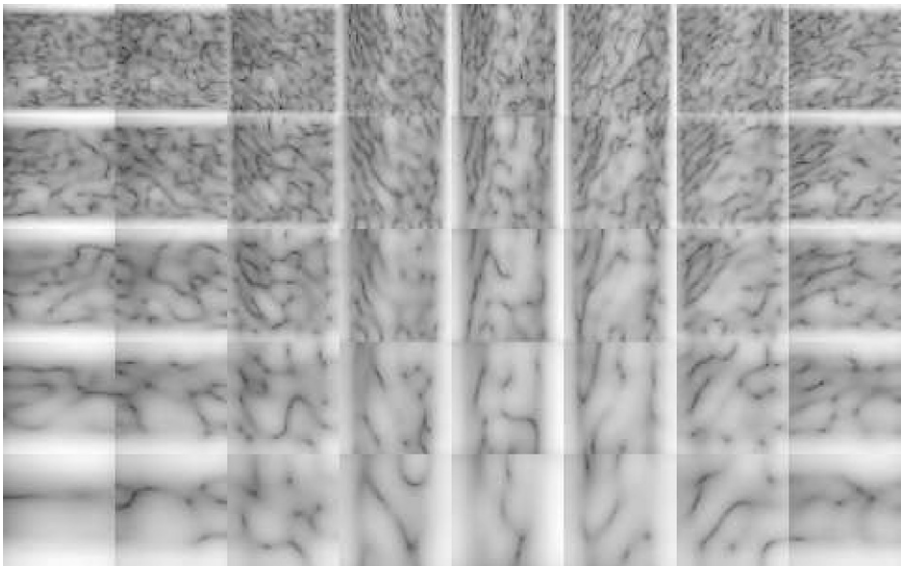

Figure 3. Magnitude of Gabor features for one noise-free mammogram image convolved with 40 Gabor filters for $\nu = \{0, 1, 2, 3, 4\}$ and $\mu = \left\{0, \frac{pi}{8}, \frac{\pi}{4}, \frac{3\pi}{8}, \frac{\pi}{2}, \frac{5\pi}{4}, \frac{3\pi}{4}, \frac{7\pi}{8}\right\}$. The sample corresponds to the left most bottom image from Figure 2.

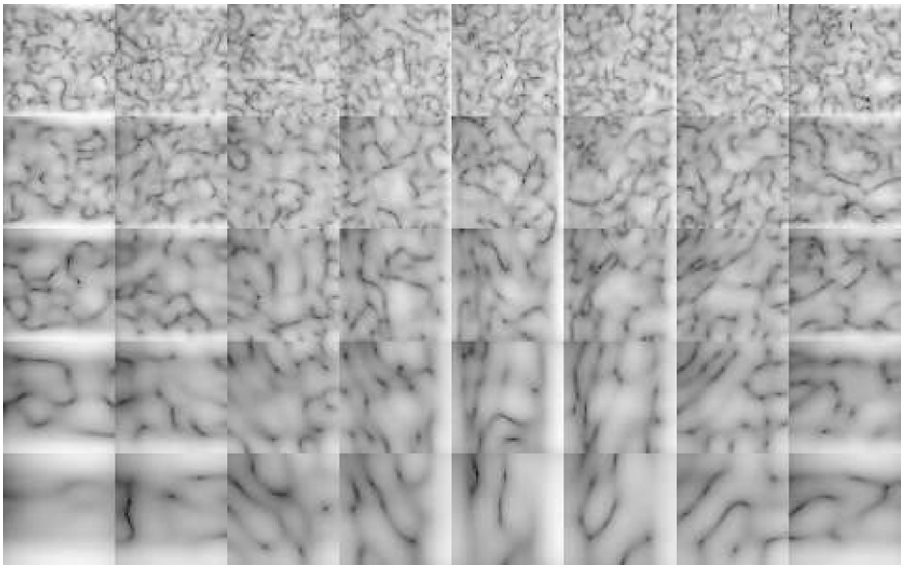

Figure 4. Magnitude of Gabor features for one noisy mammogram image with $\lambda = 0.1$, convolved with 40 Gabor filters.

magnitude for one sample corresponding to the left most bottom image from Figure 2, with eight orientations and five frequencies Gabor filters. Horizontal, vertical and oblique oriented features are emphasized after applying Gabor filtering. Low frequencies tend to stress smooth image areas while high frequencies outline fine details in image.

Gabor features were extracted from noisy mammogram patches as well. Figures 4 and 5 respectively, illustrates the output for the convolution of 40 Gabor filters with one noisy image corresponding to different λ. Although intense quantum noise is generated, the Gabor wavelets were able to intrinsically filter the noise, and emphasize discriminant features in the same time. This can be noticed by visual comparison of Figure 3 with Figures 4–5, where insignificant modifications are observed.

Once Gabor features were extracted, two novel Gabor based feature matrices are formed. Let us denote with $\mathbf{X}^{kc}$ and $\mathbf{X}^{uc}$ the matrix containing the Gabor features with known and unknown classes (test matrix) in their columns, respectively. $\mathbf{X}^{kc}$ is a p×258 matrix and $\mathbf{X}^{uc}$ is a $p \times 64$ matrix, where p denotes the feature vector size after concatenating the Gabor features. Due to high feature vector dimension p (recall that, each column of this matrix is a 86400 - dimensional Gabor feature vector when patch size is 60×60 and 24 Gabor filters are used), PCA was next applied to $\mathbf{X}^{kc}$ matrix for dimensionality reduction. In a large-scale processing the matrix $\mathbf{X}^{kc}$ should be much bigger (its number of columns increases as more samples are added). Gabor + PCA feature vectors (denoted here by "*GabPCA*") that are actually used for classification, were formed by projecting the zero mean and unit variance data into the PCA eigenvectors V_r, i.e. $\mathbf{F}^{kc}_{GabPCA} = \mathbf{V}_r^{'}\mathbf{X}^{kc}$, where V_r is the r - rank PCA projection matrix ansents the transpose operator. Employing PCA projection, the dimension was reduced from p to r, where $r \ll p$. The experiments were run for the set r={5, 10, 20, 30, ..., 150}. Whenever a new zero mean unseen patch from the test matrix $\mathbf{X}^{uc}$ comes, its corresponding feature vector was formed in a similar way, i.e., $\mathbf{f}^{uc}_{GabPCA} = \mathbf{V}^{'}\mathbf{x}^{uc}$. For comparison purpose we carried out experiments were PCA was solely applied to the initial unfiltered patch images. Proximal Support Vector Machines [30] with polynomial kernel of degree 1,2 and 3 were used to classify $\mathbf{F}^{uc}_{GabPCA}$ (or F^{uc}_{PCA}, for unfiltered images) corresponding to unknown patch samples. As PSVM is typically designed for two - class problem, a "one

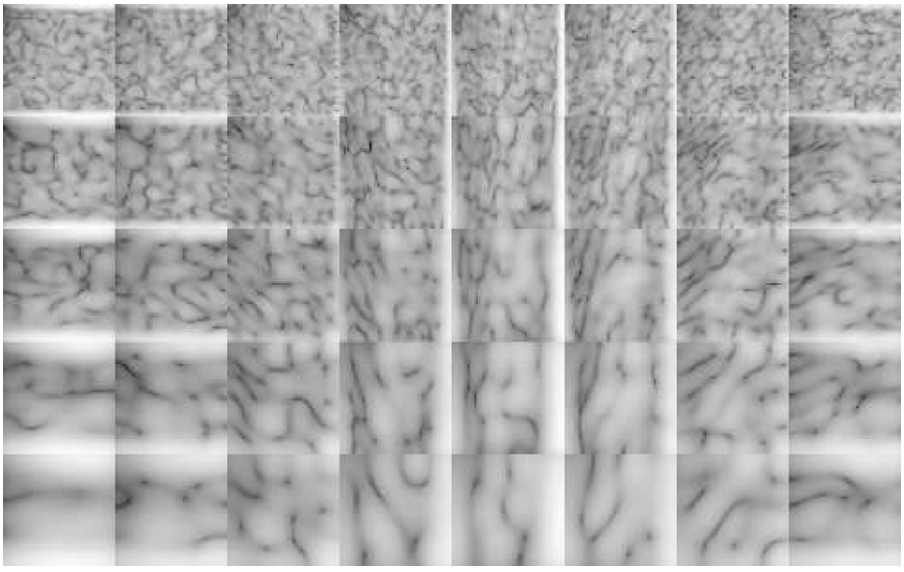

Figure 5. Magnitude of Gabor features for one noisy mammogram image with $\lambda = 1$, convolved with 40 Gabor filters.

Table 1
The recognition rate expressed in percentage corresponding to noise-free patches and for different patch size, Gabor features and PCA. The highest RR is in bold. The rank (r) corresponding to the best RR is given in parenthesis.

Patch size	Features	Frequency range	Orientation	RR (%)
			0	73.43 (100)
			$\pi/8$	67.18 (5)
			$\pi/4$	67.18 (20)
			$3\pi/8$	68.75 (20)
		lfr	$\pi/2$	68.75 (5)
			$5\pi/4$	68.75 (20)
			$3\pi/4$	64.06 (5)
			$7\pi/8$	**75** (80)
	GabPCA		all	**75** (20)
			0	65.62 (10)
			$\pi/8$	68.75 (50)
			$\pi/4$	64.06 (5)
30×30			$3\pi/8$	67.18 (30)
		hfr	$\pi/2$	71.87 (40)
			$5\pi/4$	67.18 (50)
			$3\pi/4$	64.06 (5)
			$7\pi/8$	68.75 (20)
			all	71.87 (30)
	PCA	-	-	65.62
			0	65.62 (100)
			$\pi/8$	67.18 (140)
			$\pi/4$	67.18 (110)
			$3\pi/8$	67.18 (30)
		lfr	$\pi/2$	**68.75** (110)
			$5\pi/4$	67.18 (5)
			$3\pi/4$	64.06 (20)
			$7\pi/8$	**68.75** (40)
	GabPCA		all	68.75 (30)
			0	64.06 (5)
			$\pi/8$	62.50 (30)
			$\pi/4$	65.62 (5)
60×60			$3\pi/8$	**68.75** (140)
		hfr	$\pi/2$	65.62 (5)
			$5\pi/4$	64.062 (50
			$3\pi/4$	65.62 (10)
			$7\pi/8$	67.18 (110)
			all	**68.75** (1100
	PCA	-	-	67.11

against all" strategy was applied for our three - class problem. Prior to the test phase, PSVMs were trained using $\mathbf{F}^{kc}_{GabPCA}$ (or $\mathbf{F}^{kc}_{PCA}$) from known class data set. We should note that each SVM was separately trained (to acquire its optimum separating hyperplane) for each case.

The classification performance was measured in terms of *recognition rate* (RR), defined as the percentage of correctly classified samples, i.e., $\{\tilde{l}(\mathbf{f}^{uc}_{GabPCA}) = l(\mathbf{f}^{uc}_{GabPCA})\}$, where $\tilde{l}$ is the predicted label (classifier output) and l is the actual label for the test mammogram image. Tables 1–3 summarize the results for different Gabor features coupled with PCA based dimensionality reduction corresponding to 30×30 and 60×60 pixel patches, respectively, and for noise-free and noisy mammogram images. The results attained when PCA was directly applied to the unfiltered images are also reported in those Tables.

The recognition rate is an overall measurement of the approach performance without pointing out to the error type. On the contrary, the confusion matrix provides us information with respect to false positive and false negatives as well as the type of misclassifications. More precisely, the confusion matrix is the matrix whose diagonal entries are the number of mammogram patches that are correctly classified, while the off-diagonal entries correspond to misclassification. The rows of the matrix describe the actual (correct) class labels and its columns the predicted ones. False negatives (FN) could lead to death while false positives (FP) have a high cost and could cause detrimental effects on patients. For automatic medical image classification, the rate of false negatives has to be very low if not zero. A false positive detection causes an unnecessary biopsy. In a false negative detection, an actual tumor remains undetected and could lead to higher costs or, ultimately, to the cost of the life of the patient. Tables 4–6 tabulates the confusion matrix for the results corresponding to the optimum parameters with noise-free and noisy images. For the normal case we have one FP where the sample was misclassified as benign while it's normal. Eight FNs correspond to the case when the samples were wrongly classified as normal while their correct class is associated to the benign case. The most crucial FNs are the ones corresponding to the first column last row of Table 4 where three malign samples were misclassified as normal tissue.

As Table 5 indicates, 5 more samples are misclassified as normal tissue while actually being malign, for λ=0.1.

3.4. Experimental Results for the Normal versus Tumor Case

When running the normal versus tumor case, the results depicted in Table 7 are obtained. Apart from the overall recognition rate, three more classification measures are reported, namely, sensitivity, specificity and area under the receiver operator characteristic (AUC). Sensitivity and specificity are described as $Sn = TP/(TP + FN)$, and $Sp = TN/(TN + FP)$, where TP denotes true positive, and TN stands for true negative.

The optimum result corresponds to the GabPCA features with a recognition rate of 84.37%, 97.56% as sensitivity and 60.86% as specificity. The highest AUC corresponds to 0.79%. As noise is added to the image samples, the RR decreases for all cases. However, GabPCA features seem to be more robust to noise compared to the

Table 2
The recognition rate expressed in percentage corresponding to noisy patches with λ = 0.1 and for different patch size, Gabor features and PCA. The highest RR is in bold. The rank (r) corresponding to the best RR is given in parenthesis.

Patch size	Features	Frequency range	Orientation	RR (%)
			0	65.62 (50)
			$\pi/8$	65.62 (40)
			$\pi/4$	64.06 (5)
			$3\pi/8$	65.62 (30)
		lfr	$\pi/2$	64.06 (5)
			$5\pi/4$	64.06 (5)
			$3\pi/4$	**67.18** (5)
			$7\pi/8$	64.06 (5)
	GabPCA		all	64.06 (5)
			0	64.06 (5)
			$\pi/8$	64.06 (5)
			$\pi/4$	64.06 (5)
30×30			$3\pi/8$	64.06 (5)
		hfr	$\pi/2$	65.62 (5)
			$5\pi/4$	64.06 (5)
			$3\pi/4$	64.06 (5)
			$7\pi/8$	64.06 (5)
			all	64.06 (5)
	PCA	-	-	62.05
			0	64.06 (5)
			$\pi/8$	64.06 (5)
			$\pi/4$	64.06 (5)
			$3\pi/8$	64.06 (30)
		lfr	$\pi/2$	**66.65** (5)
			$5\pi/4$	64.06 (5)
			$3\pi/4$	64.06 (5)
			$7\pi/8$	64.06 (5)
	GabPCA		all	64.06 (5)
			0	64.06 (5)
			$\pi/8$	64.06 (5)
			$\pi/4$	64.06 (5)
60×60			$3\pi/8$	64.06 (5)
		hfr	$\pi/2$	64.06 (5)
			$5\pi/4$	64.06 (5)
			$3\pi/4$	64.06 (5)
			$7\pi/8$	64.06 (5)
			all	64.06 (5)
	PCA	-	-	65.02

Table 3
The recognition rate expressed in percentage corresponding to noisy patches with $\lambda = 1$ and for different patch size, Gabor features and PCA. The highest RR is in bold. The rank (r) corresponding to the best RR is given in parenthesis.

Patch size	Features	Frequency range	Orientation	RR (%)
			0	**68.75** (10)
			$\pi/8$	67.18 (5)
			$\pi/4$	64.06 (5)
			$3\pi/8$	64.06 (5)
		lfr	$\pi/2$	64.06 (5)
			$5\pi/4$	64.06 (5)
			$3\pi/4$	65.62 (5)
			$7\pi/8$	65.62 (30)
	GabPCA		all	64.06 (5)
			0	64.06 (5)
			$\pi/8$	64.06 (5)
			$\pi/4$	64.06 (5)
30×30			$3\pi/8$	64.06 (5)
		hfr	$\pi/2$	67.18 (5)
			$5\pi/4$	64.06 (5)
			$3\pi/4$	64.06 (5)
			$7\pi/8$	64.06 (5)
			all	64.06 (5)
	PCA	-	-	62.05
			0	64.06 (5)
			$\pi/8$	64.06 (5)
			$\pi/4$	64.06 (5)
			$3\pi/8$	64.06 (5)
		lfr	$\pi/2$	**66.65** (30)
			$5\pi/4$	64.06 (5)
			$3\pi/4$	64.06
			$7\pi/8$	64.06 (5)
	GabPCA		all	64.06 (5)
			0	64.06 (5)
			$\pi/8$	64.06 (5)
			$\pi/4$	64.06 (5)
60×60			$3\pi/8$	64.06 (5)
		hfr	$\pi/2$	65.62 (10)
			$5\pi/4$	64.06 (5)
			$3\pi/4$	64.06 (5)
			$7\pi/8$	64.06 (5)
			all	64.06 (5)
	PCA	-	-	64.06

Table 4
Confusion matrix for noise-free 30×30 image patches, *lfr*, all orientations and 20 PCs.

		predicted class normal	benign	malign
true class	normal	40	1	0
	benign	8	2	3
	malign	3	1	6

Table 5
Confusion matrix for noisy 30×30 image patches ($\lambda = 0.1$), *lfr*, orientation = $3\pi/4$ and 5 PCs.

		predicted class normal	benign	malign
true class	normal	41	0	0
	benign	11	1	1
	malign	9	0	1

Table 6
Confusion matrix for noisy 30×30 image patches ($\lambda = 1$), *lfr*, orientation = 0 and 10 PCs.

		predicted class normal	benign	malign
true class	normal	41	0	0
	benign	10	1	2
	malign	7	1	2

Table 7
Sensitivity (Se), specificity (Sp) and area under the ROC curve (AUC) for the 30×30 and 60×60 image patches, respectively, corresponding to the case normal versus tumor.

Image size	Image type	Method	RR	Sn	Sp	AUC
	noise-free	GabPCA	**84.37** (10)	97.56	60.86	0.79
		PCA	73.43 (120)	97.56	30.43	0.77
30×30	$\lambda = 0.1$	GabPCA	73.43 (90)	75.60	69.56	0.77
		PCA	66.75 (80)	70.48	43.47	0.69
	$\lambda = 1$	GabPCA	76.56 (130)	75.60	78.26	0.78
		PCA	71.87 (5)	60.97	39.13	0.51
	noise-free	GabPCA	75 (90)	92.68	40	0.70
		PCA	75 (150)	97.56	34.78	0.78
60×60	$\lambda = 0.1$	GabPCA	67.18 (10)	88.05	70.23	0.69
		PCA	73.43 (140)	78.04	65.21	0.77
	λ=1	GabPCA	67.18 (10)	78.04	65.21	0.77
		PCA	75.00 (100)	80.48	65.21	0.77

Highest recognition rate is in bold.

pure PCA features for the 30×30 image patches. Interestingly, a larger patch size (60×60) weakens the RR for the GaborPCA.

3.5. Experimental Results for the Benign versus Malign Case

Table 8 shows the output for the the benign versus malign case. The results are in trend with the other two cases, i.e., GabPCA features turn out to be more robust against noise compared to pure PCA features and seem to possess more discriminative power.

3.6. Discussions

Considering the experimental results, the following discussions can be drawn:

- Increasing the patch size from 30×30 to 60×60 does not help the classification performance for the *GabPCA* features. Rather, an increase in image resolution conducts to a decrease in the recognition rate, at least for the GabPCA features.
- Low frequency range possesses more discriminant power than the high frequency range for all cases, when *GabPCA* features are employed to 30×30 image patches. As image patch size increases to 60×60, the difference between *RR* for low and high frequency range shrinks, emphasizing the contribution of details (comprised in the high frequency range). However, overall, the highest recognition rate is due to the low frequency range.
- Generally, using all Gabor wavelet feature orientations improved the recognition rate. This is not surprising as many features are oriented towards many spatial directions.

Table 8
Sensitivity (Se), specificity (Sp) and area under the ROC curve (AUC) for the 30×30 and 60×60 image patches, respectively, corresponding to the case benign versus malign.

Image size	Image type	Method	RR	Sn	Sp	AUC
	noise-free	GabPCA	**78.26** (80)	84.61	80	0.78
		PCA	56.52 (80)	76.92	30	0.57
30×30	$\lambda = 0.1$	GabPCA	69.56 (30)	46.15	70	0.60
		PCA	65.21 (50)	76.92	50	0.64
	λ=1	GabPCA	69.56 (20)	76.92	50	0.67
		PCA	65.21 (40)	76.92	50	0.63
	noise-free	GabPCA	69.56 (10)	30.07	80	0.70
		PCA	56.52 (40)	74.22	10	0.55
60×60	λ=0.1	GabPCA	69.56 (10)	54.61	40	0.66
		PCA	60.86 (70)	53.84	70	0.63
	λ=1	GabPCA	69.56 (10)	75.92	40	0.68
		PCA	69.56 (50)	69.23	70	0.65

Highest recognition rate is in bold.

- Considering the normal versus tumor case, promising results were accomplished. We should stress here that, adding more representative tumor-free and tumor samples in the SVM training procedure would likely to increase the specificity associated to this case.
- As far as the benign versus malign case is concerned, the recognition rate would definitely get higher with an increase in the samples number. Similar benign and malign patterns made the recognition quite difficult. This was reflected into the SVM kernel's degree. More precisely, the best results correspond to the polynomial or cubic kernel, indicating a nonlinear separating hyperplane between benign and malign features.
- Overall, *GabPCA* features tend to lead to superior classification performance compared to the features solely based on *PCA* decomposition.
- The experimental results revealed the fact that the GabPCA features are less sensitive to the noise level compared to the PCA features extracted from noisy row image patches. In addition, low frequency range is more robust to noise, due to the fact that, as expected, noise mainly affects the high frequency image components.

4. Conclusions

Digital mammograms are among the most difficult medical images to be read due to the low tissue image contrast and slight perceptible differences. Gabor features obtained by convolving image patches representing tumor or tumor-free areas with several Gabor filters are employed in this paper for recognizing three tissue types (normal, benign and malign) in automatic way, from digital mammograms. This task is very challenging even for medical specialists. From the human observer (physician) point of view, the distinction between benign and malign tumor types is very ill-defined in terms of visual inspection, since what usually a physician does is to ask for further analysis including other more relevant tests (biopsy) for characterizing the tumor as benign or malign. It was reported that 65 to 90% of the biopsies of suspected cancers turn out to be benign [31,32]. Two recent papers that are closely related to our work and involve de same database are worth to mention here. The first one is the work of Verma et al. [33], where a novel classification scheme is proposed to distinguish between malign and benign cases. However, the work did not focus the feature extraction procedure and rather dealt with the learning part. The second work attributed to Islam et al. [34] reported a value of 60.91% for the sensitivity and 83.87% for the specificity for the malign versus benign case. Both reported values are higher than the ones yielded from our experiments. It is difficult to have a direct and fair comparison as the experiment setup differs. For instance, they only used a subset of 69 correctly segmented samples. Also, different samples might have been picked up to form the training and test set. However, our results also outperform the radiologist sensitivity reported as being 75%. For the normal versus tumor case, though the specificity obtained is relatively low, a very promising value for the sensitivity as achieved. We believe that, by introducing more (representative) tumor samples would improve the specificity value.

Summarizing, two main conclusions can be drawn from our experiments. Firstly, Gabor features seem to possess more discriminative power than PCA features which are largely used for this task in the literature. Unlike PCA features (derived from unfiltered images), Gabor features are more robust to quantum noise, noise that typically may occur when X-ray images are acquired. Technically, as indicated by our experiments, Gabor features have proven to be less insensitive to noise with Poisson distribution even for large noise level.

Acknowledgment

This work was supported by CNCSIS - UEFISCSU, project number PNII - IDEI 668/2008.

References

[1] R.N. Strickland, H. Hahn, Wavelet Transforms for Detecting Microcalcifications in Mammograms", IEEE Trans. on Medical Imaging 15 (1996) 218–229.
[2] R. Swiniarski, T. Luu, A. Swiniarska, H. Tanto, Data Mining and On-line Recognition of Mammographic Images Based on Haar Wavelets, Principal Component Analysis and Rough Sets Methods, International SPIE Symposium Medical Imaging (2001) 17–23.
[3] R. Swiniarski, H.K. Lim, J.H. Shin, A. Skowron, Independent Component Analysis, Principal Component Analysis and Rough Sets in Hybrid Mammogram Classification, Proc. of the 2006 International Conference on Image Processing, Computer Vision, Pattern Recognition Vol. II (2006) 640–645.
[4] Y. Sun, C.F. Babbs, E.J. Delp, Normal Mammogram Classification based on Regional Analysis, The 2002 45th Midwest Symposium on Circuits and Systems 2 (2002) 375–378.
[5] O.R. Zaiane, M.L. Antonie, A. Coman, Mammography Classification by an Association Rule-based Classifier", in: Third International ACM SIGKDD workshop on multimedia data mining (MDM/KDD'2002) in conjunction with eighth ACM SIGKDD, 2002, pp. 62–69.
[6] G. Lemaur, K. Drouiche, J. DeConinck, Highly Regular Wavelets for the Detection of Clustered Microcalcifications in Mammograms, IEEE Trans. on Medical Imaging 22 (2003) 393–401.
[7] C.B.R. Ferreira, D.L. Borges, Analysis of Mammogram Classification Using a Wavelet Transform Decomposition, Pattern Recognition Letters 24 (2003) 973–982.
[8] C.B.R. Ferreira, D.L. Borges, An Evaluation of Wavelet Features Subsets for Mammogram Classification", *Progress in Pattern Recognition*, Image Analysis and Applications vol.3773 (2005) 620–630.
[9] L. Wei, Y. Yang, R.M. Nishikawa, Y. Jiang, A Study on Several Machine-Learning Methods for Classification of Malignant and Benign Clustered Microcalcifications, IEEE Trans. on Medical Imaging 24 (3) (2005) 371–380.
[10] H.S. Sheshadri, A. Kandaswamy, Breast Tissue Classification Using Statistical Feature Extraction Of Mammograms, Medical Imaging and Information Sciences 23 (3) (2006) 105–107.
[11] F. A. Amara and I. A. Qader, "Hybrid Mammogram Classification using Rough Set and Fuzzy Classifier", *International Journal of Biomedical Imaging*, Vol. 2009, Article ID 680508, 12 pages.
[12] L.O. Martins, G.B. Junior, A.C. Silva, A.C. de Paiva, M. Gattass, Detection of Masses in Digital Mammograms using K-means and Support Vector Machine, Electronic Letters on Computer Vision and Image Analysis 8 (2) (2009) 39–50.
[13] J. Bozek, M. Mustra, K. Delac, M. Grgic, in: M. Grgic, K. Delac, M. Ghanbari (Eds.), A Survey of Image Processing Algorithms in Digital Mammography Recent Advances in Multimedia Signal Processing and Communications, Springer-Verlag, Heidelberg, October 2009, pp. 631–657.
[14] M. Analoui, Radiographic digital image enhancement. Part II: transform domain techniques, Dentomax. Radiol. 30 (2001) 65–77.
[15] S. Webb, The Physics of Medical Imaging, Adam Hilger, Bristol, U.K, 1988.
[16] P.M. Goebel, A.N. Belbachir, M. Truppe, Noise estimation in panoramic x-rays images: an application analysis approach, in: In: Proc. of the 13th IEEE Workshop on Statistical Signal Processing, 2005, pp. 996-1001.
[17] T. Aach, D. Kunz, R. Florent, S. Makram-Ebeid, Noise Reduction and Image Enhancement Algorithms for Low-Dose X-Ray Fluoroscopy, *Proc. BVM* (1996) 95–100.
[18] P. Gravel, G. Beaudoin, J.A. De Guise, A Method for Modeling Noise in Medical Images, IEEE Trans. on Medical Imaging 23 (10) (2004) 1221–1232.
[19] I. Frosio, M. Lucchese, N.A. Borghese, A new and reliable Poisson noise estimator for radiographic images, in: Proc. of the 14th International Conference on Image Analysis and Processing, 2007, pp. 725-730.
[20] A. Toprak, Quantum Noise Suppression in X-ray Images using Fuzzy 2-D Wiener Filter, Journal of Medical Systems 31 (5) (Oct 2007) 351–355.
[21] A. Mencattini, M. Salmeri, R. Lojacono, M. Frigerio, F. Caselli, Mammographic images enhancement and denoising for breast cancer detection using dyadic wavelet processing, IEEE Trans. on Instrumentation and Measurement 57 (7) (2008) 1422–1430.
[22] R.M. Rangayyan, R.J. Ferrari, J.E.L. Desautels, A.F. Frère, Directional analysis of images with Gabor wavelets, in: Proc. of XIII Brazilian Symposium on Computer Graphics and Image SIBGRAPI, 2000, pp. 170-177.
[23] I. Buciu, A. Gacsadi, Gabor Wavelet Based Features for Medical Image Analysis and Classification, in: 2nd International Symposium on Applied Sciences in Biomedical and Communication Technologies, 2009, pp. 1- 4.
[24] Mammographic Image Analysis Society, http://www.wiau.man.ac.uk/services/MIAS/MIASweb.html.
[25] T. Lee, Image Representation using 2d Gabor Wavelets, IEEE Trans. Pattern Analysis and Machine Intelligence vol.18 (10) (1996) 959–971.
[26] M.R. Turner, Texture Discrimination by Gabor functions, Biological Cybernetics vol.55 (1986) 71–82.
[27] S.E. Grigorescu, N. Petkov, P. Kruizinga, Comparison of Texture Features Based on Gabor Filters, IEEE Trans. on Image Processing 11 (10) (2002) 1160–1167.

[28] J. Han, K.-K. Ma, Rotation-invariant and Scale-invariant Gabor Features for Texture Image Retrieval, Image and Vision Computing 25 (9) (2007) 1474–1481.
[29] I.T. Jolliffe, Principal Component Analysis Series: Springer Series on Statistics, 2nd Ed., Springer, NY, 2002.
[30] G. Fung, O.L. Mangasarian, Proximal Support Vector Machine Classifiers, in: Proceedings KDD-2001: Knowledge Discovery and Data Mining, 2001, pp. 77-86.
[31] H.D. Cheng, X.J. Shi, R. Min, L.M.L.M. Hu, X.P. Cai, H.N. Du, Approaches for Automated Detection and Classification of Masses in Mammograms", Pattern Recognition 39 (4) (2006) 646–668.
[32] J. Billingsley, "Radiologists' Mammogram Accuracy Varies Widely" http://news.healingwell.com/index.php?p=news1&id=526229.
[33] B. Verma, P. McLeod, A. Klevansy, Classification of benign and malignant patterns in digital mammograms for the diagnosis of breast cancer, Expert Systems with Applications 37 (4) (2010) 3344–3351.
[34] M.J. Islam, M. Ahmadi, M.A. Sid-Ahmed, An Efficient Automatic Mass Classification Method In Digitized Mammograms Using Artificial Neural Network, International Journal of Artificial Intelligence & Applications 1 (3) (2010) 1–13.