Computer Aided Classification of Mammographic Tissue Using Shapelets and Support Vector Machines

George Apostolopoulos[1], Athanasios Koutras[1,2],
Ioanna Christoyianni[1], and Evaggelos Dermatas[1]

[1] Wired Communications Lab.,
Electrical and Computer Engineering Department, University of Patras, Greece
{gapost,ioanna,dermatas}@george.wcl2.ee.upatras.gr
[2] Informatics and Mass Media Department,
Technical Educational Institute of Western Greece, Greece
koutras@teipat.gr

Abstract. In this paper a robust regions-of-suspicion (ROS) diagnosis system on mammograms, recognizing all types of abnormalities is presented and evaluated. A new type of descriptors, based on Shapelet decomposition, derive the source images that generate the observed ROS in mammograms. The Shapelet decomposition coefficients can be used efficiently to detect ROS areas using Support-Vector-Machines (SVMs) with radial basis function kernels. Extensive experiments using the Mammographic Image Analysis Society (MIAS) database have shown high recognition accuracy above 86% for all kinds of breast abnormalities that exceeds the performance of similar decomposition methods based on Zernike moments presented in the literature by more than 8%.

Keywords: Shapelets, support vector machines, mammogram, breast cancer, computer-aided diagnosis (CAD), feature extraction.

1 Introduction

Breast cancer has been a major cause of fatality among all cancers for women [1]. However, mortality rates have been decreasing during the last years due to better diagnostic facilities and effective treatments [2]. Screening mammography using X-ray imaging of the breast is the most effective, low-cost, and highly sensitive technique for detecting early, clinically unsuspected breast cancer [3]. The radiographs are searched for signs of abnormality by expert radiologists but complex structures in appearance and signs of early disease are often small or subtle. That's the main cause of many missed diagnoses that can be mainly attributed to human factors [3,4]. Studies have shown an error rate between 10% - 30% for detection of cancer in screening studies [5], [6]. Of these, a percentage of 52% can be attributed to breast cancer signs misinterpretation while another 43% is mainly due to sheer overlook of signs in abnormal scans [6] by expert radiologists. The consequences of errors in detection or classification are costly, so there has been a considerable interest in developing

A. Likas, K. Blekas, and D. Kalles (Eds.): SETN 2014, LNAI 8445, pp. 510–520, 2014.

methods for automatically classifying suspicious areas of mammography tissue, as a means of aiding radiologists in double reading screening mammograms and improving the efficacy of screening programs thus avoiding unnecessary biopsies.

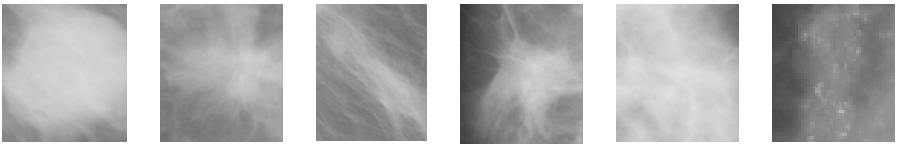

Fig. 1. Types of breast cancer on mammograms

Among the various types of breast abnormalities, clustered microcalcifications and mass lesions are the most important ones. Masses and clustered microcalcifications often characterize early breast cancer [7] that can be detectable before a woman or the physician can palp them. Masses appear as dense regions of varying sizes and properties and can be characterized as circumscribed (Fig 1a), spiculated (Fig 1b), or ill defined (Fig 1c). On the other hand, microcalcifications (Fig 1f), appear as small bright arbitrarily shaped regions on the large variety of breast texture background. Finally, asymmetry and architectural distortion (Fig 1e-d), are also very important yet difficult abnormalities to detect.

Computer-aided methods in the field of digital mammography are divided into two main categories; *computer aided detection* methods [8-11] that are capable of pinpointing regions of suspicion (ROS) in mammograms for further analysis from an expert radiologist and *computer aided diagnosis* methods [12-16] which are capable of making a decision whether the examined ROS consist of abnormal or healthy tissue. Sampat et al. [17], and Rangayyan et al. [18], provide an extensive review on different stages of CAD methodology for breast cancer.

In this study, we propose a method to classify regions of suspicion (ROS) that contain abnormal or healthy tissue using a novel feature extraction technique based on *shapelets* and a SVM classifier. The proposed method linearly decomposes each tested ROS into a series of localized basis functions with different shapes, which are called *shapelets*. The basis set that is chosen consists of the weighted Hermite polynomials which correspond to the perturbations about a circular Gaussian and in their asymptotic form to the Edgeworth expansion in several dimensions [19-20]. The coefficients of the linear decomposition are used as features that are fed into the classifier. Extensive experiments have shown great accuracy of over 86% in recognizing normal and abnormal breast tissue in mammograms which outperforms by 8% similar features based on Zernike moments [21].

The structure of this paper is as follows: In the next section, a detailed description of the features extracted from mammograms using *Shapelet* analysis as well as Zernike moments is given. Additionally a description of the utilized classifier based on Support Vector Machines is also given. In section 3 we present the data set and the experimental results and finally, in section 4 some conclusions are drawn.

2 The Proposed Method

The proposed system consists of a feature extraction module that decomposes each examined ROS into a linear composition of a set of localized basis functions with different shapes. The feature vectors which are composed by the calculated shapelet coefficients are used in the neural classifier module which is comprised by a scheme of support vectors machine (SVM) [22].

2.1 Feature Extraction

Shapelet Based Feature Extraction

Shapelets in the Cartesian Domain

The Shapelet image decomposition method was introduced in [19], providing an efficient method for the estimation of discrimination data for accurate detection of ROS areas, based on several expressions of the spatial pixels distribution of an object as a linear sum of orthogonal 2D functions,

$$f(x,y)=\sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} f_{n_1,n_2}\cdot\Phi_{n_1,n_2}(x,y;\beta) \tag{1}$$

where f_{n_1,n_2} are the Shapelet coefficients to be determined, whereas the Shapelet basis functions $\Phi_{n_1,n_2}(x,y;\beta)$ are

$$\Phi_{n_1,n_2}(x,y;\beta)=\frac{\mathrm{H}_{n_1}\left(x\cdot\beta^{-1}\right)\cdot\mathrm{H}_{n_2}\left(y\cdot\beta^{-1}\right)\cdot e^{-\frac{|x|^2+|y|^2}{2\cdot\beta^2}}}{\beta\cdot 2^n\cdot\sqrt{\pi\cdot n_1!n_2!}} \tag{2}$$

where $\mathrm{H}_n(.)$ is a Hermite polynomial of order n, and β is the scaling factor of each Shapelet. The Hermite polynomials form an orthonormal basis set, ensuring that the features which are extracted from any image can be determined by:

$$f_{n_1,n_2}=\iint_{\mathbb{R}} f(x,y)\cdot\phi_{n_1,n_2}(x,y;\beta)\cdot dx\cdot dy \tag{3}$$

Shapelets in the Polar Domain

Polar shapelets have been introduced in [19]. They include all the major properties of the Cartesian Shapelets with a scaling β in spite of the polar shapelets are separable in r and θ. For this reason, the polar shapelets coefficients are easier to comprehend in terms of rotational symmetries, are simpler and more intuitive. A function $f(r,\theta)$ in polar coordinates is decomposed as a weighted sum:

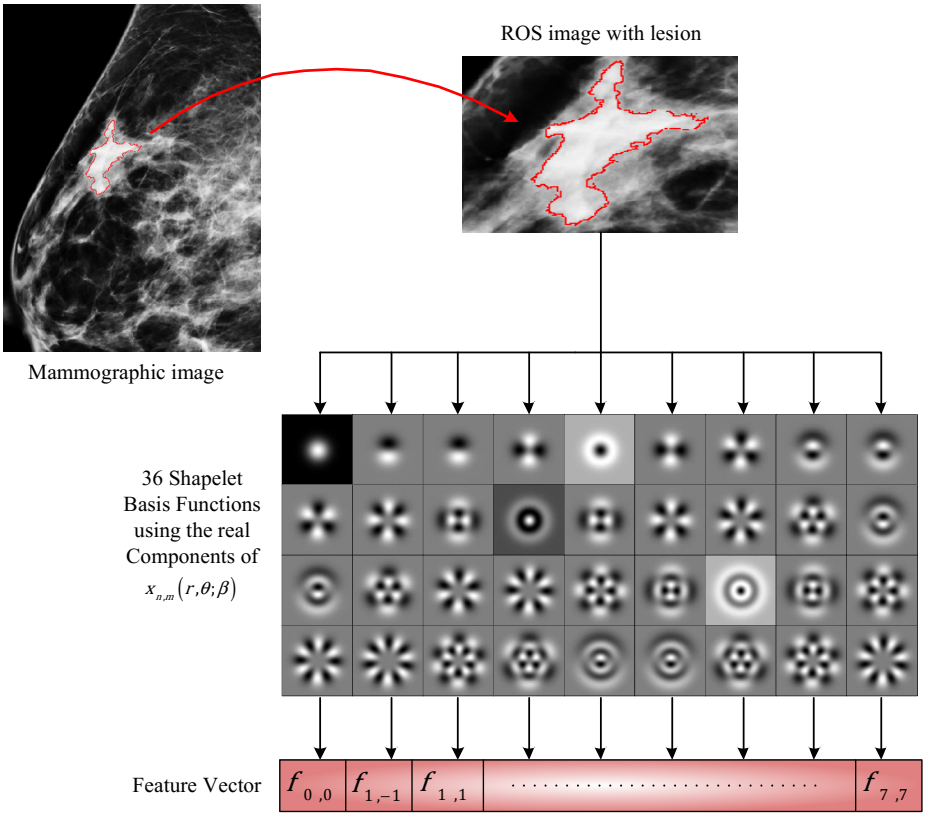

Fig. 2. The feature extraction process using the real parts of the polar shapelets basis functions

$$f(r,\theta)=\sum_{n=0}^{\infty}\sum_{m=-n}^{+n} f_{n,m}\cdot x_{n,m}(r,\theta;\beta) \tag{4}$$

where $x_{n,m}(r,\theta;\beta)$ are the polar basis functions related to Laguerre polynomials $L_{\frac{n-|m|}{2}}^{|m|}$

$$x_{n,m}(r,\theta;\beta)=\frac{(-1)^{\frac{n-|m|}{2}}}{\beta^{|m|+1}}\cdot\left[\frac{\left[(n-|m|)/2\right]!}{\pi\left[(n+|m|)/2\right]!}\right]^{1/2}\cdot r^{|m|}\cdot L_{\frac{n-|m|}{2}}^{|m|}\left(\frac{r^2}{\beta^2}\right)\cdot e^{-\frac{r^2}{2\cdot\beta^2}}\cdot e^{-i\cdot m\cdot\theta} \tag{5}$$

with n and m both even or odd, respectively.

The polar shapelets coefficients of order n and m are calculated using the overlap integral

$$f_{n,m}=\iint_{\mathbb{R}} f(r,\theta)\cdot x_{n,m}(r,\theta;\beta)\cdot r\cdot dr\cdot d\theta \tag{6}$$

In Fig. 2 we present the feature extraction process with the 36 estimated shapelets in the polar domain.

Zernike Moments Based Feature Extraction

In optical systems, which are constructed using lenses, optical fibers or other optical components are often circular, therefore when the structure of the measured deformations and aberrations is required to be characterized efficiently the Zernike functions forming a complete, orthogonal basis set over the unit circle. The Zernike moments firstly introduced in digital image analysis problems in 1980 [23, 24] and evaluated for many types of images [25-27]. In digital mammography, Zernike moments have been already been proposed and used in Computer Aided Diagnosis Systems (CAD) for the diagnosis of breast masses as descriptors (features) of shape and marginal characteristics [21].

The Zernike moments defined as a family of orthogonal functions over the unit disk, ensuring that there is minimum correlation among the moments and consequently, minimum redundancy of information, invariant both to rotation and displacement. In polar coordinates the Zernike function $Z_n^m(r,\theta)$ is defined by:

$$Z_n^{\pm m}(r,\theta) = R_n^m(r)\cdot e^{jm\theta} \tag{7}$$

where $R_n^m(r)$ are the Zernike polynomials. The index n is the degree of the polynomial, while m is the polynomial order. The Zernike polynomials are defined as a finite sum of powers of r^2:

$$R_n^m(r) = \sum_{k=0}^{(n-|m|)/2} \frac{(-1)^k\cdot(n-k)!}{k!\left(\frac{n+|m|}{2}-k\right)!\left(\frac{n-|m|}{2}-k\right)!}\cdot r^{n-2k} \tag{8}$$

where $n = 0,1,2,\ldots,\infty$ and $m = -n,-n+2,-n+4,\ldots,+n$, with $n-|m| = even$ and $|m| \le n$.

2.2 The SVM Classifier

The main idea behind SVMs, when dealing with a real life pattern classification problem, is to find an *optimal* hyperplane in a vector space called, feature space, where the original data are embedded via a nonlinear mapping. By the term *optimal*, it is suggested that for a separable classification task (linear separable in feature space), the hyperplane (w, b) with the maximum margin from the closest data points belonging to the different classes is selected among the infinite choices of hyperplanes.

Using mathematical notation, having a data set $D = \{(\boldsymbol{x}_i, y_i)\}_{i=1}^n$ of labeled examples $y_i \in \{-1,1\}$ and a nonlinear mapping from the input space into a high dimensional feature space $\boldsymbol{\varphi}(.)$ where the data samples are linearly separable we are seeking for the vector $\mathbf{w}_0$ that minimizes $\frac{1}{2}\|\mathbf{w}\|^2$ subject to the constraints

$$y_i\left(\mathbf{w}^T\varphi(\mathbf{x}_i)+b\right)\geq 1 \quad i=1,...,n\,. \tag{9}$$

However, the mapping into a higher feature space through a nonlinear function does not guaranty perfect separation of the classes for many real-life problems. Therefore we have to introduce slack-variables ξ_i that measure the deviation of a data point from the ideal condition of pattern separability and relax the hard margin constraints as follows:

$$y_i\left(\mathbf{w}^T\varphi(\mathbf{x}_i)+b\right)\geq 1-\xi_i\,,\ \ \xi_i\geq 0\,,\ \ i=1,...,n\,. \tag{10}$$

From the above formulation it is obvious that data with $1\geq\xi_i\geq 0$ are correctly classified while data with $\xi_i>1$ are classified incorrectly. With the introduction of the slack variables the goal is to maximize the margin and at the same time to keep the number of data samples with $\xi>0$ as small as possible (for $\xi_i=0$ for all i we have the linear separable case and all the data have at least maximum margin from the separating hyperplane). Therefore the quantity that has to be minimized is

$$\frac{1}{2}\|\mathbf{w}\|^2+C\sum_{i=1}^{n}g(\xi_i) \tag{11}$$

$$\text{where}\quad g(\xi_i)=\begin{cases}0, & \xi_i=0\\ 1, & \xi_i>0\end{cases}$$

The above problem is a combinatorial problem, which is difficult to solve and as a result an alternative approach is required. A mathematical tractable implementation of the previous two demands is given by minimizing

$$\frac{1}{2}\|\mathbf{w}\|^2+C\sum_{i=1}^{n}\xi_i \tag{12}$$

which corresponds to the so called L1 soft margin SVMs. The parameter C is a positive constant that controls the relative influence of the two competing terms. The solution to this optimization problem subject to the constraints is given by the saddle point of the primal Lagrangian equation:

$$L_p(\mathbf{w},b,\boldsymbol{\xi},\boldsymbol{\alpha},\boldsymbol{\beta})=\frac{1}{2}\|\mathbf{w}\|^2+C\sum_{i=1}^{n}\xi_i-\sum_{i=1}^{n}\alpha_i\left(y_i\left(\mathbf{w}^T\boldsymbol{\varphi}(\mathbf{x}_i)+b\right)-1+\xi_i\right)-\sum_{i=1}^{n}\beta\xi_i \tag{13}$$

This leads to the dual maximization problem of the dual Langrangian equation:

$$L_d(\boldsymbol{\alpha})=\sum_{i=1}^{n}\alpha_i-\frac{1}{2}\sum_{i,j=1}^{n}\alpha_i\alpha_j y_i y_j\left(\boldsymbol{\varphi}^T(\mathbf{x}_i)\boldsymbol{\varphi}(\mathbf{x}_i)\right) \tag{14}$$

subject to the constraints

$$\sum_{i=1}^{n} \alpha_i y_i = 0 \quad \text{and} \quad C \geq \alpha_i \geq 0 \quad , \quad i = 1,...,n \tag{15}$$

The solution of the above optimization problem leads to the *optimal* discriminating function:

$$f(\mathbf{x}) = sign\left(\sum_{i=1}^{n} y_i a_i \left(\boldsymbol{\varphi}^T (\mathbf{x}_i) \boldsymbol{\varphi}(\mathbf{x}) \right) + b \right) \tag{16}$$

with $b = \frac{1}{|I|} \sum_{i \in I} \left(y_i - \sum_{j=1}^{n} y_j a_j \left(\phi^{\mathrm{T}} (x_i) \phi(x_j) \right) \right)$, $i \in I := \{ i : 0 < \alpha_i < C \}$.

The points for which $a_i > 0$ are called Support Vectors and they are usually a small portion of the original data set. Especially those support vectors that have $a_i = C$ are those that lie within the margin or in the wrong side of the separating hyperplane have the strongest influence on the solution $\mathbf{w}_0$. The choice of C which is done a-priori by the user heavily affects the width of the margin.

In this formulation, if the nonlinear mapping function is chosen properly, the inner product in the feature space can be written in the following form

$$\phi^T (x_i) \cdot \phi(x_j) = K(x_i, x_j) \tag{17}$$

where K is called the inner-product kernel. A kernel function is a function in input space and, therefore, we do not explicitly perform the nonlinear mapping $\varphi(.)$. Instead of calculating the inner product in a feature space $\varphi^T(x_i) \cdot \varphi(x_j)$, one can indirectly calculate it using the kernel function $K(x_i, x_j)$. Different kernels produce different learning machines and different discriminating hypersurfaces.

Among others the most popular are the polynomial learning machines, the radial basis function networks and the two-layer perceptrons. In our experimental procedure we have employed radial basis function machines (the width σ^2, which is common to all kernels is specified also a priori by the user)

$$K(\mathbf{x}, \mathbf{x}_i) = \mathrm{e}^{\left(-\frac{1}{2\sigma^2} \|\mathbf{x} - \mathbf{x}_i\|^2 \right)}, \quad i = 1,...n \tag{18}$$

for recognizing healthy and abnormal mammographic tissue testing all kinds of abnormalities.

3 Experimental Results

3.1 The MIAS Data Set

In our experiments the MIAS MiniMammographic Database [28], provided by the Mammographic Image Analysis Society (MIAS), was used. The mammograms are digitized at 200-micron pixel edge, resulting to a 1024×1024 pixel resolution. In the MIAS Database there is a total of 119 ROS containing all kinds of existing

abnormal tissue from masses to clustered microcalcifications. The smallest abnormality extends to 3 pixels in radius, while the largest one to 197 pixels. These 119 ROS along with another 119 randomly selected sub-images from entirely normal mammograms were used throughout our experiments. A database of 238 ROSs of 35×35 pixels size is used. The image database has been designed to include all types of different ROS areas abnormalities, i.e. circumscribed, spiculated, ill defined masses, microcalcifications, asymmetry and architectural distortion as well as regions with normal (healthy) tissue.

3.2 Training and Evaluation of the SVM Classifier

From a total number of 238 ROS included in the MIAS database, 119 regions are used for the training procedure: 60 groundtruthed abnormal regions along with 59 randomly selected normal ones. In the evaluation procedure, the remaining 119 regions are used that contain 59 groundtruthed abnormal regions together with 60 entirely normal regions. Therefore, no ROS was used both in the training and testing procedure. The above procedure was repeated 10-times using randomly chosen ROSs in the training and testing sets in order to get the unbiased classification performance of the SVM.

Each ROS image is converted in polar coordinates and the inner product among images and each mask of the polar Shapelets basis functions is obtained. In this research, basis functions of order $n = 7$ and $-n \le m \le n$ have been chosen. The feature vector is composed by the real parts of the polar Shapelets resulting to 36 coefficients (SH-Features), defining the image information used for detection of ROS areas. Therefore, each ROS image is described by a feature vector with 36-dimensionality. In the classification experiments, the features derived by the polar Shapelets are used to train the SVM scheme. In order to find the SVM configuration for maximum recognition accuracy the values of (σ^2, C) for the RBF machine, a *grid-search* approach was used in a systematic manner with different values for the parameters followed by a cross validation.

For comparison reasons we have also estimated the Zernike moments of an image $I(r,\theta)$ in polar coordinates, by projecting the image onto the orthogonal basis of the Zernike functions, thus resulting to the following complex numbers:

$$a_n^{\pm m} = \frac{n+1}{\pi} \cdot \sum_{n=0}^{\infty} \sum_{\substack{m=-n, \\ n-|m|=even, \\ |m| \le n}}^{+n} I(r,\theta) \cdot Z_n^m(r,\theta) \tag{19}$$

The factor $\frac{n+1}{\pi}$ is used to normalize the moment's expression. In this research, Zernike functions of degree $n = 6$ and order $m = -n, -n+2, -n+4, \cdots, +n$ have been chosen. The feature vector is composed by the real parts of the Zernike moments $a_n^{\pm m}$ resulting to 28 coefficients.

In Fig. 3 we present the recognition accuracy when different values of the shapelet scaling factor β where used. It is clear that the best results were achieved when using a scaling factor equal to 0.2 that results to the best overall true positive identification rate of over 86% for RBF machines.

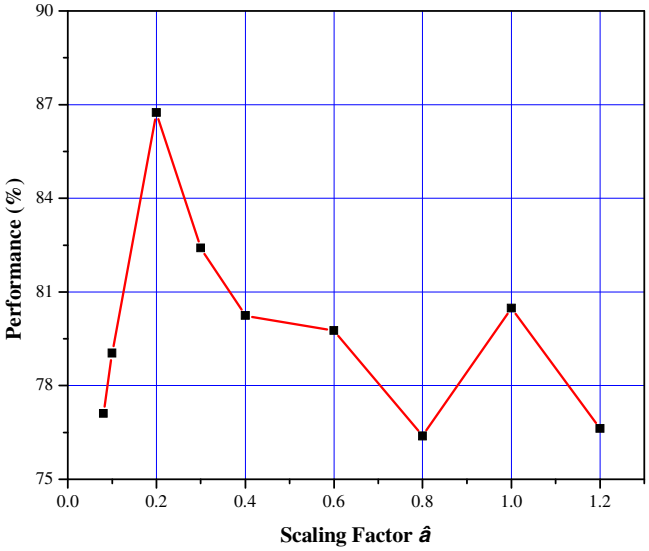

Fig. 3. The performance of the Shapelet features using different values of scaling factor β

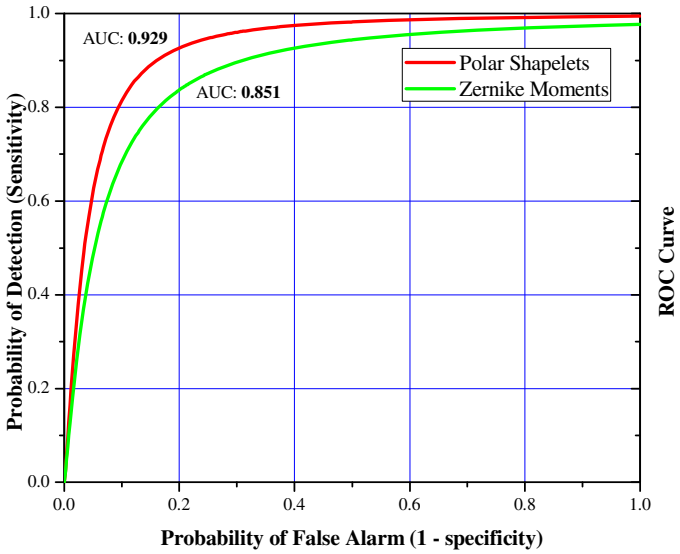

Fig. 4. The ROC curve of the proposed method

In Fig. 4 a ROC curve, is presented visualizing the relative trade-offs between benefits (true positives) and costs (false positives). As it is shown, the closer the ROC curve follows the left-hand and the top border of the ROC space, the more accurate the classification result will be. Additionally, a quantitative comparison of the proposed feature extraction method's accuracy is shown in Fig. 4 where the ROC curve

is compared with the corresponding ROC calculated using Zernike moments (ZE-features) based features described in the previous section. As it is shown, the proposed features outperform Zernike based features resulting in 86% classification accuracy compared to 78.5% obtained by ZE-Features for RBF support machines.

4 Conclusions

In this paper we investigated the performance of a classifier based on Support Vector Machines and a novel set of features based on Shapelet image decomposition, in the problem of recognizing breast cancer in ROS of digital mammograms. It is well known that the disease diagnosis on mammograms is a very difficult task even for experienced radiologists due to the great variability of the mass appearance. The experimental results showed superior performance and accuracy of the proposed feature set compared to similar features that have already been proposed when used to recognize all different types of breast abnormalities. However further studies will address the problem of automatic selection of the shapelet scaling β and order in more complex tasks such as the discrimination between different types of abnormalities.

References

1. Feuer, E.J., Wun, L.M., Boring, C.C., Flanders, W.D., Timmel, M.J., Tong, T.: The lifetime risk of developing breast cancer. Journal of the National Cancer Institute 85(11), 892–897 (1993)
2. Sirovich, B.E., Sox, H.C.: Breast cancer screening. Surgical Clinics of North America 79(5), 961–990 (1999)
3. Martin, J.E., Moskowitz, M., Milbrath, J.R.: Breast cancer missed by mammography. American Journal of Roentgenology 132(5), 737–739 (1979)
4. Kalisher, L.: Factors influencing false negative rates in xero-mammography. Radiology 133, 297 (1979)
5. Kerlikowske, K., Carney, P.A., Geller, B., Mandelson, M.T., Taplin, S.H., Malvin, K., Ballard-Barbash, R.: Performance of screening mammography among women with and without a first-degree relative with breast cancer. Annals of Internal Medicine 133(11), 855–863 (2000)
6. Bird, R.E., Wallace, T.W., Yankaskas, B.C.: Analysis of cancers missed at screening mammography. Radiology 184(3), 613–617 (1992)
7. Tabar, L., Dean, B.P.: Teaching Atlas of Mammography, 2nd edn. Thieme, NY (1985)
8. Sampaio, W.B., Diniz, E.M., Silva, A.C., Cardoso de Paiva, A., Gattass, M.: Detection of masses in mammogram images using CNN, geostatistic functions and SVM. Computers in Biology and Medicine 41(8), 653–664 (2011)
9. Pereira, D.C., Ramos, R.P., Zanchetta do Nascimento, M.: Segmentation and detection of breast cancer in mammograms combining wavelet analysis and genetic algorithm. Computer Methods and Programs in Biomedicine (2014)
10. Oliver, A., Freixenet, J., Martí, J., Pérez, E., Pont, J., Denton, E., Zwiggelaar, R.: A review of automatic mass detection and segmentation in mammographic images. Medical Image Analysis 14(2), 87–110 (2010)

11. Wang, D., Shi, L., Heng, P.A.: Automatic detection of breast cancers in mammograms using structured support vector machines. Neurocomputing 72(13-15), 3296–3302 (2009)
12. Eltoukhy, M.M., Faye, I., Samir, B.B.: A statistical based feature extraction method for breast cancer diagnosis in digital mammogram using multiresolution representation. Computers in Biology and Medicine 42(1), 123–128 (2012)
13. Eltoukhy, M.M., Faye, I., Samir, B.B.: A comparison of wavelet and curvelet for breast cancer diagnosis in digital mammogram. Computers in Biology and Medicine 40(4), 384–391 (2010)
14. Buciu, I., Gacsadi, A.: Directional features for automatic tumor classification of mammogram images. Biomedical Signal Processing and Control 6(4), 370–378 (2011)
15. Ren, J.: ANN vs. SVM: Which one performs better in classification of MCCs in mammogram imaging. Knowledge-Based Systems 26, 144–153 (2012)
16. Zanchetta do Nascimento, M., Martins, A.S., Neves, A.L., Ramos, R.P., Flores, E.L., Carrijo, G.A.: Classification of masses in mammographic image using wavelet domain features and polynomial classifier. Expert Systems with Applications 40(15), 6213–6221 (2013)
17. Sampat, M.P., Markey, M.K., Bovik, A.C.: Computer-Aided Detection and Diagnosis in Mammography. In: Handbook of Image and Video Processing. Elsevier, London (2003)
18. Rangayyan, R.M., Ayres, F.J., Leo Desautels, J.E.: A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs. Journal of the Franklin Institute 344(3), 312–348 (2007)
19. Refregier, A.: Shapelets–I. A method for image analysis. Monthly Notices of the Royal Astronomical Society 338(1), 35–47 (2003)
20. Refregier, A., Bacon, D.: Shapelets—II. A method for weak lensing measurements. Monthly Notices of the Royal Astronomical Society 338(1), 48–56 (2003)
21. Tahmasbi, A., Saki, F., Shokouhi, S.B.: Classification of benign and malignant masses based on Zernike moments. Computers in Biology and Medicine 41(8), 726–735 (2011)
22. Scholkopf, B., Burges, C.J.C., Smola, A.J.: Advances in Kernel Methods. Support Vector Learning. The MIT Press, London (1999)
23. Zernike, F.: Diffraction theory of the cut and its improved form, the phase contrast method. Physica 1, 689–704 (1934)
24. Teague, M.: Image analysis via the general theory of moments. J. Opt. Soc. Am. 70, 920–930 (1980)
25. Teh, C., Chin, R.: On image analysis by the methods of moments. IEEE Transactions on Pattern Analysis and Machine Intelligence 10, 496–513 (1988)
26. Bailey, R., Srinath, M.: Orthogonal moment feature for use with parametric and non-parametric classifiers. IEEE Transactions on Pattern Analysis and Machine Intelligence 18, 389–399 (1988)
27. Khotanzad, A., Hong, Y.: Classification of invariant image representations using a neural network. IEEE Transactions on Acoustics and Speech Signal Processing 38, 1028–1038 (1990)
28. http://peipa.essex.ac.uk/info/mias.html