Accepted refereed manuscript of:

Wajid S & Hussain A (2015) Local energy-based shape histogram feature extraction technique for breast cancer diagnosis, *Expert Systems with Applications*, 42 (20), pp. 6990-6999.

Local Energy-based Shape Histogram Feature Extraction Technique for Breast Cancer Diagnosis

Summrina Kanwal Wajid* and Amir Hussain

Division of Computing Science & Maths, School of Natural Sciences, University of Stirling, Stirling SCOTLAND, UK

Abstract—This paper proposes a novel local energy-based shape histogram (LESH) as the feature set for recognition of abnormalities in mammograms. It investigates the implication of this technique on mammogram datasets of the Mammographic Image Analysis Society and INbreast. In the evaluation, regions of interest were extracted from the mammograms, their LESH features were calculated, and they were fed to support vector machine (SVM) classifiers. In addition, the impact of selecting a subset of LESH features on classification performance was also observed and benchmarked against a state-of-the-art wavelet based feature extraction method. The proposed method achieved a higher classification accuracy of range 99.00±0.50 as well as an A_z value of 0.9900±0.0050 with multiple SVM kernels where linear kernel performs with 100% accuracy for distinguishing between the abnormalities (masses vs. microcalcifications). Hence, the general capability of the proposed method was established, in which it not only distinguishes between malignant and benign cases for any type of abnormality but also among different types of abnormalities. It is therefore concluded that LESH features are an excellent choice for extracting significant clinical information from mammogram images with significant potential for application to 3-D MRI images.

Keywords—Computer-aided decision support system (CADSS), local energy-based shape histogram (LESH), support vector machine (SVM), local energy model, receiver operating characteristic (ROC) curve.

1   Introduction

Breast cancer is a fatal disease that originates in breast tissue. It primarily affects women; however, men can also develop it. Breast cancer caused more deaths than any other cancer in women in the US in 2011, when it was the second-most diagnosed cancer after skin cancer. In the UK, breast cancer accounts for 30% of all female cancers; almost one in nine women is estimated to develop it. In the EU, a woman is diagnosed with breast cancer every 2.5 minutes (American Cancer Society, 2011). Although the risk of breast cancer is lowest in less-developed countries, it is increasing every year (Ahmedin et al., 2011; Hollander, D. 2002; Shulman et al., 2010).

The fatality rate of breast cancer can be reduced by early diagnosis and treatment. Computer-aided decision support systems (CADSSs) combined with mammography help doctors diagnose breast cancer at an early stage. CADSSs function in four key steps: preprocessing, feature extraction, feature selection, and classification. The preprocessing step includes breast image segmentation and filtering, which is followed by image normalization to improve image quality and reduce noise. In feature extraction, images of lesions are extracted from enhanced images by using various techniques. In feature selection, an additional set of features is selected. In classification, the selected feature set is classified to separate false signals from true ones.

The essence of an accurate diagnosis exists in the selection of suitable features that can differentiate between normal and abnormal cases. The literature reports many techniques for feature extraction. (Yu et al., 2006) used a combination of model-based and statistical texture features to detect microcalcifications. This approach first detects the region of interest (ROI)—in this case, the area that contains the microcalcifications—by using wavelets and thresholds. It then extracts texture features from the ROI by using Markov random fields, fractal models, and statistical features. The performance of this approach was evaluated by using the area under the free-response receiver operating characteristic curve (FROC). A true positive rate of 94% was achieved with a 1.0 false positive per image rate.

Multi-resolution methods greatly interest researchers in image processing, analysis, biology, and other fields. (Eltoukhy et al., 2010) applied curvelet transforms to mammograms. They used the largest coefficients of the curvelet as feature vectors. The method resulted in a 98.59% classification accuracy when using SVM classifiers. (Lladó et al., 2009) used extended local binary pattern (LBP) histogram features. These features were extracted from regions that showed masses on the mammogram.

The method resulted in an A_z value of 0.94 ±0.02. (Wang et al., 2009) classified malignant and benign masses by using SVM for features selected from the ROI. The feature set included curvilinear, texture, Gabor, and multi-resolution features. Performance was measured using by ROC and reached 0.97 with a maximum accuracy of 91.4%. (Karahaliou et al., 2008) found that the texture properties of the area surrounding the microcalcification were significant when detecting malignancies in mammograms. This approach considered features such as grey-level first-order statistics, grey-level co-occurrence matrices, and Laws' texture energy measures, which were extracted from the surrounding tissue of regions of interest (ST-ROI). A redundant discrete wavelet transform (RWT) was then applied to the image. The wavelet coefficient first-order statistics and wavelet coefficient co-occurrence matrices for ST-ROI were then used as features. This combination of features resulted in an A_z performance of 0.9989. (Delogu et al., 2007) extracted 16 different features, including mass perimeter, circularity, mean of the normalized radial length, and others, from segmented masses and studied their various combinations. These features, which were further selected by using feature discriminating power and linear correlation interplay techniques, were fed to a multi-layer perceptron neural network for classification. It achieved 97.8% accuracy. (Rashed et al., 2007) experimented with different wavelet transforms to analyse their ability to discriminate among different classes of abnormalities, such as clusters, speculated lesions, circumscribed masses, and ill-defined lesions. (Moussa et al., 2005) used wavelet (Daubechies) coefficients as feature vectors. In this approach, the horizontal, vertical, and diagonal detailed coefficients were first extracted from the image decomposition via Daubechies wavelets. The coefficients were then normalized and the energy for each feature vector was calculated. This output was used for classification purposes. The features were reduced depending on the level of energy required.

(Verma et al., 2010) used density and shape features, such as mass margins, abnormality assessment rank, patient age, and subtlety values as features and then applied soft clustering using k-mean to separate each class further. After clustering within malignant and benign cases, neural network was applied with two additional layers other than the conventional three layer neural network architecture. One of these additional layers was a fixed weight layer while the other one is the cluster layer which prescribes the natural tendency of the output towards a malignant or benign class. The algorithm achieved a maximum accuracy of 97.5% on test set. (Diaz-Huerta et al., 2014) applied contrast enhancement with extended maxima threshold for image enhancement next it extracted special, texture, and spectral domain features from it. It further applied support vector machine classifier to distinguish between malignant and benign microcalcifications. The result reached sensitivity of 85.9% over all. (Rahimeh et al. 2015) Applied artificial neural network (ANN) and cellular neural network (CNN) based segmentation upon mammograms. Further it extracted texture and shape features from segmentation area and fed it to artificial neural network for classification between malignant and benign masses. The result reached accuracy of 96.87% at maximum. (Soltanian-Zadeh et al., 2004) Experimented with texture: multiwavelet, wavelet, Haralick, and shape features to classify between malignant and benign microcalcifications. Features were further selected using genetic algorithm. The result was maximum 0.89 ROC value. (Marcelo et al., 2013) considered wavelets (Daubechies-8 (db8), Symlet 8 (sym8) and bi-orthogonal 3.7 (bior3.7)) as features and applied polynomial classifier to discriminate abnormal cases from nomal ones. They achieved A_z value of 0.98±0.03 as classification performance on DDSM mammogram dataset.

Most of the above mentioned algorithms focus on one type of abnormality (either mass or microcalcification) and try to diagnose their malignancy whereas this paper proposes a method which can efficiently deals with different types of abnormalities at the same time while discriminating between malignant and benign cases with higher accuracy rate. The algorithm also produces good results in discriminating between different types of abnormalities as can been seen in results section.

Main objective of this research is to evaluate the implication of a novel LESH feature extraction technique for diagnosing malignancies in mammograms and this algorithm generates more accurate results than other mentioned algorithms above. LESH has been successfully applied in pattern matching applications with promising results. (Zakir et al., 2011) applied LESH to automatically detect and recognize different road signs. (Sarfraz & Hellwich, 2008a; 2008b; 2009) experimented with LESH in face recognition systems for different face and head poses. LESH works by calculating a histogram of the local energy pattern within the image. A histogram is a simple technique that forms the basis of many spatial domain image processing techniques (Gonzalez & Woods, 2002). It provides useful image statistics that can be used to further analyse and process the image, and is described later in this paper.

The remainder of this paper is organized as follows. Section 2 introduces the proposed LESH scheme. Section 3 describes the experimental work; the results and discussion are presented in Section 4. Conclusions and future work are presented in Section 5.

2 Local Energy-based Shape Histogram Feature Extraction

LESH works by converting an image into a combination of local energies along different orientations. (Morrone, M.C. & Owens, R., 1987) suggested that features extracted at the points of maximum phase congruency can be helpful in image analysis. The type of phase and amplitude of local maxima of the energy function determines the type, sign, and contrast of a

feature. This framework to calculate the phase congruency in two-dimensional images while using a high-pass filter to obtain image features at different scales is given as (Kovesi, P. D., 1999):

$$PC(z) = \frac{\sum_n W(z) \lfloor A_n(z) \Delta\Phi_n(z) - T \rfloor}{\sum_n A_n + \epsilon} \qquad (1)$$

where

$$\Delta\Phi_n(z) = \cos\left(\phi_n(z) - \bar{\phi}_n(z)\right) - \left|\sin(\phi_n(z) - \bar{\phi}_n(z))\right| \quad (2)$$

and T is the noise cancellation factor, while $W(z)$ is the weighting of the frequency spread. A_n and ϕ_n represent the amplitude and phase angle, respectively, of local complex value Fourier components at location z in the image of size n. ϵ is a constant value incorporated to avoid division by zero (Kovesi, P. D., 1999).

Here, A_n and ϕ_n are calculated using the logarithmic Gabor wavelets filter (Morrone, M.C. & Owens, R., 1987). It detects low-level features that are invariant to image illumination, contrast, and image magnification (Kovesi, P. D., 2000). These one-dimensional symmetric/anti-symmetric filters are transformed into a two-dimensional form by applying a Gaussian spreading function across the direction that is perpendicular to its orientation. The image is convolved with a bank of Gabor kernels at each of the 16 sub-regions along 8 different orientations as defined by (Zhang et al., 2007; Zakir et al., 2011):

$$G_{u,v}(z) = I(z) * \Psi_{u,v}(z) \qquad (3)$$

where $z = (x, y)$ represents the image position, the asterisk '*' is the convolution operator, and $G_{u,v}(z)$ is the convolution result of the Gabor kernel at orientation u and scale v. The Gabor wavelet kernel $\Psi_{u,v}(z)$ can be calculated as follows ((Zhang et al., 2007) :

$$\Psi_{u,v}(z) = \frac{\|k_{u,v}\|^2}{\sigma^2} e^{-\|k_{u,v}\|^2 \|z\|^2 / 2\sigma^2} \left[e^{ik_{u,v}z} - e^{-\sigma^2/2}\right] \quad (4)$$

Where $\|.\|$ is a norm operator and wave vector $k_{u,v} = k_v e^{i\emptyset_u}$ with $k_v = k_{max}/f^v$ and $\emptyset_u = \pi\mu/8$ and k_{max} is the maximum frequency and and f is the space factor between the kernels in the frequency domain and σ=2π (Liu, C., 2004). These filters are designed to detect features in all orientations because they uniformly tile the frequency plane. Energy, $PC(z)$, at each orientation is calculated and then summed as a whole. The sum of energies for each orientation is normalized by overall sum and scales of amplitude of individual wavelet responses at a specific location z. Hence, the calculated two-dimensional PC is given as:

$$PC_2(z) = \frac{\sum_0 \sum_n W_0(z) \lfloor A_{n,0}(z) \Delta\Phi_{n,0}(z) - T_0 \rfloor}{\sum_0 \sum_n A_{n,0} + \epsilon} \qquad (5)$$

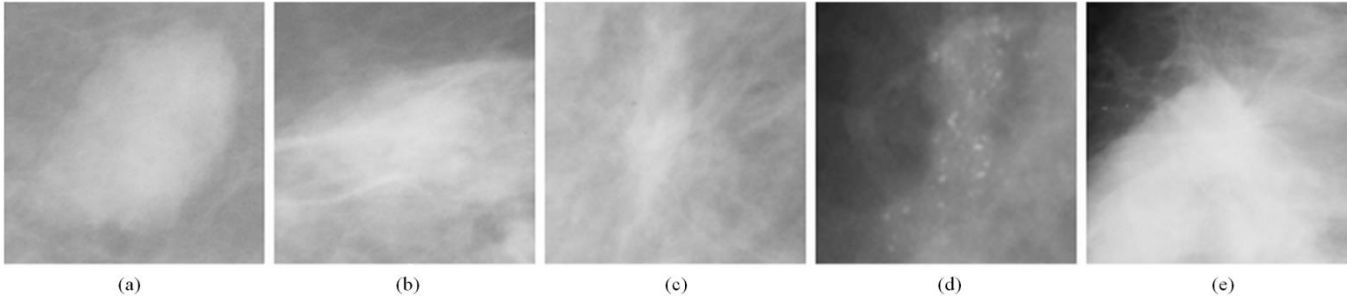where *0* denotes the index of orientation, as detailed in (Kovesi, P. D., 2000).

Fig. 1. Different samples of abnormalities in the MIAS database: (a) circumscribed masses (CIRC) mdb015; (b) ill-defined masses (MISC) mdb032; (c) speculated masses (SPIC) mdb145; (d) calcification (CALC) mdb241; (e) architectural distortion (ARCH) mdb115.

3   Experimental Work

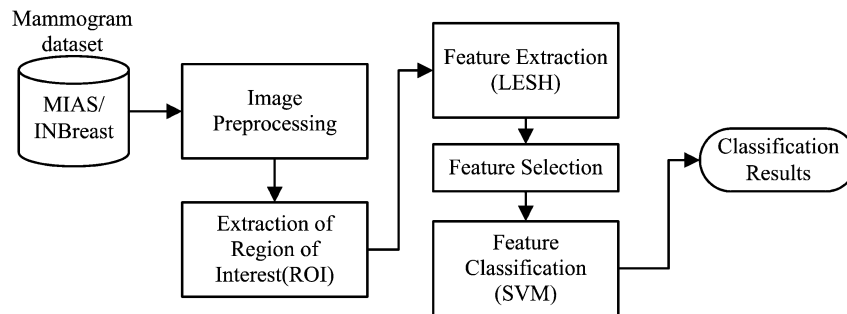The experiment proceeds through specific steps, as shown in Fig. 2.

3.1   Mammogram Datasets

We experimented with two publically available mammogram datasets: the Mammographic Image Analysis Society (MIAS) and INbreast datasets. The MIAS dataset (Suckling J. et al., 1994) contains images in 1024x1024-pixel resolution, which have been digitized with 50-micron pixel edges; they are available in Portable Gray Map (PGM) format. The dataset contains 322 mammogram cases from 161 patients. Of these 322 mammograms, 207 are normal, 51 are malignant, and 64 are benign. The images have been labelled by radiologists and provide a variety of abnormal cases naming: microcalcification (25 cases); circumscribed masses (23 cases); architectural distortion (19 cases); speculated masses (19 cases) and miscellaneous (15 cases). Almost 50% of the images are comprised of backgrounds with different types of noise.

The INbreast dataset contains 115 cases (410 images) with different types of abnormalities (Moreira I. C. et al., 2012). The images were acquired from the Centro Hospitalar de S. João [CHSJ], Breast Centre, Porto, with the permission of the Portuguese National Committee of Data Protection and Hospital's Ethics Committee. The images are in Digital Imaging and Communications in Medicine (DICOM) format with annotations and metadata available. The prominent abnormalities are masses (108 cases) and calcifications (308 cases) with only 3 cases of architectural distortion and 14 cases of asymmetry. 396 ROI were gathered from the dataset for experimentation out of those 117 were masses (with some cases had multiple masses) and 279 were calcifications.

3.2 Image Preprocessing

Mammograms are inherently noisy images. This noise hinders the true detection of small micro calcifications and masses. Mammograms were pre-processed to suppress the noise and enhance important image features. The preprocessing includes two steps: image normalization and then application of Contrast Limited Adaptive Histogram Equalization (CLAHE). The steps are summarized below:

3.2.1   Image Normalization

Mammograms obtained from the MIAS dataset have intensity values in the range of 0 to 255, whereas INbreast mammograms have intensity values in the 0 to 32,767 range. Image normalization was applied to the mammograms from both datasets to adjust the intensity within a common range of 0 to 1 values. Let z_0 be the original mammogram image then after normalization image z is given as:

$$z = (z_0 - Min)\frac{newMax - newMin}{Max - Min} + newMin \quad (6)$$

Where original image z_0 has intensity range (Min, Max) and new image z has intensity range (newMin, newMax) (Gonzalez & Woods, 2002).

Further, contrast limited adaptive histogram equalization (CLAHE) (Derek T Puff et al., 1994) was applied for image enhancement.

3.2.2   Contrast Limited Adaptive Histogram Equalization (CLAHE)

CLAHE applies histogram equalization(HE) to sub-regions of an image by first dividing an image into contextual blocks (tiles). It then creates a histogram for each block using a specific number of bins and clips the histogram at a certain threshold. It maps each region according to the new histogram results. Finally, it interpolates gray-level mapping to reconstruct the final CLAHE image (Sundaram et al., 2011).

This approach is the most widely used technique for enhancing medical images and is based on histogram equalization.

3.2.2.1   Histogram Equalization (HE)

It is a classical technique which uses histogram of an image for contrast adjustment. It increases the contrast of low intensity areas in the image by spreading out the most frequent intensity values. For an image z with pixel intensity values ranging from 0 to L-1, where L being 256 here, a normalized histogram p with a bin for each possible intensity level, the probability of occurrence of intensity value k can be written as (Gonzalez & Woods, 2002):

$$p(k) = \frac{number\ of\ pixels\ with\ intensity\ k}{total\ number\ of\ pixels\ in\ the\ image} = \frac{n_k}{n} \qquad k: 0\ to\ L - 1 (7)$$

The Cumulative Distribution Function corresponding to p is given as:

$$cdf(k) = \sum_{j=0}^{k} p(j) \quad (8)$$

Then histogram equalization is an image transformation such that the transformed image has flat Cumulative Frequency Distribution (CDF) given as:

$$df(m) = i * m \quad (9)$$

where *m* is a contrast within the range 0 to L-1.

3.3 Region of Interest Extraction

After preprocessing step, ROIs were extracted from 113 abnormal cases of the MIAS dataset and 396 abnormal cases from the INbreast dataset. The abnormalities included:

1. Circumscribed masses (CIRC): These are the most commonly detected abnormalities in mammograms. They are dense and have circumscribed margins. Malignant masses are markedly dense with irregular borders (de Paredes E. S., 2007).
2. Ill-defined masses (MISC): These masses have poor structure and are predominantly malignant. They are typically surrounded by fine tendrils (de Paredes E. S, 2007).
3. Speculated masses (SPIC): These masses have an irregular appearance with speculation (de Paredes E. S., 2007).
4. Calcification (CALC): These can be defined as benign or malignant depending on their sizes, shapes, and distribution patterns. Their sizes may vary from minute to 3 mm in diameter (de Paredes E. S., 2007).
5. Architectural Distortion (ARCH): These are speculated masses without a central dense mass (de Paredes E. S., 2007).

Fig.1 depicts examples of these abnormalities. The ROIs were extracted according to the information available with the dataset.

3.4 Feature Extraction

Histogram of the local energy information, LESH, was obtained along 16 sub-regions of the ROI along 8 different orientations as the feature set. This resulted in a 16 x 8 = 128 dimensional feature vector, as suggested by (Zakir et al., 2011). The histogram of the local energy of each sub-region along each filter was generated and combined to preserve the relationship between different regions. Next, an orientation map was generated by assigning each pixel an orientation that had the largest energy across all scales. The local histogram was calculated as follows:

$$h_{r,b} = \sum W_r \times PC_2(z) \times \delta_{r-b} \tag{10}$$

$$W_r = \frac{1}{\sqrt{2\pi}\sigma} e^{\left[(x-r_{x0})^2+(y-r_{y0})^2\right]/\sigma^2} \tag{11}$$

where *Wr* is the Gaussian weighting function of region r in the image z, PC_2 represents the local energy computed by the equation, and δ_{r-b}, represents Kronecker's delta of the orientation label map *L* and current bin *b* (Zakir et al., 2011). Fig.3 shows the effect of enhancement and the extracted LESH histogram of the ROI.

3.5 Feature Selection

Impact of selecting a subset of *N* largest coefficients say h_N from LESH feature vectors, $h_{r,b}$ and feeding them to the SVM classifier has been studied in detail. (Cristiane et al., 2003) applied the same technique for dimensionality reduction and classification improvement. They experimented with different numbers of largest wavelet coefficients as features for classification and reported the results.

3.6 Classification with Support Vector Machine

The selected LESH features, h_N were fed to the SVM classifier with different kernel functions. SVMs are supervised learning models that are used for classification and regression analysis. SVMs take a set of features as classes input data and predict to which of the possible classes each feature vector belongs. For feature vectors of training set in h_N, SVM classifier draws a hyper plane which separates them into two distinct classes: Malignant (class label being 1) and Benign (class label being 0). The hyper plane is given as below:

$$g(x) = w^T h_N + w_0 = 0 \tag{12}$$

Here *w* is a normal vector to the hyper plane. SVM is trained on different sample feature sets. This training algorithm generates a model that can assign a specific class to a new unknown feature vector. SVM can perform linear and nonlinear classifications. For nonlinear classifications, it first maps features into a high-dimensional feature space using different kernels. SVM performance can generally be improved by selecting appropriate kernel methods. Experiments were performed with the four kernels outlined below.

3.6.1 Linear Kernel

The linear kernel is given as follows:

$$k(h_N, h_{N'}) = h_N{}^T h_{N'} + c \tag{13}$$

where c is the optimal constant.

3.6.2 Polynomial Kernel

The polynomial kernel with degree d can be written as:

$$k(h_N, h_{N'}) = \left(\gamma h_N{}^T h_{N'} + c\right)^d \tag{14}$$

where γ and c are adjustable constants and d is the degree of the polynomial.

3.6.3 Gaussian, Radial Basis Function (RBF) Kernel

$$k(h_N, h_{N'}) = exp(-\gamma \|h_N - h_{N'}\|^2) \quad \gamma > 0 \tag{15}$$

where γ is a positive parameter to control the radius. It uses the Euclidean distance to avoid outliers.

3.6.4 Hyperbolic Tangent (Sigmoid) Kernel or Multi-Layer Perceptron (MLP)

It was originally derived from neural networks as:

$$k(h_N, h_{N'}) = tanh\left(\gamma h_N{}^T h_{N'} + c\right) \quad \gamma > 0 \tag{16}$$

where γ and c are adjustable constants.

4 Results and Discussion

Experiments have been conducted with MIAS and INbreast datasets to distinguish malignant and benign cases and to classify abnormalities according to their types. The results were stated using classification accuracy and the ROC curve as measures of performance. The classification accuracy was computed as:

$$Accuracy(T) = \frac{\sum_{i=1}^{T} Assess(t_i)}{|T|} \quad t_i \in T \tag{17}$$

$$Access(t) = \begin{cases} 1 \; if \; classify(t) = correct \; classification \\ 0 \; if \; classify(t) = incorrect \; classification \end{cases} \tag{18}$$

The ROC curve plots the fraction of true positives out of positives versus the fraction of false positives out of negatives. The ROC plots in Fig.4, 5 and 6 depict classification performance for MIAS and INBreast datasets for different SVM kernels. The area under the curve (AUC), A_z, signifies the classification performance and ranges between 0 and 1, with 1 being the highest performance. All reported results were averaged with ten-fold cross validation; they are presented in the tables below.

The results show that almost all SVM classifiers performed better with the LESH feature set. The prominent performance was achieved by the SVM linear classifier, which performed at a stable rate of 99.73% over all multiple feature selection sets (Table 1) and with an A_z value of 0.9975 ±0.0010 (Table 2) while using the INbreast dataset. The results were verified for the MIAS dataset when an accuracy of 100% was achieved by the SVM linear classifier (Table 3) and an A_z value of 1.0000 for 70 selected LESH features (Table 4). The second dominant classifiers were the polynomial and RBF; MLP showed an overall low performance accuracy.

The results in Table 5 verify that LESH is a good choice when discriminating among different types of abnormalities. The SVM linear classifier was capable of classifying mass or calcifications with no errors (Tables 5 and 6) for the INbreast dataset. The polynomial classifier was second, followed by RBF, in the performance score; MLP again showed a low performance. For the MIAS dataset, a few cases   existed   for   different   abnormalities.
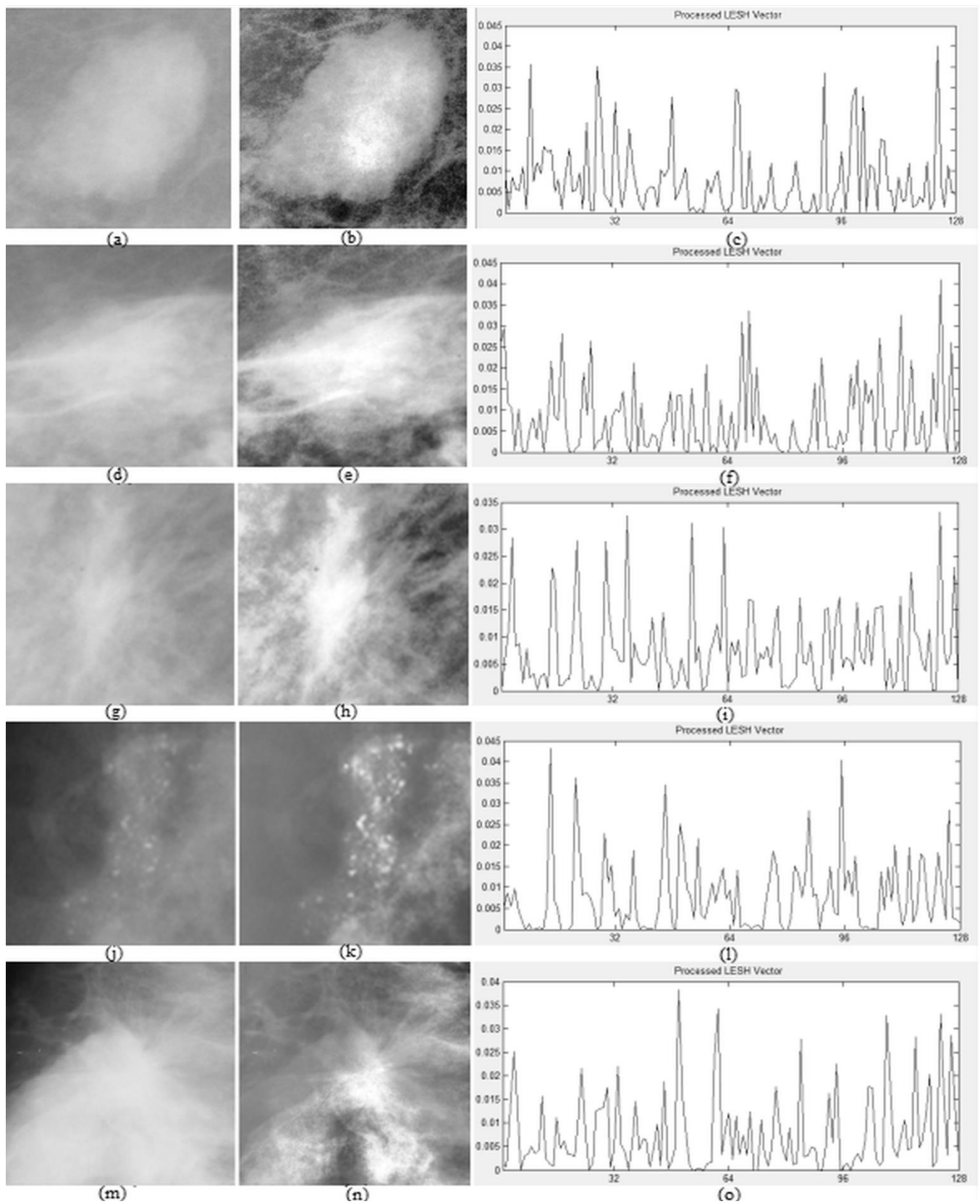
Fig. 3. LESH vector for different abnormalities obtained from software developed by (Zakir et al., 2011) (a) circumscribed masses (CIRC); (b) CIRC after enhancement; (c) LESH vector for CIRC; (d) ill-defined masses (MISC); (e) MISC after enhancement; (f) LESH vector for MISC; (g) speculated masses (SPIC); (h) SPIC after enhancement; (i) LESH vector for SPIC; (j) calcification (CALC); (k) CALC after enhancement; (l) LESH vector for CALC; (m) architectural distortion (ARCH); (n) ARCH after enhancement; (o) LESH vector for ARCH.

Multiclass SVM classification was performed based on the one-for-all scheme to distinguish between different types of abnormalities, as noted in Fig. 2. Application of the proposed feature set and SVM classifier still appropriately performed (Table 7) and affirmed the importance of LESH features.

In summary, the results reported in Tables 1-7 suggest that LESH features have the power to emphasize important features in the images and that these features can be used to extract useful information. These features perform well with different kernel methods; selecting a subset from all available features does not deteriorate classification performance.

Table 1 LESH-based Classification Accuracy % (Malignant/Benign) for INbreast Dataset

Features Selected	SVM with RBF	SVM with Linear	SVM with MLP	SVM with Polynomial
50	99.46	99.73	93.28	99.46
70	99.73	99.73	93.61	99.46
100	99.73	99.73	92.80	99.46
All	99.45	99.73	92.78	99.73

Table 2 LESH-based Classification Az Value (Malignant/Benign) for INbreast Dataset

Features Selected	SVM with RBF	SVM with Linear	SVM with MLP	SVM with Polynomial
50	0.9955	0.9964	0.9472	0.9940
70	0.9975	0.9970	0.9487	0.9937
100	0.9980	0.9982	0.9440	0.9936
All	0.9955	0.9975	0.9448	0.9975

Table 3 LESH-based Classification Accuracy % (Malignant/Benign) for MIAS Dataset

Features Selected	SVM with RBF	SVM with Linear	SVM with MLP	SVM with Polynomial
50	99.09	99.09	94.44	99.17
70	99.09	100.00	94.54	98.18
100	99.09	99.09	93.63	96.36
All	99.09	99.09	94.62	96.51

Table 4 LESH-based Classification Az Value (Malignant/Benign) for MIAS Dataset

Features Selected	SVM with RBF	SVM with Linear	SVM with MLP	SVM with Polynomial
50	0.9944	0.9944	0.9430	0.9900
70	0.9917	1.0000	0.9444	0.9733
100	0.9929	0.9929	0.9279	0.9616
All	0.9929	0.9936	0.9470	0.9583

Table 5 LESH-based Classification Accuracy % (Mass vs. Micro Calcification) for INbreast Dataset

Features Selected	SVM with RBF	SVM with Linear	SVM with MLP	SVM with Polynomial
50	99.50	100.00	86.61	99.75
70	99.49	100.00	87.62	99.74
100	99.50	100.00	91.56	99.75
All	99.44	100.00	87.09	100.00

Table 6 LESH-based Classification Az Value (Mass vs. Micro Calcification) for INbreast Dataset

Features Selected	SVM with RBF	SVM with Linear	SVM with MLP	SVM with Polynomial
50	0.9967	1.0000	0.8990	0.9982
70	0.9964	0.9964	0.9172	0.9983
100	0.9965	1.0000	0.9057	0.9982
All	0.9960	1.0000	0.9087	1.0000

Table 7 LESH Feature with Multiclass (One-vs.-All) SVM Classifier Performance Accuracy % for MIAS Dataset

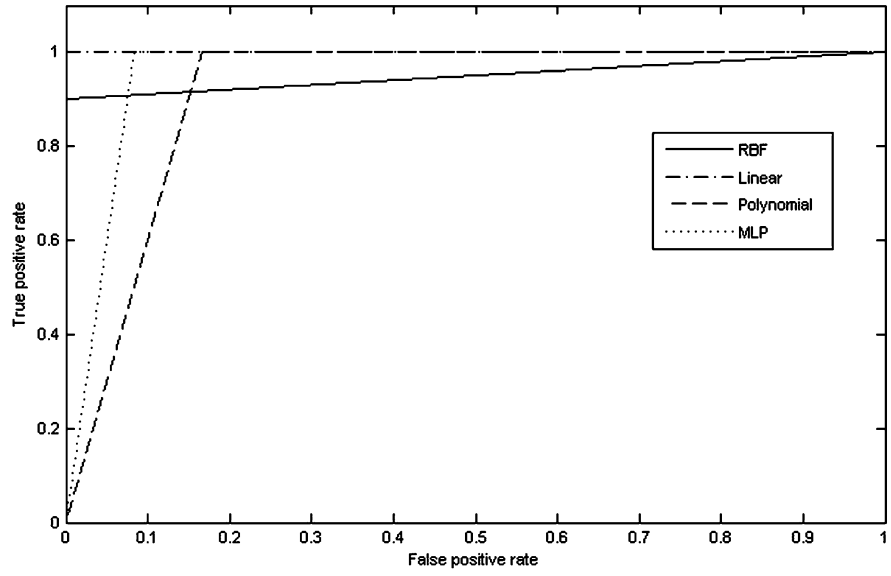| Abnormality Type	SVM with RBF	SVM with Linear	SVM with MLP	SVM with Polynomial
ARCH	85.45	100.00	79.90	100.00
MISC	100.00	95.45	79.90	100.00
SPIC	95.67	82.17	90.00	95.45
CIRC	97.56	100.00	72.15	92.20
CALC	97.56	97.10	77.89	97.56

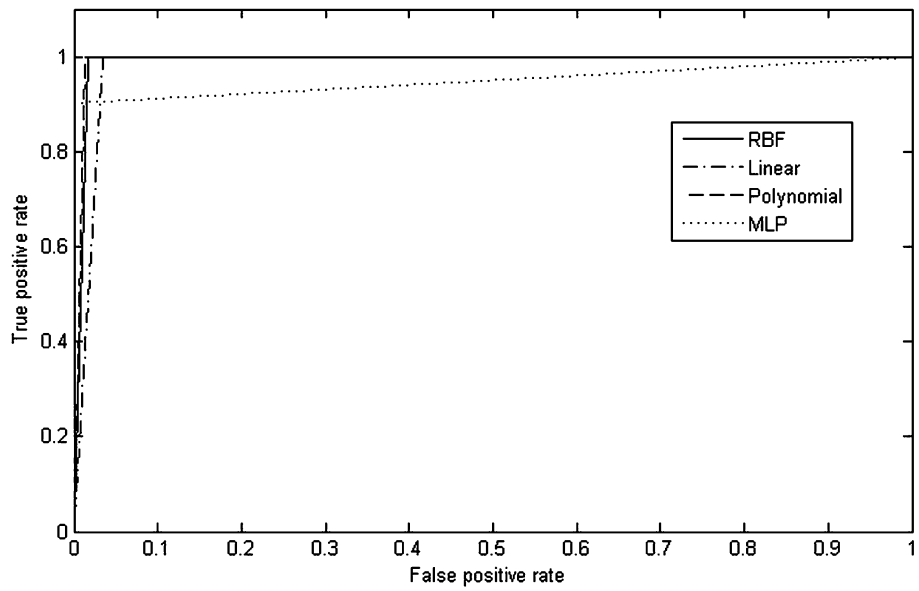Fig.4. ROC for classification between malignant/benign cases for MIAS dataset.

Fig.5. ROC for classification between malignant/benign cases for INBreast dataset.
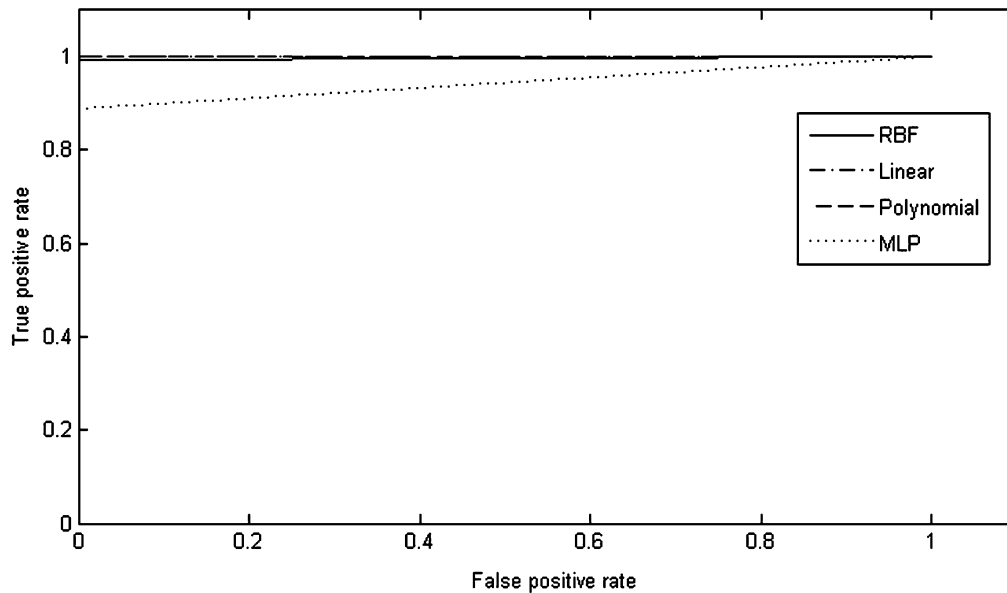
Fig.6. ROC for classification between masses/calcifications cases for INBreast dataset.

4.1 Comparative Analysis of LESH and Wavelet-based Feature Extraction for Classification Performance

In this study, LESH features were compared with wavelet transform-based feature extraction technique, which was used by (Moussa et al., 2005; Cristiane et al., 2003. In (Cristiane et al., 2003), the classification results of applying different types of wavelets to extract features from mammograms were compared. *N* largest wavelet coefficients were selected as feature vectors; later the SVM classifier was applied. These results were reproduced using Daubechies wavelet (Daubechies, I., 1992) for the 100 largest coefficients, as recommended by (Cristiane et al., 2003). These results are shown in Tables 8, 9, and 10. The results indicate superior performance using LESH features in both binary and multiclass classification.

Table 8 Wavelet (Daubechies) Feature Extraction-based Classification Results for Binary Classification between Malignant and Benign Cases for INbreast Dataset

Measure	SVM with RBF	SVM with Linear	SVM with MLP	SVM with Polynomial
ACCURACY %	99.19	99.46	94.56	99.46
A_z	0.9928	0.9919	0.9433	0.9919

Table 9 Wavelet (Daubechies) Feature Extraction-based Classification (Mass vs. Micro Calcifications) Results for INbreast Dataset

Measure	SVM with RBF	SVM with Linear	SVM with MLP	SVM with Polynomial
ACCURACY %	98.73	100.00	95.67	99.75
A_z	0.9898	1.0000	0.9600	0.9980

Table 10 Wavelet-based Feature Extraction for Multiclass (One-vs.-All) SVM Classifier Performance Accuracy % for MIAS Dataset

Abnormality Type	SVM with RBF	SVM with Linear	SVM with MLP	SVM with Polynomial
ARCH	80.10	94.10	80.10	92.44
MISC	78.34	76.67	86.56	71.89
SPIC	73.56	71.56	86.56	67.12
CIRC	67.12	79.90	46.15	98.00
CALC	76.67	100.00	96.10	100.00

4.2 Statistical Analysis

The analysis of variance (ANOVA) statistical inference test (Creech, S., 2003) was applied to calculate the significance of the improvement in accuracy by using LESH features compared to the wavelet texture features. The null hypothesis was defined as follows:

1) H0: LESH features are more discriminative in classifying differences in mammograms compared to wavelet features. This hypothesis was tested against an alternate hypothesis.

2) H1: LESH features are not discriminative in classifying differences in mammograms.

A 0.05 level of confidence was used for experiments, which is standard for statistical testing. The results for binary and multiclass classification are shown in Tables 6 through 9. The probability value (p-value) determined whether the null hypothesis should be rejected. In both binary and multiclass cases, the p-value was less than the 0.05 confidence level, which suggests that the difference between the classification accuracy was insignificant. Because the accurate detection of malignancies is critical in breast cancer diagnosis—to avoid unnecessary surgery—these results may be considered an improvement (Verma et al., 2010).

Table 11 ANOVA Feature Extraction Information for Binary Classification between Malignant and Benign Cases

SUMMARY		ANOVA single factor test		
Groups	Count	Sum	Average	Variance
LESH Features	8	7.90030	0.98754	0.00022
Wavelet Features	8	7.65630	0.95704	0.00012

Table 12 ANOVA Feature Extraction Analysis for Different Feature Selection Methods for Binary Classification

ANOVA

Source of Variation	SS	Df	MS	F	P-value	F crit
Between Groups	0.0037	1	0.0037	21.931	0.0004	4.60011
Within Groups	0.00237	14	0.0002			
Total	0.00609	15				

Table 13 ANOVA Feature Extraction Information for Multiclass Classification

SUMMARY		ANOVA Single Factor Test		
Groups	Count	Sum	Average	Variance
LESH Features	8	7.9003	0.987538	0.000216
Wavelet Features	8	7.6563	0.957038	0.000123

Table 14 ANOVA Feature Extraction Analysis for Different Feature Selection Methods for Multiclass Algorithm

ANOVA

Source of Variation	SS	Df	MS	F	P-value	F crit
Between Groups	0.00371	1	0.0040	20.916	0.0005	5.0011
Within Groups	0.00238	14	0.0002			
Total	0.00609	15				

5   Conclusion and future work

As mentioned earlier, LESH is a histogram of local energy, which is at a maximum level at the abrupt change of image intensity. Accordingly, it marks significant the texture variations in the local area. The highest degree coefficients correspond to the most significant and prominent set of features in the local area within an image. A subset of these features can be selected with some compromise of classification accuracy, which hence reduces the dimensionality. Experiments were performed with different numbers of highest degree coefficients and determined that $N = 70$ is the most appropriate number for improving classification accuracy. Overall, the results suggest that LESH generates an effective set of features that improves classification performance compared to a state-of-the-art wavelet based feature selection approach which is also evident from plots in Fig. 4,5 & 6.

The reason being that although wavelets have efficient image representation but fail to represent discontinuities along curves

and edges. Other competitive features are curvelets and contourlets. Curvelets on the other hand are superior to wavelets to sustain edges. Contourlets are closer to curvelets and are counted as discrete form of curvelets transform but have less clear directional features (Jianwei Ma and Plonka, G., 2010). Other than wavelets, many feature extraction techniques has been mentioned in section 1. Most of these techniques focus their study on a specific type of abnormality and try to detect their malignancy, whereas LESH based methodology provides the advantage of detecting malignancy of any type of abnormality as well as can differentiate among different types of abnormalities with quite efficiency.

Note that while the preliminary results reported in this paper should be taken with care, they do demonstrate the capability of exploiting LESH features for breast cancer detection, and a range of contributions and potential impact is envisaged from this work, both for clinical practice and further research into employing such models in other clinical applications.

For future work, more extensive evaluation, and clinical validation is required using additional clinical datasets benchmarked against other state-of-the-art feature selection and classification approaches. We also intend to extend cancer detection methodologies and experiments to breast magnetic resonance imaging (MRI) and 3-D breast imaging. A hybrid approach to adaptively optimise an ensembled feature set based on combination of LESH and other state of the art feature extraction technique while selecting significant features can be evaluated for enhancing classification performance. Further, an intensive study of clinical significance of LESH features when applied to real clinical dataset may lead to potential uselful intimations.

Acknowledgment

References

Ahmedin Jemal et al., (2011), Global Cancer Statistics, CA: A Cancer Journal for Clinicians, 61(2).

American Cancer Society. (2011). Breast Cancer Facts & Figures 2011-2012. Atlanta: American Cancer Society, Inc.

Cristiane Bastos Rocha Ferreira, & Díbio Leandro Borges (2003) .Analysis of mammogram classification using a wavelet transform decomposition, Pattern Recognition Letters, 24(7), 973-982.

Creech, S. (2003).ANOVA, http://www.statisticallysignificantconsulting.com/Anova.htm. Accessed March 2014.

Daubechies, I. (1992).Ten Lectures on Wavelets, PA: SIAM.

de Paredes E. S. (2007). Atlas of Mammography, (3rd ed.), Lippincott Williams and Wilkins, Philadelphia, USA.

Delogu, P., Fantacci, M. E., Kasae, P., & Retico, A.(2007). Characterization of mammographic masses using a gradient-based segmentation algorithm and a neural classifier, Computers in Biology and Medicine, 37, 1479-1491.

Derek T Puff et al. (1994). A method for determination of optimal image enhancement for the detection of mammographic abnormalities. Journal of Digital Imaging, 7(4),161-171.

Diaz-Huerta et al. (2014). Quantitative analysis of morphological techniques for automatic classification of micro-calcifications in digitized mammograms, Expert Systems with Applications, 41(16), 7361-7369, ISSN 0957-4174.

Eltoukhy, M. M., & Faye, I., Samir, B. B. (2010). Breast cancer diagnosis in digital mammogram using multi scale curvelet transform. Computerized Medical Imaging and Graphics, 34, 269–276.

Gonzalez, R. C., Woods, R. E. (2002). Digital image processing. (2nd ed.). ISBN 0-201-18075-8.

Hollander, D. (2002). In developed and developing countries, breast cancer risk is reduced by 4% for each year of breastfeeding. Perspectives on Sexual and Reproductive Health, 34, 318–321.

Jianwei Ma and Plonka, G. (2010). The Curvelet Transform, Signal Processing Magazine, IEEE, 27(2), 118-133.

Karahaliou, Anna N. et al. (2008). Breast Cancer Diagnosis: Analyzing Texture of Tissue Surrounding Microcalcifications. Information Technology in Biomedicine, IEEE Transactions, 12,731-738.

Kovesi, P. D. (1999). Image features from phase congruency. Videre: Journal of Computer Vision Research, MIT Press 1, 3, 1-26.

Kovesi, P. D. (2000). Phase congruency: a low-level image invariant, Psychological Research , 64, 136–148.

Lladó, X., Oliver, A., Freixenet, J., Martí, R., & Martí, J. (2009). A textural approach for mass false positive reduction in mammography. Computerized Medical Imaging and Graphics: The Official Journal of the Computerized Medical Imaging Society, 33,415-422.

Marcelo Zanchetta do Nascimento, Alessandro Santana Martins, Leandro Alves Neves, Rodrigo Pereira Ramos, Edna Lúcia Flores, Gilberto Arantes Carrijo. (2013). Classification of masses in mammographic image using wavelet domain features and polynomial classifier. Expert Systems with Applications, 40(15), 6213-6221, ISSN 0957-4174.

Moreira I. C. et al. (2012). INbreast: toward a full-field digital mammographic database, Academic Radiology, 19(2), 236-248, ISSN 1076-6332, http://dx.doi.org/10.1016/j.acra.2011.09.014. (http://www.sciencedirect.com/science/article/pii/S107663321100451X).

Morrone, M.C. & Owens, R. (1987). Feature detection from local energy. Pattern Rec. Letters. 1 103-113.

Mousa, R., Munib, Q. & Moussa, A. (2005). Breast cancer diagnosis system based on wavelet analysis and fuzzy-neural, Expert Syst. Appl., 28, 713-723.

Rahimeh Rouhi et al. (2015). Benign and malignant breast tumours classification based on region growing and CNN segmentation, Expert Systems with Applications, 42(3), 990-1002.

Rashed, E. A., Ismail, I. A. & Zaki, S. I. (2007). Multiresolution mammogram analysis in multilevel decomposition, Pattern Recognition Letters, 28, 286-292.

Sarfraz, M. S. & Hellwich, O. (2008a). Head Pose Estimation in Face Recognition across Pose Scenarios. Int. Conference on Computer Vision Theory and Applications VISAPP, Portugal, 1, 235 -242.

Sarfraz, M. S. & Hellwich, O. (2008b). An efficient front-end facial pose estimation system for face recognition. International Journal of Pattern Recognition and Image Analysis, Springer, 18, 434–441.

Sarfraz, M. S., & Hellwich, O. (2009). On Head Pose Estimation in Face Recognition. Computer Vision And Computer Graphics. Theory And Applications, Lecture Notes CCIS, 24,162-175.

Shulman, L. N., Willett, W., Sievers, A., & Knaul, F. M. (2010). Breast Cancer in Developing Countries: Opportunities for Improved Survival. Journal of Oncology, 2010, 595167. doi:10.1155/2010/595167.

Soltanian-Zadeh et al. (2004). Comparison of multiwavelet, wavelet, haralick, and shape features for microcalcification classification in mammograms, Pattern Recognit, 37, 1973–1986.

Suckling J. et al. (1994). The Mammographic Image Analysis Society Digital Mammogram Database Exerpta. Medical International Congress Series, 1069 ,375-378.

Sundaram, M. et al. (2011). Histogram based contrast enhancement for mammogram images, Signal Processing, Communication, Computing and Networking Technologies (ICSCCN), 2011 International Conference, 842-846.

Verma, B., McLeod, P. & Klevansky, A. (2010). Classification of benign and malignant patterns in digital mammograms for the diagnosis of breast cancer. Expert Systems with Applications, 37, 3344-3351.

Wang, D. F., Shi, L., Heng, P. A. (2009). Automatic detection of breast cancers in mammograms using structured support vector machines, Neurocomputing, 72, 3296–3302.

Yu, S.-N., Li, K.-Y., Huang, Y.-K. (2006). Detection of microcalcifications in digital mammograms using wavelet filter and Markov random field model, Computerized Medical Imaging and Graphics, 30,163–173.

Zakir, U., Zafar, I., Edirisinghe, A. E. (2011). Road Sign Detection and Recognition by using Local Energy Based Shape Histogram (LESH). International Journal of Image Processing, 4, 566–582.

Zhang, B., Shang, S., Chen, X., Gao, W. (2007). Histogram of Gabor Phase Pattern (HGPP): A novel object representation approach for face recognition. IEEE Trans. on Image Processing, 16, 57-68.