Title: Effective combination of Zernike moments and LBP operators on Curvelet coefficients for the diagnosis of breast tumors

Authors: Maedeh Abbasi[1,*] , Mohammad Hosein Parvizi[2], Keivan Maghooli[3]

[1] MSc, Department of Medical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.
E-mail address: maede.abbasi69@gmail.com Tel: +98 9128430764.

[2] MSc, Department of Medical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.
E-mail address: mhparvizi1990@gmail.com Tel: +98 9352865353.

[3] PhD, Department of Medical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.
E-mail address: k_m_iau@yahoo.com Tel: +98 21 44474330.

*** Corresponding Author:** Maedeh Abbasi
MSc, Department of Medical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.
Unit 27, Floor 8, No. 6, Jamshidi St., Kaj Sq., Sa'adat Abad district, Tehran, Iran. Tel.: +98 9128430764.
E-mail: maede.abbasi69@gmail.com

Effective combination of Zernike moments and LBP operators on Curvelet coefficients for diagnosis of breast tumors

Maedeh Abbasi[a], Mohammad Hosein Parvizi [a], Keivan Maghooli[a]

[a] Department of Medical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

ARTICLE INFO

Keywords:

Tumor classification
Breast cancer
LBP operators
Zernike moments
Curvelet transform

ABSTRACT

Breast cancer is one of the most common cancer among women. This paper has been focused on the enhancement of a Computer-Aided Diagnosis (CADx) system for breast cancer diagnosis. The extraction of features is a critical step in CADx systems. We aimed to make use of extracting of features by two different methods. Those images which are taken by X-ray have detailed information about breast and its texture. So, texture analysis methods have been used in this research. According to BI_RADS, three different features have been introduced to distinct all kinds of breast masses, therefore we also used shape and margin characteristics. After applying the Homomorphic filter on mammogram images from DDSM database in order to increase contrast, the tumor region of interest was extracted manually. In this proposed method, we utilized two series of different features, Zernike moments as a shape and margin descriptor and LBP operators applied on Discrete Curvelet Transform as a texture descriptor. Finally, features were reduced using Principal Component Analysis (PCA) technique and are given to the Support Vector Machine for both learning and testing procedures. In experimental results, the accuracy and precision are 95.26 and 96.89 respectively that these results show superior performance.

1. Introduction

One of the most common cancers observed among women in the age range of 40-55 years is breast cancer, in developed and developing countries. It has been considered as the most common cause of death after lung cancer. Early detection of breast cancer increases the chances of survival and treatment options (Nascimento et al., 2013; Rouhi, Jafari, Kasaei, & Keshavarzian, 2015). Among the different tools, X-ray mammography is a common method for early detection of cancerous lumps which has very low-cost and is available, and plays an important role in reducing deaths (Braz Junior, da Rocha, Gattass, Silva, & Paiva, 2013; Zhang, Silber, & Mazurowski, 2015). Despite the fact that current mammography is the best screening method for breast cancer, its interpretation is difficult, and even though the recent development in digital mammography technology to help in better detection of breast disorders has dramatically increased the survival chances of patients (Guzmán-Cabrera et al., 2012), it is not perfect and has limitations. For example, in the case of dense breast tissue in young women, it can be seen that the result is not always correct and has an error. The physiological characteristics and the overlap of the patient's breast tissue will create a mammography image with low contrast which can cause medical errors. Furthermore, this screening method will create a huge amount of images (Lei Zhen & Chan, 2001). So, radiologists do not recognize 30% of cancer cases, depending on the density of breast tissue (Dheeba, Albert Singh, & Tamil Selvi, 2014). That is why in order to improve the accuracy and efficiency of screening mammography in the detection, CAD systems were introduced.

In general, there are two methods for computer-aided detection systems: The first is computer-aided detection method to determine the presence or absence of lesions focused on the identification of suspicious area (CADe), the second category is used to diagnose the type of lesion as benign or malignant and interpret their characteristics (CADx) (Tambasco Bruno et al., 2016) .CADx systems are used to classify lesions and complete the analysis by calculating the probability of malignancy in the suspicious area (Elter & Horsch, 2009).

Mammogram lesions are categorized into masses, microcalcifications and architectural distortion in CAD systems. In mammograms, microcalcifications appear as small spots because the high contrast in images are easily realizable but the diagnosis of masses is difficult because they are inseparable from adjacent tissues and also A) they are varied in size, shape, and density; B) have low contrast; C) their densities are very close to the density of the surrounding tissue; D) are surrounded by non-uniformity tissue that has the same properties. Finally, sensitivity and specificity of tumor diagnosis from mammography images is not optimal (Byung-Woo Hong & Bong-Soo Sohn, 2010; Nascimento et al., 2013). The diagnostic process in CAD systems performs such as a professional radiologist and checks the features that a radiologist considers and then classifies them.

A CAD system consists of four basic steps: (1) Pre-processing (2) segmentation (3) feature extraction and (4) classification (Saraswathi & Srinivasan, 2017). One of the fundamental steps in CADx to classify the suspicious regions is the feature extraction.

In general, in these systems, two categories of features are extracted from mammography which includes: morphological and non-morphological features. Morphological features provide information such as the size and shape of the lesion and non-morphological features work on the properties of gray levels that are not interpretable by humans simply. An important class of such features is feature sets that describe the texture (Gupta & Markey, 2005).

The masses are determined by shape, size, marginal characteristics and density. Similarly, microcalcifications are characterized by size, morphology, distribution, and heterogeneity. Based on these characteristics, a tumor in a mammogram image can be classified as benign or malignant (Surendiran & Vadivel, 2011; Vadivel & Surendiran, 2013).

A smooth visual texture includes minor changes in the values of digital numbers (DN) on a surface, while a rough visual texture includes many abrupt changes in the values of DN on a surface (Coburn & Roberts, 2004). The texture is always a sign of visual pattern recognition and gives a visual perception of roughness and smoothness of image features (Tang & He, 2013).

There are several approaches to pattern recognition, which includes global and local features. A local binary pattern is a local approach in the domain of structural operators and the root of this operator is from the co-occurrence statistics, n-tuple method and textons or texels (Mäenpää, 2003). The LBP operator was originally introduced as a complementary measure for the local contrast of the image (Harwood, Ojala, Pietikäinen, Kelman, & Davis, 1995). This method produces very useful histograms for describing the texture of features that are not sensitive to brightness or rotation of the image and are resistant to such changes (Jae Young Choi, Dae Hoe Kim, Plataniotis, & Yong Man Ro, 2012).

In this research, we also used Zernike moments that are a function that map image into a set of complex Zernike polynomials (Haddadnia, Ahmadi, & Faez, 2003; Wang, Mottershead, & Mares, 2009). Zernike moments were used to extract shape and mass margin features in this study. Orthogonality makes no redundancy and overlap of information between the moments of different order and repetition. This property helps each moment to be independent and unique in the information on the image.

There are two main approaches to enhance the operation of these systems include: (1) model-based approaches; (2) frequency-based approaches (Tambasco Bruno et al., 2016). Model-based approaches specify texture by creating stochastic properties. Some systems that use this approach are such as Markov random fields (Yu & Huang, 2010), autoregressive model (Mayerhoefer et al., 2010) and fractal features (Mohammed et al., 2018). Frequency-based models like wavelet transform (Beura, Majhi, & Dash, 2015), the ridgelet transform (Ramos, Nascimento, & Pereira, 2012) and the curvelet transform (Eltoukhy, Faye, & Samir, 2010) can be obtained from local operators and statistical attributes.

A vast variety of feature extraction method in combination with different feature selection and reduction methods and classification methods have been used to develop CADx systems like the work done by (Nascimento et al., 2013) which used the features of the wavelet domain together with the polynomial classifier for the classification of masses; in another work, the texture analysis was performed based on the Curvelet transform and local binary pattern and it was continued using the ANOVA feature selection method, and reached 85% accuracy on the DDSM database (Tambasco Bruno et al., 2016). In (Moura & Guevara López, 2013), a new descriptor invariant of rotation, Histogram of Invariant Divergence (HGD), was used to on rounded-shaped objects like masses, and they could develop patterns of gradients that are variable in rotation. (Vyas & Rege, 2007) introduced kinds of malignant lesions were classified into circumscribed, speculated, architectural distortion, miscellaneous, asymmetry, and calcification using the combination of Chebyshev moments which are discrete moments with Logpolar coordinate system and compared with Zernike moment which is continuous. In the work of (Venkatalakshmi & Janet, 2017), Pseudo-Zernike moments were used for the feature extraction step and finally, classified into benign and malignant classes using SVM. Two methods for classifying three classes: normal, malign and benign, were proposed; the first method is called convolutional

neural network-discrete wavelet (CNN-DW) which improves mammography images split into four sub-bands by means of a two-dimensional discrete wavelet transform (2D-DWT); the second method is the convolutional neural network-Curvelet transform (CNN-CT), in which the discrete Curvelet transform (DCT) is used. In both methods, the dense scale-invariant feature (DSIFT) is extracted for all sub-bands and used as a convolutional neural network input (CNN), and Softmax layer and support vector machine (SVM) layer are used for CNN training (Jadoon, Zhang, Haq, Butt, & Jadoon, 2017). In the study of (Gardezi, Faye, Adjed, Kamel, & Eltoukhy, 2017), the texture features were extracted and the geometric invariant shift transform and the statistics feature from the Curvelet gray level co-occurrence matrices were used and combined and the SVM classifier with a polynomial kernel was used for classification into normal and non-normal classes. In (Saraswathi & Srinivasan, 2017)'s study, Curvelet was used to analyze the input image, that is, separation of suspicious areas from normal tissue and then, Fractal features of the Curvelet Scale Layers were extracted for analysis and classification.

This article focuses on the development of the CADx system in order to recognize cancerous masses. The rest of the article is organized as follows: Section 2 provides an introduction of the Homomorphic filter, Zernike moments, local binary pattern, PCA and SVM in detail. Section 3 discusses the experimental work; Section 4 summarizes the results and Section 5 is the conclusion.

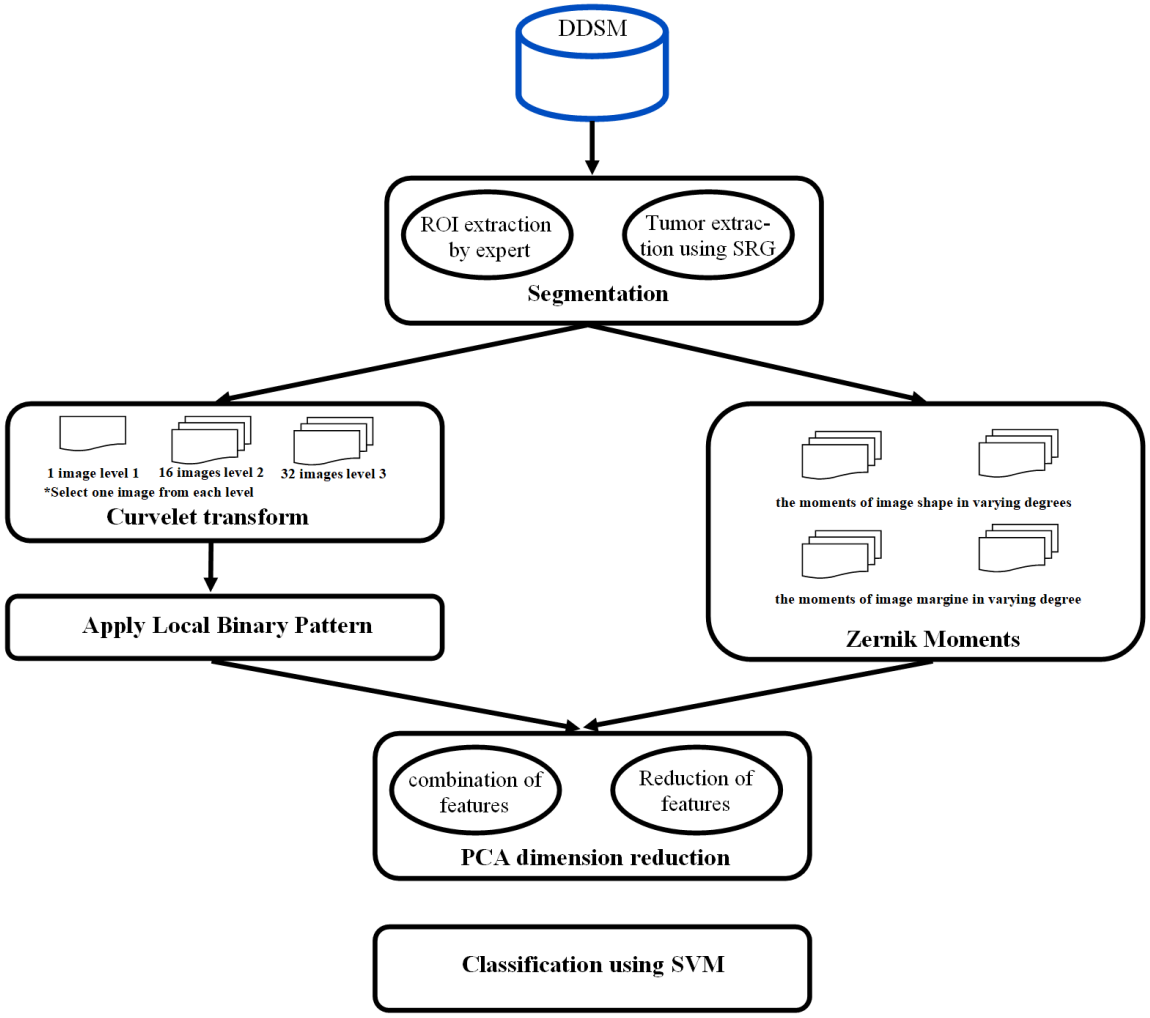

Fig.1. Procedures of the proposed methodology

2. Methods and materials

In this work, the combination of two feature extraction methods, Zernike moments and LBP operators that are applied to Curvelet coefficients, were used. The reduction of features was done using PCA and SVM classifier was then applied for classification. The proposed CAD system is shown in the block diagram in Figure 1.

2.1. Contrast enhancement using homomorphic filter

Using the homomorphic filter, we would be able to normalize the brightness across an image and increase a contrast simultaneously. As the tumor tissue is denser, by increasing the contrast of the images, the tumor can be easily recognized.

the damages of image $f(x\,.\,y)$ are due to components intensity $i(x\,.\,y)$ or reflection $r(x\,.\,y)$ and homomorphic filter has good control over the intensity and reflection of images. (Görgel, Sertbas, & Ucan, 2013) utilized a homomorphic filter to enhance the contrast of mammograms. For this purpose, a filter function $H(u\,.\,v)$ was used to give a different effect on high and low-frequency components of the Fourier transform.

Generally, pixels in an image according to the intensity of the relevant point in scene and reflection of that point are valued. Therefore, the image $f(x\,.\,y)$ can be expressed by the following relationship:

$$f(x\,.\,y) = i(x\,.\,y)r(x\,.\,y) \tag{1}$$

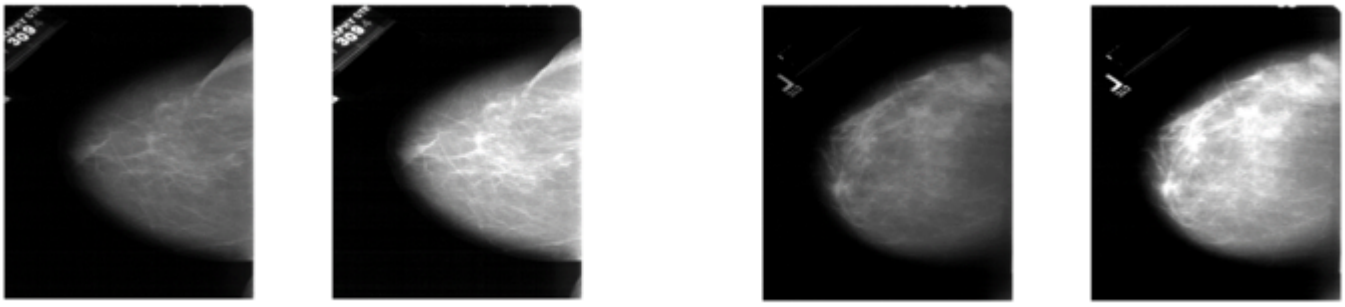

Fig.2. Images are taken from applying the Homomorphic filter. The left images are original ones; the right images are the output images after applying the Homomorphic filter.

$z(x\,.\,y)$ by performing logarithm on both sides of Equation (1) is defined as:

$$z(x\,.\,y) = \ln f(x\,.\,y) = \ln i(x\,.\,y) + \ln r(x\,.\,y) \tag{2}$$

Then, the Fourier Transform of both sides of the equation (2) is taken:

$$Z(u\,.\,v) = F_i(u\,.\,v) + F_r(u\,.\,v) \tag{3}$$

This work will continue by applying a filter function $H(u\,.\,v)$

$$S(u\,.\,v) = H(u\,.\,v)Z(u\,.\,v) = H(u\,.\,v)F_i(u\,.\,v) + H(u\,.\,v)F_r(u\,.\,v) \tag{4}$$

Function $S(u\ .\ v)$ is the result of performing a Homomorphic filter. Inverse Fourier Transform of the equation above gives Equation (5)

$$s(x\ .\ y) = i'(x\ .\ y) + r'(x\ .\ y) \tag{5}$$

Where in relations above, the filter function $H(u\ .\ v)$ can be expressed as follows:

$$H(u\ .\ v) = (\gamma_H - \gamma_L)\left[1 - e^{c(D^2(u\ .\ v)/D_0^2)}\right] + \gamma_L \tag{6}$$

Where, D_0 is the distance from the origin of the transform; the distance from the point $(u\ .\ v)$ to the center of the frequency rectangle; using constant c, the sharpness of the slope of the filter function can be adjusted. Finally, the enhanced image is calculated using the following equation. Figure 2 illustrates the mammogram images before and after applying the Homomorphic filter.

$$g(x\ .\ y) = e^{s(x.y)} = e^{i'(x\ .\ y)} * e^{r'(x\ .\ y)} = i_E(x\ .\ y)r_E(x\ .\ y) \tag{7}$$

2.2. Zernike Moments

Zernike moments map an image into a set of complex Zernike polynomials. Zernike moments are invariant to rotation, non-redundant, resistant to noise and shape and have a multi-level display (Haddadnia et al., 2003; Wang et al., 2009). In this study, Zernike moments were used for shape and mass margin feature extraction.

In order to eliminate sensitivity to scale and translation of Zernike moments, some pre-processing steps such as segmentation, scaling and translation were performed on the input ROI and then the mass shape was obtained in binary format. The appropriate numbers of Zernike moments that can fully interpret the masses were extracted.

For extracting margin features, histogram equalization, translation and scaling were performed as a pre-processing step on the input ROI and the edge of the masses in the form of gray levels were obtained. Then, Zernike moments used to describe edge of the masses were extracted (Tahmasbi et al., 2011).

Three steps are used in calculating the Zernike moments. A) Calculation of radial polynomials: The following equation gives the real value of a one-dimensional polynomial.

$$\mathrm{R_{n.m}} = \sum_{s=0}^{\frac{(n-|m|)}{2}} (-1)^s \frac{(n-s)!}{s!\left(\left(\frac{n+|m|}{2}\right)-s\right)!\left(\left(\frac{(n-|m|)}{2}\right)-s\right)!}\ \rho^{\mathrm{n-2\,s}} \tag{8}$$

Where n is a non-negative integer which determines the order of polynomial and m is a non-zero integer that represents the repetition of Azimuthal angle as (1) $n - |m|\, is\, even$; (2) and |m|<= n.

ρ Represents the length of the vector from the origin to the point $(x\ .\ y)$. Also, it can be proven that $R_{n.-m}(\rho) = R_{n.m}(\rho)$. B) Calculation of basic functions of Zernike moments: Based on the amount of radial polynomial obtained in the last stage, Zernike two dimensional basic functions were calculated by the following equation.

$$\mathrm{V_{n.m}}(\rho\ .\ \theta) = R_{n.m}(\rho)e^{jm\theta} \quad |\rho| \le 1 \tag{9}$$

Where $j = \sqrt{-1}$: θ is the Azimuthal angle between the vectors ρ and the x-axis in the counterclockwise direction.

C) Calculation of Zernike moments: Zernike moments of the image $f(x\,.\,y)$ state its projection on the space of orthogonal Zernike polynomials of order n by repeating m.

$$Z_{n.m} = \frac{< f(x\,.\,y)\,.\,V_{n.m} >}{< V_{n.m}\,.\,V_{n.m} >} \tag{10}$$

Where $< k\,.\,\mathrm{g} >$ express the inner product of k and g, meaning:

$$< k\,.\,\mathrm{g} \geq \iint k(x\,.\,y)g^{*}(x\,.\,y)\,dx\,dy \tag{11}$$

According to the orthogonality of Zernike polynomials, different Zernike moments with a m repetition and order of n can be obtained using inner product of image function and basic function of Zernike by the following equation.

$$\mathrm{Z_{n.m}} = \frac{\mathrm{n}+1}{\pi} \iint_{0\leq x^2+y^2\leq 1} f(x\,.\,y)\mathrm{V^{*}_{n.m}}(x\,.\,y)dx\,dy \tag{12}$$

In the form of polar coordinates, it can be expressed as follows:

Where is known as an even part of real value of Zernike polynomials and is known as an odd part of it.
And

$$\mathrm{Z_{n.-m}} = \mathrm{Z^{*}_{n.m}} \text{ و } \left|\mathrm{Z_{n.-m}}\right| = \left|\mathrm{Z_{n.m}}\right| \text{ یا } \arg(\mathrm{Z_{n.-m}}) = -\arg(\mathrm{Z_{n.m}})$$

In the above relation, $f(\rho\,.\,\theta)$ and * shows image function and the complex conjugate, respectively. Since the images are discrete and defined as a matrix, the discrete form will be obtained by replacing an integral with Sigma. In addition, image is normal in the range of [0.1] by mapping function shown in the figure below.

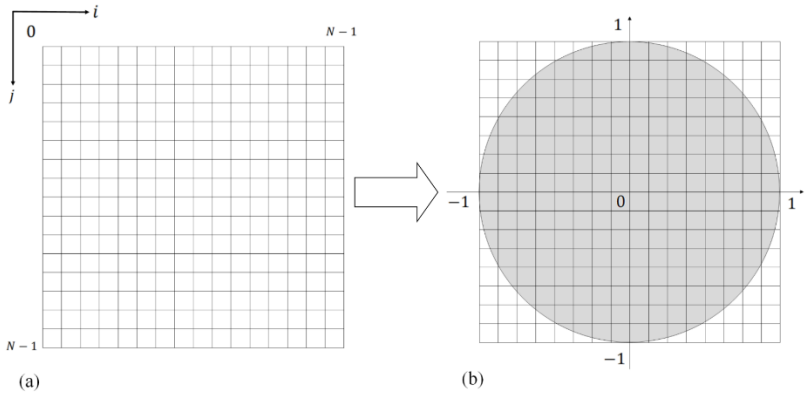

Fig.3.General mapping function (a) N×N image; (b) the image mapped onto the unit circle

The figure above shows the general form of the mapping function. Pixels that are located outside the unit circle are not involved in the calculation of Zernike moments. So, Zernike moments calculated using this mapping function, do not describe the characteristics outside the unit circle.

The final form of discrete Zernike moment for an image of size N * N is obtained as follows:

$$Z_{n.m} = \frac{n+1}{\lambda_N} \sum_{c=0}^{N-1} \sum_{r=0}^{N-1} f(c\,.\,r)\, V^{*}_{n.m}(c\,.\,r) = \frac{n+1}{\lambda_N} \sum_{c=0}^{N-1} \sum_{r=0}^{N-1} f(c\,.\,r)\, R_{n.m}\left(\rho_{cr}\right) e^{-jm\theta_{cr}} \tag{14}$$

Here, $0 \le \rho_{cr} \le 1$, and λ_N is the number of pixels inside the unit circle of the mapping transform used to normalize what relates to the unit circle area, π, in continuous domain.

$$\rho_{cr} = \frac{\sqrt{(2\,c - N + 1)^2 + (2\,r - N + 1)^2}}{N}$$

$$\theta_{cr} = \tan^{-1}\left(\frac{N - 1 - 2\,r}{2\,c - N + 1}\right) \tag{15}$$

c and r are rows and columns of images. To get rid of the problems related to mapping functions, Equation (14) and (15) were combined, in which, we have the relationship as follows. The above relationships will be usable until the size of image is an even number because by choosing an odd number for the size of an image, relations will be as follows:

$$(c\,.\,r) = \left(\frac{N-1}{2}\,.\,\frac{N-1}{2}\right) \tag{16}$$

$$\rho_{cr} = 0$$

$$\theta_{cr} = \tan^{-1}\frac{0}{0} = Nan$$

Where, $(c\,.\,r)$ is the central point. As shown in the above relations, when the size of the image is even number, the phase will be undefined. By choosing the odd number for image size, the image will not have one center, so the problem will be solved without any redundancy.

The high sensitivity of high order Zernike moment to noise and their high computational complexity, these moments may decrease the performance of the algorithm. But, when they are chosen exactly, will be proper descriptors for shape and margin features than low-order moments. Therefore, to evaluate the performance of our algorithm, the impact of orders of Zernike moments were analyzed. Two groups of Zernike moments were extracted: high-order and low orders Zernike moments of images of geometric shapes and mass margin were extracted. The first group included 32 low-order moments that satisfy the following conditions:

$$Group\ 1 = \left\{Z_{n.m}\right\} \forall \begin{cases} 3 \le n \le 10 \\ |m| \le n \\ n - |m| = 2\,k \\ k \in N \end{cases}$$

The second group included 32 high-order moments that satisfy the following conditions:

$$Group\,2 = \{Z_{n.m}\}\forall \begin{cases} 10 \le n \le 17 \\ |m| \le n \\ n - |m| = 4\,k \\ k \in N \end{cases}$$

Four different groups of features from the magnitude of low-order and high-order Zernike moments which are extracted from margin and shape were considered.

1) Extraction the magnitude of low-order Zernike moments from mass shape mammogram images (LMSI)
2) Extraction the magnitude of low-order Zernike moments from mass margin mammogram images (LMMI)
3) Extraction the magnitude of high-order Zernike moments from mass shape mammogram images (HMSI)
4) Extraction the magnitude of high-order Zernike moments from mass margin mammogram images (HMSI)

In this work, all these groups applied to the next stage of the algorithm. Fig 4 shows the shape and margin of the benign and malignant masses for extraction of Zernike moments features.

Fig.4. Image of the shape and margin of mass for extracting Zernike moments; (a) malignant mass, (b) benign mass

2.3. The Fast Discrete Curvelet Transform (FDCT)

In 2006, *Fast Discrete Curvelet Transform* (FDCT) model Based on changes in Curvelet transform function was introduced. The FDCT is superior than Curvelet transform because it uses *fast Fourier transform* (FFT) which was less complicated in computation and time. Two methods used in performing FDCT, named Curvelet transform via unequally spaced fast Fourier transform and Curvelet transform via wrapping function, were explained (Candès, Demanet, Donoho, & Ying, 2006). In this study, to extract texture features of breast tissue images, the wrapping method of FDCT was used and the works based on translation and wrapping of Fourier transform, are so fast and reduce the information redundancy.

The following relation shows how to calculate the wrapping method of FDCT. The two-dimensional input image is considered as an array $f[m\,.\,n]$, where $0 \le m < M$ and $0 \le n < N$ are dimensional array. After scanning the input image by wedge data, a subset of Curvelet coefficients $c^D(j\,.\,l\,.\,k_1\,.\,k_2)$ will be obtained on the scale j, an orientation l and located in the two spatial locations of the image k_1 and k_2. D index shows the digital coefficients and is applicable in software implementation. Also, Figure 3 shows calculation of the Curvelet coefficients in a discrete space (AlZubi, Islam, & Abbod, 2011; Tambasco Bruno et al., 2016)

$$\mathrm{c^D(j\,.1\,.k1.k2)} = \sum_{n=1}^{N}\sum_{m=1}^{M} f[m\,.\,n]\varphi^{D}{}_{j.l.k1.k2}[m\,.\,n] \qquad (17)$$

Finally, in order to produce the image resulting from the Curvelet transform, the Fourier transform of Curvelet coefficients must be product for Fourier transform of the image and to get a return from the frequency domain to the spatial domain, the inverse Fourier is taken. Final output of the discrete Curvelet includes 1 image in the first level, 16 images in the second level, 32 images in the third and fourth levels and 64 images in the fifth level, and each one is obtained in a special rotation and scale.

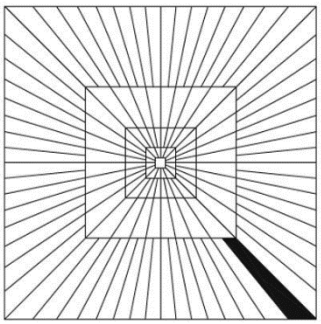

Fig.5. Digital image tiling in the frequency domain (Hwang & Kim, 2006)

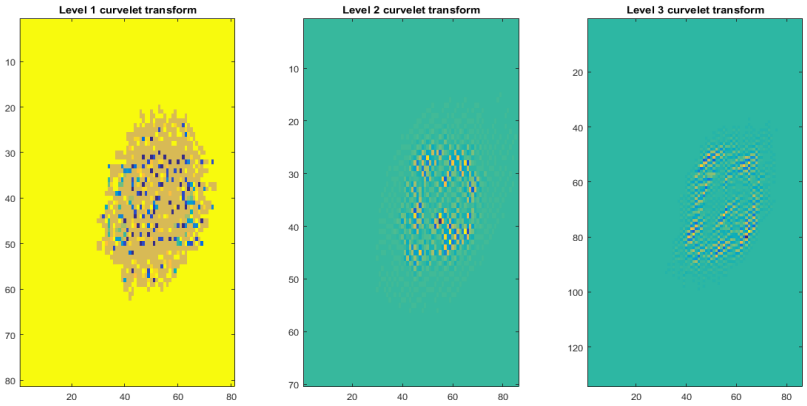

Fig.6. Discrete Curvelet transform output for three different levels

In general, Curvelet coefficients are defined for several levels and each contains useful information to classify images. In our proposed model, the first three levels of the Curvelet coefficients were used to cover more, the extraction information. We used 16 degrees for orientation l and scale of 3. Using three levels will create 49 output images that are a result of 1 output image at level 1, 16 output images at level 2 and 32 output images at level 3. The size of the tumor in the images is usually smaller than a 256 * 256 matrix. So, the size of the input image for Curvelet transform is considered as 256 * 256. Because of huge number of features we obtained by performing FDCT, using all 49 images does not seem logical when each level has similar valuable information in the output images. So, as suggested by Bruno and colleagues, we will use one output image at each level.

2.4. Local Binary Pattern

The local binary pattern technique was introduced in 1994 by Ojala and then followed by Pietikainen (Ojala, Pietikainen, & Maenpaa, 2002; Pietikäinen, Hadid, Zhao, & Ahonen, 2011). The local binary pattern contains important information on the texture of the image based on its spatial structure and contrast. In mammograms, separating the breast tissue and recognizing similar patterns in benign and malignant tumors will greatly help in categorizing them. Generally, a local binary pattern computes the value of the local binary pattern of the central

pixel C, by taking each pixel of the image as the central pixel and the value of the threshold relative to its P pixel in each frame and comparing the value of the central pixel with other neighboring pixels. The Equation (20) is used to calculate the local binary pattern of each pixel with the radius of neighboring R and with the number of pixels P(Ojala et al., 2002; Pietikäinen et al., 2011; Tambasco Bruno et al., 2016).

$$\mathrm{LBP_{P.R}} = \sum_{p=0}^{P-1} s\left(g_p - g_c\right)2^p \tag{18}$$

The function $s(x)$ for $x \geq 0$ has a value of 1 and for$x < 0$, the zero value is taken as the output. Also, the value of g_c is the intensity of the central pixel (grayscale of the central pixel) and g_p is the amount of intensity of the neighboring pixels.

In 2002, Ojala et al. developed a new model of local binary pattern method called the uniform local binary pattern; the difference is in how to calculate the central pixel value and its mapping. With the following terms, the uniform local binary pattern can be calculated and used(Ojala et al., 2002; Pietikäinen et al., 2011; Tambasco Bruno et al., 2016).

$$\mathrm{LBP}^{u}_{P.R} = \begin{cases} \sum_{p=0}^{P-1} s & if(U(LBP_{p,r}) \leq 2 \\ P+1 & otherwise \end{cases} \tag{19}$$

$$U\left(\mathrm{LBP_{P.R}}\right) = \left|\mathrm{s}\left(\mathrm{g_{p-1}} - \mathrm{g_c}\right) - \mathrm{s}\left(\mathrm{g_0} - \mathrm{g_c}\right)\right| + \sum_{p=1}^{P-1}\left|s\left(g_p - g_c\right) - s\left(g_{p-1} - g_c\right)\right| \tag{20}$$

In the presented uniform local binary pattern model, instead of calculating the difference between the central pixels and neighboring pixels and then multiplying the threshold values in a pixel number, the difference of the central pixel and the first and last pixels of the frame was considered and then, the difference between the two pixels adjacent to the central pixel was added to the effect of image uniformity in calculating the local binary pattern (Ojala et al., 2002; Pietikäinen et al., 2011; Tambasco Bruno et al., 2016).

Note that the absolute value of the coefficients is used to compute LBP and the parameters are chosen p=8 and r=2 (Nanni, Brahnam, & Lumini, 2012). Figure 7 shows the result of applying a local binary pattern on Curvelet coefficients.

Fig.7.The output of application of LBP on Curvelet coefficients

3. Experiments and results

3.1. Performance of the proposed algorithm

In this research, the images from the DDSM database were used and in order to evaluate the performance of the proposed algorithm the images were divided into benign and malignant groups.

3.1.1. Reduction using PCA

After extracting all features described in section 2, a huge number of features were produced that make the algorithm time-consuming. So, we use the PCA algorithm to reduce the number of features.

The existence of almost twenty thousand features, in addition to making the training process slower, certainly have much useless information, such as the black parts of the image margins and this can cause a tangible decrease in classification accuracy. For this reason, the use of principal component analysis for the combination of features and the reduction of feature dimensions, in addition to increasing the speed of training of the classifier, will likely greatly increase the accuracy of classification.

3.1.2. Classification of tumors using SVM

Support vector machine is one of the most common data classification methods. The main idea behind the support vector machine is the separation of data based on the maximum distance between the different data classes.

Suppose that we have a training dataset $\left\{ \left(x_i . y_i \right) :\ x_i \in \boldsymbol{R}^d \, y_i \in \ \{-1. +1\} \right\}$ where $i\,th$ sample vector is shown by x_i and its corresponding label is y_i. Then, a hyperplane is found to separate the features that satisfy the blow constraints:

$$y_i . \left(w . x + b \right) \geq 0 \tag{21}$$

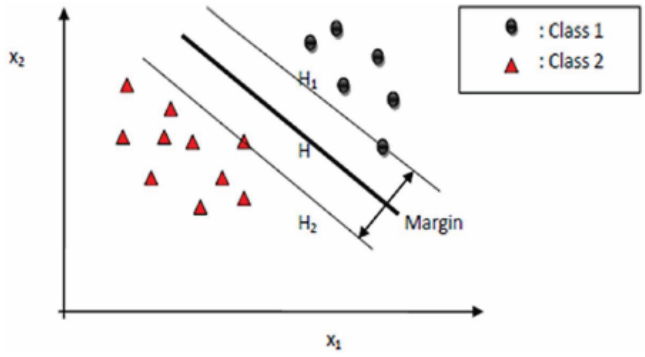

Fig.8.Support vector machine separating two classes [chapter 9-support vector-Bn Yousef]

SVM recognizes the features that are the nearest one to the hyperplane named support vector. SVM aims to find a hyperplane that should be optimal. For this purpose, the objective function must be minimized:

$$\min_{w} \left(\frac{1}{2} \|w\|^2 \right) \quad subject\,to \ y_i . \left(w . x + b \right) \geq 1 \quad . \quad i = 1 \ldots . . n \tag{22}$$

By using Lagrangian formalism, the constrained problem converts into an unconstrained problem. We use Lagrange multipliers $\left\{ \alpha_i \right\}_{i=1}^{n}$ to minimize the objective function:

$$Q(\alpha) = \frac{1}{2} - \sum_{i=1}^{n} \alpha_i + \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j K\left(x_i . x_j\right) \quad (23)$$

$$\sum_{i=1}^{n} \alpha_i y_i = 0 \qquad 0 \leq \alpha_i \leq C \quad . \quad \forall i = 1.2 \ldots . n \quad (24)$$

Where, K is a kernel function and C is a positive constant that the user determines. After obtaining an optimal solution from the above relations, the decision function will be:

$$\beta_{SVM}(x) = b + \sum_{i=1}^{n} \alpha_i y_i K(x_i . x) \quad (25)$$

In this paper, we used polynomial kernel function as below [chapter 9-suuport vector-Bn Yousef]

$$K\left(x_i . x_j\right) = (x_i x_j + 1)^2 \quad (26)$$

3.1.3. Cross-validation

The 10-fold model is used for testing and training. In this model, the existing data (238 images) are divided into 10 different folds and each time, 9 folds are given to the support vector machine and one fold is tested. This process is used 10 times and the final accuracy will be calculated from the average accuracy obtained at each time. Using this technique will make the combination of data to have no effect on the amount of accuracy and in fact, the samples are not biased for training and testing. Also, variance and average accuracy show the performance of the model in terms of the presence of different tests and teaching samples.

3.1.4. Measurement metrics

The results of the classification system are false positive if the positive class system is labeled positive, false negative if the positive system is labeled negative, and true positive and true negative when the class system is correctly labeled (Görgel et al., 2013).

The confusion matrix provides the rate of TP, TN, P and N, and is expressed in the following table.

The feature vectors were combined with each other and in order to assess the proposed method, the accuracy (AC) and area under the ROC curve (AUC) are considered (Dua, Singh, & Thompson, 2009; Fawcett, 2006).

The accuracy criterion is the result of dividing the correct predictions into the total number of items.

$$AC = \frac{TP + TN}{P + N} \quad (27)$$

Other validation criteria in addition to accuracy are precision, sensitivity, and F-measure, which can be used for comparison. According to the confusion matrix, sensitivity and specificity are expressed as follows (Cook, 2017).

$$Sensitivity = \frac{TP}{P} \quad (28)$$

$$Specificity = \frac{TN}{N} \quad (29)$$

Finally, in order to combine and evaluate the performance of the model with respect to both classes with the help of the expression (28), the F-measure criterion can also be calculated (Ameta, 2017).

$$\mathrm{F-measure} = \frac{2*(sensitivity*\mathrm{precision})}{(sensitivity + \mathrm{precision})} \tag{30}$$

$$AUC = \frac{1 + TP - FP}{2} \tag{31}$$

3.2. Results

Table 1 shows the confusion matrix of this method and Table 2 provides the four criteria besides the accuracy range in comparison with the three proposed models. It is clear that all the feature extraction models for benign diagnosis (sensitivity criterion) provide a similar performance of 94.69%, but the superiority of the proposed method is in identifying malignant tumors. Also, based on the accuracy range in this table, it is known that in the worst-case scenario, the accuracy of the proposed model using PCA and without it will be more than 81%.

Table1
Confusion matrix

		Predicted class	
		Benign	Malignant
Real class	Benign	125	7
	Malignant	4	102

Table 2
Comparison of criteria of the proposed method and two other methods

Method	F-measure	Sensitivity	Specificity	Precision	Accuracy range
LBP	92.93	94.69	88.67	91,24	77.27-100
LBP+ Zernike moments	93.98	94.69	91.5	93.28	81.81-100
PCA on LBP+ Zernike	95.78	94.69	96.22	96.89	81.81-100

Table 3
Average accuracy to compare three methods

Method	Average accuracy (%)	superiority of proposed method (%)
LBP features	91.89 ± 0.075	3.37
LBP+ Zernike moments	93.18 ± 0.059	2.08
PCA on LBP+ Zernike	95.26 ± 0.063	0

Table 3 shows average accuracy for the three methods and also the superiority of the proposed method than the other two methods. The use of features derived from LBP has the lowest level of accuracy than the other two models. In fact, it seems that using LBP alone is not enough to obtain accuracy of the classification above 95%. The main reason for this may be the resemblance of the tissue of some malignant and benign masses, and the Curvelet coefficients and LBP operators are not able to differentiate these two groups.

Therefore, with the addition of features of Zernike moments, the shape and margin of masses are also considered as mass features in addition to the texture features. Finally, with the use of PCA and the combination of the features of Zernike moments and the application of LBP on the Curvelet coefficients, not only is the number of features of the model reduced significantly and the speed of training and testing increased, but also the best combination of features is created by choosing and defining the coefficients which eventually yield more than 3% progress and more than 95% accuracy.

Finally, to complete the explanation of the behavior of the proposed model, a statistical T-test was used to compare the effect of PCA on the extracted features. In this test, the result of the prediction of the model with extracted features with the PCA implementation model was compared with the rejection or confirmation of the behavioral theory of the proposed model by the T-test standard. The T-test output has a value of 1, which in fact eliminates the randomness hypothesis of the proposed model behavior and its p-value is 0.045, indicating the correctness of the analysis.

As previously mentioned, many studies and articles on the diagnosis of benign and malignant tumors have been presented and each of them has achieved success, depending on how the feature extraction and classification are used and of course, they also have weaknesses.

Table 4 shows a comparison of some similar approaches using AC values and shows that the proposed model has better results than other similar techniques.

Table 4

Comparison of different approach by AC values

Reference	Dataset	ROIs	Feature extraction	Classifier	AC
M. M. Eltoukhy et al. (2010b) (Eltoukhy et al., 2010)			Curvelet transform and Haralick features		91.68
S. Dhahbi et al. (2015) (Dhahbi, Barhoumi, & Zagrouba, 2015)		252	Curvelet transform and moment theory	K-NN	91.27
P. Görgel et al. (2015) (Görgel, Sertbas, & Uçan, 2015)	MIAS	60	Spherical wavelet transform		83.3
D. O. Tambasco Bruno et al. (2016) (Tambasco Bruno et al., 2016)	DDSM	240	Curvelet transform, LBP and ANOVA	SVM	85
Our approach	DDSM	238	Curvelet transform, LBP and PCA	SVM	95

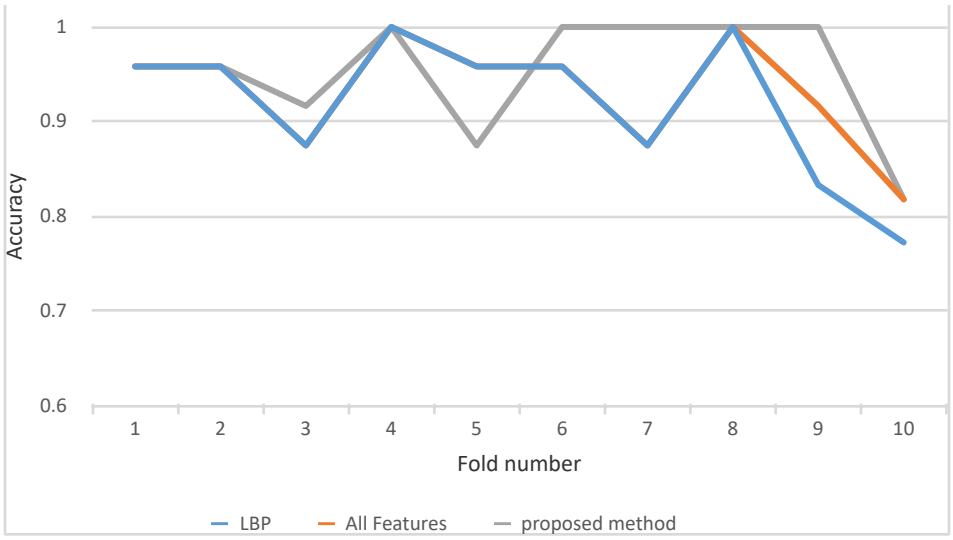

Fig.9. Chart showing a comparison of the accuracy in the implementation of 10-fold

Figure 9 shows the accuracy taken from the implementation of the 10-fold model, the three evaluated models had a high accuracy of 85% for most categories but the challenge categories are 9 and 10. The use of LBP alone yielded much lower accuracy than the other two models. Another point in the diagram is the tangible advantage of applying PCA in the extraction of a combination of LBP and Zernike moment in all the categories, except Category 5, which seems to make the proposed model considerably more appropriate. In categories 4 and 6 to 9, the proposed method showed 100% accuracy, which includes 5 different categories, while the method of using all the features together in at most 2 categories have 100% accuracy.

`

4. Conclusion

A new approach based on the combination of LBP applied to Curvelet transform and Zernike moments are presented in this paper. The reason for choosing this combination of feature extraction method is the existence of different strengths in these two models. By using Zernike moments applied to the mass shape and margin, masses can be simply categorized into benign and malignant classes. On the other hand, discrete Curvelet coefficients are appropriate for extracting a texture model of the image. These coefficients which are extracted at several levels provide useful information on the texture. Because changes in texture have important information, after extracting the features of the discrete Curvelet coefficients, the local binary pattern was applied. After extracting both feature categories, due to the very large number of features extracted from LBP applied on the Curvelet coefficients, to reduce the features, Principle Component Analysis was used.

After preparing all the features and applying PCA, in order to classify the final features, they were sent to support vector classifier for training and testing and the results were compared with all the features of LBP. The superiority of the proposed method to all the features without applying PCA was about 3%, which is more than 90% detectable.

As expected, the proposed method has a significant improvement as compared to the successful and widely used approaches in diagnosing mammogram masses. On the other hand, the proposed model is based on two models with high-performance to extract features which are less complex than the other methods.

It seems that improving the detection rate of the tumor by using other features and adding them to the current features is expected. Since frequency domain features such as the Curvelet coefficients provide valuable information on the image texture, one can expect that the use of other models such as wavelet transforms can generate appropriate information as well. Wavelet transformation features have already been used in other works and the use of these features also has the appropriate precision for classification (Saraswathi, Dharani, &

Srinivasan, 2016). In addition to Zernike moments, DCT Cosinus coefficients can also be used in the feature extraction step. These coefficients will provide relevant information on the problem.

Finally, it seems that to design and fabricate an integrated software capable of using new and old mammograms that can detect tumor area and automatically extracts it and report the outcome and type of tumor by itself is essential, to save the lives of people.

References

AlZubi, S., Islam, N., & Abbod, M. (2011). Multiresolution Analysis Using Wavelet, Ridgelet, and Curvelet Transforms for Medical Image Segmentation. *International Journal of Biomedical Imaging*, *2011*, 1–18. https://doi.org/10.1155/2011/136034

Ameta, D. (2017). Ensemble classifier approach in breast cancer detection and malignancy grading- A review. *International Journal of Managing Public Sector Information and Communication Technologies*, *8*. https://doi.org/10.5121/ijmpict.2017.8102

Beura, S., Majhi, B., & Dash, R. (2015). Mammogram classification using two dimensional discrete wavelet transform and gray-level co-occurrence matrix for detection of breast cancer. *Neurocomputing*, *154*, 1–14. https://doi.org/10.1016/J.NEUCOM.2014.12.032

Braz Junior, G., da Rocha, S. V., Gattass, M., Silva, A. C., & Paiva, A. C. de. (2013). A mass classification using spatial diversity approaches in mammography images for false positive reduction. *Expert Systems with Applications*, *40*(18), 7534–7543. https://doi.org/10.1016/J.ESWA.2013.07.034

Byung-Woo Hong, & Bong-Soo Sohn. (2010). Segmentation of Regions of Interest in Mammograms in a Topographic Approach. *IEEE Transactions on Information Technology in Biomedicine*, *14*(1), 129–139. https://doi.org/10.1109/TITB.2009.2033269

Candès, E., Demanet, L., Donoho, D., & Ying, L. (2006). Fast Discrete Curvelet Transforms. *Multiscale Modeling & Simulation*, *5*(3), 861–899. https://doi.org/10.1137/05064182X

Coburn, C. A., & Roberts, A. C. B. (2004). A multiscale texture analysis procedure for improved forest stand classification. *International Journal of Remote Sensing*, *25*(20), 4287–4308. https://doi.org/10.1080/0143116042000192367

Cook, J. A. (2017). ROC curves and nonrandom data. *Pattern Recognition Letters*, *85*, 35–41. https://doi.org/10.1016/J.PATREC.2016.11.015

Dhahbi, S., Barhoumi, W., & Zagrouba, E. (2015). Breast cancer diagnosis in digitized mammograms using curvelet moments. *Computers in Biology and Medicine*, *64*, 79–90. https://doi.org/10.1016/j.compbiomed.2015.06.012

Dheeba, J., Albert Singh, N., & Tamil Selvi, S. (2014). Computer-aided detection of breast cancer on mammograms: A swarm intelligence optimized wavelet neural network approach. *Journal of Biomedical Informatics*, *49*, 45–52. https://doi.org/10.1016/j.jbi.2014.01.010

Dua, S., Singh, H., & Thompson, H. W. (2009). Associative classification of mammograms using weighted rules. *Expert Systems with Applications*, *36*(5), 9250–9259. https://doi.org/10.1016/j.eswa.2008.12.050

Elter, M., & Horsch, A. (2009). CADx of mammographic masses and clustered microcalcifications: A review. *Medical Physics*, *36*(6), 2052–2068. https://doi.org/10.1118/1.3121511

Eltoukhy, M. M., Faye, I., & Samir, B. B. (2010). Curvelet based feature extraction method for breast cancer diagnosis in digital mammogram. In *2010 International Conference on Intelligent and Advanced Systems* (pp. 1–5). IEEE. https://doi.org/10.1109/ICIAS.2010.5716125

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*(8), 861–874. https://doi.org/10.1016/J.PATREC.2005.10.010

Gardezi, S. J. S., Faye, I., Adjed, F., Kamel, N., & Eltoukhy, M. M. (2017). Mammogram Classification Using Curvelet GLCM Texture Features and GIST Features (pp. 705–713). Springer, Cham. https://doi.org/10.1007/978-3-319-48308-5_67

Görgel, P., Sertbas, A., & Ucan, O. N. (2013). Mammographical mass detection and classification using Local Seed Region Growing-Spherical Wavelet Transform (LSRG-SWT) hybrid scheme. *Computers in Biology and Medicine*, *43*(6), 765–774. https://doi.org/10.1016/j.compbiomed.2013.03.008

Görgel, P., Sertbas, A., & Uçan, O. N. (2015). Computer-aided classification of breast masses in mammogram images based on spherical wavelet transform and support vector machines. *Expert Systems*, *32*(1), 155–164. https://doi.org/10.1111/exsy.12073

Gupta, S., & Markey, M. K. (2005). Correspondence in texture features between two mammographic views. *Medical Physics*, *32*(6), 1598. https://doi.org/10.1118/1.1915013

Guzmán-Cabrera, R., Guzmán-Sepúlveda, J. R., Torres-Cisneros, M., May-Arrioja, D. a., Ruiz-Pinales, J., Ibarra-Manzano, O. G., … Parada, a. G. (2012). Digital Image Processing Technique for Breast Cancer Detection. *International Journal of Thermophysics*, *34*(8–9), 1519–1531. https://doi.org/10.1007/s10765-012-1328-4

Haddadnia, J., Ahmadi, M., & Faez, K. (2003). An Efficient Feature Extraction Method with Pseudo-Zernike Moment in RBF Neural Network-Based Human Face Recognition System. *EURASIP Journal on Advances in Signal Processing*, *2003*(9), 890–901. https://doi.org/10.1155/

S1110865703305128

Harwood, D., Ojala, T., Pietikäinen, M., Kelman, S., & Davis, L. (1995). Texture classification by center-symmetric auto-correlation, using Kullback discrimination of distributions. *Pattern Recognition Letters*, *16*(1), 1–10. https://doi.org/10.1016/0167-8655(94)00061-7

Hwang, S.-K., & Kim, W.-Y. (2006). A novel approach to the fast computation of Zernike moments. *Pattern Recognition*, *39*(11), 2065–2076. https://doi.org/10.1016/j.patcog.2006.03.004

Jadoon, M. M., Zhang, Q., Haq, I. U., Butt, S., & Jadoon, A. (2017). Three-Class Mammogram Classification Based on Descriptive CNN Features. *BioMed Research International*, *2017*, 1–11. https://doi.org/10.1155/2017/3640901

Jae Young Choi, Dae Hoe Kim, Plataniotis, K. N., & Yong Man Ro. (2012). Combining multiple feature representations and AdaBoost ensemble learning for reducing false-positive detections in Computer-aided Detection of masses on mammograms. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 4394–4397). IEEE. https://doi.org/10.1109/EMBC.2012.6346940

Lei Zhen, & Chan, A. K. (2001). An artificial intelligent algorithm for tumor detection in screening mammogram. *IEEE Transactions on Medical Imaging*, *20*(7), 559–567. https://doi.org/10.1109/42.932741

Mäenpää, T. (2003). *The local binary pattern approach to texture analysis : extenxions and applications*. Dissertation. Acta Univ Oul C 187, 78 p + App.: Oulun yliopisto. Retrieved from http://herkules.oulu.fi/isbn9514270762/

Mayerhoefer, M., Perry, N., Milner, S., Mokbel, K., Duffy, S., & Pinker, K. (2010). Texture analysis applied to full-field digital mammography: ability to discriminate between invasive ductal and invasive lobular breast cancer - preliminary results. *Breast Cancer Research*, *12*(S3), P12. https://doi.org/10.1186/bcr2665

Mohammed, M. A., Al-Khateeb, B., Rashid, A. N., Ibrahim, D. A., Abd Ghani, M. K., & Mostafa, S. A. (2018). Neural network and multi-fractal dimension features for breast cancer classification from ultrasound images. *Computers & Electrical Engineering*. https://doi.org/10.1016/J.COMPELECENG.2018.01.033

Moura, D. C., & Guevara López, M. A. (2013). An evaluation of image descriptors combined with clinical data for breast cancer diagnosis. *International Journal of Computer Assisted Radiology and Surgery*, *8*(4), 561–574. https://doi.org/10.1007/s11548-013-0838-2

Nanni, L., Brahnam, S., & Lumini, A. (2012). A very high performing system to discriminate tissues in mammograms as benign and malignant. *Expert Systems with Applications*, *39*(2), 1968–1971. https://doi.org/10.1016/J.ESWA.2011.08.050

Nascimento, M. Z. Do, Martins, A. S., Neves, L. A., Ramos, R. P., Flores, E. L., & Carrijo, G. A. (2013). Classification of masses in mammographic image using wavelet domain features and polynomial classifier. *Expert Systems with Applications*, *40*(15), 6213–6221. https://doi.org/10.1016/j.eswa.2013.04.036

Ojala, T., Pietikainen, M., & Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*(7), 971–987. https://doi.org/10.1109/TPAMI.2002.1017623

Pietikäinen, M., Hadid, A., Zhao, G., & Ahonen, T. (2011). Local Binary Patterns for Still Images (pp. 13–47). Springer London. https://doi.org/10.1007/978-0-85729-748-8_2

Ramos, R. P., Nascimento, M. Z. do, & Pereira, D. C. (2012). Texture extraction: An evaluation of ridgelet, wavelet and co-occurrence based methods applied to mammograms. *Expert Systems with Applications*, *39*(12), 11036–11047. https://doi.org/10.1016/J.ESWA.2012.03.020

Rouhi, R., Jafari, M., Kasaei, S., & Keshavarzian, P. (2015). Benign and malignant breast tumors classification based on region growing and CNN segmentation. *Expert Systems with Applications*, *42*(3), 990–1002. https://doi.org/10.1016/J.ESWA.2014.09.020

Saraswathi, D., Dharani, D., & Srinivasan, E. (2016). An efficient feature extraction technique for breast cancer diagnosis using curvelet transform and swarm intelligence. In *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)* (pp. 441–445). IEEE. https://doi.org/10.1109/WiSPNET.2016.7566172

Saraswathi, D., & Srinivasan, E. (2017). A high-sensitivity computer-aided system for detecting microcalcifications in digital mammograms using curvelet fractal texture features. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, *5*(4), 263–273. https://doi.org/10.1080/21681163.2015.1089793

Surendiran, B., & Vadivel, A. (2011). Feature Selection using Stepwise ANOVA Discriminant Analysis for Mammogram Mass Classification. *ACEEE Int . J . on Signal & Image Processing*, *2*(1), 17–19. Retrieved from https://www.researchgate.net/profile/Surendiran_Balasubramaniam/publication/258052973_Feature_selection_using_stepwise_ANOVA_discriminant_analysis_for_mammogram_mass_classification/links/0c96052942c3e97cda000000.pdf

Tahmasbi, A., Saki, F., Shokouhi, S. B., Santiago, S., Guang-wei, X., Du, H. N., … Boas, D. A. (2011). Classification of benign and malignant

masses based on Zernike moments. *Computers in Biology and Medicine*, *41*(8), 726–35. https://doi.org/10.1016/j.compbiomed.2011.06.009

Tambasco Bruno, D. O., Do Nascimento, M. Z., Ramos, R. P., Batista, V. R., Neves, L. A., & Martins, A. S. (2016). LBP operators on curvelet coefficients as an algorithm to describe texture in breast cancer tissues. *Expert Systems with Applications*, *55*, 329–340. https://doi.org/10.1016/j.eswa.2016.02.019

Tang, L., & He, C. (2013). Multiscale Texture Extraction with Hierarchical (BV,G p ,L 2) Decomposition. *Journal of Mathematical Imaging and Vision*, *45*(2), 148–163. https://doi.org/10.1007/s10851-012-0351-1

Vadivel, A., & Surendiran, B. (2013). A fuzzy rule-based approach for characterization of mammogram masses into BI-RADS shape categories. *Computers in Biology and Medicine*, *43*(4), 259–267. https://doi.org/10.1016/j.compbiomed.2013.01.004

Venkatalakshmi, S., & Janet, J. (2017). Classification of Mammogram Abnormalities Using Pseudo Zernike Moments and SVM. *I.J. Image, Graphics and Signal Processing*, *4*(4), 30–36. https://doi.org/10.5815/ijigsp.2017.04.04

Vyas, V. S., & Rege, P. P. (2007). Malignancy Texture Classification in Digital Mammograms based on Chebyshev Moments and Logpolar Transformation. *ICGST-BIME Journal*, *7*(1), 29–35. Retrieved from https://pdfs.semanticscholar.org/2fa9/10f3d41e81f1647c5b8e3b8931cf6f644cfd.pdf

Wang, W., Mottershead, J. E., & Mares, C. (2009). Mode-shape recognition and finite element model updating using the Zernike moment descriptor. *Mechanical Systems and Signal Processing*, *23*(7), 2088–2112. https://doi.org/10.1016/j.ymssp.2009.03.015

Yu, S.-N., & Huang, Y.-K. (2010). Detection of microcalcifications in digital mammograms using combined model-based and statistical textural features. *Expert Systems with Applications*, *37*(7), 5461–5469. https://doi.org/10.1016/J.ESWA.2010.02.066

Zhang, J., Silber, J. I., & Mazurowski, M. A. (2015). Modeling false positive error making patterns in radiology trainees for improved mammography education. *Journal of Biomedical Informatics*, *54*, 50–57. https://doi.org/10.1016/J.JBI.2015.01.007