World Applied Sciences Journal 13 (5): 1067-1076, 2011
ISSN 1818-4952

Breast Cancer Diagnosis System Based on Contourlet Analysis and Support Vector Machine

Sara Dehghani and Mashallah Abbasi Dezfooli

Department of Computer, Science and Research Branch,
Islamic Azad University, Khuzestan, Iran

Abstract: Breast cancer is one of the most epidemic diseases among women and the second cause of death among them. Because of non-existence of a method for prevention of catching breast cancer and also non-existence of a definite therapy for this disease, the upcoming and in-time diagnosis in preventing progress of this kind of cancer has a noticeable effect. Unfortunately these tumors aren't visible in the first stage and emerge suddenly, so in this area, we need CAD system as the second idea in diagnosing tumors of breast cancer and in helping medical technicians. Along with supplying this need, in this project, a system has been designed and complemented for diagnosis and classification of breast cancer tumors by using process of image. Firstly, this system is able to extract features of the texture from ROI areas by statistical method and signal process method and secondly, the system can classify the samples into two salubrious classes by using classifier based on SVM.

Key words: Breast Cancer・Mammography・Contourlet coefficient・Weighted SVM classification・Texture Features・Image Processing

INTRODUCTION

The interpretation and analysis of medical images represent an important and exciting part of computer vision and pattern recognition. Developing a computer-aided diagnosis system for cancer diseases, such as breast cancer, to assist physicians in hospitals is becoming of high importance and priority for many researchers and clinical centers. It is a complex process to develop a computer vision system to perform such tasks. The high incidence of breast cancer in women has increased significantly in the recent years. It is the cause of the most common cancer death in women. It is a leading cause of fatality in women, with approximately 1 in 12 women affected by the disease during their lifetime [1]. In Australia, approximately 1 of 13 women develops the disease [2]. A report from the National Cancer Institute (NCI) estimates that about one in eight women in the United States (approximately 12.5%) will develop breast cancer during their lifetime [3]. Early detection plays a very important factor in cancer treatment and allows better recovery for most patients. The required medical image for the diagnosing process of breast cancer, mammogram (breast X-ray), is considered the most reliable method in early detection [2,3]. Due to the high volume of images to be analyzed by radiologists and since senior radiologists are rare, reliable radiological diagnosis is not always available and the accuracy rate tends to decrease. A statistics shows that only 20-30% of breast biopsies are proved cancerous [4] and 10% of all cases of breast cancer go undetected by mammography [5]. Moreover, digital mammograms are among the most difficult medical images to be read according to the differences in the types of tissues and their low contrasts. Important visual clues of breast cancer include preliminary signs of masses and microcalcification clusters [6]. Unfortunately, at the early stages of breast cancer, these signs are very subtle and varied in appearance, making diagnosis even difficult to specialists. Therefore, automatic reading of digital medical images becomes highly desirable. It has proven that double reading of the mammogram, by two radiologists, increases the accuracy, but at high costs [7]. Therefore, the motivation of the computer-aided diagnosis systems [8-11] is to assist medical staffs to achieve high efficiency and accuracy.

Corresponding Author: Sara Dehghani, Department of Computer, Science and Research Branch, Islamic Azad University, Khuzestan, Iran. E-mail: s_dehghani61@yahoo.com.

Many studies have been made on the problem of breast cancer diagnosing based on digital mammograms [2,10,12,13].

[7] used neural network and data mining techniques for detection and classification of digital mammograms. Histogram equalization are used to enhance the images. The proposed methods classified the digital mammograms in two categories: normal and abnormal. The data collection they used in their experiments was taken from MIAS [14]. The extracted features that used are two existing features (type of the tissue and position of the breast) and four statistical parameters. In their experiments they used 90% of the dataset-322 images-for training the systems and 10% for testing them. The success rate obtained using the neural network (backpropagation algorithm) is 81% on average. On the other hand, it is 69% on average for association rule classifier. In the following research for [4] the data mining classifier is enhanced by applying two pruning methods of rules. They are eliminating the specific rules and keep only those that are general and with high confidence and prune some rules that could introduce errors at the classification stage. All the extracted features presented in [4] have been computed over smaller windows of the original image. The classification rate increased to 80%.

[15] proposed a new generalization of the rank nearest neighbor (RNN) rule for multivariate data for diagnosis of breast cancer. The performance of this rule using two well known databases and compare the results with the conventional k-NN rule is studied. The two wellknown databases are (i) Wisconsin diagnostics breast cancer (WDBC) database; (ii) Wisconsin breast cancer (WBC) database. They observed that this rule performed remarkably well and the computational complexity of the proposed k-RNN is much less than the conventional k-NN rules. This approach suffers from a major drawback which is the unavailability of dataset comparable to Wisconsin breast cancer (WBC).

Data Sources: It is difficult to access real medical images for experimentation due to privacy issue. The data collection that was used in our experiments was taken from the Mammographic Image Analysis Society (MIAS) [14]. This same collection has been used in other studies of automatic mammography classification. It consists of 322 images, which belong to three categories: normal, benign and malign, which are considered abnormal. In addition, the abnormal cases are further divided into six categories: circumscribed masses, spiculated masses, microcalcifications, ill-defined masses, architectural distortion and asymmetry. All images are digitized at a resolution of 1024!1024 pixels and eight-bit accuracy (gray level). They also include the locations of any abnormalities that may be present. The existing data in the collection consists of the location of the abnormality (like the center of a circle surrounding the tumor), its radius, breast position (left or right), type of breast tissues (fatty, fatty-glandular and dense) and tumor type if exists (benign or malign).

Methodology: In suggestive system, extracting features based on texture features has been done by using statistic and signal process method and classifying the samples has been done by using a new classification based on SVM named Weighted SVM.
In this part, each method is explained.

Texture Features: There isn't a clear and accurate definition that what is texture and how to describe. But, generally, texture features aren't usually the features which give us the information about order and arrangement of points in space in respect of color, brightness and so on…..

Many researches are done about analyzing the texture and its features.

In general, extract method of features based on texture can be classified into four different classes [16]:

- Statistical method
- Methods based on modeling
- Methods based on signal process
- Geometric methods

In this essay, we have also used statistic and signal process methods in extracting the features.

Statistical Methods in Extracting Texture Features: In this method, extracting texture features is based on space diffusion of gray surfaces. The most current extract method of features in this class which is common in mammography is also accidental based on matrix that has been defined by Harlick [17].

This matrix is a two-dimension histogram of accident of two gray surfaces with attention to certain distance and angle.

Members of this matrix can be shown with p_{ij} (d, é) in which p expresses numbers of accident of two gray surfaces i and j with distance d and angle é. This matrix is a collection of features that is extracted like above. Four common usable angles include: $é \in (0°, 45°, 90°, 135°)$ and distance which is usually one or two pixels.

Table 1: Some features which are extracted from co-occurrence matrix

Feature No.	Feature Name	Formula
1	Contrast(CON)	$\Phi_1 = \sum_{i,j} (i-j)^2 c(i,j)$
2	Correlation(COR)	$\Phi_2 = \sum_{i,j} \frac{(i-\mu_x)(j-\mu_y)}{\sigma_x \sigma_y}$
3	Dissimilarity(DIS)	$\Phi_3 = \sum_{i,j} (\lvert i-j \rvert^k) c(i,j);$ $where, k = 1 and i \neq j$
4	Entropy(ENT)	$\Phi_4 = -\sum_{i,j} c(i,j) \log c(i,j)$
5	Inverse Difference Moment(IDM)	$\Phi_5 = \sum_{i,j} \frac{c(i,j)}{1+(i-j)^2}$
6	Normalized Inverse Difference Moment(IDN)	$\Phi_6 = \sum_{i,j} \frac{c(i,j)}{1+((i-j)^2/\sigma^2)}$
7	Inverse Difference(INV)	$\Phi_7 = \sum_{i,j} \frac{c(i,j)}{1+\lvert i-j \rvert}$
8	Inverse Difference Normalized(INN)	$\Phi_8 = \sum_{i,j} \frac{c(i,j)}{1+(\lvert i-j \rvert/\sigma)}$
9	Uniformity(UNI)	$\Phi_9 = \sum_{i,j} c(i,j)^2$
10	Maximum Probability(MAX)	$\Phi_{10} = \max C$

Some features of Harlick matrix which are used in this route include: energy, correlation, Inertia, entropy, inverse difference in size of movement, total of variance, total of entropy, average of difference, variance of difference, entropy of difference and size of data in correlation. In Table 1, some features which are extracted from co-occurrence matrix have been mentioned.

Signal Process Methods in Extracting Texture Features: In signal process methods, extracting features of filtered image is done. This method can be classified into two categories:

- Filters in space area
- Filters in frequency area

In space area, the filtered image is obtained with convolution of main image with a collection of filters.

The most current method in this case is set of energy filters named Laws [16]. This set of filters consists of 25 masks which are gained by using convolution of a core surface of zero surface with a column of zero core surface.

Bellow, 5 Zero Core Surfaces Are Showed:

- Level: L5=[1,4,6,4,1]
- Edge: E5=[-1,-2,0,2,1]
- Spot: S5=[-1,0,2,0,-1]
- Wave: W5=[-1,2,0,-2,1]
- Ripple: R5=[1,-4,6,-4,1]

In area of frequency, the first transformations are Fourier transform (FFT) and also, conversion of cosines'(DCT) that generally analyzes the frequencies in the image.

To do this action, one of the ways is inserting a window on Fourier transform.

$$F_w(u,\psi) = \int_{-\infty}^{\infty} f(x) w(x-\psi) e^{-j2\pi ux} dx \tag{1}$$

If window function w(x) is a Gaussian function, this transformation w ill be Gabor transform. Limitation of transformation window whether in space area or in frequency area is expressed by bellow inequality which is called Gabor-Heisenberg inequality.

$$\Delta x \Delta u \geq \frac{1}{4\pi} \tag{2}$$

By selecting a Fourier transform window, space-frequency accuracy would be constant. Wavelet transform has over come this limitation.

This transformation acts based on variable Δ_u,Δ_x. Theoretically, as much the central frequency increases in analyzing filter, the space accuracy decreases [16].

In other transformations of this area, structural methods are used. In structural methods, texture is defined by a set of primal elements and a hierarchic Configuration of these primal elements.

To describe the texture, we should define the primal elements and positioning rules. S electing a primal elements (from a set of primal elements) and the possibility of being that element in a specific place can be a function of place or primal elements near that place.

The appropriate and effective display and configuration of visible information is one of the important tasks in processing the image that consists of compression, denoise and extracting the features.

Efficiency and display of an image is related to the ability of getting useful and meaningful information from an object in the image by using little descriptions.

For example, for empirical samples, an appropriate display is gained by structural transformation and pyramid algorithms.

Radon Transform: Radon transform is able to transmit a two-dimension image with different lines to other apace with parameters of line.

In this space, each line in major image is analogous to a peak in parameters of line in new space. Several definitions of this transformation are available, but Radon transform can depict an image in different directions [18]. Depicting can be done under each angle é. In general, Radon transform of function f (x.y) is equal to total of lines f, parallel to axis y that is defined like bellow:

$$R_\theta(x') = \int_{-\infty}^{\infty} f\left(x'\cos\theta - y'\sin\theta + y'\cos\theta\right)dy'$$
$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} \tag{3}$$

This method in image process is used in diagnosing line and vibration.

Ridgelet Transform: Ridgelet theory was stated in Emmanuel Candes`s doctorate thesis in 1998 [19]. He demonstrated that they can set up a system of decomposition and analysis based on function of Ridgelet.

$$\psi_{a,b,\theta}(x_1,x_2) = a^{-1/2}\psi\left(\left(x_1\cos(\theta) + x_2\sin(\theta) - b\right)/a\right) \tag{4}$$

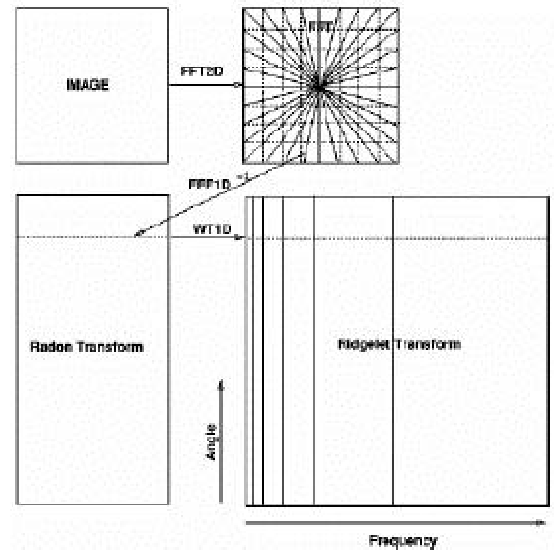

Fig. 1: General structure of Ridgelet transforms [20]

First in Ridgelet transform, a two-dimension Fourier transform photo is taken on from the image. Then, we separate the lines which pass from the destination and take a one-dimension Fourier transform photo from each of them. To this stage, Radon transform has been done.

At the end, a one dimension Wavelet transform is taken from gained lines. In figure 1, the stages of actions are showed.

With this general introduction of Ridgelet Transform, we compare that with Wavelet transform. In pyramid decomposition method, each image breaks into different surfaces with different qualities. This structure also makes the magnification (zoom in) and thumbnail (zoom out) of formation of main texture possible. Hence, texture extract isn't under the main effect of size of neighbor pixels.

This property helps us to be able to use wavelet transform for image compression, denoise and image classification.

Wavelet transform has been used and examined in many researches.

This transform is useful in diagnosing the ridge and getting texture information. Wavelet transform gives the details in 3 horizontal, vertical and diameter directions by decomposing an image into some high-pass and low-pass canals. However, these 3 directions are so limited and may not be able to get sufficient information in the denoising images which don't have strong horizontal, vertical and diameter lines.

Ridgelet transform like wavelet transform demonstrates the texture information in pyramid shape with different qualities.

This transform presents the structural information of an image based on several circular directions in frequency area for us. I research [21] has showed that this transform in isolating texture acts to stronger Wavelet transform. One of the limitations in this method is responding well in diagnosing circular-linear structure, but these kinds of lines aren't main and inherent lines in medical images.

Curvelet Transform: In most transformations which nowadays are used in processing signal and image, pyramid model is used. Frequently, like finding lines in an image, for in-bits one-dimension signals, Wavelet transforms which is a pattern of pyramid model is utilized as a suitable tool. For this transformation presents an optimal display of signal in this case and is considered as a fast transform and a suitable tree-like structure.

These are the reasons of success for Wavelet transform in examining signals and similar usages. For example, Wavelet transform has been accepted as anew compression method of image (JPEG-2000).

But natural image like in-bits flat one-dimension lines aren't simple. As, there are flat borders in many physical images, disconnected points like ridges are placed between flat curves.

Natural images have their own inherent geometric structures which are the key of finding information and their features.

The results of former researches [22, 23] have showed that Wavelet transform is suitable for finding ridges of image, but isn't successful in diagnosing contours and curves. In addition, isolating by using Wavelet gives information only in some limited directions and however, information in different directions is one of the most important characteristics of multi-dimension signals. So there is need to a stronger tool for meaningful display in higher dimension.

Suppose a series of $\{\varphi_n\}_{n=1}^{\infty}$ or Wavelet transform for signal f:

$$f = \sum_{n=1}^{\infty} C_n \varphi_n \tag{5}$$

The error which the best sentence M gives first of these functions is a structure for measuring efficiency. A non-linear estimate of sentence M is obtained like bellow:

$$\hat{f}_M = \sum_{n \in I_M} C_n \varphi_n \tag{6}$$

As according to f_M is set of the biggest M. $|C_n|$ quality of estimate f_M is computed so that how this expansion can compress energy of function f in the least numbers of coefficients.

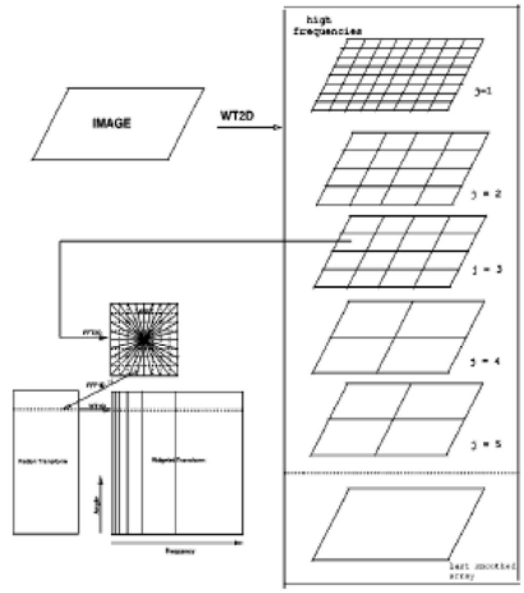

Fig. 2: General structure of Curvelet transforms [20]

Donoho Candes invented a new method in cohesive two-dimension space named Curvelet. This expansion has obtained an optimal estimate of in-bits flat cohesive functions, but for disconnected function with this intensity isn't good. The best estimate of sentence M which is obtained by error square $(L_2 - norm\,square) \| f - \hat{f}_M \|_2^2$ has rank $O((\log M)^3 M^{-2})$ in Curvelet transform. While this rank is in Wavelet transform $O(M^{-1/2})$ and Fourier transform $O(M^{-1})$ [24]. So for images with flat contours by using Curvelet with regard to Wavelet transform, we expect improvement.

As Wavelet transform with regard to Fourier transform has better result for in-bits one-dimension flat signals.

At first, Curvelet transform was done for cohesive areas by multi-scales filters and by transforming a block Ridgelet on each middle-pass image.

In figure 2, general structure of Curvelet transform has been showed. Later, its author invented the second generation of Curvelet transform. This copy had been defined by using segmentation frequencies and without Ridgelet transform directly. In both copies, we need rotation operators and analogous to that segmentation frequencies like one-dimensional in polar space. This characteristic causes that Curvelet transform would be simple in cohesive space and difficult in some cases in disconnected space.

This transform is difficult in disconnected space because rectangular sampling for disconnected images cause to impose horizontal and vertical directions to that.

This fact became a motivation for improvement in directional hierarchic transformation like Curvelet in disconnected area.

Do and his colleagues [22] have surveyed this enterprise and introduced Contourlet.

Contourlet transform is new expansions of Wavelet transform that use directional filters in different scales. Contourlet has been shaped based on image rotation in different directions in several scales with convertible rates. With this strong collection of images, this transformation can always have flat contours which are the main and inherent characteristics if image. Analyzing images are usually based on statistical models. Natural images tend to have a series of definite and current characteristic that makes them natural.

The purpose of statistical modeling is to save these characteristic in a few numbers of parameters to be able to use as series of former information in different parts of image process such as compression, denoise and feature extract.

Images (pictures) can be a model by wavelet transform which is a multi scale transform and frequency time. In the beginning, it is supposed that wavelet transform is a suitable transform for omitting the image's dependence. So by the usage of the wavelet transform's coefficient which are independent of each other, we can consider a good model for Marginal statistical data statistic data. Next researches show that there is implicit dependency between wavelet transform's coefficient in various scales and also between neighborhoods coefficient in one level and in one under-image especially around the edges. This finding performs an important contribution in progressing compression in the wavelet domain by the usage of statistical method.

The main restriction in transforming the two-dimension wavelet is not receiving information from different directions. As you know the wavelet transform just can extract information in vertical, horizontal and diametrical direction. To conquer this problem researcher use the multi-direction scales method which can conclude the main structure of picture's geometry like Contour, some of these methods are as follow[22]:

- Steerable Pyramid
- Brushlet
- Complex wavelet transform
- Contourlet

Contourlet transform in addition to have main features of wavelet transform which has multi scale and time frequency it also have high degree of direction and anisotropy. This transform presents different and changeable number of various directions in different scales of the image. Contourlet transform uses filters repeatedly, in a way to have sufficient calculation and suitable complexity. For processing point N in this transform we need O(N) Operator.. Contourlet transform can make model from the dependency between various directions. In the other way give this power to us to model three main visible parameters of scale, place and direction. [22] The aim of Contourlet transform is to find a small expansion for one smooth and Pieces image.in two-dimension wavelet the direction is lost. This transform is just good for separated points but geometrical smoothness does not cover this Contour. Contourlet transform is a progressed version of wavelet transform in this field. In comparison with wavelet transform, Contourlet keep a less amount of correlation for presenting a Contour. Recently for suitable presenting of the image which has the geometrical features various methods have been used like wedgelet, bandele and Code tree square Forums Read. In all cases, there is a need for an edge diagnosis process which has compatibility with its presenting method. Generally Contourlet should have important features and characteristics which are in wavelet transform but since there is a high correlation between these coefficients in a smooth image we should omit these extra coefficients in the new transform.

At first we should use one transform the same as wavelet for the edge diagnosing. Then we do a local directional transform for Contour separation. It is interesting since the other stages are the same as Hough transform in the polar space for line diagnosis. Two filter's banks are used for small expansion of Contour. (In the filter's bank, filters are vertical over eachother. These filters decrease the extra coefficient). At first these two filters Laplacian Pyramid (LP) are used for covering cut points. Then we use the directional filter's bank for connecting this cut to the linear structure. In this structure we use whole to details in various scales which their size are depended on apparent form. The research results show that [22] this expansion can calculate Contour in various qualities with total capability. This transform in the frequency domain acts in the form of multi scales and directional analyzing.

One of the directional analyzing methods is the Laplace pyramids usage which has been introduced by Adelson and Burt.[25]in it,the pyramids transforms of each phase is one sample of the lower rank of the original image which has been filtered by the Low-pass filters. The difference of main and predictive image has been delivered in the form of between-pass filter.this action has been done repeatedly over the general canal. In multi dimensions filter's banks sampling has been done by the sampling matrix. As an example x[n] low rank sampling by the M filter is as follow:

$X_d[n] = x[M_n]$ In it M is a correct matrix.

One of the LP problem is that has various model that put on each other incorrectly. In any way in LP comparison with the wavelet sampling, LP has specific characteristics that in each pyramidal stage produce just one between-pass image(even in multi dimensions case) which this image does not have the scramble frequencies of image. But the frequency variances in the wavelet filter's bank is happened in up passing canal after low rank sampling, which put down in the frequency band and so their spectrum have been reflected. But in LP by low ran sampling which has just taken from the low passing canal, prevent from this problem. In the research which has done by DO and *et al,* [22] they paid to the LP framework. In this research, it is showed that LP produces the strong framework with vertical filters over each other which their analyzing is in the form of inversion time. We use the Contourlet transform for extracting features, with attention to mentioned subjects in this research.

Classification: After choosing the proper characteristics from the existing samples, we use classifier for their classification.

SVM (Support Vector Machine): (Support Vector Machine)SVM is a closed classification for a large amount of researchers for solving various classified problems in real world [26].in this part a complete definition of SVM has been presented. One classified binary problem with an observer usually is expressed as follow: by the usage of n instruction sample $(\langle x_i \rangle, y)_i$ in which $\langle x_1 \rangle = \langle x_{i1}, x_{i2}, \dots x_{im} \rangle$ is the vector of input features and $yi \in \{-1,+1\}$_is the label of classified class. The work of segmented or classified function is learning the existing samples in the instructional samples in a way that in the next stages it can reliably specified an unknown x_i input, a suitable y_i. Originally SVM is designed for this time it means binary classification, but it can vast to the multi classification's capability. Like the other linear classification, SVM tries to create a linear decision boundary (suppose that data can have linear separation) or a linear hyper plane between two classes.

Usually the belief is that the points which belonged to the two classes, put in a way that there is distance between them.SVM tries to maximize the distance between two classes with consider to the problem as a second class of programmed problem.[27,28]. In the case of non-linear condition, a mapping function $Ø(x_i)$ for writing the input space to the feature's space with higher dimensions in a way that non-linear hyper plane change to linear, has been used. To prevent the augmentation in calculation complexity and working problems in high dimension, a *K(xi, xj)* kernel has been used that originally kernel's amount which is equal to the input space is calculated in a way that there is no need for any Explicit mapping.

Usually Kernels Are as Follow:

Linear : $xi^T x_j$

Radial basis function RBF: $e^{-\gamma \left\| x_{i-x_j} \right\|^2}, \gamma > 0$

$\left(\gamma x_i^T x_j + a\right)^d, \gamma > 0.$

Polynomial of degree d: $\tanh\left(\gamma x_i^T x_j + a\right)^d$,MLP

In all of this relations *g,* γ and *d* are the classified parameters.

Weighted SVM: In this research we use a classifier for classifying healthy and diseased texture which is based on the strong classifier to land SVM. (Weighted SVM) weighted Support Vector Machine is a classifier in this research.

Usually the numbers of healthy samples are more than diseased one. The importance of diseased class is more important than healthy class since if a person is patient but diagnose as a healthy one he will face with irreparable harms. Because of these two reasons we should give value to the patient's data. With attention to have less number of diseased samples in relation to healthy one we give them specified weight to equal their value approximately. This valuing is calculated with attention to the number of two classes. For weighting data or valuing we use the (7) equation.

$$\begin{cases} \dfrac{W_1}{L_2} = \dfrac{W_2}{L_1} \\ W_1 + W_2 = 1 \end{cases} \quad (7)$$

L1,L2,W1,W2 in order, there are the numbers of each class and their corresponding weights.

Until now each two classes participates equally in each classification. Then by the Cross Validation method, 99 percent of samples are used for instruction and the remained one for testing (examination).

RESULTS AND DISCUSSION

We can refer to Ridgelet transform and the family of wavelet transform from the previous transforms in the time frequency domain, that each one of these transforms has some restrictions. All three Contourlet transform, Ridgelet and wavelet transforms are from hierarchical transform in the pyramid form. Data wavelet transform is just obtained in 3 vertical, horizontal and diametrical directions. And there is also some restriction in scale's amount. Contourlet transform in addition to have the main features of wavelet transform which is multi scales and time frequency it covers high degree of direction and Anisotropy. This transform presents various and changeable number of various directions in different scales of the image. Two dimensions of wavelet transform are just good for separated points but it does not express the geometrical smoothness of a Contour. Contourlet transform is a prepared version of wavelet transform in this domain. In comparison with wavelet transform the Contourlet transform keep the short amounts of correlations for expressing a Contour.

Energy, correlation, enercy Inertia, entropy, Inverse Difference movement, total average, total variance, total entropy, difference average, difference variance, difference entropy and information size in co relation.

For extraction of texture feature with signal processing method, filters in the frequency domain are used.

In this domain Contourlet is used which is a new version of the Curvelet transform. Four levels of Contourlet have been taken from the doubtful region. The number of the obtained images in each area (level, surface) of the body is as follow: the first surface shows one image, the second one show four, the third shows eight and the fourth surface shows sixteen images in different scales and directions.

Considered features from Contourlet are as follow: a complex of max K of the first coefficients, average and Standard deviation of, energy, entropy and esque

Features classification is performed according to the strong SVM classifier with RBF kernel with the name of SVM Weighted .in this classifier two classes of healthy and diseased classification are considered. And with attention to the number of the samples in each two classes, we weight the classes. With attention to have twenty eight diseased sample and thirty two healthy samples, the weight of healthy and diseased samples is orderly calculated 55% and 45% and then with the Cross Validation method [29] approximately we use ninety percents of samples for instruction and the remained one for testing.

Searching and results show that the surprising power of Contourlet coefficients has an effective role in compounding with the mentioned classifier in diagnosing and classifying breast cancer tumors.

	SVM				Weighted SVM			
Measures	Co-occurrence	Wavelet	Contourlet	Contourlet +Co-occurrence	Co-occurrence	Wavelet	Contourlet	Contourlet +Co-occurrence
Mean ClassRate	0.441	0.442	0.446	0.453	0.512	0.524	0.537	0.554

CONCLUSION

Features extraction is according to the texture characteristics and is performed by statistical method and signal processing methods.

Statistical method in extraction of texture features base on simultaneously matrix.

A complex of simultaneously matrix which has used in this project is as follow:

REFERENCES

1. Spence, D., L. Parra and P. Sajda, 2001. Detection, synthesis and compression in mammographic image analysis using a hierarchical image probability model. Artificial Intelligence in Medicine, 25(31): 365-371.
2. Verma, K. and J. Zakos, 2000. A computer-aided diagnosis system for digital mammograms based on

fuzzy-neural and feature extraction techniques. IEEE Transactions on Information Technology in Biomedicine, 16: 219-223.
3. Arun, K., 2001. Computer vision fuzzy-neural systems. Englewood Cliffs, NJ: Prentice-Hall.
4. Zaiane, O., A. Maria-Luiza and C. Alexandru, 2002. Mammograph classification by an association rule-based classifier. In Proceedings o second international workshop on multimedia data minin (MDM/KDD') in conjunction with seventh ACM SIGKDD, USA.
5. Bird, R., T. Wallace and B. Yankaskas, 1992. Analysis of cancers missed at screening mammography. Radiology, Gonzalez, R. C. and Woods, R. E. (2002). Digital image processing (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall. 178: 234-247.
6. Roberts, M., E. Kahn and P. Haddawy, 1995. Development of a Bayesian network for diagnosis of breast cancer. IJCAI-95 workshop on building Probabilistic Networks.
7. Zaiane, O., A. Maria-Luiza and C. Alexandru, 2001. Application of data mining techniques for medical image classification. In Proceedings of second international workshop on multimedia data mining (MDM/KDD') in conjunction with seventh ACM SIGKDDS, USA.
8. Mendez, A., P. Tahoces, M. Lado and J. Souto, 1998. Computer-aided diagnosis: Automatic detection of malignant masses in digitized mammograms. Medical Physics, 25: 109-131.
9. Taylor, P., 1995. Decision support for image interpretation: A mammography workstation Information processing in medical imaging. Dordrecht: Kluwer Academic Publishers.
10. Woods, S.K., 1994. Automated image analysis techniques for digital mammography. PhD Thesis. University of South Florida.
11. Yin, F., M. Giger, C. Vyborny, K. Doi and R. Schmidt, 1993. Comparison of bilateral-subtraction and single imageprocessing tech- niques in the computerised detection of mammographic masses. Investigative Radiol., 28(142): 179-183.
12. Qian, W., P. Sunden, H. Sjostrom, H. Fenger-Krog and U. Brodin, 2002. Comparison of image quality for different digital mammogram image processing parameter settings versus analogue film. Electromedica, 71(1):57-73.
13. Sameti, M. and R. Ward, 1996. A fuzzy segmentation algorithm for mammogram, partitioning. Third international workshop on digital mammography, Chicago.
14. Suckling, J., J. Parker, D. Dance, S. Astley, I. Hutt, C. Boggis, *et al.* 1994. The mammographic images analysis society digital mammogram database. Exerpta Medical International Congress Series, 1069: 375-378.
15. Baguia, C., 2003. Breast cancer detection using rank-nearest neighbor classification rules. Pattern Recognition, 36: 367-381.
16. Malagelada, I. and A. Oliver, 2007. Automatic Mass Segmentation in Mammographic Images, PHD-thesis, Universitat de Girona.
17. Haralick, R.M., K. Shanmugan and I. Dinstein, 1973, Textural features for image classification, IEEE Transactions on Systems, Man and Cybernetics, 3(6): 610- 621.
18. Unsalan, C., 1998, pattern recognition methods for texture analysis case study : steel surface classification, B.S. thesis in E.E., Hacettepe University,
19. Candes, E.J. and D.L. Donoho, 2000, Curvelets - A Surprisingly Effective Nonadaptive Representation for Objects with Edges, Saint-Malo Proceedings.
20. Starck, J.L., E.J. Candés and David L. Donoho, The Curvelet Transform for Image Denoising, 2002, IEEE Transactions on Image Processing, 11:6.
21. emler, L. and L. Dettori, 2006, A Comparison of Wavelet-Based and Ridgelet-Based Texture Classification of Tissues in Computed tomography, proceedings of International Conference on Computer Vision Theory and Applications.
22. Do, M.N. and M. Vetterli, 2005. The Contourlet Transform: An Efficient Directional Multi - resolution Image Representation, IEEE Trans. 14(12): 2091-2106.
23. I- Donoho, D.L. and M.R. Duncan, 1999. Digital Curvelet Transform Strategy, Implementation and Experiments, Department of Statistics Stanford University .
24. Po, D.Y. and N. Do, 2006. Directional Multi - Scale Modeling of Images using the Contourlet Transform, IEEE Trans. pp: 1-11.
25. Burt, P.J. and E.H. Adelson, 1983, The Laplacian pyamid as a compact image code, IEEE Trans. Commun., COM - 31(4): 532- 540.

26. Rajpoot, K.M. and N.M. Rajpoot, 2004. Wavelets and Support Vector Machines for Texture Classification. Proceedings 8th IEEE International Multitopic Conference (INMIC), Lahore, Pakistan.
27. Cristianini, N. and J. Shawe-Taylor, 2000. An Introduction to Support Vector Machines and other kernel-based learning methods. Cambridge University Press, Cam-bridge, UK.
28. Vapnik, V., 1998. Statistical Learning Theory, Wiley, New York,
29. Introduction to pattern Analysis, Ricardo Gutierrez Osuna, Texas AandM University, cross validation, Lecture, pp: 133.