Volume: 04 Issue: 01 June 2015 Page No.1-4
ISSN: 2278-2419

Microcalcification Classification in Digital Mammogram using Moment based Statistical Texture Feature Extraction and SVM

K. Sankar[1], K.Nirmala[2]
[1]ResearchScholar, ManonmaniamSundaranarUniversity,India.
[2]Associate Professor, Dept. of Computer Science,Quaid-e-Millet College,Chennai,India
Email: rkmvcks@gmail.com, .nimimca@gmail.com

Abstract-The digital mammogram is a reliable technique to detect early breast cancer without any symptoms. The main aim objective is to classify the mammogram microcalcifications images either benign or malignant. This system consist of three stage that is mammogram enhancement, statistical texture feature extraction and classification. The mammogram images are enhanced by shift-invariant transform which consist of shift-invariant multi-scale, multi-direction property and classify mammogram pixels into strong edges, weak edges and noise edges. It clearly distinguishes weak edges and noise edges. The moment based statistical texture features are extracted from enhanced images and stored. Finally, these features are fed into SVM classifier to classify the mammogram images.

Keywords—Micocalcification,shift-invariant Transform,Moments,SVM.

I. INTRODUCTION

Breast cancer arises due to uncontrollably of breast cells which produce a breast tumor. The breast tumor can be normal and abnormal. The normal tumor represents no cancerous. The abnormal consists of two classes such as benign and malignant. The benign considered as a non-cancerous which is close to normal in appearance. They grow gradually and do not spread or invade nearby tissues to other parts of the body. The malignant is cancerous that spreads beyond the original tumor to the other parts of the body. There are several kinds of abnormalities revealed such as asymmetrical breast tissue, asymmetrical density, architectural distortion, mass, microcalcification, interval changes compared with previous films, adenopathy and other miscellaneous. In the literature, among several abnormalities we use mammogram microcalcifications are used for experiments. The microcalcifications associated small calcium deposits in the breast tissues. This can be categorized based on shape, size, distribution and density.

This can be benign microcalcification and malignant microcalcification. Benign micocalcifications are distributed in diffuse or bilateral arrangement in the acini or with a round or punctate shape or scattered in dense breast tissue. Malignant microcalcifications are in a brachning or linear pattern and with irregular borders, or with variable density, or distributed in a segmental way.Several techniques have been used to detect the breast abnormalities such as mammography, ultrasound, MRI, and nuclear medicine. The mammography is a reliable method to detect breast cancer at the early stage without no symptoms. The early detection saves many lives and reduces mortality rates. There are several well-known features related to shape, size, and texture (histogram, Haralick's texture features, and moment-based features) are extracted from the mammograms images [1]. Texture is a significant feature that has been used in wide range of application such as automated inspection, medical image processing, document processing, remote sensing and content-based image retrieval. There are four types of texture feature extraction namely structural, statistical, model-based and transform domain [2].This work aims to Mammogram enhancement, extract moments based statistical texture features and SVD features and classification by SVM. This paper is organized as follows. Section II discusses about the background work of some papers through literature survey. Microcalcification enhancement is explained in section III. Section IV explores the concept of feature extraction. The proposed system using the support vector machines is discussed in section V. Finally, Section VI concludes the research work.

II. REVIEW OF LITERATURE

There are numerous researchers have been proposed to detect the breast cancer. Moment based statistical features are extracted and neural network classifiers have been proposed to detect breast cancer [3]. Ultrasound imaging is one of the most frequently used diagnosis tools to detect and classify abnormalities of the breast. In this method the masses are classified as either benign or malignant [4]. Microcalcification based statistical features and Stochastic Neighbor Embedding (SNE) is proposed to extract features such as mean, standard deviation, skewness and kurtosis. Then the extracted features are given as an input to the robust K-Nearest Neighbor (KNN) classifier to classify the mammogram images into normal or abnormal, and the abnormal into benign or malignant [5]. Vermaet al. (2009) proposed a novel soft cluster neural network. The highest classification accuracy obtained by this approach was 93% on mammograms from the DDSM [6]. Jacobi moments are utilized to extract features mammogram features and SVM classifier is used to classify the images into normal and abnormal then the abnormal images are classified into benign or malignant [7].

III. MICROCALCIFICATION ENHANCEMENT

The mammogram images are noise, low-contrast and blur due limitations of X-ray hardware systems. The detection of microcalcificationsare difficult due their small shapes and size and also exhibits poor contrast. So we have to enhance mammogram images using Shift-invariant Transform [8]-[11]. The shift-invariant transform consist of two parts: NonsubsampledNonSubsampled Pyramid (NSP) ensures multi-scale property and NonSubsampled Directional Filter Bank (NSDFB) handling multi-directionality as shown in fig 1.

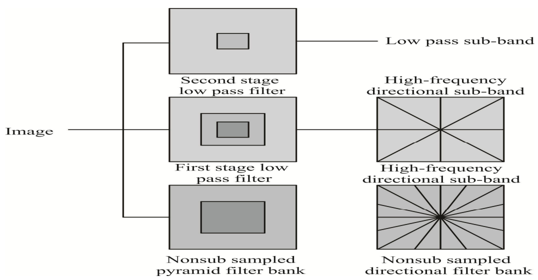

Fig. 1: NSCT Process - NSP decomposition followed by NSDFB

In this transform, the pixels are classified into three types: strong edges, weak edges and noise. The strong edges contains large magnitude coefficients in all subbands, weak edges contains larrge magnitude coefficients in some directional subbands and small magnitude coefficients in other sunbands and noise edges contains only low magnitude coefficients. Based on the observation the pixels are classified by

$$\begin{cases} \text{strong edges}, & \text{if mean} \geq c\sigma \\ \text{weak edges}, & \text{if mean} < c\sigma, \max \geq c\sigma \\ noise, & \text{if mean} < c\sigma, \max < c\sigma \end{cases} \quad (1)$$

Where c is a parameter ranging from 1 to 5 and σ is the noise standard deviation of the sub-bands at a specific pyramidal level.The main aim of proposed method is to amplify weak edges and to suppress noise. We have to modify the NSCT coefficients according to the category of each pixel.

$$y(x) = \begin{cases} x, & \text{strong edges pixels} \\ \max\left(\left(\frac{c\sigma}{|x|}\right)^p, 1\right) x, & \text{weak edges pixel} \\ 0, & noise \end{cases} \quad (2)$$

Where the input x is the original coefficient, and 0 < p < 1 is the amplifying gain. This function keeps the coefficients of strong edges, amplifies the coefficients of weak edges, and zeros the noise coefficients.

IV. FEATURE EXTRACTION

A. Moment Based Statistical Texture Features

The moments based statistical texture features are extracted from digital mammogram images. This features carries information about mammogram like mean, standard deviation, smoothness, entropy skewness and kurtosis.The n^{th} central moment is defined as

$$\mu_n = \sum_{i=0}^{L-1} (\lambda_i - m)^n p(\lambda_i) \quad (3)$$

Where $\lambda_i$is a random discrete variable,$p(\lambda_i)$ repsents histogram intensity levels and L is the number of intensity levels.

$$m = \sum_{i=0}^{L-1} \lambda_i \, p(\lambda_i) \quad (4)$$

The first order moments of central moment gives average intensity values. The second order moment represents standard deviation that gives average contrast of mammogram images. The third high-order moments denotes skewness that measures the degree of asymmetry of distribution. Skewed means thedistribution is not symmetrical. In a symmetrical distribution, the values of the variable equally distant from their mean. The mean, median and mode are coincided in a perfect symmetrical distribution. When the distribution is skewed to the right, mean is greater than mode. When the distribution is skewed to the left, mean is less than mode. The fourth high-order moments expresses kurtosis that describes flatness or peakness of a distribution.The smoothness computes the relative smoothness of intensity in a histogram. It can be computed as

$$Smoothness = 1 - \frac{1}{1 + \sigma^2} \quad (5)$$

The entropy represents randomness of the mammogram images. It defined as

$$Entropy = -\sum_{i=0}^{L-1} p(\lambda_i) \, log \, p(\lambda_i) \quad (6)$$

The energy measures uniformity of the mammogram images, it can be computed as

$$Energy = \sum_{i=0}^{L-1} p^2(\lambda_i) \quad (7)$$

B. SVD Features

The Singular Value Decomposition (SVD) is a powerful technique in many matrix computations and analysis. Using the SVD of a matrix in computations, rather than the original matrix, has the advantage of being more robust to numerical error. Additionally, the SVD exposes the geometric structure of a matrix, an important aspect of many matrix calculations. The SVD is employed in a variety applications, from least square to solving system linear equations. Each of these application exploits key properties of the SVD, its relation to the rank of a matrix and ability to approximate matrices of a given matrix.A singular value decomposition of an M × N matrix A is any factorization of the form

$$A = U \sum V^T \quad (8)$$

Where U is an M × M orthogonal matrix, V is an N × N orthogonal matrix, and $\sum$ is an M × N diagonal matrix with $S_{ij} = 0$ if $i \neq j$ and $S_{ii} = 0 \; S_i \geq 0$.

V. PROPOSED sYSTETEM USING sUPPORT VECTOR MACHINES

Support vector machines (SVMs) are one of the more popular approaches to data modelling and classification, more recently subsumed within kernel methods. The SVM has the ability to classify correctly samples that are not within feature space used for training. This can used wide range applications such as text and hypertext categorizations, human face detection, pattern recognition, texture classification,especially in the medical domain. The main aim of SVM is to classify the mammogram images are normal or abnormal, then the abnormal images are classified either benign or malignant.An overview proposed mammogram microcalcificationis adopted which consistsofmicrocalcificationenhancement, feature extraction and classification as shown in fig 2.

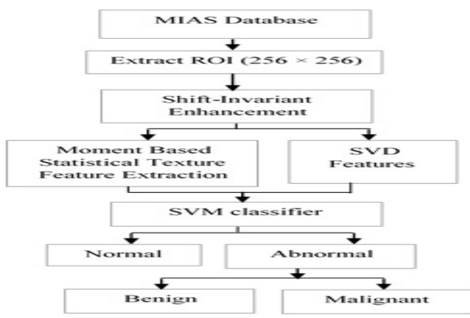

Fig 2: .Proposed mammogram microcalcification classification system

VI. EXPERIMENTAL RESULTS

In this research work, the MIAS database is used for experiments which consist of a total of 322 digital mammogram images both normal and abnormal. First, the enhancement pre-processing experiments were conducted over (100 Normal, 8 Benign Microcalcifications, and 10 MalignantMicrocalcifications) image by proposed shift-invariant transform as shown fromFig 3 to Fig 5. Second, moment based statistical features and SVD features from 120 Normal, 8 Benign Microcalcifications, and 10 MalignantMicrocalcifications images are extracted. The sample result are shown from Tables I to Table VI. Finally, these features are fed into SVM classifier to classify either normal or abnormal, then the abnormal images are classified either benign or malignant.

Table 1: Moment Based Statistical Texture Features for Normal Mammogram images

Features	Image ID				
	mdb006	***mdb007***	***mdb014***	***mdb016***	***mdb020***
Mean	143.9475	158.5528	193.0389	157.1544	157.1231
Variance	30.5061	143.7972	136.3357	264.7423	97.1343
Skewness	0.0030	0.0002	0.0002	0.0000	0.0007
Kurtosis	0.0023	0.0001	0.0001	0.0000	0.0003
Smoothness	0.9682	0.9931	0.9927	0.9962	0.9898
Energy	0.0523	0.0236	0.0237	0.0168	0.0304
Entropy	4.4613	5.6119	5.5171	6.0505	5.2412

Table 2: Moment Based Statistical Texture Features for Benign Microcalcification images

Features	Image ID				
	Mdb252	***Mdb222***	***Mdb226***	***Mdb236***	***Mdb227***
Mean	150.2569	173.3930	170.7828	195.0608	171.5282
Variance	277.3134	468.2192	571.4393	338.0761	301.7835
Skewness	-0.0004	0.0000	0.0000	-0.0001	-0.0004
Kurtosis	0.0001	0.0000	0.0000	0.0000	0.0001
Smoothness	0.9964	0.9979	0.9983	0.9971	0.9967
Energy	0.0239	0.0136	0.0119	0.0177	0.0207
Entropy	5.7843	6.4101	6.4901	6.0523	5.9235

Table 3: Moment Based Statistical Texture Features for Malignant Microcalcification images

Features	Image ID				
	Mdb209	***Mdb211***	***Mdb213***	***Mdb216***	***Mdb231***
Mean	157.8415	172.2273	132.5669	211.4736	130.6192
Variance	858.0987	372.7025	750.8552	160.9373	118.0120
Skewness	0.0000	0.0001	0.0000	-0.0005	0.0003
Kurtosis	0.0000	0.0000	0.0000	0.0002	0.0002
Smoothness	0.9988	0.9973	0.9987	0.9938	0.9916
Energy	0.0106	0.0150	0.0108	0.0252	0.0261
Entropy	6.6575	6.1860	6.7234	5.5573	5.4348

Table 4: SVD Features for Normal Mammogram images

Image ID				
mdb006	mdb007	mdb014	mdb016	mdb020
36864.3995	40633.0912	49480.7382	40318.0566	40262.2951
408.5342	1404.5736	1119.2819	1918.5631	1022.3996
372.0636	1093.1044	594.5213	1827.6877	711.5510
311.8370	671.5103	451.4491	920.0903	594.3169
263.6510	616.3089	371.4866	742.0006	484.1747
228.6806	589.8388	347.2063	583.3193	384.8078
210.8154	452.1648	296.2909	523.0148	371.3417
204.6966	426.7935	281.7656	448.3690	332.7626
180.4221	386.0630	241.4179	375.3236	297.4562
162.3241	368.8901	223.4207	367.5789	255.7541

Table 5: SVD Features for Benign Microcalcification images

Image ID				
mdb252	mdb222	mdb226	mdb236	mdb227
38575.1558	44639.0816	44081.2689	50118.2330	44056.7669
2593.7221	1912.5218	1414.6865	1122.9370	1844.9779
875.2290	1155.8294	1320.3499	989.0030	929.4223
650.3027	954.9926	766.3159	824.1783	757.7109
587.9459	857.8789	538.3219	422.2911	685.8577
495.2127	677.2515	466.8432	347.6494	543.4836
405.8960	527.3957	356.7434	278.5032	446.3302
326.7466	411.4764	349.7444	267.4861	412.7891
322.2812	328.2563	306.9024	251.2630	382.2359
304.7974	299.1059	268.8535	240.9955	333.7675

Table 6: SVD Features Malignant Microcalcification images

Image ID				
mdb209	mdb211	mdb213	mdb216	mdb231
40872.1575	44294.2847	34469.2227	54198.2604	33513.7587
3312.0855	1312.3209	3006.1196	1243.3136	859.1925
1657.4817	1155.2130	1242.8730	924.9878	598.8261
1199.5201	801.4471	747.9701	651.9572	503.9472
845.9325	653.8005	623.5661	494.5887	480.3885
687.5055	584.0844	494.1277	415.1999	410.8642
603.9160	560.9895	342.1732	333.5813	362.9551
515.2406	530.2370	331.2750	294.7645	337.6098
494.2237	440.4890	302.9584	270.8016	287.9222
414.2007	365.3820	269.7974	256.8182	221.2257

VII. CONCLUSIONS

In this research work, a new method is proposed to enhance the mammogram microcalcification images which provide perfect reconstruction after modifying coefficient, faster implementation and also clearly distinguishes noise edges and weak edges. The implementation shows better visual quality mammogram image. Moment based statistical texture features mean, variance, skewness, kurtosis, smoothness, energy and entropy are extracted and SVD features are extracted from normal, benign and malignant images.This system is classified

International Journal of Data Mining Techniques and Applications
Volume: 04 Issue: 01 June 2015 Page No.1-4
ISSN: 2278-2419

95% normal, abnormal 89% benign 88% malignant 87% for the mammogram images. This moment is not orthogonal and the image reconstruction is difficult.

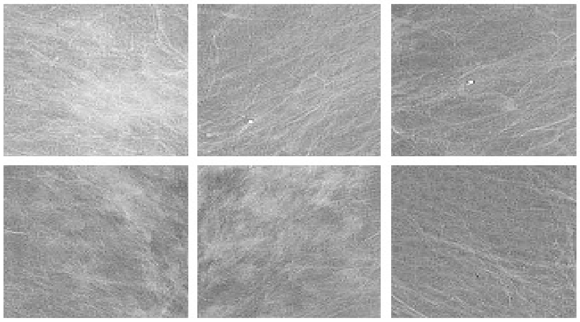

Fig 3. Sample Normal Images: Enhanced Shift-invariant Transform images (256× 256)

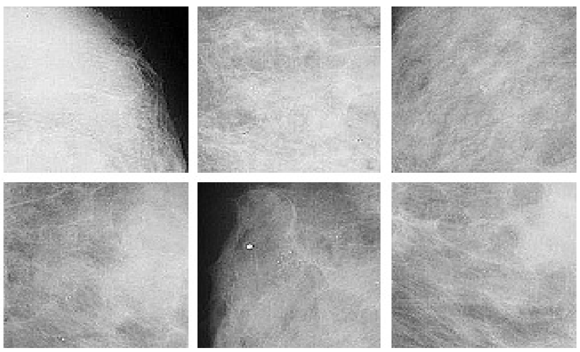

Fig 4. Sample Benign Images: Enhanced Shift-invariant Transform images (256× 256)

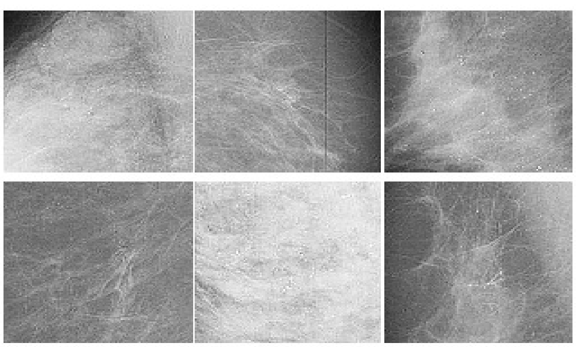

Fig 5. Sample Malignant Images: Enhanced Shift-invariant Transform images (256× 256)

REFERENCES

[1] Sergio Koodi Kinoshita, Paulo Mazzoncini de Azevedo-Marques, Roberto Rodrigues Pereira Jr, Jośe Antônio Heisinger Rodrigues and Rangaraj Mandayam Rangayyan, “Content-based Retrieval of Mammograms Using Visual Features Related to Breast Density Patterns”, June 2007, Volume 20, Issue 2, pp 172-190.

[2] M.Tuceryan and A. K. Jain, Texture analysis, In The Handbook of Pattern Recognition and Computer Vision (1998), 207-248.

[3] R. Nithya and B. Santhi, “Breast Cancer Diagnosis in Digital Mammogram using Statistical Features and Neural Network”, Research Journal of Applied Sciences, Engineering and Technology 4(24): 5480-5483, 2012.

[4] K.Subashini and K.Jeyanthi, “Masses detection and classification in ultrasound images”, IOSR Journal of Pharmacy and Biological Sciences, e-ISSN: 2278-3008, p-ISSN:2319-7676. Volume 9, Issue 3 Ver. II (May -Jun. 2014), PP 48-51.

[5] S Mohan Kumar and G. Balakrishnan,” Statistical Features Based Classification of Microcalcification in Digital Mammogram Using Stochastic Neighbor Embedding”, International Journal of Advanced Information Science and Technology (IJAIST) ISSN: 2319:2682 Vol.7, No.7, November 2012.

[6] Verma, B., P. McLeod and A. Klevansky, “A novel soft cluster neural network for the classification of suspicious areas in digital mammograms”. Pattern Recogn., 42: 1845-1852. 2009.

[7] N.V.S.Sree Rathna Lakshmi and C.Manoharan, “An Automated System for Classification of MicroCalcification in Mammogram Based on Jacobi Moments”, International Journal of Computer Theory and Engineering, Vol. 3, No. 3, June 2011.

[8] A. Cunha, J. Zhou, and M. Do , “ The nonsubsampled contourlet transform: theory, design, and applications ”, IEEE Transactions on Image Processing, Vol. 15, pp. 3089-3101, Sep. 2006.

[9] Hong Zhang, Chenxi Zhangand Mingui Sun," Infrared image enhancement based on NSCT and neighborhood information ", Proc. SPIE 8768, International Conference on Graphic and Image Processing (ICGIP 2012), volume 9069, 2013.

[10] Sangeetha, T. A. and Dr. Saradha, A, “An efficient way to enhance mammogram image using nonsubsampled contourlet transform”, International Journal of Current Research, Vol. 4, Issue, 12, pp. 385-390, December, 2012.

[11] Yunlan Tan, Chao Li, Guangyao Li, WenlangLuo and Weidong Tang, “A Novel Approach for Image Enhancement via NonsubsampledContourlet Transform”, Volume 23, Issue 3, Pages 345–355, ISSN (Online) 2191-026X, ISSN (Print) 0334-1860, April 2014