Tumour in Mammogram using Digital Image Processing Techniques

Er.Kanchan Sharma[1], Dr. Deepti Mittal[2]
[1]Research Scholar Thapar University,Patiala,
[2]Assistant Professor Thapar University,Patiala,

***Abstract*—Breast cancer is the second most common cancer.Mostly it founds in females as compared to males.In this proposed work,features(energy,correlation,contrat and homogeneity) are extracted using GLCM(Gray Level Co-Occurrence Matrix) and compare the feature values of original image with the corresponding image having cancer.DDSM/MIAS(Digital Database for screening Mammography/The Mammographic Image Analysis Society) images have been taken to detect benign and malignant cancer using digital image processing techniques.**

***IndexTerms*—Benign,DDSM/MIAS,GLCM Features,Malignant Cancer.**

I. Introduction

GLCM stands for Gray level Co-occurrence matrix. It is of 2nd order statistics, so information about pixels of pairs are collected by GLCM.GLCM shows how the pixel brightness in an image occurs..A matrix is constructed at a distance d=1 and at angles in degrees(0,45,90,135) .Haralick also proposed different measures i.e. entropy,energy,contrast,correlation etc.These measures calculate at different angles.

GLCM is texture feature and texture mention to touch i.e. smooth, silky and rough etc.The order of texture statics are:

1. First order texture measures are statistics determined from the original image values, like variance, and pixel neighbour relationship are not taken.
2. Second order measures the relationship between groups of two (usually neighbouring) pixels in the original image.
3. Third and higher order textures (noting the relationships among three or more pixels) are theoretically possible but practically/ commonly not implemented due to calculation time and interpretation difficulty. [1]

time and interpretation difficulty.

GLCM texture takes the relation between two pixels at a time, called the reference and the neighbour pixel.GLCM illustrates the distance and angular spatial relationship over an image sub- region of specific size. GLCM is created from gray scale values.It is calculated how often a pixel with gray level(gray scale intensityor gray tone) values come either horizontally,vertically and diagonally to alongside the pixels with the value j.GLCM directions are:

1. Horizontal(0)
2. Vertical(90)
3. Diagonal

Manuscript received Aug 15, 2012.
***Kanchan Sharma**, EIC, Thapar University, Patiala, India*
***Dr.Deepti Mittal**, ECE / Thapar University Patiala, / India.,*

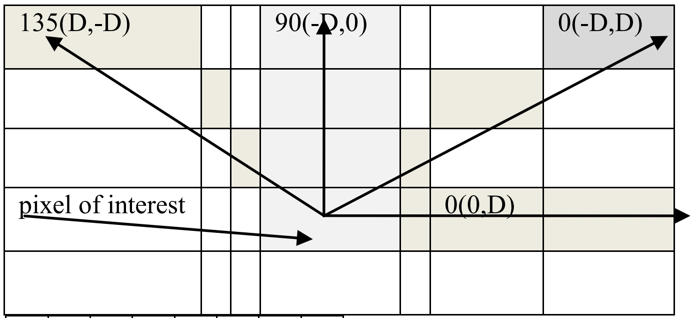
a) bottom left to top right(-45)
b) top left to bottom right (-135)

They are denoted as P0,P45,P90 and P135 respectively.P0(i,j)-Its direction is horizontal(0,D) means i and j are gray level values of the image.It depends on resolution of the image(image consists of 8 gray tones or 256)[1]

A. GLCM CALCULATION:

GLCM

1	2	0	0	1	0	0	0
0	0	1	0	1	0	0	0
0	0	0	0	1	0	0	0
0	0	0	0	1	0	0	0
1	0	0	0	0	1	2	0
0	0	0	0	0	0	0	1
2	0	0	0	0	0		0
0	0	0	0	1	0	0	0



SUB-REGION/IMAGE

1	1	5	6	8
2	3	5	7	1
4	5	1	1	2
8	5	1	2	5

Firstly take angle at 0 degree(horizontal).Here from the image ,the values at 0 degree angle .In the GLCM output,the element(1,1) has value1 because in the input image, there is only 1 occurrence where two horizontally near to pixels of distance 1 having values 1 and1.GLCM(1,2) has value 2 because in the input image there are two occurrences where two horizontally near to pixels of distance 1 having value 1 and 2. GLCM(1,3) has value 0 because in the input image there are no occurrence where two horizontally near to pixels of distance 1 having value 1 and 3.the proceure is repeated for the whole GLCM matrix at different angles.[2],[6]

ISSN: 2278 – 909X

B . Properties of GLCM

The properties of GLCM are:

1. GLCM is of square in shape because the reference and neighbouring pixels have same range of values.
2. Number of rows and columns equal to the quantization level of the image.
3. The test image consists of four gray level values that is 0,1,2 and 3.Eight bit data consists 256(2^8) possible values,256 X 256 matrix would be obtained,65536 cells.16 bit data having matrix of 65536 X 65536,having cells 429,496,720.
4. It is symmetrical about the diagonal.
5. The diagonal elements pairs having no gray level difference(0-0,1-1,2-2,3-3etc).Most pixels are identical to their neighbouring cells,very less contrast is there in the image.If there is a difference of 1 cell away from the diagonal,one level gray difference is there(0-1,1-2,1-3 etc).More the distance from the diagonal,more the gray level difference.

C . GLCM Features

The texture measures according to the weight of the equation.The texture is classified according to the degree.Square term means second order equation it is,Cube term means third order equation it is.[6,7]The features extracted are:

1. Contrast: In short form, it is called CON.'Sum of Square Variance' is the second name of Contrast.It submits the calculation of the intensity contrast linking pixel and its neighbor over the whole image.For constant image contrast value is 0.In contrast measure,weight increases exponentially(0,1,4,9) as proceed away from the diagonal.Range=[0,size(GLCM,1)-1)^2] [1],[2],[6],[7]

$$\sum_{i,j=0}^{N-1} Pi,j\,(i-j)^{\wedge}2$$

As (i-j) increases contrast continue to increase exponentiallyWhen i and j are equal i.e. i-j=0.When i and j are differ by 1,small contrast is there is 1.when i and j differ by 2, the contrast is increasing and weight is 4.[1],[2],[6],[7]

For horizontal GLCM,

Contrast is=contrast weight X horizontal GLCM.

Correlation:It gives the calculation of the correlation of a pixel and its neighbor over the whole image means it calculates the linear dependency of gray levels on those of neighbouring pixels..For a perfectly positively or negative correlated image,the correlation value is 1 and -1.For constant image its value is NaN..Range=[-1,1] and the formula is

$$\sum_{i,j=0}^{N-1} Pi,j\,\frac{(i-\mu i)(j-\mu j)}{\sqrt{(\sigma i^2)(\sigma j^{\wedge}2)}}$$

2. Energy:As we know that energy is used for doing work,hence orderliness.It is used for the texture that calculates orders in an image.It gives the sum of square elements in GLCM.It is completely different from entropy.When the window is perfect orderly,energy value is high .The square root of ASM(Angular Second Moment) texture feature is used as Energy.Its range is[0 1].For constant image its value is 1.[1],[2],[6],[7].The equation of energy is

$$\sum_{I,J=0} P(i,j)^{\wedge}2$$

3. Homogeneity:In short term it is called HOM.It givesthe value that calculates the closeness of distribution of the elements in the GLCM to the GLCM diagonal.For diagonal GLCM its value is 1 and its range is [0,1].Inverse of contrast weight is homogeneity weight values,with weight decreases exponentially away from the diagonal.The weight used in contrast is (i-j)^2 and in homogeneity ,it is 1/1+(i-j)^2.[1],[2],[6],[7].The equation is

$$\sum_{i,j=0}^{N-1} P(i,j)\,/R$$

Here R=1/1+(i-j)^2.Here four features are extracted i.e. Contrast,Correlation,Energy and Homogeneity.[2],[6],[7]The feature values of original image and its corresponding image having cancer are compared and result is shown in Table 1.

II PROCEDURE

1.The original image has been taken(Digital database for screening mammography/The Mammographic Image Analysis Society).

2.The DDSM/MIAS image is converted into Gray Scale

3.By using cropping method ,remove the noise like artifacts.

4.Filtering is done.

5.Thresholding is done on filtered image.

6.Segmenttaion is done after thresholding of the image.[4],[5]

Here PN stands for Patient.The results by using digital image processing techniques are shown below:

A. DATABASE MAMMOGRAM IMAGE(PN1)

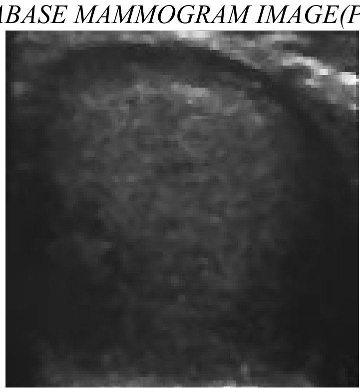


Fig 1(a)

ISSN: 2278 – 909X

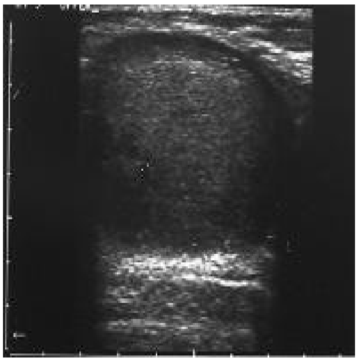


Fig 1(b)

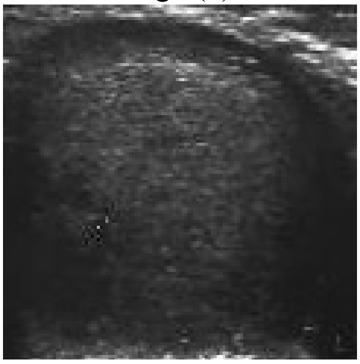


Fig 1(c)

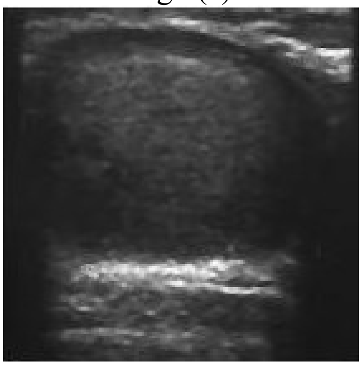


Fig 1(d)

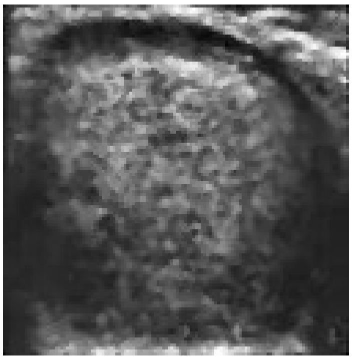


Fig 1(e)

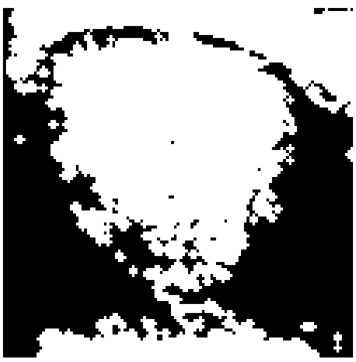


Fig 1(f)

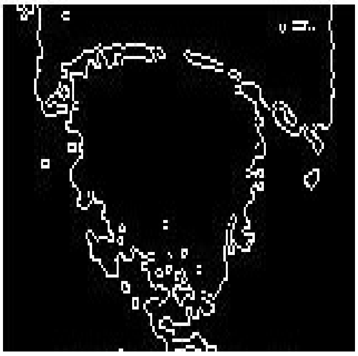


Fig 1(g)

Fig 1(a,b,c,d,e.f.g) shows the original, gray scale, cropped, filtered, contrast and cancerous image respectively.

B. DATABASE MAMMOGRAM IMAGE(PN2)

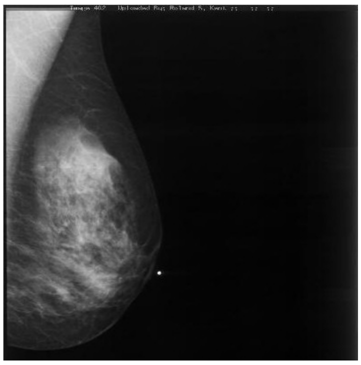


Fig 2(a)

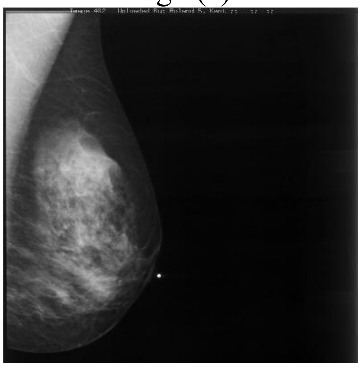


Fig 2(b)

Fig 2(c)

ISSN: 2278 – 909X

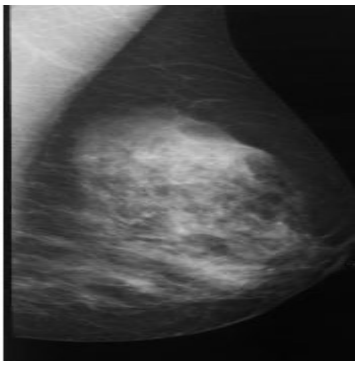


Fig 2(d)

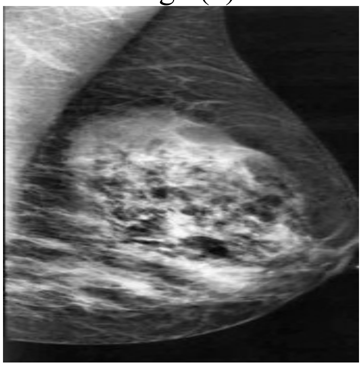


Fig 2(e)

Fig 2(f)

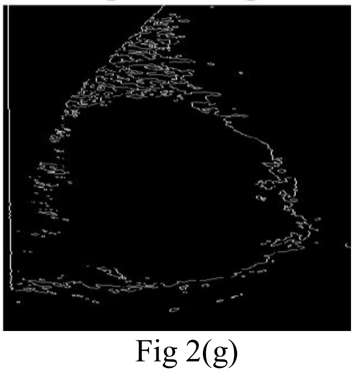


Fig 2(g)

Fig 2(a,b,c,d,e,f.g) shows the original, gray scale, cropped, filtered , contrast, cancerous and segmented image respectively.

C. DATABASE MAMMOGRAM IMAGE(PN3)

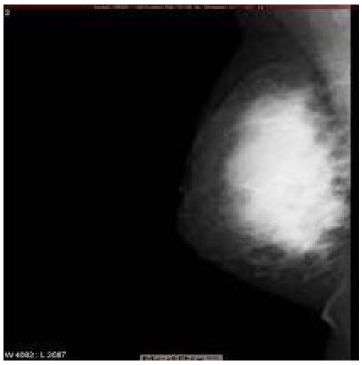


Fig3(a)

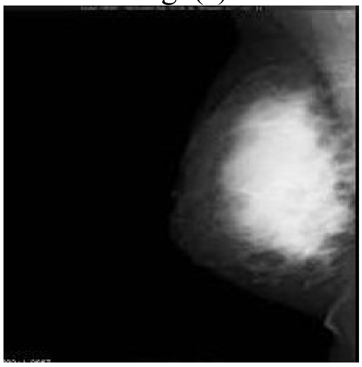


Fig3(b)

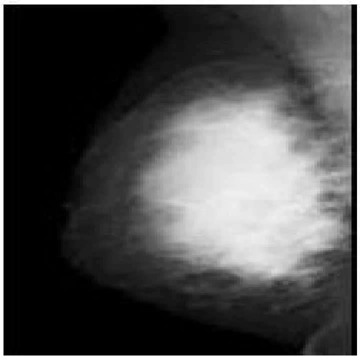


Fig3(c)

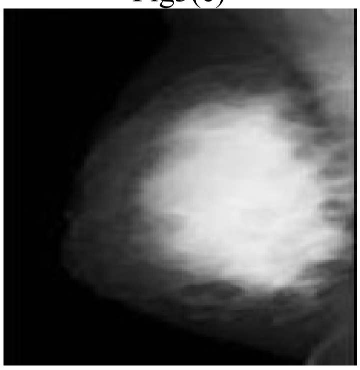


Fig3(d)

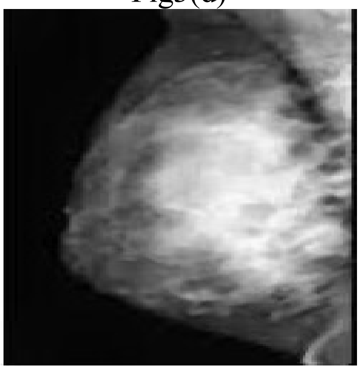


Fig3(e)

ISSN: 2278 – 909X

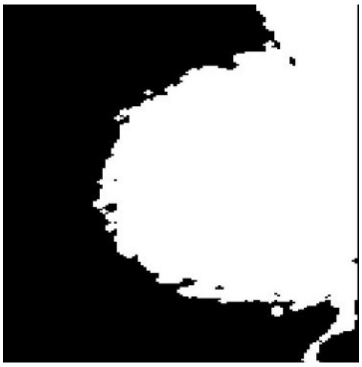


Fig3(f)

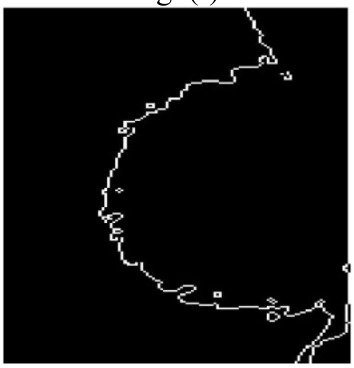


Fig3(g)

Fig3(a,b,c,d,e,f,g) shows the original,gray scale,croppes,filtered , contrast, cancerous and segmented image respectively.

D. DATABASE MAMMOGRAM IMAGE(PN4)

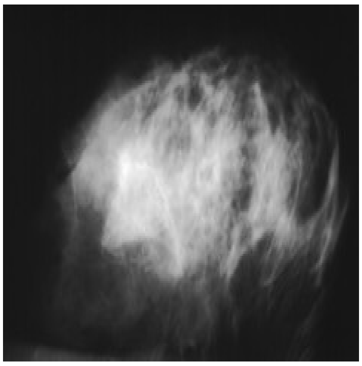


Fig4(a)

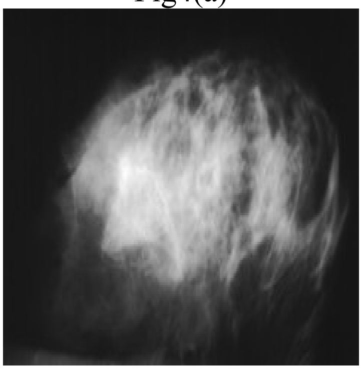


Fig4(b)

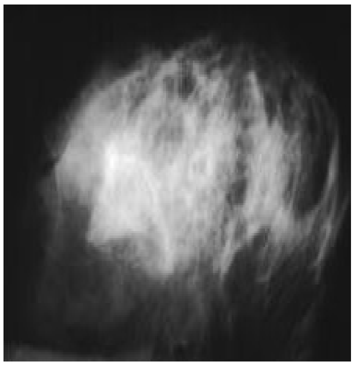


Fig4(c)

Fig4(d)

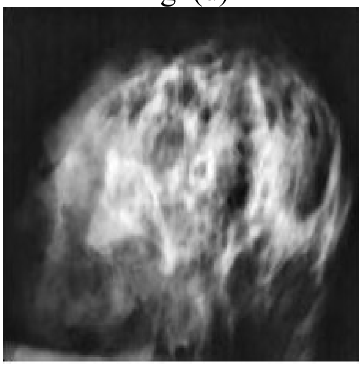


Fig4(e)

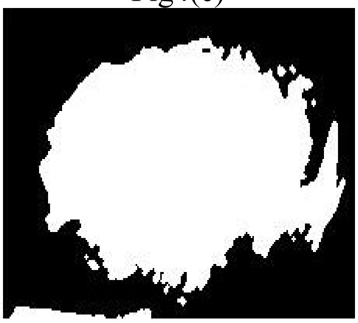


Fig4(f)

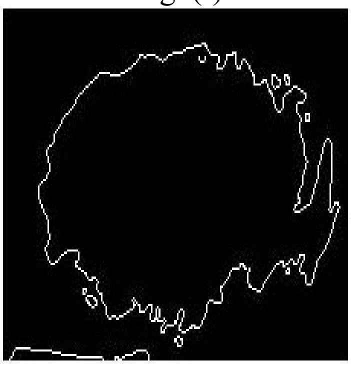


Fig4(g)

ISSN: 2278 – 909X

Fig4(a,b,c,d,e,f,g) shows the original,gray scale,cropped,filtering,contrast,cancerous image and segmented image respectively.

Table 1: GLCM Texture Features Values

Texture Features	PN1 OIV	PN1 CIV	PN2 OIV	PN2 CIV	PN3 OIV	PN3 CIV	PN4 OIV	PN4 CIV
Offset[0 1]								
Contrast	0.4112	0.4859	.1200	.7626	.1093	1.0467	.0692	.9315
Correlation	0.9494	0.842	.9685	.9588	.9829	.9569	.9867	.9620
Energy	0.2473	0.9275	.4099	.4865	.4499	.4835	.3747	.4816
Homogeneity	0.9423	0.9913	.9756	.9864	.9717	.9813	.9654	.9834
Offset[-1 1]								
Contrast	0.983	0.9856	.1832	1.0501	.1422	1.1304	.0845	1.1065
Correlation	0.8727	0.7004	.9519	.9570	.9778	.9535	.9838	.9548
Energy	0.2318	0.9176	.4056	.4809	.4452	.4816	.3690	.4781
Homogeneity	0.8724	0.9833	.9665	.9812	.9627	.9798	.9580	.9802
Offset[-1 0]								
Contrast	0.6646	0.4983	.1111	.7403	.0730	.5996	.0631	.9404
Correlation	0.9155	0.8781	.9709	.9697	.9886	.9753	.9879	.9616
Energy	0.2353	0.4065	.4092	.4868	.4547	.4927	.3756	.4815
Homogeneity	0.886	0.9911	.9763	.9868	.9752	.9893	.9684	.9832
Offset[-1-1]								
Contrast	0.9919	1.0906	.1794	.9676	.1428	1.2481	.0995	1.3849
Correlation	0.8716	0.7375	.9529	.9604	.9778	.9486	.9809	.9435
Energy	0.2316	0.8935	.4059	.4825	.4455	.4793	.3652	.4727
Homogeneity	0.8716	0.9805	.9680	.9827	.9626	.9777	.9507	.9753

OIV-Original Image Value
CIV-Cancerous Image Value
PN-Patient

III CONCLUSION

It was found that the values of features(contrast,energy and homogeneity) of original image are less than their corresponding cancerous images but only correlation feature value of original image is more as compared to corresponding cancerous image.(Table 1).Benign and malignant cancer which is detected using Digital image processing techniques is shown in Fig1,Fig2,Fig3,Fig4.

REFERENCES

[1]http://www.fp.ucalgary.ca/mhallbey/what_is_texture.html
[2]http://matlab.izmiran.ru/help/toolbox/images/enhanc15.html
[3] Blaine Martinez," Digital Image Processing Final Project" and Nicholas Sia Pik Kong, Theam Foo Ng ,"Simple Adaptive Median Filter for the Removal of Impulse Noise from Highly Corrupted Images HaidiIbrahim", IEEE Transactions on Consumer Electronics, Vol 54, No.4 ,Novermber 2008
[4] Robert M. Haralick ,K. Shanmugam and Its'hak Dinstein,"Textural Features for Image Classification", IEEE Transcations on systems,man and cybernetics,vol.SMC-3 pp. 610-621,6-Nov 1973.
[5] Pawar B.V., Patil Sharma,"Early Stage Detection ofTumors in Mammograms",Electronic Dept.,W.C.E. Sangli.
[6] Aswini Kumar Mohanty, Swapnastikta Beberta,S aroj Kumar Lenka," Classifying Benign and Malignant Mass using GLCM and GLRLM based Features from Mammogram", International Journal of Engineering Research and Applications,col.1,pp.6876-693
[7] P. Mohanaiah, P. Sathyanarayana, L. Guru Kumar," Image Texture Feature Extraction Using GLCM Approach", International Journal of Scientific and Research Publications, Volume 3, Issue 5, May 2013.

ISSN: 2278 – 909X