THE IIOAB JOURNAL

ARTICLE OPEN ACCESS

DETECTION OF CALCIFICATION IN MAMMOGRAM USING NEAREST NEIGHBOUR ALGORITHM

Usha[1*] and Arumugam[2]

[1]*PARK College of Engineering and Technology, Coimbatore, TN, INDIA*

[2]*Dept of Computer Science and Engineering, Nandha Engineering College, Erode, TN, INDIA*

ABSTRACT

Breast cancer is a dangerous and it increases the death rate among women cancer detection in early stage is not an easy task. The reason of the cancer is uncontrollable cells growth. In this paper an automatic mammogram classification techniques using symlet wavelet, gabor filter and nearest neighbour algorithm are used for getting better result.

Published on: 10th– August-2016

KEY WORDS

Breast Cancer, Digital mammogram Wavelets, gabor, Feature Selection, nearest neighbour algorithm.

***Corresponding author: Email:** usha.samiappan@gmail.com, arumugamdote@yahoo.co.in

INTRODUCTION

Memography is a useful method for early breast carcinomas detection [1].Diagnosis of Breast cancer using xray mammography is very good method. But many studies reveal that radiologists can be done misdetection of abnormalities in addition to having higher rates of false positive.75% is the estimated sensitivity of radiologist in Breast cancer screening. To avoid this double reading was suggested to be effective to improve sensitivity but it is costly and it takes time to read [2].cost effectiveness is important for mass screening programs to succeed. The diagnosis may be affected due to human factor of the abnormality indicators usually varied in shape, size and brightness. Difficult task to find out tumours and ordinary calcification [3,4]. In this paper, an automatic mammogram classification techniques using wavelets, gabor,feature selection is used to know the relationship between the features and classifier used to classify benign and malignant.

METHODOLOGY

MIAS is a digital mammogram database where films from U.K.'s National Breast Screening Programme have been digitized to 50 micron pixel edge with a Joyce-Loebl scanning microdensitometer, a linear device in a 0-3.2 optical density range representing pixels with an 8-bit word [5]. The database includes 322 digitized films and radiologist "truth"-markings on locations of abnormalities detected. The database was reduced to 200 micron pixel edge with padding/clipping to ensure that all images are 1024x1024 pixels at 8 bits per pixel. Erosion followed by dilatation has a similar structuring element, completing the opening function.

The MIAS database [6] though no longer supported, is old used much in literature. MIAS annotations are insufficient for some studies as all circumscribed/speculated lesions are to bemanually segmented [7]. Another drawback is its digitized resolution which renders it unsuitable for micro-calcification detection experiments. Regarding calcifications, healthier tissue is found in the ground truth region which is justified through calcifications shape as the latter are small lesions spread over a large area with all this being included in the annotation. Cross validation ensures higher formalism in entry data division considered necessary due to limited images with calcifications available in the database. The database Mini MIAS, prevents excessive network training and so a better system generalization.

THE IIOAB JOURNAL
ISSN: 0976-3104

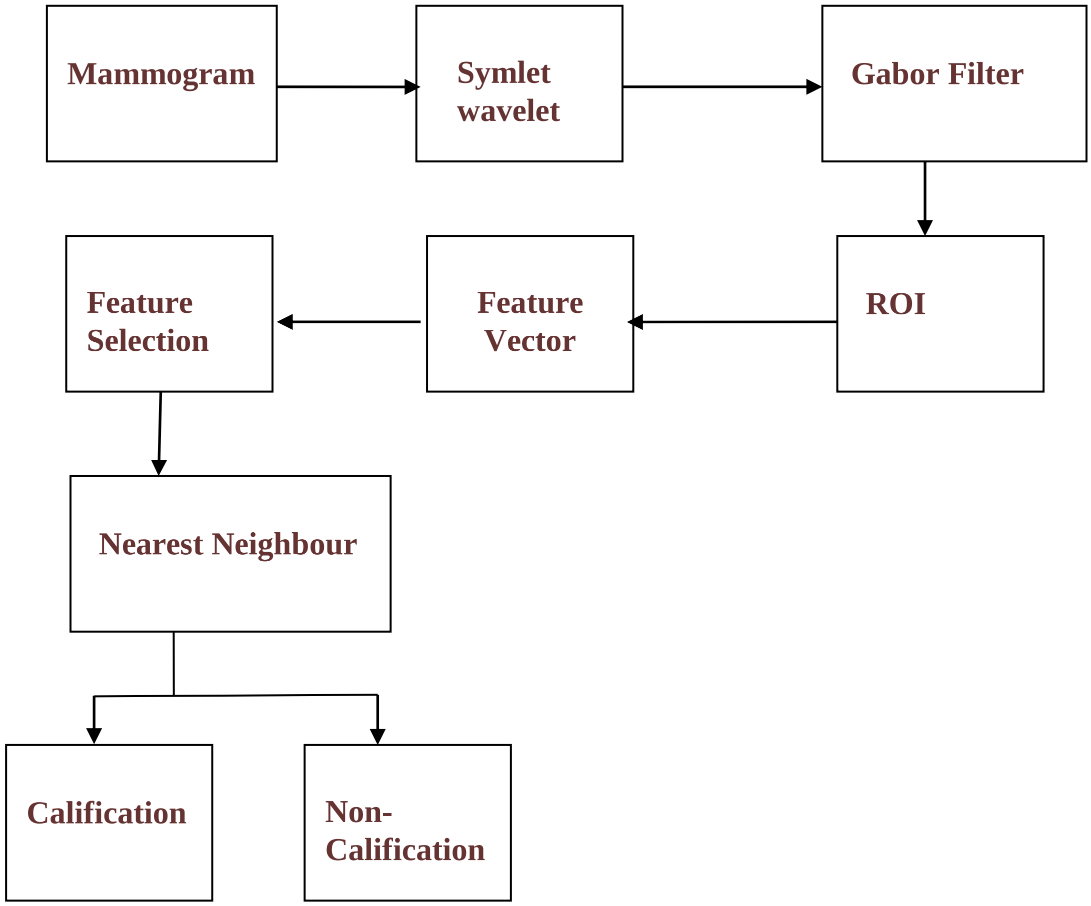

Fig: 1. Proposed system for calcification of Mammogram

....

ANALYSIS USING MULTI RESOLUTION APPROACH

Multi resolution is a scaling function.It is used to create a series of approximation of a function or image[8]. Each differing by a factor of 2 in resolution from its nearest neighbouring approximation [Gonalez et al 2004]. In this system image scaling is used in order to reduce computing time. The wavelet transform have been used for micro calcification detection using symlet wavelet, a function of a continous variable into a sequence of coefficient like sequence of number usually called as detail and approximation [9]. The wavelets are functions and those functions are basis for other functions called mother wavelet. The series of function generated by tranlation and dilation of mother function.

The four different coefficient are produced in each level of decomposition. The decomposition takes place while applying 2D wavelet transform on image. Those are horizontal, vertical,diaganol and approximation coefficient.DWT analysis the image by decomposing into coarse approximation via lowpass filtering and detailed information via highpass filtering. The decomposition is performed recursively on lowpass approximation coefficient at each level until the necessary iteration are reached.Micro calcification appear in the mammogram image as fine and bright grains in the tissue.so the detailed coefficient having micro calcification in the mammogram image using wavelet decomposition. In this research three level discrete wavelet decomposition by using symmetric daubechies of order 2[10]. The preprocessing time can be reduced using region of interest(ROI). The ROI extraction ignored the dark areas at particular frequency and orientation gabor filter can be viewed as a sinusoidal plane[11]. Symlet wavelet along with gabor filter for detecting mass easily. Gabor filter bank applied in different frequency and orientation on HH high frequency subband image obtained using symlet and extracted statistical features like standard deviation and mean in the form of feature vector.

ROI selected the centers of the abnormality. After getting the feature vector, the feature can be selected using searching method[12]. The number of features obtained after applying the feature selection method[13]. The classifier used to classify the mass into calcification andnon-calcification. The classification can be done using Euclidian distance as a measurement between the coefficient[14][15]. In a class,there is a N of images. These images are used to create class core vector and the core vector calculated by

THE IIOAB JOURNAL ISSN: 0976-3104

$$\mathrm{CCV}^{\mathrm{i}} = \frac{1}{n}\sum_{j=1}^{j=n} \mathrm{CCVj}^{\mathrm{i}}$$

$$\text{Where } \mathrm{j} = 1,2,3, \ldots\ldots \mathrm{m}$$

The Euclidian distance can be calculated by

$$D = \sqrt{\sum_{i=1}^{m} (CCV^{\mathrm{i}} - V_{Test}^{i})^2}$$

It is used to calculate the distance between the tested image and class core vector.

DISTRIBUTION OF THE MIAS DATA BASE

In this research MIAS (Mammographic Image Analysis Society) data set is used. This data set was investigated & labeled by the experts. This data set having 322 mammogram image of right and left breast. From 161 patients, 51 diagnosed as maligant, 64 has benign and 207 as normal. This result shown in **[Table -1]**.

Table: 1. Distribution of the Mias data base

Cases	B	M
Circ	19	04
Ill	06	08
Spic	11	08
Arch	09	10
Ass	06	09
Norm	-----	----

Abnormality Class
Norm - normal tissue
Calc - microcalcification clusters
Circ - circumscribed masses
Ill – ill-defined masses
Spic – spiculated lesions
Arch – architectural
Asym – asymmentry
Type of cancer
B – benign
M – malignant

RESULT AND ANALYSIS

To distingush between the types of tumors based on the physical properties and level of risk. The six abnormality cases are used as important classs. Those are microcalcification, circumscribed masses, ill-defined masses, spiculated lesions, architectural, asymmentry. Next,classifying wheather those cases are benign or malignant tumors. The following table shows the classification rate with accuracy based on 10 fold cross validation. In each fold,the average rate calculated and total average of 10 fold can be calculated (Table-1).

Table-1: Successful rate of calcification and non-calcification.

Class	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	AVG
B	100	90.20	84.33	92.35	80.20	94.3	85.60	90.00	86.78	93.50	89.72
M	100	100	98.50	98.40	98.04	96.20	100	97.30	100	96.30	98.47
Avg	100	95.1	91.41	95.37	89.12	95.25	92.8	93.65	93.39	94.9	94.09

B : Bennign, M : Maligant, F: Fold

COMPUTER SCIENCE

CONCLUSION

The diagnosis of breast cancer in a digital mammogram is a practical field of investigation. In this study the concept of using wavelet coefficients and gabor cofficients are used to form future vector and multi resolution analysis used in future extraction. Experiment which applied on real data shows the above results. The accuracy rate of classification achieved to distinguish between bennign ang maligant is 94.09 %. The results shows a special concentration on the wavelet coefficient that gives high percentage of success to find the tumour in each class.

CONFLICT OF INTEREST
The authors declare no conflict of interests.

ACKNOWLEDGEMENT
None

FINANCIAL DISCLOSURE
The authors report no financial interests or potential conflicts of interest.

REFERENCES

[1] Curtis C, Shah SP, et al. [2012] The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*, 486(7403): 346-352.

[2] Tang J, Rangayyan RM, Xu J, El Naqa I, Yang Y. [2009]Computer-aided detection and diagnosis of breast cancer with mammography: recent advances. Information Technology in Biomedicine, *IEEE Transactions on*, 13(2):236-251.

[3] Virnig BA, Tuttle TM, Shamliyan T, Kane RL. [2010] Ductal carcinoma in situ of the breast: a systematic review of incidence, treatment, and outcomes. *Journal of the National Cancer Institute*, 102(3): 170-178.

[4] Wasif N, Maggard MA, Ko CY, Giuliano AE. [2010] Invasive lobular vs. ductal breast cancer: a stage-matched comparison of outcomes. *Annals of surgical oncology*, 17(7):1862-1869.

[5] Suckling J et al. [1994] The mammographic image analysis society digital mammogramdatabase, *International Congress Series* 1069:375–378.

[6] HD Cheng, X cia, X chen, LH lou. C.[2003]omputer-aided detection and classification of microcalcification in mammogram: a survey. *Pattern Recognition Letters* 36:2967-2991.

[7] Dean J, Ilvento V.[2006] Improved cancer detection using computer-aided detection with diagnostic and screening mammography: Prospective study of 104 cancers. *American Journal of Roentology*,187:20-28.

[8] S Liu, CF Babbs, E Delp.[2001]Multiresolution detection of spiculated lesions in digital mammograms, *IEEE Transactions on Image Processing* 10 (6):874–884.

[9] EA Rashed, IA, Ismail SI Zaki. [2007] Multiresolution mammogram analysis in multilevel decomposition. *Pattern Recognition Letters* 28:286–292.

[10] Hajar Moradmand,Saeed Setayeshi,Ali Reza Karimian,Mehri Sirous,Mohammad Esmaeil Akbari"Comparing the performance of image Enhancement methods to detect microcalcification clusters in digital Mammography", *Spring* ,5(2)

[11] Yufeng Zheng[2010] Breast cancer detection with gabor features from digital mammograms. *Algorithms*, ,3:44-62.

[12] Wei CH, Li Y, Li CT. [2007] Effective extraction of Gabor features for adaptive mammogram retrieval. In Multimedia and Expo, 2007 IEEE International Conference on (pp. 1503-1506). *IEEE*

[13] Mohammed Meselhy Eltoukhy,Ibrahima Faye,Brahim Belhaouari Samir"Curvelet based feature extraction method for breast cancer diagnosis in digital mammogram.

[14] Min-Ling Zhang; National Lab. for Novel Software Technol., Nanjing Univ., China ; Zhi-Hua Zhou.A k-nearest neighbor based algorithm for multi-label classification.

[15] Keller JM. Gray, M.R.; Givens, J.A.. A fuzzy K-nearest neighbor algorithm

COMPUTER SCIENCE