Early Detection of Breast Cancer Using Grey Level Co-Occurrence Matrix Features

P.Sathia [1], K.K.Kavitha [2]

Research Scholar, Dept. of Computer Science, Selvamm Arts & Science College, Tamilnadu, India[1]

HOD & Vice Principal, Dept. of Computer Science, Selvamm Arts & Science College (Autonomous),

Tamilnadu, India [2]

ABSTRACT: Breast cancer is that the most typical malignancy of ladies and cause several deaths. The first detection is that the best resolution so as to avoid excision cut back the possibilities to come back, and reduce the speed of death. This can be through with the assistance of CAD system. CAD system has been developed to assist radiotherapist for detection cancer in girls breast. CAD system examines mammographic pictures of a breast. The aim of this paper is to examine the accuracy of GLCM options and improve system's accuracy victimization totally different options of GLCM that area unit being employed less range of times. This can be done by victimization GLCM (Grey level co-occurrence matrix) for extracting options and Neural Network for coaching and testing. The developed techniques are enforced in MATLAB and tested on real case studies.

KEYWORDS: Gray-level Co-Occurrence Matrix, Graylevel Run Length Matrix, mammograms, benign mass, malignant mass, texture features, textures analysis, association rule mining, Receiver operating characteristics.

I. INTRODUCTION

Breast cancer is one amongst the frequent and leading reason for mortality among lady, especial in developed countries. Age is one amongst the danger issue of carcinoma. Ladies at intervals the age of 40-69 have additional risk of carcinoma. In western countries concerning 53%-92% of the population has this sickness. During a Phillipine study a X-ray picture screening was done to 151,198 women. Out of that 3479 ladies had this sickness and was referred for designation. Though carcinoma results in death, early detection can increase the survival rate. The present diagnostic methodology for early detection of carcinoma is diagnostic technique. Mammographies square measure low dose X-ray projections of the breast, and it's the most effective methodology for detecting cancer at the first stage. Micro-calcifications (MC) square measure quiet small bits of metal, and will show up in clusters or in patterns and square measure related to additional cell activity in breast tissue. typically the additional cell growth isn't cancerous, however typically tight clusters of micro-calcification will indicate early breast cancer. Scattered micro-calcifications square measure typically an indication of benign carcinoma. 80% of the megahertz is benign. megahertz within the breast shows up as white speckles on breast X-rays. The   calcifications area unit small; sometimes varied from a hundred micrometer to three hundred micrometer, but in reality could also be as giant as 2mm. It's terribly tough to observe the calcifications per se, when over ten calcifications area unit clustered along, it becomes potential to diagnose malignant illness. However the survival depends on however early the cancer is detected. So, any MC formation ought to be detected at the benign stage. Hence, a pc motor-assisted designation (CAD) system is employed to observe megacycle per second clusters.

II. EXISTING SYSTEM

Mammogram–breast x-ray is taken into account the foremost effective, low cost, and reliable methodology in early detection of carcinoma. Though general rules for the differentiation between benign and malignant breast lesion exist, solely fifteen to half-hour of plenty referred for surgical diagnostic test are literally malignant. During this work, Associate in Nursing approach is planned to develop a computer-aided system for cancer detection from digital

mammograms. The planned system consists of 3 major steps. The primary step is region of interest (ROI) extraction of 256×256 pixels size. The second step is that the feature extraction; we have a tendency to used a collection of nineteen GLCM Associate in Nursing GLRLM options and therefore the nineteen (nineteen) options extracted from gray level run-length matrix and grey level co-occurrence matrix may distinctive malignant plenty from benign mass with an accuracy ninety 4.9%. Further analysis dispensed by involving solely twelve of the nineteen options extracted, that consists of five options extracted from GLCM matrix and seven options extracted from GLRL matrix.

DRAWBACKS OF EXISTING SYSTEM

- Texture options supported GLCM are often accustomed distinguish between malignant plenty and benign plenty on roentgenogram pictures, with accuracy levels above texture options supported GLRLM, however still below texture options supported combined GLRLM and GLCM.
- The main drawback facing extracting the visual options of the mammographic pictures is noise, the various resolution, quality and therefore the weak distinction of the mammograms.

III. PROPOSED SYSTEM

With the target of obtaining additional accuracy in benign and malignant tissues we offer this work to investigate the performance of GLCM (textural feature extraction) with RBFNN classifier. Numerous techniques are used for computing textural options. GLCM is proven to be a decent method for extracting textural options from numerous pictures. Grey level element distribution delineate by statistics like chance of 2 pixels having specific grey level at specific abstraction relationships. This abstraction info is provided as 2 dimensional grey level matrices. Radial Basis perform Neural Network (RBFNN) could be a variety of computing (AI) system is employed for tissue classification . This work is projected for analyzing the accuracy of GLCM options by passing these options for the classification of carcinoma from mammographic pictures as either benign or malignant.

ADVANTAGES OF PROPOSED SYSTEM

- The classification rates obtained make sure the superior performances obtained by the planned secret writing methodology. Being freelance on the window size of research, this secret writing has the advantage of reducing the quantity of grey level per the quantity of grey level contained within the block of research chosen.
- The level of prediction accuracy with texture options supported GLCM is 0.995, indicates that the prediction accuracy may be classified as excellent.
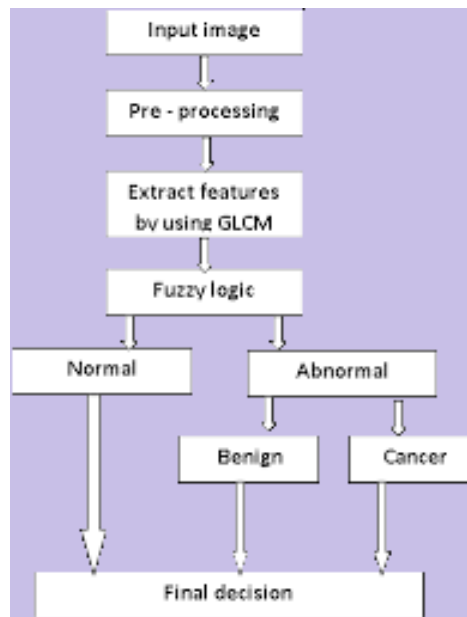
SYSTEM ARCHITECTURE

Fig 1: SYSTEM ARCHITECTURE

IV. IMPLEMENTATION

PRE-PROCESSING

In pre-processing step we tend to improve the standard of the image by removing noise. The pre- process is vital to eliminate some parts that don't seem to be connecting to the region of interest. So, the second step of pre-processing is removing the background space, removing the muscular pectorals and rib portion from the breast space as a result of we tend to ar victimization mammograms of MLO read. and therefore the last is segmentation to extract ROIs containing all lots and find the suspicious mass from the ROI. Segmentation of the suspicious regions on a mammographic image is intended to own a awfully high sensitivity and an oversized range of false positives are acceptable. This step proves to be a vital success think about the CAD system and helps in suppressing distortions and neglecting those elements that don't seem to be a part of breast and this ensures an improved classification accuracy of CAD formula.

FEATURE EXTRACTION

Feature Extraction is most significant step, uses for extracting price of different-different options and plays a key role in pattern classification. Image options will be either general options, like extraction of color, texture and form or domain specific options. X-ray photograph pictures encompass heterogeneous data that shows differing kinds of tissues, blood vessels, organ ducts, breast edges and X-ray photograph machine characteristics. Thus to own higher approach for detection traditional and abnormal tissues, we've got to settle on such a system which provides higher result with additional accuracy. Here we tend to are victimization GLCM methodology for extracting the textural options.

GLCM (GREY LEVEL
CO-OCCURRENCE MATRIX)

The GLCM is table illustration of however the no of times numerous mixtures of picture element brightness values (grey levels) occur in a picture. Counting on the quantity of pixels or dots in every combination, Feature extraction supported GLCM is that the second order statistics that may be accustomed analyzing image as a texture. In keeping with the quantity of intensity points (pixels) in every combination, statistics are classified into first-order, second order and higher-order statistics. Harlick has outlined fourteen options , in most of the CAD system solely five or seven options are used. Wherever victimization five options system provides ninety six accuracy of fifty check cases and victimization seven options system provides ninety 5.50% accuracy of 112 check cases [9, 10]. Within the projected work, we tend to use eight options particularly distinction, Correlation, Autocorrelation, add Entropy, Variance, system of measurement of Correlation, Inverse distinction Momentum, distinction Entropy to live the accuracy and different performance factors.

CLASSIFICATION

In projected CAD technique we tend to use artificial neural network for the classification of carcinoma in benign and malignant. As a result of machine vision perception there's continually a retardant in detection of lots in digital diagnostic technique. This drawback is solved to a good extent during this CAD, victimization Artificial Neural Network. Neural network classification consists of two processes: Training and Testing. Neural network is that the best tool in pattern classification application. The classifier is trained and so tested on X-ray photograph image. The classification accuracy depends on coaching as a result of it generates output supported its past expertise of coaching once associate unknown input is being passed.

PERFORMANCE ANALYSIS

Performance of CAD system is analyzed in terms of Accuracy, Sensitivity and Specificity. Wherever accuracy measures the quality of the binary classification. It takes under consideration of 2 categories these are true and false positives and negatives. Accuracy is usually calculable with balanced live. Sensitivity/Precision deals with solely positive cases and specificity deals with solely negative cases.

V. CONCLUSION

We developed a replacement methodology for the detection of micro-calcifications supported cryptography of unsmooth mammographic pictures. This cryptography consists in assignment to every peal a code and not a grey level. Through this methodology, we tend to found that cryptography preserves the structure of the pictures whereas reducing complexness because of high grey levels values. The classification rates based victimisation the GCLM applied to the coded pictures area unit more than those given by the initial pictures and even far better than the results obtained by classification performed on coded pictures by rank cryptography (82.6%). additionally, the creation of the coded pictures makes it attainable to decrease considerably the amount of the values related to every pel that is reduced from 256 to 64 at the most. This yields a way easier implementation of the strategy and reduces computation complexness. The projected methodology of cryptography the image is also applied to any unsmooth image before any process so as to cut back info knowledge to be analyzed.

VI. FUTURE WORK

1. In extraction of image object options, we'll contemplate a lot of feature vectors to attain a lot of wealthy multi-dimensional multi-level association mining therefore on increase classification exactitude additional.
2. The task of object segmentation in Blob world system leaves space for improvement. We have a tendency to shall integrate different simpler ways for up additional the accuracy. Therefore, we are able to increase classification exactitude by complete objects segmentation.
3. Since the planned multi-dimensional multi-level association mining relies on Apriori, it incurs high execution value. We'll do additional improvement on the rule improve the execution performance.

REFERENCES

[1]   Majid AS, de Paredes ES, Doherty RD, Sharma N Salvador X. "Missed breast carcinoma: pitfalls and Pearls". Radiographics 23(2003)881-895.

[2]   Osmar R. Zaïane,M-L. Antonie, A. Coman "Mammography Classification by Association Rulebased Classifier," MDM/KDD2002 International Workshop on Multimedia Data Mining with (ACM SIGKDD 2002, Edmonton, Alberta, Canada, 17-19 July 2002,), pp.62-69.

[3]   Xie Xuanyang, Gong Yuchang, Wan Shouhong, Li Xi ,"Computer Aided Detection of SARS Based on Radiographs Data Mining ", Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference Shanghai, China, September 1-4, 2005

[4]   Shuyan Wang, Mingquan Zhou and Guohua Geng, "Application of Fuzzy Cluster analysis for Medical Image Data Mining" Proceedings of the IEEE International Conference on Mechatronics & Automation Niagara Falls, Canada • July 2005.

[5]   R.Jensen, Qiang Shen, "Semantics Preserving Dimensionality Reduction: Rough and Fuzzy-Rough Based Approaches", IEEE Transactions on Knowledge and Data Engineering, 16 (2004) 1457-1471.

[6]   Walid Erray, and Hakim Hacid, "A New Cost Sensitive Decision Tree Method Application for Mammograms Classification" IJCSNS International Journal of Computer Science and Network Security, 6 (2006) No.11.

[7]   Ying Liu, Dengsheng Zhang, Guojun Lu, "Region based image retrieval with high-level semantics using decision tree learning", Pattern Recognition, 41 (2008) 2554 – 2570

[8]   Kemal Polat , Salih Gu¨nes, A novel hybrid intelligent method based on C4.5 decision tree classifier and oneagainst- all approach for multi-class classification problems, Expert Systems with Applications,

[9]   C.Chen and G.Lee, "Image segmentation using multitiresolution wavelet analysis and Expectation Maximum(EM) algorithm for mammography" , International Journal of Imaging System and Technology, 8(5):491-504,1997

[10] T.Wang and N.Karayaiannis, "Detection of microcalcification in digital mammograms using wavelets", IEEE Trans. Medical Imaging, 17(4):498- 509,1998

[11] I.Christiyanni et al ., "Fast detection of masses in computer aided mammography", IEEE Signal processing Magazine, Pages:54- 64,2000.

[12] Deepa S. Deshpande "association rule mining based on image content" International Journal of Information Technology and Knowledge Management January-June 2011, Volume 4, No. 1, pp. 143-146

[13] L. Jaba Sheela & V.Shanthi "A Novel Texture Classification Procedure by using Association Rules "ITB J. ICT Vol. 2, No. 2, 2008,pp,103-114

[14] http://marathon.csee.usf.edu/Mammography/Database.h tml

[15] Ali Cherif chaabani, Atef boujelben, Adel mahfoudhI, Mohamed ABID, "An Automatic-Pre-processing Method For Mammographic Images ", JDCTA: International Journal of Digital Content Technology and its Applications, Vol. 4, No. 3, pp. 190 ~ 201, 2010

[16] M. Hanmandlu, V.K. Madasu and S. Vasikarla, "A fuzzy Approach to texture segmentation", Proc. International Conference on Information Technology: Coding and Computing, vol-1, pp.636-642, 2004.

[17] V. Chalana and Y. Kim, "A Methodology for Evaluation of Boundary algorithms on Medical Images," IEEE Trans. on Medical Imaging, vol. 16(5), pp. 642-52, 1997.

[18] W. A. Perkins, "Area segmentation of images using edge points." IEEE Trans. on Pattern Analysis & Machine Intelligence, vol. PAMI-2, pp. 8-15. 1980.