Application of GLCM technique on Mammograms for Early Detection of Breast Cancer

Jinan FadhilMahdi
Dept. of Physic, Foundation of technical education Bagdad, Iraq.
Jinan F @yahoo.com

Abstract

Breast cancer represents the second leading cause of cancer deaths in women today، and it is the most common type of cancer in women. It has become major health issue in the world. Mammography is the main test used for screening and early diagnosis. Early detection performed on X-ray mammography is the key to improve breast cancer prognosis. GLCM technique introduced by Haralick is mostly used in image processing to study the grey levels intensities present in the image, Haralick features also known as first order have been used for detection of malignant masses in breast tissue images, I have attempted firstly to obtain well known Haralick features as well as the second order GLCM features and have studied its effects in breast cancer image classification.

Keywords:-Breast Cancer, Mammogram, CLAHE, texturefeatures, GLC**M**

الخلاصة

باستخدام التفاصيل اللازمة نتنبئ بالأسباب الممكنة لحدوث سرطان الثدي فبالرغم من تقدم التقنيات للكشف عن سرطان الثدي مبكراً يبقى الماموكرام هو الاساس وخاصتاً في الدول طور التقدم حيث تكاليفه رخيصة وغير اقتحاميه تجعلها تستعمل بشكل واسع فعال تقنياً.

ودراستنا الحالية والدراسات التمهيدية المشار لها تجعل النتائج مؤثره وفعاله من عدة عوامل وتشمل نقاوة بكسل وكذالك خطـــوات حسابيه والاتجاه والمسافة في هذا النوع منGLCMS

الكلمات المفتاحية: سرطان الثدي، الماموجرام، CLAHE، texturefeatures، GLCM

1. Introduction

Breast cancer has emerged as the leading sites of cancer in India among males and females. It is the most common type and frequent form of cancer and one in 22 women in India is likely to suffer from breast cancer. This is the second main cause of cancer deaths among women. Recently, the estimated number of breast cancer cases for the years 2015 and 2020 will be 106,124 and 123,634 respectively, according to the National Cancer Registry Programme report of the Indian Council of Medical Research (ICMR). Detecting a cancer at an early stage can improve the cure rate from breast cancer (Ramnath *et al*., 2010).

The highest survival rates for breast cancer occur when it is detected in its earlier stages. Mammography is a special type of x-ray imaging used to create detailed images of the breast, and is the most widely used method for breast cancer detection in its early stages. Once a lump is discovered, mammography can be very useful in evaluating the lump to determine if it is cancerous.

Image feature extraction is important step in mammogram classification. These features are extracted using image processing techniques. Several features are extracted from digital mammograms including texture feature,position feature and shape feature etc. Textures are one? of the important features used for many applications. Texture features have been widely used in mammogram classification. The texture features are ability to distinguish between normal and abnormal pattern. Texture is an alteration and variation of surface of the image (NITHYA, 2011).

Some researches show that the better detection rate can be achieved by appropriate feature selection that must include in the system that may require the number of features. However, having more features increases the complexity and time used to analyse the mammogram.

2. Methodology

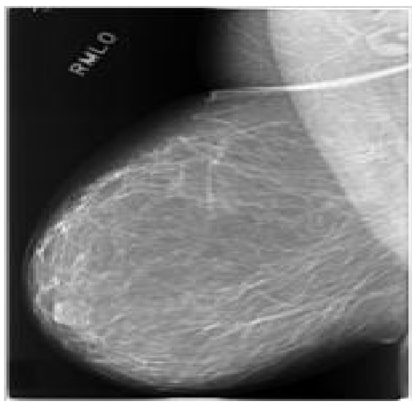
2.1 Pre-Processing

It is difficult for radiologists to identify the masses on a mammogram because they are surrounded by complicated tissues. In current breast cancer screening, radiologists often miss approximately 10-30% of tumors because of the ambiguous margins of lesions and visual fatigue resulting from long-time diagnosis.

The main objective of the pre-processing step is to eliminate those elements in the image that could negatively affect the process of microcalcification detection & make the feature extraction phase easier and reliable (Rolando,2012) and comprises the sub-steps of:

(1) Receiving the original images as input.

(2) Applying a CLAHE (Contrast Limited Adaptive Histogram Equalization).The contrast enhancement phase is done using the Contrast Limited Adaptive Histogram Equalization (CLAHE) technique, which is a special case of the histogram equalization technique that functions adaptively on the image to be enhanced.(Rolando,2012; Ramnath , 2010).



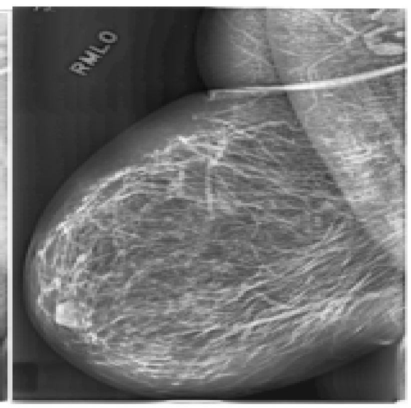


Fig.1 Original image Fig.2.CLAHE filtered image
(Ramnath , 2010;Ramnath, 2010)

CLAHE method seeks to reduce the noise and edge shadowing effect produced in homogeneous areas and was originally developed for medical imaging.

method seeks to reduce the noise produced in homogeneous areas and was originally developed for medical imaging. CLAHE operates on small regions in the image called tiles rather than the entire image. Each tile’s contrast is enhanced, so that the histogram of the output region approximately matches the uniform distribution or Rayleigh distribution or exponential distribution. Distribution is the desired histogram shape for the image tiles. The neighbouring tiles are then combined using bilinear interpolation to eliminate artificially induced boundaries. The contrast, especially in homogeneous areas, can be limited to avoid amplifying any noise that might be present in the image.

2.2 Feature Extraction

Texture analysis is important for application of computer image analysis for classification, detection segmentation of an image based on intensity and colour. Traditionally texture analysis can be broadly classified into two، they are statistical and structural approach. In structural approach the texture can be represented by primitive (micro texture) and spatial arrangement of this primitive (macro texture). In the statistical approach represents the texture indirectly by the non-deterministic properties that govern the distributions and relationships between the gray levels of an image. There have been eight statistical approaches are used for the measurement and characterization of texture analysis (Robert,1979).

In 1973, Haralick introduced 14 statistical features. These features are generated by calculating the features for each one of the co-occurrence matrices obtained by using the directions 0◦, 45◦, 90◦ , and 135◦, then averaging these four values. The symbol Δ, representing the distance parameter, can be selected as one or higher.

In general, Δ value is set to 1 as the distance parameter. A vector of these 14 statistical features is used for characterizing the co-occurrence matrix contents.

Texture analysis, using some or all of the 14 texture features proposed by Haralick *et al*. based upon gray-level co-occurrence matrices (GLCMs), is a popular approach for the analysis and classification of many medical images, including breast masses and tumors seen in mammograms (Sivaramakrishna, 2002; Bovis, 2000; Gupta, 2005).

GLCM is a statistical texture measure. GLCM collects information about pixel pairs, hence it is of second orderstatistics (Mohd ,2009; Sreeja and Ganesan,2010). GLCM is a tabulation of frequencies or how the pixel brightness values in an image occur. The matrix is constructed at a distance of d = 1 and for direction of θ given as 0°,45°,90° and 135° . A single direction might not give enough and reliable texture information. For this reason, the four directions are used to extract the texture information.

GLCMs are constructed by comparing the gray-level values for two pixels with a defined spatial relation. Combination frequencies of occurrence are calculated for each possible gray-level value. The frequencies in the GLCMs are used to calculate 13 different features like energy, entropy, contrast, local homogeneity,correlation, and shade provenance, sum of squares, sum average, sum entropy, difference entropy, sum variance anddifference variance. (Karahaliou, 2007; Al Mutaz , 2011)

Gray level co-occurrence matrices capture properties of a texture but they are not directly useful for further analysis, such as the comparison of two textures. The co-occurrence matrix computation applied to a mammogram is shown in figure 1.

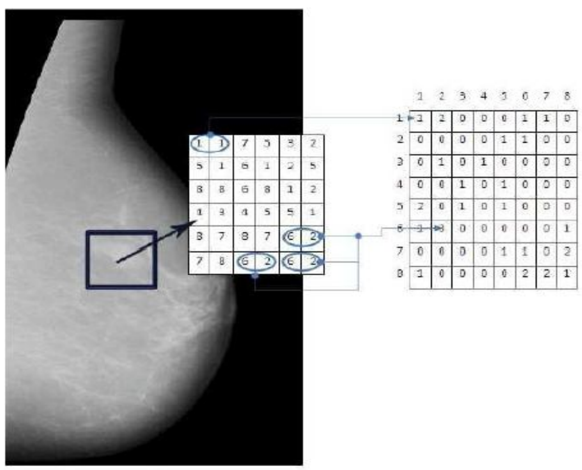


Fig. 2: Example of a co-occurrence matrix with eight grey levels, computed using one for distance between pixels and zero degrees for the direction, applied to a mammogram

Computer image processing techniques are used to enhance the image which is followed by segmentation of Region of Interest. Then the texture feature will be extracted from ROI. The texture feature is used to classify the ROI as masses or non-masses. The extraction of texture feature can be done by using grey level co-occurrence matrices.

Experimental Results

The proposed features mentioned in previous section was implemented and applied on real breast cancer mammography images selected from MIAS database .After we calculated GLCM features from images, we save the features value in database presented in Table 2.

The results indicate that benign masses and malignant tumors are not linearly separable, in the space defined bythe 14 texture features, to a high level of classification accuracy.The dataset used in the present study was prepared by combining digitized film mammographic images from MIAS dataset.

Table 1. MIAS Dataset

Class	Benign	Malignant	Total
Microcalcification	12	13	25
Circumscribed masses	19	4	23
Ill-defined masses	7	7	14
Spiculated masses	11	8	19
Architectural distortion	9	10	19
Asymmetry lesion	6	9	15
Normal tissue	-	-	207
Total	64	51	322

Table 2**.** Texture features extracted from mammogram

Mean ;Mean is the mean pixel value within the image. It is important to know the brightness of image since the tumor has high value of brightness.

Uniformity : Uniformity measures the uniformity of intensity in the histogram of a ROI. Tumor has positive low value of uniformity.

Standard Deviation : Standard Deviation is the root mean square (RMS) deviation of the values from their arithmetic mean. It is the most common measure of statistical dispersion, measuring how widely spread the values are in a data set. If the data points are all close to the mean, then the standard deviation is close to zero. If many data points are far from the mean, then the standard deviation is far from zero. If all the data values are equal, then the standard deviation is zero. Tumor is far from zero.

Smoothness :Smoothness measures the relative smoothness of the intensity in a image. It used to determine the region whether it is smooth or not. When its value is near from 0 this region is smooth and when the region is more complex than smoothness value is near from 1. Tumor is smooth, if the ROI value is near from 0, thus when the region is smooth, it could be a sancer.

Inverse: It gives high value when the high value of the entry of NH is near the main diagonal. Tumor has positive low value of inverse

Skewness :Skewness measures the asymmetry of the probability distribution of a real-valued random variable. Adistribution has positive skew (right-skewed) if the right (higher value) tail is longer or fatter and negative skew (left skewed)if the left (lower value) tail is longer or fatter. Tumor has left skewness, which takes direction to the bright gray level.

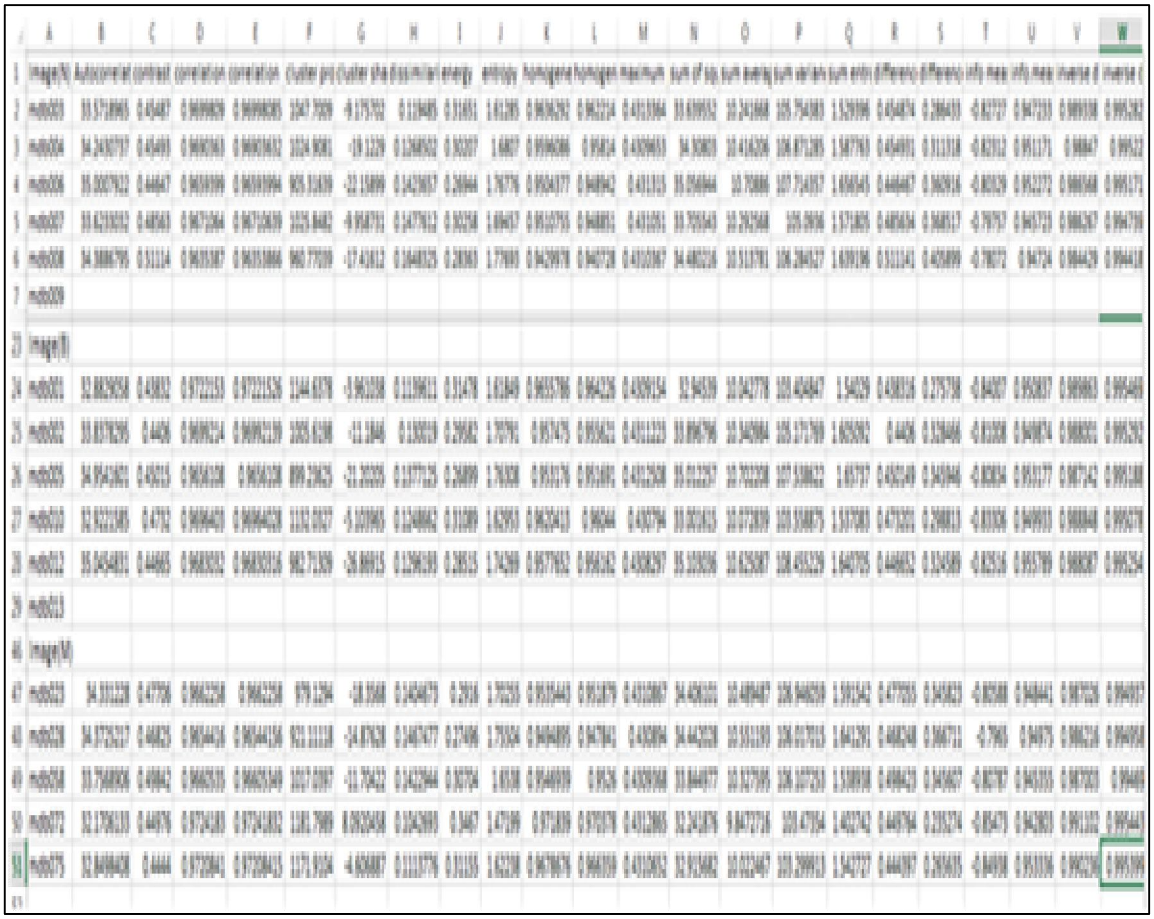
Entropy: Entropy represent the measure of disorder in an object gray level organization, large value of entropy correspond too much disorganized distribution, such as salt and paper random field. Low entropy images have very little contrast. Image that is perfect flat will have zero entropy. Tumor has negative low value of entropy.

Correlation: Correlation is the most famous statistical approach. It is popular today, by virtue of good performance. The Co-occurrence matrix contains elements that are counts of the number of pixel pairs for specific brightness levels, when separated by some distance and at some relative inclination . Moreover, correlation produces a large value if an object contains large connected subcomponents of constant gray level, and with large gray level differences is created between adjacent components. So the tumor has high value of correlation .

Conclusion:-

Using texture features to predict the risk of breast cancer appears feasible.Despite many advances in techniques for early breast cancer detection, mammography is still the main standard, especially in developing countries. Its low cost, and the fact that is not invasive, coupled with its high effectiveness make it a widely used technique.

Our present study and a related preliminary study indicate that the results are affected by several factors, including pixel resolution, preprocessing and filtering steps used, and the direction and distance used to compute the GLCMs.



References

Al Mutaz M. Abdalla1, Safaai Dress2, NazarZaki 2011. “Detection of Masses in Digital Mammogram Using Second Order Statistics and Artificial Neural Network”,International Journal of Computer Science & Information Technology (IJCSIT), Vol 3, No 3, June

Bovis K, Singh S: 2000 Detection of masses in mammograms using texture features. Proceedings of the 15th InternationalConference on Pattern Recognition; Sept 3–7, 2:267–270,

Gupta S, Markey MK: 2005. Correspondence in texture features between two mammographic views. Med Phys 36(6):1598–1606,

Karahaliou, A. ,Boniatis, I., Sakellaropoulos, P., S.Skiadopoulos, Panayiotakis, G., L. Costaridou” 2007 Can texture of tissue surrounding micro calcifications in mammography be used for breast cancer diagnosis?"Nuclear Instruments and Methods in Physics Research A 580 1071–1074

Mohd. A, Khuzi1R Besar, WMD Wan Zaki, NN Ahmad, 2009. "Identification of masses in digital mammogram using gray level co-occurrence matrices" Biomedical Imaging and Intervention Journal,

Nithya, R., Santhi, B. " comparative study on feature extraction method for breast cancer classification" 2011. journal of theoretical and applied information technology 30th November Vol. 33 No.2

RamnathTakiar, DeenuNadayil, A Nandakumar2010. "Projections of Number of Cancer Cases in India (2010-2020) by Cancer Groups", Asian Pacific Journal of Cancer Prevention, Vol 11,

Robert M. Hawlick, Senior Member, 1979 "Statistical and Structural Approaches to Texture" Procceedings of IEEE VOL. 67, NO. 5, MAY

Rolando-Rafael Hernández-Cisneros 2012 "METHOD FOR THE DETECTION AND CLASSIFICATION OF MICROCALCIFICATION CLUSTERS IN DIGITAL MAMMOGRAMS" Publication date: 2012-05-17 ,Patent application number: 0121159

Sivaramakrishna R, Powel KA, Lieber ML, 2002 ChilcoteWA,Shekhar R: Texture analysis of lesions in breast ultrasoundimages. Comput Med ImagGrap 26:303–307,

Sreeja, S.S Mole, Dr.L.Ganesan, 2010 "Unsupervised Hybrid Classification for Texture Analysis Using Fixed and Optimal Window Size", (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 09,2910-2915,