On mammograms and CAD for breast cancer

Mr S MOHAN KUMAR, Prof. Dr .G. BALAKRISHNAN ME., Ph.D.

A.P, SAMSCET & Research Scholar, Director, IGCE & Research Guide,
Karpagam University, Karpagam University,
Coimbatore, Tamil Nadu, India. Coimbatore, Tamil Nadu, India.
E-Mail: mohankumar.sugumar@gmail.com,
mohankumar@janhoo.com

Abstract: Breast cancer is the most common type of cancer in women with approximately one in nine women developing the disease in her life time. Breast cancer is a disease in which cancerous cells are found in the tissues of the breast. These cancerous cells divide and grow continuously and eventually form a lump known as tumor. Although breast is the leading causes of cancer deaths in women, the causes of breast cancer are unknown. However, heredity does play a vital role in the development of breast cancer. The DNA in your cells carry the genetic information that you receive from your parents. Mammography is a specific type of imaging that uses a low dose x-ray for examination of the breast. The images can be viewed on a film at a view box. Most experts agree that successful treatment of breast cancer often is linked to early diagnosis. Mammography plays a major part in early detection of breast cancer, because it can show changes in the breast up to two years before a patient or a physician can feel them. In this work various aspects of mammography and computer aided detections have been analyzed.
[S MOHAN KUMAR, G. BALAKRISHNAN ME. **On mammograms and CAD for breast cancer.** *Cancer Biology* 2012;2(2):1-4]. (ISSN: 2150-1041). http://www.cancerbio.net. 1

Keywords: Breast; cancer; women; disease; life; time; cell; tissue; vital role; genetic information; x-ray; examination; diagnosis; patient; physician; mammography

Introduction

Numerous fields, which range from military to commercial applications, are in need of paced and efficacious analysis of multimedia data. In the area of multimedia data, images get the place of a fortress. A plenty of knowledge can be hidden in image data and the knowledge discovery in image databases is known as image mining [3]. The tasks of image mining are mostly concerned with classification problems such as "labeling" regions of an image based on presence or absence of some characteristic patterns, and with image retrieval problems where "similar" images are identified [1]. Image mining is more than just an extension of data mining to the image domain [2]. Applications of the technology include, beyond remote sensing, geographic information systems, medical imaging, geomarketing, navigation, traffic control, environmental studies, and many other areas where spatial data are used [4].

Besides the other applications, medical imaging plays a significant role in the human's life. In medical imaging, abnormality detection in images is predicted to play an important role in many real-life applications [5]. In fact, the New England Journal of Medicine regards medical imaging as one of the most important medical developments of the past thousand years, basically du e to the fact that it provides physicians with physiology and functionality of organs and cells inside human bodies. The medical image is utilized to detect the abnormalities of tissues; the one among them is the breast cancer tissue which is the major death cause of women. Among different imaging modalities used for breast cancer detection, mammography remains the key screening tool for the detection of breast abnormalities [6]. Breast cancer is considered as one of the primary causes of women mortality [7] [10]. The mortality rate in asymptotic women can be brought down with the aid of premature diagnosis. Despite the increasing number of cancers being diagnosed, the death rate has been reduced remarkably in the past decade due to the screening programs [8].

Premature detection of breast cancer increases the prospect of survival whereas delayed diagnosis frequently confronts the patient to an unrecoverable stage and results in death [9]. The most effective way to reduce breast cancer deaths is detect it earlier. However, earlier treatment requires the ability to detect breast cancer in early stages. Early diagnosis requires an accurate and reliable diagnosis procedure that allows physicians to distinguish benign breast tumors from malignant ones [11]. Mammography is the most contemporary option for the premature detection of breast cancer in women [6].

Mammography [13], [14] is the single most effective, reliable, low cost and highly sensitive method for early detection of breast cancerous. Mammography offers high quality images at low radiation doses and is the only widely accepted imaging metho d for routine breast cancer screening. It is reco mmended that women at the ages of 40 or above should have a mammogram every one to two years. [12]. Detecting breast cancer in mammograms is challenging because the cancerous structures have many features in common with normal breast tissue [15].

Review on Related Researches

Some of the recent research works are briefly reviewed here. **Muthu Rama Krishnan et al.** [16] have aimed at designing a support vector machine (SVM)-based classifier for breast cancer detection with higher degree of accuracy. It has been introduced a best possible training scheme of the features ex tracted from the mammogram, by first selecting the kernel function and then choosing a suitable training-test partition. Prior to classification, detailed statistical analysis viz., test of significance, density estimation have been performed for identifying discriminating power of the features in between malign ant and benign classes. A comparative stud y has been performed in respect to diagnostic measures viz., confusion matrix, sensitivity and specificity. They have considered two data sets from UCI machine learning database having nine and ten dimensional feature spaces for classification. In addition, the ov erall classification accur acy obtained by using the proposed classification strategy is 99.385% f or dataset-I and 93.726% for dataset-II,respectively.

Subashini et al. [17] have compared the use of polynomial kernel of SVM and RBFNN in ascertaining the diagnostic accuracy of cytological data obtained from the Wisconsin breast cancer database. The data set has been included nine different attributes and two categories of tumors namely benign and malignant. Known sets of cytologically proven tumor data was used to train the models to categorize cancer patients according to their diagnosis. Performan ce measures such as accuracy, specificity, sensitivity, F-score and other metrics used in medical diagnosis such as Youden's index and discriminant power were evaluated to convey and compare the qualities of the classifiers. Their research has demonstrated that RBFNN outperformed the polynomial kernel of SVM for correctly classifying the tumors.

Jihene Malek e al. [18] have proposed a design of automated detection, segmentation, and classification of breast cancer nuclei using a fuzzy logic. The first step was based on segmentation using an active contour for cell tracking and isolating of the nucleus in the cytological image. Some of the textural features have been extracted from this nucleus, using the wavelet transforms to characterize image using its texture, so that malign texture could be differentiated from benign on e with the assumption that tumoral texture was different from the texture of other kinds of tissues. Finally, the obtained features will be introduced as the input vector of a fuzzy C -means (FCM) clustering algorithm to classify the images into malign and benign ones. The implementation of such algorithm has been done using a methodology based on very high speed integrated circuit, hardware description language (VHDL). The design of the circuit has been performed b y using a CMOS 0.35 μm technology.

Mehmet Faith Akay [19] has been proposed breast cancer diagnosis based on a SVM-based method combined with feature selection. Experiments have been conducted on different training-test partitions of the Wisconsin breast cancer dataset (WBCD), which has been commonly used among researchers who used machine learning methods for breast cancer diagnosis. The performance of the method has been evaluated using classification accuracy, sensitivity, specificity, positive and negative predictive values, receiver operating characteristic (ROC) curves and confusion matrix. The results have been showed that the highest classification accuracy (99.51%) has been obtained for the SVM model that contains five features, and that has been very promising when compared to the previously reported results.

Lukasz Jelen et al. [20] have presented a framework for automatic malignancy grading of fine needle aspiration biopsy tissue. The malignan cy grade was one of the most important factors taken into consideration during the prediction of cancer behavior after the treatment. Their framework was based on a classification using Support Vector Machines (SVM). The SVMs were able to assign a malignancy grade based on pre-extracted features with the accuracy up to 94.24%. They have also showed that the SVMs performed best out of four tested classifiers.

Inspiration of the Research

Recently, the image processing is an upcoming research field for the researchers. The image mining is used to extract the images from the large image database and it has been utilized in many fields. Apart from other applications, the medical image mining plays a vital role in human's life, where the abnormality has been detected. The medical image mining is utilized for the detection of a plenty of diseases. The breast cancer is one of the diseases, which has been detected with the aid of medical image mining. The breast cancer is the primary cau se for the women mortality.

The earlier detection of the breast can cer may diminish the mortality rate. For the earlier detection of breast cancer, the mammography is used. Mammography is supposed to minimize mortality from breast cancer. Early detection of breast cancer is the main objective of mammography and in general this is achieved through detection of characteristic masses and/or microcalcifications. Yet, detecting breast cancer in mammograms seems to be on demand since the cancerous structures posses numerous characteristics in common with normal breast tissue. The deadly nature of the breast cancer and the difficulties in detecting the cancer in mammograms necessitates an efficient technique for the detection of breast cancer cells. This has motivated to do the research wo rk in identifying an efficient technique.

The Anticipated Solution

The primary intent of my work is to detect the anomaly of the breast cancer tissue. To accomplish this, the digital mammography will be segmented using contour-based segmentation and the segmented mammography will be subjected under clustering so as to identify the cancer affected breast tissue locations. The cancer cells will be classified whether the cancer is benign or malign ant. To accomplish this, an extensive feature set will be extracted from the clustered breast cancer cells. The extensive feature set is comprised of shape, texture and gray intensity features. The shape features plays major role in classification of the cancer to benign or malignant. Owing to its foremost role in the identification and for the further differentiation of lesion throughout the diagnosis process, the texture feature shall thus be extracted. The ex tracted features will be used to train the Support Vector Machine (SVM). The SVM is chosen because of its capability of learning with very little samples. The well-trained SVM can effectively detect and classify the breast cancer, when a mammography is given. This paves the way for effective clinical diagnosis of the breast cancer.

References

[1] Manoranjan Dash and Deepak Kolippakkam, “Automatic View Selection: An Application to Image Mining”, Advances in Knowledge Discovery and Data Mining, p.p. 107-113, DOI 10.1007/b136725.

[2] S.P.Victor and S.John Peter, “A Novel Minimum Spanning Tree Based Clustering Algorithm for Image Mining”, European Journal of Scientific Research, Vol. 40, No, 4, pp.540-546, 2010, ISSN 1450-216X .

[3] Kun-Che Lu, Don-Lin Yang, and Ming-Chuan Hung, "Decision Trees Based Image Data Mining and Its Application on Image Segmentation", In proceedings of International Conference on Chinese Language Computing, pp.81-86, 2002.

[4] Han and Kamber, "Data mining - Concepts and techniques", pp.405-412, Morgan Kaufmann Publishers, San Diego, CA, 2001.

[5] Yu Xiaoxue, Wynne Hsu, Lee and Lozano -Perez, "Abnormality Detection in Retinal Images", MIT Dspace, 2003 .

[6] Editorial. Looking back on the millennium in medicine, New En gland journal of Medicine, Vol.342, No.1, pp.42-49, 2000 .

[7] R.A. Smith, "Epidemiology of breast cancer in a categorical course in physics”. Technical Aspects of Br east Imaging, 2nd ed., RSNA publication, Oak Book, II, pp.21, 1993.

[8] R. Peto, J. Boreham, M. Clarke, C. Davies., V. Beral, “UK and USA Breast cancer deaths down 25% in year 2000 at ages 20-69 years”, THE LANC ET, Vol.355, No. 9217, pp. 1822, May 2000.

[9] Ranadhir Ghosh, Moumita Ghosh, John Year wood, "A Modular Framework for Multi category feature selection in Digital mammography", In Proceedings of the 12th European Symposium On Artificial Neural Networks ESANN’2004, Bruges (Belgium), pp. 175-180, April 2004.

[10] Petroudi, Marias, English, Adams and Brad y, “Classification of mammogram patterns using area measurements and the Standard Mammogram Form (SMF)”, In proceedings of International Workshop on Digital Mammography, Lecture Notes in Computer Science, Springer Verlag Berlin Heidelberg, 2002.

[11] Emad A. El-Sebakhy, Kanaan Abed Faisal, Helmy, Azzedin and Al-Suhaim, "Evaluation of Breast Cancer Tumor Classification with Unconstrained Functional Networks Classifier", in proceedings of the 4th ACS/IEEE International Conference on Computer Systems and Applications (AICCSA-06), Dubai, USE, March 2006

[12] Amar Partap Singh Pharwaha and Baljit Singh, "Shannon and Non-Shannon Measures of Entropy for Statistical Texture Feature Ex traction in Digitized Mammograms", In proceedings of the World Congress on Engineering and Computer Science, Vol.2, October 2009 .

[13] Sheshadri and Kandaswamy, “Experimental investigation on breast tissue classification based on statistical feature extraction of mammogram,” Computerized medical imaging and graphics, vol. 31, pp. 46-48, 2007.

[14] Christoyianni, Koutras, Dermatas and Kokkinakis, “Computer aided diagnosis of breast cancer in digitized mammograms,” Computerized medical imaging and graphics, vol. 26, 2002, pp. 309-319.

[15] Dave Tahmoush, "Image Similarity to Improve the Classification of Breast Cancer Images", Algorithms, Vol.2,pp. 1503-1525, 2009.

[16] Muthu Rama Krishnan, Shuvo Banerjee, Chinmay Chakraborty, Chandan Chakraborty and Ajoy K. Ray, "Statistical analysis of mammographic

features and its classification using support vector machine", Expert Systems with Applications, Vol.37, No.1,pp.470-478, January 2010.
[17] Subashini, Ramalingama and Palanivela, "Breast mass classification based on cytological patterns using RBFNN and SVM ", Expert Systems with Applications, Vol.36, No. 3, pp.5284-5290, April 2009.
[18] Jihene Malek, Abderrahim Sebri, Souhir Mabrouk, Kholdoun Torki and Rached Tourki, "Automated Breast Cancer Diagnosis Based on GVF-Snake Segmentation, Wavelet Features Extraction and Fuzzy Classification", Journal of Signal Processing Systems, Vol.55, No.3, pp.49-66, April 2009.
[19] Mehmet Faith Akay, "Support vector machines combined with feature selection for breast cancer diagnosis", Expert Systems with Applications, Vol.36, No.2, pp.3240-3247, March 2009.
[20] Lukasz Jelen, Thomas Fevens and Adam Krzyzak, "Classification of Breast Cancer Malignancy Using Cytological Images of Fine Needle Aspiration Biopsies", Int. J.Appl.Math.Comput.Sci., Vol.18, No.1, pp.75-83, 2008.

4/14/2012