A Novel Approach for Detecting and Classifying Breast Cancer in Mammogram Images

S. Shanthi, Kongu Engineering College, Erode, Tamil Nadu, India
V. Murali Bhaskaran, Paavai College of Engineering, Namakkal, Tamil Nadu, India

ABSTRACT

This study uses data mining techniques for computer-aided diagnosis that involves the feature extraction for cancer detection, so as to help doctors towards making optimal decisions quickly and accurately. Features play an important role in detecting the cancer in the digital mammogram and feature extraction stage is the most vital and difficult stage. In this paper, an enhanced feature extraction method named Multiscale Surrounding Region Dependence Method (MSRDM) is proposed to be effective in classifying the mammogram images into normal or benign or malignant. This proposed system is based on a four-step procedure: Regions of Interest specification, two dimensional discrete wavelet transformation, and multiscale surrounding region dependence matrix computation and feature extraction. The performance of the proposed feature set is compared with the conventional texture-analysis methods such as gray level cooccurrence matrix features and surrounding region dependence method features. Experiments have been conducted on both real and benchmark data and the results have been proved to be progressive.

Keywords: Computer Aided Diagnosis, Mammogram, Multiscale Surrounding Region Dependence Method, Texture Features, Wavelet Transformation

1. INTRODUCTION

Advances in image acquisition and storage technology have led to tremendous growth in large and detailed text or image databases. A vast amount of medical data is daily generated in hospitals and medical centers which most of the times makes it difficult or even impossible for radiologists to discover the underlying knowledge and patterns of the data. To reduce such difficulty, computer assistance is needed for the radiologists especially in handling information overload.

Computer Aided Diagnosis (CAD) systems are being developed to automate the interpretation process of medical data. CAD is an interdisciplinary challenge that draws up on proficiency in computer vision, digital image processing, image extraction, data mining, image mining, machine learning, databases, and artificial intelligence.

In the recent years, CAD is widely used in medical decision support across a wide range of medical areas, such as cancer research (Hadjiiiski, Sahiner, Chan, Petrick, & Helvie, 1999; Mencattini, Salmeri, Rabottino, & Salicone,
Breast cancer is one of the significant public health problems in the world because it acts as the major cause of fatality among all cancers for women between the ages 35 and 55. Until now, there has not been a known way to prevent breast cancer but the earlier the cancer can be detected, the higher the chance of survival for patients. Digital mammography is one of the most suitable methods for early detection of breast cancer. Early detection plays an effective role in breast cancer treatment since the causes for cancer are still unknown. As the symptoms especially at the images taken in the early stage of breast cancer are very subtle and vary in appearance, it is very difficult for the radiologists to accurately identify the presence of breast cancer. So it has become necessary to go for CAD systems which use computer technologies to detect abnormalities such as calcifications, masses, architectural distortion, and bi-lateral symmetry in mammograms and use of this results by the radiologists can play a key role in the early detection of breast cancer (Tang, Rangayyan, Xu, Naqa, & Yang, 2009).

CAD system consists of four stages: pre-processing, identification of Region of Interest (ROI), feature extraction and classification. The efficiency and robustness of this CAD system is often largely determined by the intrinsic separability of the features available to it (Kim, Park, Song & Park, 1998). There have been great efforts that are spent on extracting the appropriate features from the ROI. Most systems extract features to detect and classify the abnormality into benign or malignant through textures, statistical properties, spatial domain, fractal domain and wavelet bases. Texture is a kind of very interesting and challenging property or feature of an image and it is characterized by a set of local statistical properties of pixel intensities. Texture analysis has been exploited extensively for various kinds of special purposes, such as for image mining, image classification, surface description, image matching and 3D-reconstruction and so on (Ali, 2007).

In mammographic image processing, texture features have been used to distinguish density patterns that indicate different levels of risk in malignant lesions. Texture analysis has been shown to be a promising technique in analyzing mammographic lesions caused by masses. Textural information is important for outlining the performance of CAD system required for the classification that distinguishes masses from normal tissues.

Similarly, multiresolution analysis has proved to be useful in mammographic image processing, image enhancement, mass detection, and feature extraction. The original image is decomposed into sub-bands that preserve high and low frequency information. Several studies have investigated the use of wavelet transform as a multiresolution analysis tool for texture analysis and classification (Tang, Liu & Sun, 2009; Tsai, Chen, & Hsu, 2011).

Initially, texture analysis was based on the first order or second order statistics of textures like Spatial Gray-Level Dependence Method (SGLDM) which is known as Gray Level Cooccurrence Matrix (GLCM) by Haralick, Shanmugan, and Dinstein (1973), Gray-Level Run-Length Method (GLRLM) by Galloway (1975), Gray-Level Difference Method (GLDM) by Weszka, Dyer and Rosenfeld (1976) and Surrounding Region-Dependence Method (SRDM) by Kim et al. (1998). All these features provide a mathematical description of a texture but it is often difficult to associate these features with the observed characteristics. In general this approach is easier to compute and is more widely used, since natural textures are made of patterns of irregular sub elements.

On the contrary Gaussian Markov Random Field (GMRF) and Gibbs random field models were proposed to characterize textures (Cross & Jain, 1983). Local linear transformations were also used to compute texture features (Unser,
17 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the product's webpage:
www.igi-global.com/article/novel-approach-detecting-classifying-breast/75544?camid=4v1

www.igi-global.com/e-resources/library-recommendation/?id=2

Related Content

Intuitionistic Fuzzy Time Series Forecasting Based on Dual Hesitant Fuzzy Set for Stock Market: DHFS-Based IFTS Model for Stock Market
www.igi-global.com/chapter/intuitionistic-fuzzy-time-series-forecasting-based-on-dual-hesitant-fuzzy-set-for-stock-market/208041?camid=4v1a

Complementary Part Detection and Reassembly of 3D Fragments
www.igi-global.com/chapter/complementary-part-detection-reassembly-fragments/64395?camid=4v1a

On Soft Graphs and Chained Soft Graphs
www.igi-global.com/article/on-soft-graphs-and-chained-soft-graphs/201560?camid=4v1a
Fuzzy Expert System to Diagnose Diabetes Using S Weights for S Fuzzy Assessment Methodology
www.igi-global.com/chapter/fuzzy-expert-system-to-diagnose-diabetes-using-s-weights-for-s-fuzzy-assessment-methodology/178406?camid=4v1a