Breast Abnormalities Detection in Digital Mammogram Using Fuzzy C-Means Clustering and Support Vector Machine-Matlab Implementation

Dr. D.Pugazhenthi
Asst. Prof, PG& Research Dept. of Computer Science, Quaid-E-Millath Govt. College for women (A), Chennai, India

I. Introduction
Breast cancer is the third causes of death for women around the world. The screening mammography is the effective way to finding the abnormalities in the breast. There are two cases of the most familiarsymptoms of breast abnormalities are masses and microcalcification. Computer-Aided Detection (CAD) increases the early detection of breast abnormalities, especially the women with dense breast masses.[1]The main improvement of a CAD system is that it can get some essential features, such as Computational, statistical features, which cannot be found visually and automatically by medical doctors and also reduces the operator-dependent nature in screening tool.[2]

Preprocessing
Digital Mammogram images are complex to process, the images are exposed to improvement and accurate extraction for additional segmentation, examination and classification are performed. [3] Preprocessing steps comprise of removal of unnecessary or irrelevant areas and to create a region of interest by enhancing the contrast manipulation. [4]

Post processing
At this level the pre-processed Digital mammogram image is segmented into the number of pixels of small matrices of 2x2, after which all pixels coordinate values of the mass are observed, and the value representing maximum amount in the block is allocated to all pixels of the block. [5]Based on the logic this value is distributed to rest of the pixels of the blocks. This ensures that the entire block consists of same pixel values. The equal blocks are combined to form a 4x4 pixel block based on the method to cover the entire digital mammogram image(DM).[6]
If the association between breast tissue density and breast risk is to be studied, a more accurate and objective method of assessing tissue density is needed. [7]In this work an attempt has been made to automate classification of breast density using statistical features. [8]

Abstract
Nowadays breast abnormalities are the most common malignancies among women in the world. The digital mammogram is the most common and familiar screening tools for breast which finding abnormal masses or lesions. Digital mammogram is the efficient and low energy x-ray for breast and helps to find the abnormalities which compare to the other screening tools. A CAD (Computer Aided Detection) system includes three main steps containing preprocessing, feature extraction, and classification. The preprocessing stage is the initial and crucial stage to remove all the noise in the digital mammogram. The feature extraction stage, extracting the statistical feature like entropy, standard deviation, etc. These features are then used in combination with SVM (Support Vector Machine) to detect the breast abnormalities. The aim of this paper predict the abnormal masses in digital mammogram, detect breast abnormalities and create awareness among the women about the risk of breast abnormalities.

Proposed methodology

A. Input image (Digital mammogram)

Preprocessing

Feature Extraction

Classification

Performance Evaluation

Fig 1: Proposed Methodology

A. Input image (Digital Mammogram)
Digital Mammography: The digital mammogram is the low cost, efficient and low energy x-ray for the breast. The Digital mammogram is the valuable screening tool and X-rays are the key tools used for capturing images and is utilized in the medical systems for mammographic images.[9] These images are specially recorded for examination the human breast. These mammographic images are useful to the radiologists in the investigation forthe breast.
B. Preprocessing

The preprocessing is the initial stage which distinguishes between normal and abnormal masses. Pre-processing is performed in the proposed method. The goal of pre-processing is to enhance the digital mammogram image data by suppressing the unwanted misrepresentations or enhances some important digital mammogram image features relevant for further process and analysis task. In this paper, Histogram manipulation is used as a preprocessing step. This is the efficient contrast enhancement method. The value in the low range of the original histogram is assigned to black and value at the high end is assigned to white. The features that were unclear on the original image will be apparent in the contrast stressed image. The pixel values are changed to new values after applying pre-processing.

C. Segmentation

Segmentation is a crucial phase of CAD systems. Both computerization and correctness of segmentation are fundamental. Computerization of separation of masses is significant because it enables the complete automation of the CAD system. A clustering method classifies data element into groups based on similarity measures. The Fuzzy c-Means clustering method is a clustering method where each data elements may belong to more than one group, where a probability scattering gives the degree of membership for each item over the clusters. It is useful when the required number of data elements clusters are pre-determined. Thus, the method efforts to put each of the data points to one of the clusters for segmenting the data elements.
D. Feature extraction

Feature extraction is a familiar part of the feature detection stage; it is very necessary to extract the information from mammographic images so that the system can correctly distinguish normal and abnormal tissues.

E. Classification

In this paper, SVM is used for classification of the extracted features. Online database of digital mammogram classifies mammographic images using two-way classifier abnormal and normal masses. The classification of normal and abnormal masses are based on the skewness of the digital mammogram.

\[
\text{Abnormal masses} = 1 \\
\text{Normal masses} = 0
\]

F. Performance Evaluation

Effective investigation of the proposed method is conducted using the online database for the digital mammogram. The proposed methodology accuracy are measures based on the feature which is extracted in the digital mammogram. The performance of the existing methods and the proposed method variation as shown in figure 12.

Conclusion

Many mammogram images of different abnormalities randomly chosen from the online database have been selected for experiment and the method applied over it have classified them in accordance with the category. Experiments results show us very clearly the presence of any abnormality in the growth of breast tissue. Also, this approach would simplify the time and computational difficulty in the examination of any given Digital mammogram image (DM).

In future other statistical methods of estimation of breast masses or lesion density like the use of Gaussian mixture model are also being considered for finding the breast cancer and reduces the unnecessary biopsy operations.

References

[3]. Senthilkumar, B., Umamaheswaru G, Kathie, J., “ A novel region growing segmentation algorithm for the detection of...
breast cancer’. In the Proceedings of the IEEE International Conference on Computational Intelligence and Computing Research, December 2010, Coimbatore, pp. 1-4.


