Impact Factor Value: 4.029 ISSN: 2349-7084
International Journal of Computer Engineering In Research Trends
Volume 4, Issue 12, December - 2017, pp. 552-555 www.ijcert.org

Cancer Detection in Mammograms by Extracting Geometry and Texture Features

[1]Pallavi P. Jadhav, [2]Prof. U. A. Nuli
D.K.T.E. Society's Textile And Engineering Institute, Ichalkaranji.

Abstract: - Breast cancer is one of the most frequently occurring diseases which cause death among women. Masses present in mammogram of breast, primarily indicates breast cancer and it is important to classify them as benign or malignant. Benign and malignant masses differ in geometry and texture characteristics. However, not every geometry and texture feature that is extracted contributes to the improvement of classification accuracy; thus, to select the best features from a set is important. Proposed new system will examine the feature selection methods for mass classification.

Index Terms— Breast cancer, mammograms, Region of Interest (ROI), Feature Extraction

1. Introduction

Cancer is an abnormal, continual multiplying of cells. The cells divide uncontrollably and may grow into adjacent tissue or spread to distant parts of body. Breast cancer remains a leading cause of cancer deaths among women. Early detection of breast cancer is necessary. Benign tumors are not cancerous. They may grow larger but do not spread to other parts of body. Malignant tumors are cancerous. They invade and destroy nearby tissue and spread to other parts of body. Tumor can be identified in mammogram because that part is highly bright (having high intensity) compared to other.

The proposed system intends to classify breast mass in mammographic image as benign or malignant. The system will start by applying image preprocessing techniques to improve image intensity so that features can be easily located and recognized. Segmentation is further performed to extract Region of Interest (ROIs). Then the geometric features or **statistical** features like concavity, concave points, fractal dimension, and other and textural features like correlation, auto correlation, contrast and other will be extracted. After extracting the features, only required features will be selected which are most important, because not all the features are going to help to improve the efficiency of the classifier. With the help of different feature selection algorithms, Like SVM/SVM-RFE/SRN. Finally, any of the classifiers like SVM, KNN, ANN or other can be used to classify the mass as benign or malignant.

- *Pallavi P. Jadhav is currently pursuing masters degree program in Computer Science and engineering department of DKTE's Textile and Engineering Institute,Ichalkaranji, India*
- *U. A. Nuli is currently working as Assitant Professor in Computer Science and engineering department of DKTE's Textile and Engineering Institute,Ichalkaranji, India*

2. Literature Review

In [1], the mass classification using gradients was investigated. In the proposed method, a mass boundary was segmented into concave and convex parts, and was polygonized. The features that quantify the extent of the sharpened nature of the boundary and the degree of narrowness of the spicules were extracted for mass classification.

In [2],[3]and [4], morphological features with manually delineated boundaries were adopted for mass classification. Pohlman et al. [4] used an adaptive region growing technique to segment masses. Six features measuring mass shape and boundary roughness were extracted and used for classification.

In [5], the mass classification based on the automated segmentation of masses was investigated. The results show that the features that are extracted from automated contours can benefit the diagnosis of breast masses. Sahiner et al. [6] utilized a rubber band straightening transform to characterize mammographic masses as malignant or benign, and Eltoukhy et al. [7] compared the wavelet and curvelet transforms for breast cancer diagnosis.

Investigation of feature selection is also important. Because not all features are useful for a typical classification problem, feature selection is very important.The use of the features selected by these methods can generally improve the classification accuracy. In this proposal, we will focus on the feature selection for mass classification.

Many methods for feature selection have been proposed [9]. One of the best methods, i.e., the support vector machine (SVM)-based recursive feature elimination (SVM-RFE), combines the RFE with an SVM classifier [10]. Although the SVM-RFE has many advantages, it also has a few disadvantages, such as the redundancy between the selected features and a high computing cost. Thus, a few improvements have been introduced [11], [12].

3. Problem Statement

Classify cancerous mass as malignant or benign by using selected geometry and textural features using feature selection method like SRN (SVM-RFE-NMIFS).

4. Proposed System

The proposed system architecture is as follows:

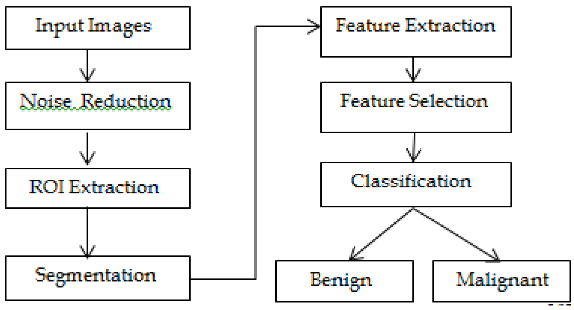

Fig -1 Block schematic of Proposed System

5. Input Images from Dataset

In this module we will take input images from Mammographic Image Analysis Society (MIAS) or Digital Database for Screening Mammography (DDSM). It contains large amount of cases of mammograms from several medical institutions. The original images are of 1024 x 1024 pixels in .pgm format (Portable Gray Map). These images contain some background noise, which needs to be removed.

Fig -2 Original Mammographic Image

6. Noise Reduction

The major limitations of mammograms are that they are low contrast, uneven illumination and presence of noise. Mammograms have existing artifacts like written labels which need to be eliminated. Due to this feature extraction and object detection in mammograms is challenging task. Image smoothing and contrast enhancement are two major steps of preprocessing. While preprocessing it is necessary to enlrge intensity difference between object and background and to reduce noise without destroying important features of the image.

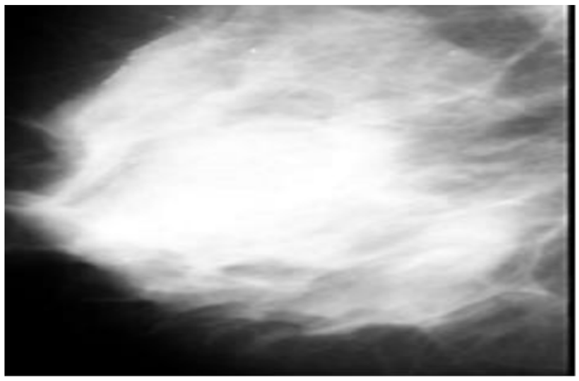

Fig- 3 Image after Noise Reduction

7. Region Of Interest (ROI) Extraction

The original images are very large and image of tumor in that mammogram is very tiny. So processing of the original image is very necessary. That is, the region where we are interested in is only used. Therefore a cropping operation is applied to the images to cut off the unwanted portion of the images. The ROIs that contain masses will be manually extracted from the original mammogram. The cropped ROI is of low contrast. The preprocessing phase of digital mammograms refers to enhancement of mammogram intensity and contrast manipulation.

8. Segmentation

This module aims in extracting images in order to search only the areas which are suspected to have abnormalities instead of searching all parts of images. Thus segmentation is the process where ROI is separated from background like pectoral tissues. Segmentation is partitioning of an image into homogeneous regions based on a set of characteristics. It extracts regions from images in order to search only the areas which are suspected to have abnormalities instead of searching all parts of image. It is to simplify or change the representation of an image into something that is more meaningful and easier to analyze. It is typically used to locate object and boundaries (lines, curves , etc). In this, a label is assigned to every pixel in an image such that pixels with the same label share certain characteristics.The fuzzy set theory has improved this process by allowing the concept of partial membership, in which an image pixel can belong to multiple clusters.

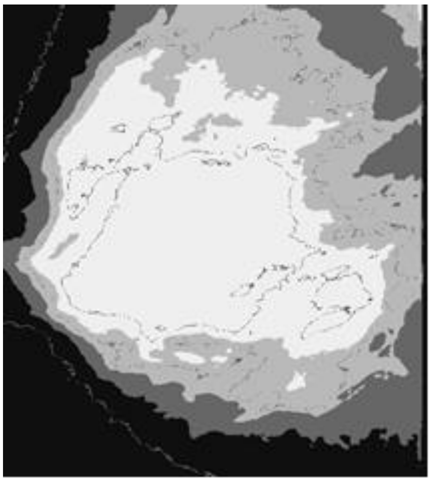

Fig - 4. Segmented image

9. Algorithm used for Segmentation

9.1 Fuzzy C-Means Clustering

Here the data point can belong to more than one clusters. With each of the data points, membership grades are associated which indicate the degree to which the data point belong to different clusters.

This algorithm tries to put each of the data point to one of the clusters. It does not decide the absolute membership of a data point to a given cluster, instead it calculates the likelihood. Each item may belong to more than one group g (hence the word fuzzy), where the degree of membership for each item is given by a probability distribution over the clusters.

Membership assignment to each data point corresponding to each cluster center on the basis of distance between the cluster center and the data point. More the data is near to the cluster center, more is its membership towards the particular cluster center. It compares the RGB value of every pixel with the value of the cluster center.

Instead of making a hard decision about which cluster the pixel should belong to, it assigns a value between 0 and 1 describing "how much this pixel belongs to that cluster" for each cluster. Fuzzy rule states that the sum of the membership value of a pixel to all clusters must be 1. The higher the membership value, the more likely that pixel is belonging to that cluster.

10. Feature extraction

After segmenting the mass from the ROI, we will compute a set of features that is related to the geometry and texture of the boundary and its neighbor regions. A typical benign mass has a round, smooth, and well-circumscribed boundary, whereas the boundary of a malignant tumor is usually speculated, rough, and blurry. Thus, we can use a boundary analysis to classify the masses into benign or malignant. We will then investigate both the geometry features and the texture features.

Geometry Features: The geometry features represent the shape of the contour of a mass. These are calculated from the boundary pixels after the segmentation. The geometry features can be any among compactness, Normalized Distance Moments, Fourier Features, Normalized Radial Length based features and relative gradient orientation based features.

Texture Features: In addition to the shape information of a mass contour, the texture information of the region surrounding the mass boundary also contains important information to discriminate the benign and malignant masses. Thus, we will also use the texture information for the mass classification. The texture features that can be extracted will be based on Gray level Co-occurrence Matrix.

11. Conclusion

Classification of benign and malignant tumor is important in initial treatment of the patient. The above discussion concludes initial preprocessing of the mammographic image is required which minimize the computational cost and improve the acuracy. After that different algorithms like FCM can be used for clustering. Geometric and texture features of the segmented image can give us mathematical data to accurately classify the image into benign and malignant.

References

[1] R. Rangayyan, N. Mudigonda, and J. Desautels, "Boundary modeling and shape analysis methods for classification of mammographic masses," Med. Biol. Eng. Comput., vol. 38, no. 5, pp. 487–496, 2000.R. Nicole, "The Last Word on Decision Theory," *J. Computer Vision*, submitted for publication. (Pending publication)

[2] N. Mudigonda, R. Rangayyan, and J. Desautels, "Gradient and texture analysis for the classification of mammographic masses," IEEE Trans.Med. Imag., vol. 19, no. 10, pp. 1032–1043, Oct. 2000.

[3] J. Kilday, F. Palmieri, and M. Fox, "Classifying mammo-

graphic lesionsusing computerized image analysis," IEEE Trans. Med. Imag., vol. 12,no. 4, pp. 664–669, Dec. 1993.S.P. Bingulac, "On the Compatibility of Adaptive Controllers," *Proc. Fourth Ann. Allerton Conf. Circuits and Systems Theory*, pp. 8-16, 1994. (Conference proceedings)

[4] S. Pohlman, K. Powell, N. Obuchowski, W. Chilcote, andS. Grundfest-Broniatowski, "Quantitative classification of breast tumors in digitized mammograms," Med. Phys., vol. 23, no. 8, pp. 1337–1345,Aug. 1996.

[5] A. Rojas Dominguez and A. Nandi, "Toward breast cancer diagnosis based on automated segmentation of masses in mammograms,"Pattern Recog., vol. 42, no. 6, pp. 1138–1148, Jun. 2009

[6] B. Sahiner, H. Chan, N. Petrick, M. Helvie, and M. Goodsitt, "Computerized characterization of masses on mammograms: The rubber bandstraightening transform and texture analysis," Med. Phys., vol. 25, no. 4,pp. 516–526, Apr. 1998.

[7] M .MeselhyEltoukhy, I. Faye, and B. Belhaouari Samir, "A comparison of wavelet and curvelet for breast cancer diagnosis in digital mammogram,"Comput. Biol. Med., vol. 40, no. 4, pp. 384–391, Apr. 2010.

[8]] R. Kohavi and G. H. John, "Wrappers for feature subset selection," Artif.Intell., vol. 97, no. 1/2, pp. 273–324, Dec. 1997.

[9] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," Mach. Learn., vol. 46,no. 1, pp. 389–422, 2002.

[10] P. A. Estévez, M. Tesmer, C. A. Perez, and J. M. Zurada, "Normalized mutual information feature selection," IEEE Trans. Neural Netw., vol. 20,no. 2, pp. 189–201, Feb. 2009.

[11] P. A. Mundra and J. C. Rajapakse, "SVM-RFE with MRMR filter for gene selection," IEEE Trans. NanoBiosci., vol. 9, no. 1, pp. 31–37,Mar. 2010

[12] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy, "IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 8, pp. 1226–1238,Aug. 2005.

