IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

A Hybrid Scheme for Mass Detection and Classification in Mammogram

Anusha D S
PG Full Time Student, CSE Department, Karunya University, Coimbatore, India
anudeepa125@gmail.com

Abstract

Mammography is used to detect breast cancer in women. This paper provides methodology for breast mass detection and classification. It consists of four steps. First step is enhancement of mammogram image using homomorphic filtering. Second step is segment the enhanced image using local seed region growing algorithm. Third step is feature extraction using shearlet transform. Finally, the fourth step is classification of mass as either benign or malignant using support vector machine.

Keywords: Mammogram, image segmentation, feature extraction, mass classification

Introduction

Mammography is used to examine human breast to detect breast cancer. It facilitates early detection of breast masses. . It is difficult in some cases to interpret suspicious region of interest (ROIs). Region of interest segmentation is the process of segmenting the suspicious regions to be analyzed for abnormalities. . The purpose here is to detect the region of interest in mammograms and classify the region of interest as either benign or malignant. Computer aided methods used here including mammographic image enhancement, segmentation and diagnosis via filtering, mass detection and classification.

Methodology

First, enhance the mammogram images by using homomorphic filtering. Next, the suspicious regions such as masses are extracted using LSRG algorithm. The detected ROIs are not always true positive masses; some of them are non-mass breast tissue. To decrement false positives and improve true positives detection, Shearlet Transform is implemented prior to feature extraction. First classification determines whether ROI is mass or non-mass and the second classification determines whether the mass is benign or malignant, which provides the breast cancer diagnosis.

Mammogram Image Enhancement Using Homomorphic Filtering

Homomorphic Filtering simultaneously normalizes the brightness across an image and increases contrast. It provides a good deal of control over the components of illumination and reflectance. It requires low and high frequency components of Fourier transform.

Image Segmentation Using Local Seed Region Growing Algorithm

Segmentation is an essential process where an image is taken as input and some detailed description of the scene or object is found for output. It divides the spatial domain pixels into meaningful non-overlapping, constituent regions that are homogeneous with respect to some characteristics. Local Seed Region Growing (LSRG) algorithm depends on the traditional similarity-based Seed Region Growing (SRG) segmentation algorithm that partitions an image directly into regions via some similarity measurements, without searching for boundaries or thresholds. LSRG determines similarity criterion and seed selection that are carried out according to both global and local conditions.

Feature Extraction Using Shearlet Transform

The shearlet transform is unlike the traditional wavelet transform which does not possess the ability to detect directionality, since it is merely associated with two parameters, the scaling parameter a and the translation parameter t. Traditional wavelet methods do not perform as well with multidimensional data. Indeed wavelets are very efficient in dealing with point wise singularities only. The idea now is to define a transform, which overcomes this vice, while retaining most aspects of the mathematical framework of wavelets. The shearlet shows optimal behavior with respect to the detection of directional information. Shearlets are a multiscale framework which can efficiently encode anisotropic features. The basic idea for the definition of continuous shearlets is the usage of a 2-parameter dilation group, which consists of products of parabolic scaling matrices and shear matrices. Shearlet transform provides a framework for analyzing and representing data with anisotropic information at multiple scales. Signal singularities such as edges can be precisely detected and located in images.

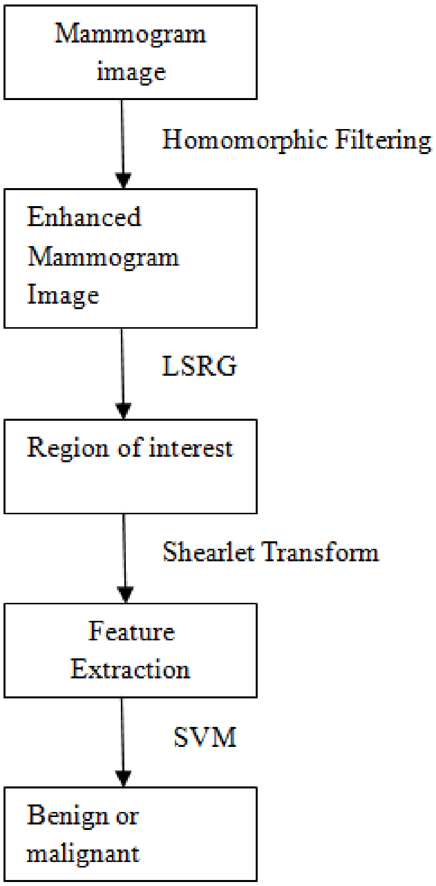

Fig 1 Flow chart of Mammogram Mass classification

Classification Using Support Vector Machine

Support Vector Machine (SVM) is a method used to estimate data classification function. The basic idea of an SVM is to construct a hyperplane as the decision surface in such a way that the margin of separation between positive and negative examples is maximized. The feature matrices are given as input to support vector machine. A classification task usually involves separating data into training and test sets. Each instance in the training set contains one target value and several attributes. The goal of SVM is to generate a model that can predict the target values of test data even the attributes are given only. An SVM uses a kernel function, in which nonlinear mapping is implicitly embedded. SVM classifier classifies the ROIs as either benign or malignant.

Experimental Work

Image Dataset

Mammogram images have been taken from the Mammographic Image Analysis Society (MIAS). And it contains 25 benign and 25 malignant masses.

Image Enhancement

Enhancement of mammogram image is implemented using homomorphic filtering. First read the input image. Then apply log domain to the image to convert image to mathematical function. Then apply fourier transform. Apply high pass filter function. Then apply inverse fourier transform. Then apply exponential function to get the enhanced image.

Image Segmentation

For segmentation of enhanced mammogram image, local seed region growing method is used. Based on threshold and seed criterion seeds are selected. Then grow the region for all seeds. The suspicious regions are detected by using this method.

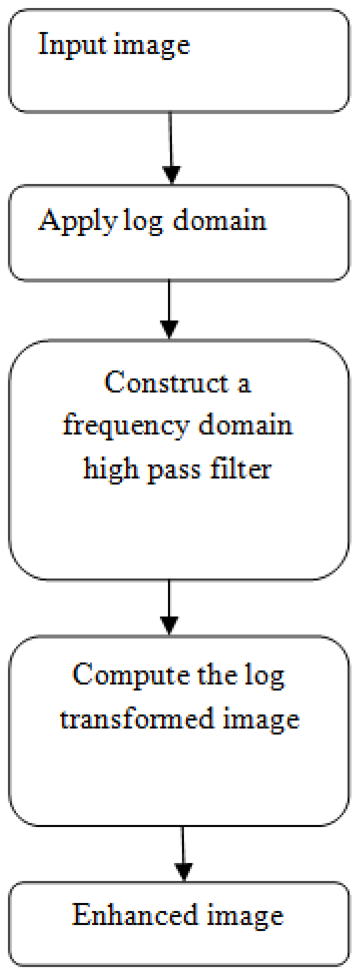

Fig 2 Flow diagram for homomorphic filtering

Feature Extraction

Features are extracted from the segmented region of interest using shearlet transform. Shearlet transform provides a framework for analyzing and representing data with anisotropic information at multiple scales. Signal singularities such as edges can be precisely detected and located in images. Shearlet transform gives approximation image on each level. From the final approximation image features are extracted. The features extracted are calculated to provide feature matrices for each ROI. Those matrices are given as input vectors to the support vector machine.

Classification

Classification of mass is done using support vector machine method. Extracted feature is given as input to the SVM classifier. It classifies the mass as either benign or malignant.

Performance Metrics

The following performance metrics are calculated.

The sensitivity is defined as the ratio between the number of true positive predictions and the number of regions in the test set. It is defined as follows:

$$\text{Sensitivity} = \frac{\text{TP}}{(\text{TP+FN})} \times 100\% \qquad (1)$$

the specificity is defined as the ratio between the number of false positive predictions and the number of regions in the test set. It is defined as follows:

$$\text{Specificity} = \frac{\text{TN}}{(\text{TN+FP})} \times 100\% \qquad (2)$$

the overall accuracy is the ratio between the total number of correctly classified regions and the test set size(total number of regions). It is defined as follows:

$$\text{Accuracy} = \left(\frac{\text{N}_\text{R}}{\text{N}}\right) \times 100\% \qquad (3)$$

where N_R is the number of correctly classified regions during the test run and N is the total number of test set. False positive fraction (FPF) gives the numbers of FPs per case (mammogram) while true positive fraction (TPF) gives the true positive detection rate.

$$\text{FPF} = \frac{\text{FP}}{\text{Total case number}} \qquad (4)$$

$$\text{TPF} = \frac{\text{TP}}{\text{TP+FN}} \qquad (5)$$

Mammograms have been taken from the Mammographic Image Analysis Society. 50 images have been taken to calculate the system performance. Among which 25 are malignant and 25 are benign.

Table 1 Confusion matrix of classification

	Malignant	Benign
Malignant	24	1
Benign	2	23

As seen in Table 1, the system performance using Shearlet transform represents 96% sensitivity, 92% specificity and 94% accuracy. FPF is 0.08 and TPF is 0.96.

Conclusion

The hybrid system segments the ROIs, detects the masses and classifies them. The ROIs are segmented using Local Seed Region Growing algorithm. The segmented regions are then given to shearlet transform. Shearlet transform provides multiresolution and produces benign/malignant classification accuracy of 94% with the MIAS database. It is also desirable to decrease the number of requested biopsy tests due to false positive detection. It provides optimum multi resolution and produces malignant/benign classification.

References

[1] *P. Görgel et al., Mammographical mass detection and classification using Local Seed Region Growing–Spherical Wavelet Transform (LSRG–SWT) hybrid scheme, Computers in Biology and Medicine 43 (2013) 765–774*

[2] *G. Kom, A. Tiedeu, M. Kom, Automated detection of masses in mammograms by local adaptive thresholding, Computers in Biology and Medicine 37 (1) (2007) 37–48.*

[3] *F. Moayedi , Z. Azimifar, R. Boostani,S. Katebi, Contourlet-based mammography mass classification using the SVM family, Computers in Biology and Medicine 40 (2010) 373–383.*

[4] *F. Moayedi , Z. Azimifar, R. Boostani,S. Katebi, Contourlet-based mammography mass classification using the SVM family, Computers in Biology and Medicine 40 (2010) 373–383.*

[5] *P. Gorgel, A. Sertbas, N. Kilic, O.N. Ucan, O. Osman, Mammographic mass classification using wavelet based support vector machine, Journal of Electrical and Electronics Engineering 9 (1) (2009) 867–875.*

[6] *G.B. Junior, A.C. Paiva, et al., Classification of breast tissues using Moran's index and Geary's coefficient as texture signatures and SVM,Comput.Biol.Med.39 (2009)1063–1072.*

[7] *G. Kutyniok et al., Introduction to shearlets, 2012*