SURFACELET TRANSFORM BASED MAMMOGRAM CLASSIFICATION SYSTEM

Leena Jasmine J S,
Department of Electronics and Communication Engineering,
Velammal Engineering College,
Chennai, India
leena_victor@yahoo.co.in

Abstract: Computer Aided Diagnosis (CAD) system plays an important role in the medical field. It helps to reduce the mortality rate due to the early diagnosis of cancers. Photographing the changes in the internal breast structure due to the formation of masses and MicroCalcifications (MC) for the detection of breast cancer is known as mammography. It uses X-rays to capture the breast tissues. In this paper, the breast tumour in the mammogram is classified into benign or malignant classes using surfacelet transform. First, the Region Of Interest (ROI) is extracted and then enhanced using histogram equalization. The enhanced mammogram ROI is subjected to surfacelet transform and features are extracted using surfacelet coefficients. Then the features are fed to Decision Tree (DT) classifier for two class prediction; benign or malignant.

Keywords: *Mammogram*, CAD, Surfacelet Transform, Decision Tree.

I. INTRODUCTION

Breast cancer is dangerous disease which affects mostly women in the world. Different kernels in Support Vector Machine (SVM) are analyzed for the detection of breast cancer in [1]. Also, they are compared with a neural network using multi-layer perception. Before classification, the effect of selecting feature subsets is also examined. Genetic Algorithm (GA) is used for selecting features subset, and 5 X 2 cross validation test is employed.

A comparison of feature extraction algorithms for mammography images is described in [2]. Feature extraction methods are compared, and different types of classifiers are used to test them. Local binary pattern, grey level run length, gray level difference, Haralick, and Gabor texture features are the five different feature extraction methods used. Three classification algorithms are used during the experiments; SVM, c4.5 algorithm, and K-Nearest Neighbor (K-NN).

The classification of masses in digital mammograms using implementation of practical CAD system is discussed in [3]. Image enhancement algorithm is used in the preprocessing stage and ROI is cropped in order to extract the tissues. A group of statistical and texture features from the ROIs is used as features. Then sequential floating forward selection is used for feature selection. Finally, testing is done by a leave-one-out method using quadratic discriminant analysis, K-NN, Linear Discriminant Analysis (LDA), and SVM classifiers.

Automatic mass detection in mammograms using hierarchical matching is explained in [4]. Template matching is used for mass detection where three types of templates are available. Based on the variety of features, a set of rules are formed to reduce the false positives. A mass detection approach by feature analysis is discussed in [5] using mammography. Noises are removed in the preprocessing stage, and then, the mass is detected by applying segmentation

approach. To find the abnormality (benign/malignant tissue), post processing is applied to the segmented region.

The detection of breast cancers in mammograms is discussed in [6] using feature selection approaches such as GA and adaptive sequential floating search. A set of features such as texture features, curvilinear features, Gabor features and multi-resolution features are extracted. Then GA and adaptive floating search based feature selection is done, and the classification of cancer regions is detected by LDA.

Vicinal SVM (VSVM) based mammographic mass detection is described in [7]. At first, SVM classifier is trained by using only normal mammograms and tested with abnormal detection. Then, malignant cases are detected by investigating the VSVM. Deterministic annealing method based on SVM kernel is used for clustering the feature space of entire training data into different soft vicinal areas.

A case-adaptive decision rule for the detection of clustered MC is presented in [8]. The accuracy of detection can be interrupted by various factors such as noises and homogeneity in breast tissue. The decision rule in the detector output is defined by using Bayes' risk approach. Translation-Invariant (TI) wavelet transform based mammogram classification is explained in [9]. At first, TI wavelet transform is employed on the mammogram for multiresolution analysis. To achieve better classification, a set of features is employed by the combination of features from TI wavelet and co-occurrence features.

Wavelet based feature extraction method for digital mammograms classification is presented in [10]. Wavelet coefficients are used for feature extraction to classify the digital mammograms. A matrix whose rows contain the wavelet coefficients of every image is constructed. Features of selected columns are used for classification. Dual-Tree Complex Wavelet Transform (DTCWT) and SVM based classification of MC is described in [11]. First feature extraction is done using DTMBWT and then classification using SVM classifier is discussed.

Gabor filter banks and hybrid Discrete Wavelet Transform (DWT) for mammogram features extraction is discussed in [12]. At first, DWT is employed on the mammogram for multiresolution analysis. Then, Gabor filter bank is applied only on the high-frequency sub-bands in order to diverse spatial orientations and frequencies. Finally, SVM classifier with the polynomial kernel is used to classify normal versus abnormal mammograms.

Statistical features based classification of MC in digital mammogram using Stochastic Neighbour Embedding (SNE) is explained [13]. SVM and DWT based breast mass classification CAD system is discussed in [14]. Categorization (benign/malignant) of mammograms with a mass tumour by SNE and DWT is explained in [15]. SVM, SNE, and DWT based breast CAD system for MC classification is explained in [16].

II. METHODS AND MATERIALS

The mammogram classification system consists of three stages. Image denoising is done in the first stage called preprocessing stage. The second stage is feature extraction in which features are extracted by the surfacelet transform. The input to this stage is the de-noised image. The last stage is classification stage where DT classifier is used for image classification into benign or malignant. A detailed summary of these stages is discussed below.

A.   Pre processing

Preprocessing is one of the essential steps in any CAD system. In this step, the given mammogram is preprocessed using de-noising and enhancement approaches so that the extracted features can be easily distinguishable for each category. In this study, the contrast of mammogram image is enhanced by applying histogram equalization method. Figure 1 shows an original mammogram and its enhanced version.

(a)   *(b)*
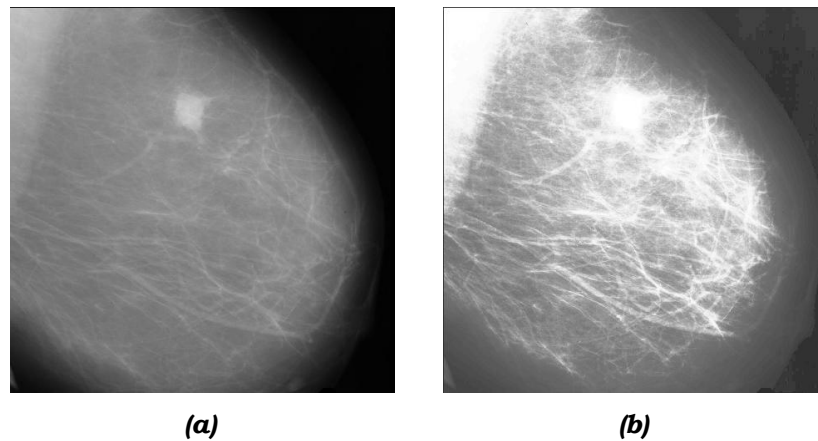
Fig. 1 (a) Abnormal mammogram (b) Contrast enhanced mammogram

B.   Feature Extraction

The performance of any CAD system can be affected by the features extracted in this stage. The character of an object is described using its features. It can be any mathematical description which is used to differentiate between two classes (benign/malignant) and more than two classes. The identification system uses them to identify the area of abnormalities and provides a final decision. To classify the given mammogram into benign or malignant, features are extracted using surfacelet transform from the ROI which contains the suspicious region of tumours. Figure 2 shows the extracted ROI from the enhanced image.


(a)   *(b)*
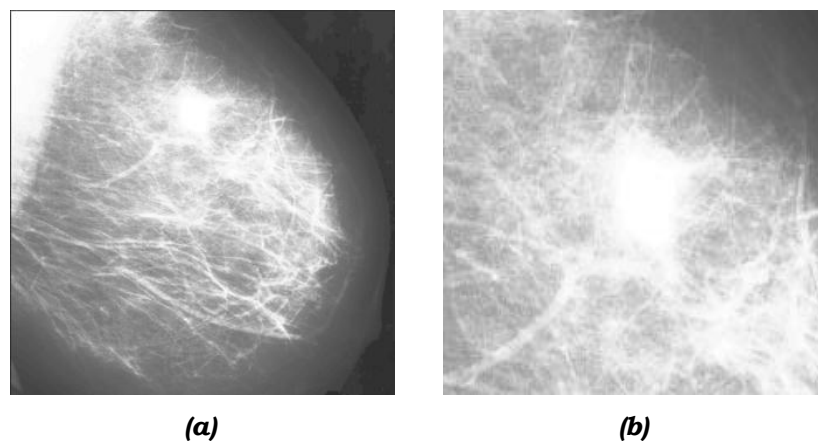
Fig. 2 (a) Enhanced mammogram (b) ROI region

B1.   Surfacelet Transform

Multi-directional as well as multiresolution analyzes are the surfacelet transform characteristics. The extension of the 2-D dimensional filter bank is the first step in the construction of the surfacelet transform to higher dimensions. For example, frequency partitioning achievement in 3-D is necessary where the ideal pass bands of the component filters are pyramids based on rectangular radiating out from the source at different orientations and tiling the full frequency space and this is a normal extension from the wedge-shaped frequency partitioning in 2D.

Surfacelet transform is initially designed to obtain related frequency partitioning like in curvelets. To achieve this partitioning both are using two different techniques. In surfacelet transform, aliasing is permitted in the directional filter bank. Carefully designed filters cancel the aliasing at the end. Surfacelets are extensively less redundant than curvelets. Surfacelet transform uses filters with rapid spatial decay. Thus, the redundancy is retained efficiently. Also, the filters do not need to strictly band limited. Figure 3 (a) and (b) show the 3-D signal resolution reduced in each dimension and frequency partitioning of the directional filter bank.

(a)   **(b)**
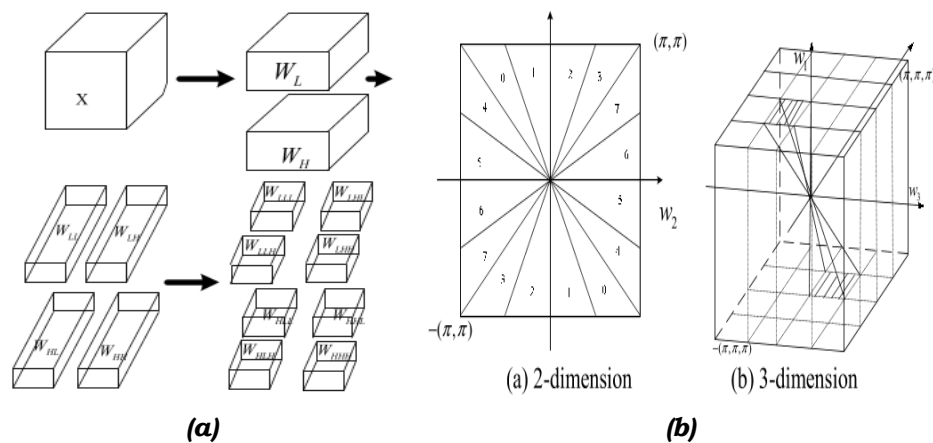
Fig. 3 (a) 3-D signal resolution reduced in each dimension (b) Frequency partitioning of the directional filter bank

The fundamental properties of the surfacelet transforms are listed below:

➤ Redundant is provided by the 2-D surfacelet transform just like the curvelet transform.
➤ Curvelets are localized to one angular wedge exactly. Surfacelet transform extends the localization across a complete scale. The parabolic scaling relation of width ∝ length obeys in spatial domain surfacelet as the oscillations occur only in the minor axis.
➤ In the 2-D case the surfacelet transform is about five times redundant.
➤ The wedges double at all scale in the surfacelet transform whereas, in curvelet, it doubles at each other scale.

C.   *Classification*

In this stage, an efficient classification approach based on DT is constructed. Before the construction of DT, the feature vector is formed by combining surfacelet transform features of each band for the classification and stored in the database. In DT, feature or data is represented by a hierarchical data structure through a divide and conquer strategy. It is also expressed in terms of cyclic partition of feature space. DT consists of a number of nodes depends on the feature space. It starts with the main node called as root, and the remaining is called as leaves. For a given training dataset or samples, best DT is constructed so that the error is minimal. The cost function is optimized in DT so that a DT is constructed. In a tree structure, a series of test questions and conditions are prepared by DT.

III. RESULTS AND DISCUSSION

MIAS database images are used to investigate the performance of mammogram classification system [17]. It has many types of mammograms which include MCs, spiculated masses, circumscribed masses, architectural distortion and normal mammograms. All benign (51 images) and malignant (64 images) images are considered and separated from all the classes of mammograms and used for the analysis. Figure 4 shows a sample benign and malignant image in MIAS database.


(a)   *(b)*
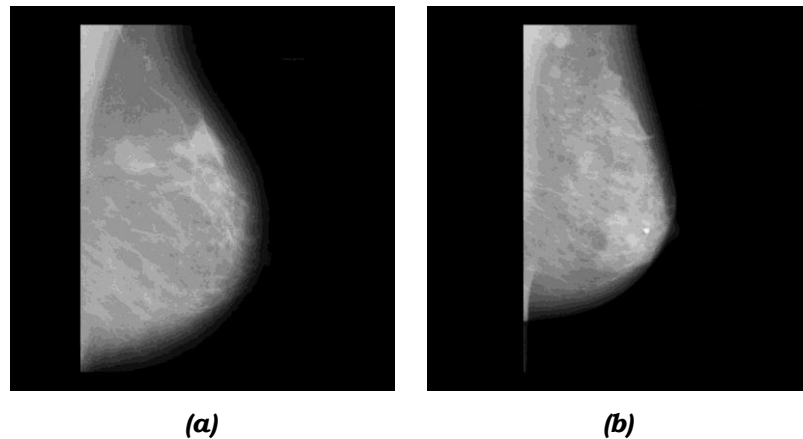
Fig. 4 (a) Benign mammogram (b) Malignant mammogram

All the images are decomposed using surfacelet transform, and the obtained surfacelet coefficients are used as features and are stored for classification stage. Finally, DT classifier which classifies the image into benign or malignant is used for the classification. Table 1 shows various parameters evaluated for the mammogram classification system.

TABLE 1 Evaluation of Mammogram Classification system

Type of Images	Total Images	Correctly classified image	Misclassified image	Sensitivity (%)	Specificity (%)	Over all Accuracy (%)
Benign	64	60	4			
Malignant	51	49	2	96.1	93.8	94.8

It is inferred from table 1 that the mammogram classification system gives better result in terms of sensitivity (96.15), specificity (93.8), and accuracy (94.8). Among the 51 malignant images, only two images are misclassified. Receiver Operating Characteristic (ROC) curve of mammogram classification system is shown in Figure 5.
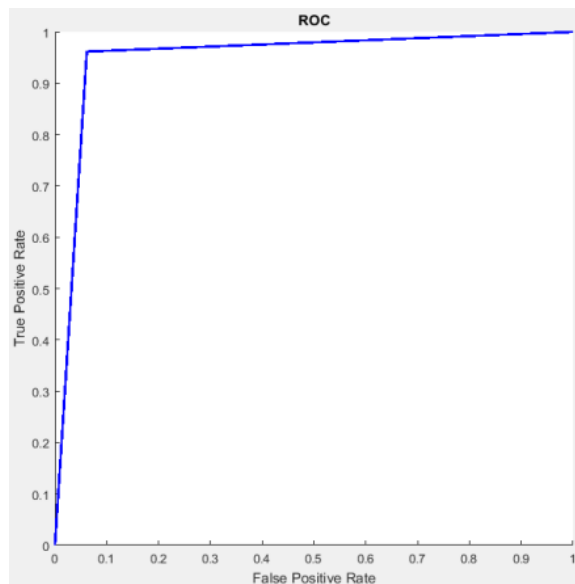
Fig. 5 ROC curve of mammogram classification system

IV. CONCLUSION

In this study, surfacelet transform is employed to classify the mammogram images. After preprocessing using histogram equalization, ROI image is cropped and given as an input to the feature extraction stage. The obtained coefficients of the surfacelet transform are used to extract the mammogram features. The mammogram images are categorized into benign or malignant by DT classifier using surfacelet transform features. The obtained accuracy is 94.9% whereas the sensitivity and specificity are 93.8% and 96.1% respectively. Results show that the mammogram classification system provides fast and accurate results for MIAS database breast images.

REFERENCES

[1]. M. Hussain, S.K. Wajid, A. Elzaart, and M. Berbar, "A comparison of SVM kernel functions for breast cancer detection", IEEE Eighth International Conference on Computer Graphics, Imaging and Visualization, 2011, pp. 145-150.

[2]. I. Kitanovski, B. Jankulovski, I. Dimitrovski, and S. Loskovska, "Comparison of feature extraction algorithms for mammography images", IEEE 4th International Congress on Image and Signal Processing, Vol. 2, 2011, pp. 888-892.

[3]. M.E. Elmanna, and Y.M. Kadah, "Implementation of practical computer aided diagnosis system for classification of masses in digital mammograms", IEEE International Conference on Computing, Control, Networking, Electronics and Embedded Systems Engineering, 2015, pp. 336-341.

[4]. S. Xu, and C. Pei, "Hierarchical matching for automatic detection of masses in mammograms", IEEE International Conference on Electrical and Control Engineering, 2011, pp. 4523-4526.

[5]. B.C. Patel, and G.R. Sinha, "Mammography feature analysis and mass detection in breast cancer images", IEEE International Conference on Electronic Systems, Signal Processing and Computing Technologies, 2014, pp. 474-478.

[6]. Y. Sun, C.F. Babbs, and E.J. Delp, "A comparison of feature selection methods for the detection of breast cancers in mammograms: adaptive sequential floating search vs. genetic algorithm", IEEE 27th Annual International Conference of the Engineering in Medicine and Biology Society, 2005, pp. 6532-6535.

[7]. A. Cao, Q. Song, X. Yang, S. Liu, and C. Guo, "Mammographic mass detection by vicinal support vector machine", IEEE International Joint Conference on Neural Networks, Vol. 3, 2004, pp. 1953-1958.

[8]. M.V.S. de Cea, and Y. Yang, "Case-adaptive decision rule for detection of clustered microcalcifications in mammograms", IEEE 12th International Symposium on Biomedical Imaging, 2015, pp. 1147-1150.

[9]. L. Zhang, and X. Gao, "Research on translation-invariant wavelet transform for classification in mammograms", IEEE 3rd International Conference on Natural Computation, Vol. 3, 2007, pp. 571-575.

[10]. I. Faye, B.B. Samir, and M.M. Eltoukhy, "Digital mammograms classification using a wavelet based feature extraction method", IEEE Second International Conference on Computer and Electrical Engineering, Vol. 2, 2009, pp. 318-322.

[11]. A. Tirtajaya, and D.D. Santika, "Classification of microcalcification using dual-tree complex wavelet transform and support vector machine", IEEE 2nd International Conference on Advances in Computing, Control and Telecommunication Technologies, 2010, pp. 164-166.

[12]. S, Lahmiri, and M. Boukadoum, "Hybrid discrete wavelet transform and Gabor filter banks processing for mammogram features extraction", IEEE 9th International New Circuits and Systems Conference, 2011, pp. 53-56.

[13]. S. Mohan Kumar, and G. Balakrishnan, "Statistical Features Based Classification of Micro calcification in Digital Mammogram using Stochastic Neighbour Embedding", International Journal of Advanced Information Science and Technology, Vol. 7, No. 7, 2012, pp. 20-26.

[14]. S. Mohan Kumar, and G. Balakrishnan, "The Performance Evaluation of the Breast Mass classification CAD System Based on DWT, SNE AND SVM",

International Journal of Emerging Technology and Advanced Engineering, Vol. 3, No. 10, 2013, pp. 581-587.

[15]. S. Mohan Kumar, and G. Balakrishnan, "Categorization of Benign And Malignant Digital Mammograms Using Mass Classification SNE and DWT", Karpagam Journal of Computer Science, Vol. 7, No. 4, 2013, pp. 237-243.

[16]. S. Mohan Kumar, and G. Balakrishnan, "The Performance Evaluation of the Breast Microcalcification CAD System Based on DWT, SNE AND SVM", CiiT International Journal of Digital Image Processing, Vol. 5, No. 11, 2013, pp. 483- 487.

[17]. J. Suckling, "The Mammographic Image Analysis Society Digital Mammogram Database", Exerpta Medica. International Congress Series 1069, 1994, pp. 375-378.