International Journal of Engineering Research in Computer Science and Engineering (IJERCSE) Vol 3, Issue 3, March 2016

# Classification of Local Binary Patterns in Mammogram Using SVM

[1] S.Venkata Lakshmi [2] V.Divya [3] B.Tamilarasi [4] S.Sayeesudha
[1]Assistant Professor, [2][3][4] Fourth year Student,
[1][2][3][4] Department of Computer Science and Engineering,
Panimalar Institute of Technology, Chennai.
[1] svlcsepit@gmail.com, [2] divyavenkattt@gmail.com [3] taamilarasibaskar18@gmail.com
[4] sayeesudha7@gmail.com

*Abstract:* **Mammogram is one of the most commonly used radiology tool for the detection of breast cancer at the earlier stage, as it helps to reveal abnormalities such as masses, micro-calcification, asymmetries and architectural distortions. In this paper, we propose a technique for diagnosing breast cancer by using SVM classifier, which segregates on the basis of LBP features. SVM (Support Vector Machine)is a supervised learning models to analyze the given data and then recognize their pattern formats; extracted from the mammographic images by using LBP technique. Further Feature Extraction is performed by HOG technique. The HOG technique helps to specify the magnitude, phase and angle value of the scanned LBP regions. This proposed method using mammograms are solicited for a different set of asymmetric cases and normal cases in the mini-MIAS (Mammographic Image Analysis Society) database, from which their data were analyzed to obtain their effective estimations with the accuracy rate of 0.85 and above.**

***Key terms*: Mammogram, Local Binary Patterns (LBP), Support Vector Machines (SVM)**.

## I. INTRODUCTION

Breast cancer is the second most common cancer overall and the most prevalent cancer among women. Based on recent statistics, conducted in the year 2014 by the international Agency for Research on cancer, breast cancer accounts for 20.9% of all cancers diagnosed and ranks as the fifth cause of cancer death in the world [8]. Deaths caused by cancer are expected to increase in future, with an estimated 12 million people dying from cancer in 2030 [2]. At present, there are no effective ways to prevent breast cancer, because its cause remains unknown. However, efficient diagnosis of breast cancer in its early stages can give a woman a better chance of good recovery. Therefore, early detection of breast cancer can play an important role in reducing the mortality rates.

The first step towards breast cancer detection is mammography. Mammography is the process where a patient's breast is exposed to low x-rays thus producing a mammographic image. It is most frequently used because is a cheap and non invasive method for investigating patient's health status and allows early stage cancer detection. The mammographic images are usually analyzed by doctors or radiologist in-order to reduce the breast cancer mortality.

In-order to improve the efficiency of mammography and also help the radiologists to achieve higher diagnostic accuracy, the Computer aided Diagnosis tool is used [2].
In this paper we are going to gather inputs to mini-MIAS (Mammographic Image Analysis Society) database and for the database mammographic images of different patients are given. These mammographic images have artifacts and digital noises. Preprocessing technique is required for obtaining a better improvement. In this process, unnecessary noise is eliminated and the contraction of poor quality images are enhanced [6].The images will be converted to gray scale. Later segment the given mammographic image into different partitions, and thus it helps to make landmark for certain regions of the breast [4].From the segmented image the features of the tissues are extracted using Local Binary Patterns (LBP). Local Binary Pattern is a powerful methodology that is used for texture classification. In this paper, the use of LBP operators helps to characterize micro-patterns (i.e. edges, lines, spots, flat areas). So that each patterns are used to cover different regions of the breast. Histogram of Oriented Gradient (HOG) is used for extracting details from the patterns that have been scanned using LBP. These

details are expressed in-terms of angle, magnitude and gradient values.

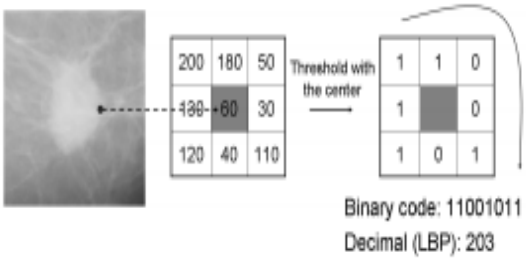


**Fig. 1.** Example of the basic LBP operator

The features that have been extracted were all classified by using SVM. Support Vector machine is a supervised learning models which have an association with the set of learning algorithms that aid to analyze the given data which provided in the datasets and then recognize their pattern formats from which the affected regions of the breast in the mammographic image can be identified.Obtained results and the comparison with previous works demonstrate thevalidity of this paper approach in terms of their accuracy rate value.

## II. RELATED WORK:

It is possible to use several imaging techniques for the examination of breast cancer, including magnetic resonance imaging, ultrasound imaging, X-ray imaging and mammography. Several surveys are done for understanding the techniques and working of the mammography device. The discussions involved in each of these surveys are determined below.

ShantanuBanik*et al* [9] submitted a paper that explains the methods for detection of sites of architectural distortion in prior mammograms of interval-cancer cases in a screening program. It is hypothesized that screening mammograms obtained prior to the detection of breast cancer might contain subtle signs of early stages of breast cancer, in particular to architectural distortion. The methods used are Gabor filters, phase portrait analysis, analysis of the angular spread of power, fractal analysis, Laws of texture energy measures and Haralick's texture feature. The use of Gabor filter and linear phase analysis leads to automatic detection of the location of nodes in the mammographic image. Then Power spectral analysis is used for identifying fractal dimensions from which fractal analysis is performed. This helps to determine the location of architectural distortion. Extraction of pattern textures is done with the help of Laws of texture energy measures. Finally, statistical analysis is done via Haralick's measures. Arianna Mencattini*et al* [7] indicated a method to analyze bilateral asymmetry in mammograms based upon systematic comparison of paired mammographic strips of the right and left breast of a patient. Their procedure is composed of the following steps: digital masking or computerized masking is performed for the paired mammograms via algorithms used for land marking and matching, bilateral analysis is used as directional components of mammograms by using Gabor filter and a novel angular similarity index. Finally, classification of the paired mammographic images is done in-order to identify whether the specified breast region is found to be normal or asymmetric. The leave-one-patient-out method was used for cross validation of the results.

R. J. Ferrari *et la* [3] proposed a new scheme based upon bank of self similar Gabor functions and the Karthunen- Leove (KL) transform to analyze directional components of mammographic images. The method is applied to detect global signs of asymmetry in fibro glandular discs of the left and right mammograms of a given subject. The Gabor filters is mainly used for determining the linear directional components. The filter responses for different scales and orientations are analyzed by using KL transform and Otsu method of thresholding. KL transform is applied to select the principal components of the filter response, preserving only the most relevant directional components of the image. Later the selected principal components are threshold by using Otsu's method that helps to determine the magnitude and the phase values of the image. At last Rose diagram is used for analyzing these value patterns.

M. Salmeri*et al* [8] presented a work on automated procedure of bilateral asymmetry detection of breast cancer in which initially mammographic densities were analyzed and fibro glandular disc detection in done through adaptive clustering techniques. Secondly, analysis and implementation of bilateral asymmetric detection algorithm is done by using Gabor filter technique. Later, use linear Bayes classifiers with leave-one-patient-out method to assess the asymmetry degree of two breast regions. Finally, metrological evaluation is done for the whole system. Their proposed method helps to provide clues about the presence of early sign of breast cancer.

J. E. L. Desaulets*et al* [4] presented a method for identifying pectoral muscle in MLO views is based upon a multi-resolution technique using Gabor wavelet. This method helps to overcome the limitations of the straight line representation of the mammographic image that was considered in their initial investigation using Hough transform.

## III. PROPOSED METHODOLOGY

Our proposed method to diagnosis breast cancer is summarized in the flow chart which is given in Fig.2. The step involves Preprocessing, Segmentation of preprocessed image using threshold algorithm, then Feature extraction is

performed using LBP and HOG techniques. Finally, classification of extracted features is done by SVM classifier. The process involved in each steps are discussed below.

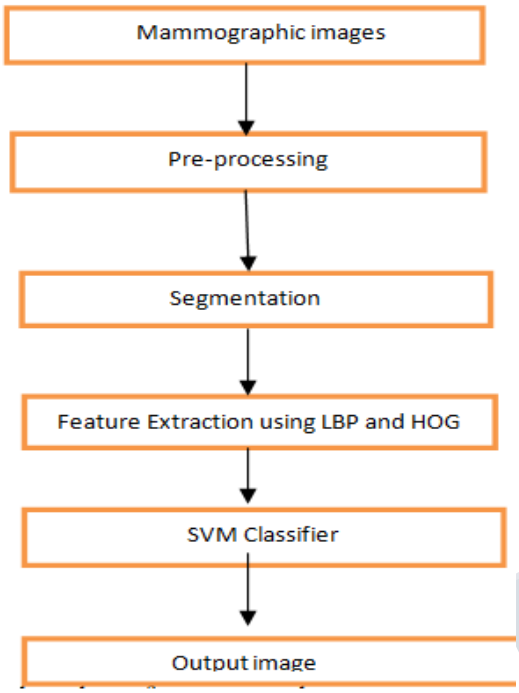


***Fig: 2 Flow chart of our proposed system***.

### *A. Preprocessing:*

Pre-processing is the initial step of every traditional image processing pipeline. This stage typically performs a variety of basic operations to eliminate distortions from the images being compared [6]. Pre-processing phase is needed to improve the image quality and make the segmentation results more accurate. The gray scale format is used for pre-processing technique. Then noise could be removed by using different filters that were used in image processing technique. For our proposed system mean filter is used for filtering. The output of this technique will provide a gray scale image as it is shown in Fig 3 and it is given as an input for the segmentation technique.

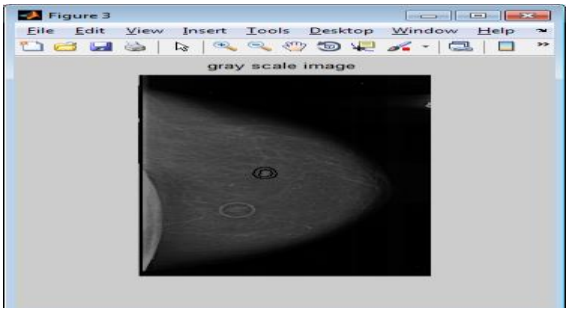


***Fig 3.a***

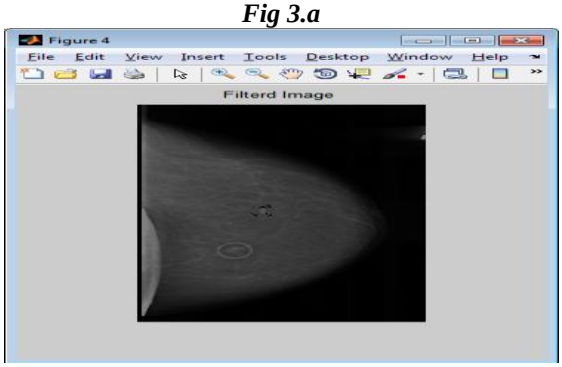


***Fig3.b***

***Fig : 3.a .Gray scale image 3.b.Filtered image***

### *B. Segmentation:*

Image segmentation is the process of partitioning a digital image into multiple segments.The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze.The thresholding in mammograms images is based on separating the histogram into background and breast tissues. Depending on the value of threshold all pixels less than the threshold are classified as background, and the reminder pixels are breast or vice versa. In the proposed system apply threshold algorithm to segment microcalcifications in mammographic pre-processed images. Finally, the output of the segmented images is shown in Fig 4 and then these segmented images are smoothened in their edges by using morphological technique.

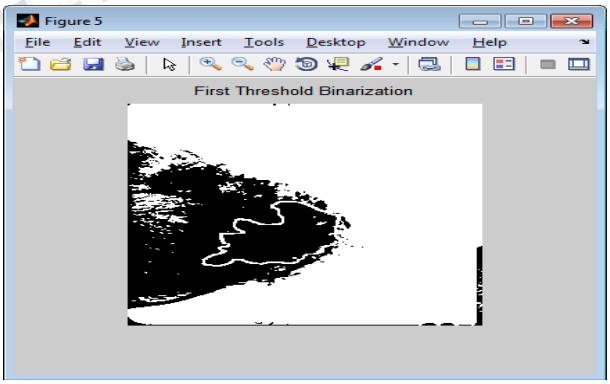


***Fig: 4: The output of the segmented image***

### *C. Feature Extraction:*

Feature extraction involves how accurately the features are extracted from the image and without loss of any information on the image. This is achieved by several image transformation techniques used in feature extraction. Local Binary Patterns (LBP) is a method for texture feature extraction very popular for face detection and recognition approaches. LBP helps to provide texture features which

are very useful for breast cancer abnormality classification in mammography's and LBP has been used to classify mammograms.

It is also possible to obtain rotation invariance by grouping the extracted features in histograms. In its original version, LBP was proposed as a histogram of binary values. Those binary values are computed if a neighbor value is higher than the central value 1 is assigned to that position, otherwise 0 is assigned.

The two main user defined parameters in the LBP method are P (number of neighbors) and R (radius of comparison). The procedure is illustrated in Figure 5. It is possible to obtain rotation invariance by grouping the extracted features in histograms.

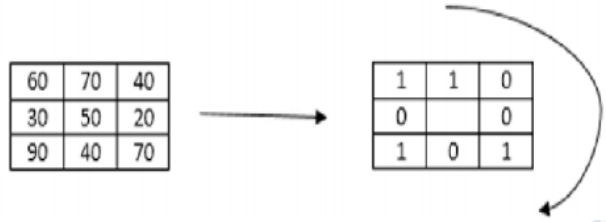


***Fig: 5. LBP operator in action.***

According to Figure 3, the binary code obtained from that image region is: (01100101)2 = (101)10. Later the further feature extraction is done by using HOG (Histogram of Oriented Gradient).Itis used in image processing for the purpose of object detection.

The technique counts occurrences of gradient orientation in localized portions of an image. It helps to specify the magnitude, phase and angle values of the scanned LBP regions.

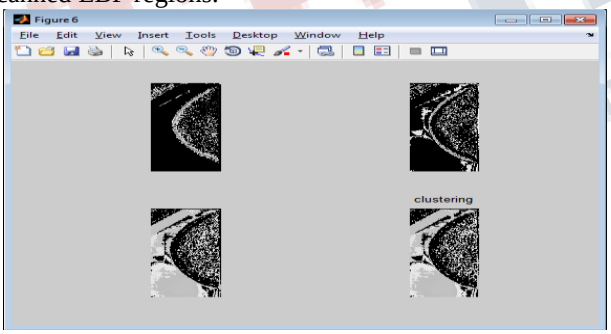


***Fig.6 Feature Extraction from segmentation phase***

From Fig.6, it is seen that the features are extracted from the segmentation phase. The features that are obtained from these techniques are compared with the trained features which are available in the database.

***D. Classifier:***

In the last phasethe selected feature is classified for normal or abnormal.In this paper we use Support Vector Machine (SVM) which is used to produce good results for medical diagnostics.Support Vector Machine(SVM) are supervised learning models which have an association with the set of learning algorithms that helps to analyze the given data and recognize their pattern formats from which the affected tissues of the breast region in the image can be identified by Support Vector machine(SVM).

Given a set of points belong to either of two classes, linear SVM finds the hyper-plane leaving the largest possible fraction of points of the same class on the same side, while maximizing the distance of either class from the hyper-plane. According to this, hyper-plane is used to minimize the risk of misclassification in the given test set.

The proper comparison with the trained features is done by the SVM classifier. From the comparison, the SVM classifier would be able to determine whether given mammographic image has severe, mild or normal issue of breast cancer.

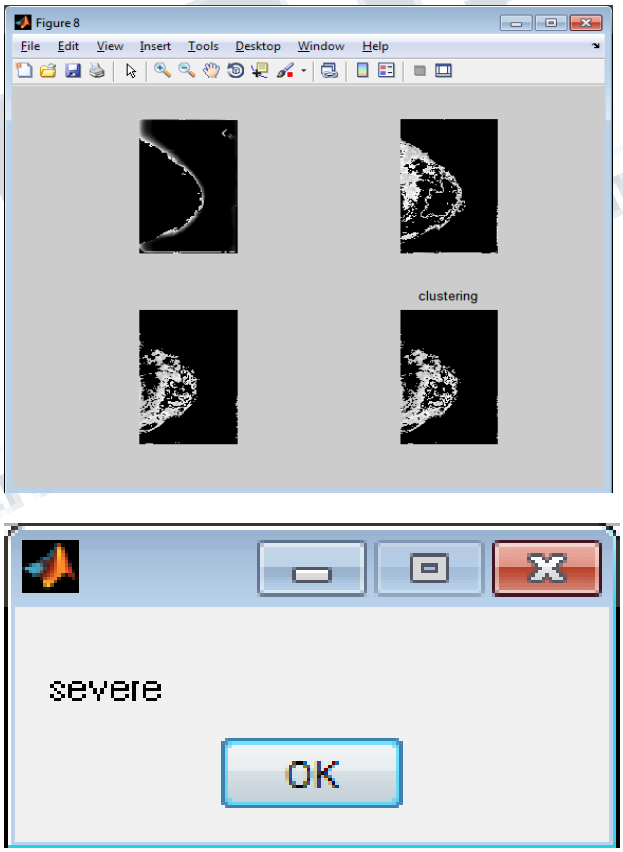


***Fig.7. Indication of presence of breast cancer.***

```
% succssive rate for svm

total_number=22;
correct_result=19;

percentage_svm=(correct_result/total_number)*100


accuracy=86.36
```

***Fig.8.Accuracy value***

## IV. CONCLUSION:

In this paper different set of asymmetric cases and normal cases for about 256 mammographic images which is stored using Mini-MIAS(MammographicImage Analysis Society)database,from which their data were analyzed to obtain their effective estimations with the accuracy rate of 0.86 and above.Analysis of LBP can analyze the neighbor pixel in order to identify the breast cancer in effective manner and the SVM classifier is used to find the images normal or abnormal. This mammographic technique could be enhanced in the feature to improve the accuracy of detection of breast carcinoma at an earlier stage.

## REFERENCE

[1.] E. A. Sickles, "Mammography: Asymmetries, masses, and architectural distortion," in Diseases of the Heart and Chest, Including Breast 2011–2014, J. Hodler, G. K. von Schulthess, and C. L. Zollikofer, Eds. New York: Springer, 2011, pp. 255–258.

[2.]J. Tang, R. M. Rangayyan, X. Jun, I. El Naqa, and Y. Yang, "Computeraided detection and diagnosis of breast cancer with mammography: Recent advances," IEEE Trans. Inf. Technol. Biomed., vol. 13, no. 2, pp. 236–251, Mar. 2009.

[3.]R. J. Ferrari, R. M. Rangayyan, J. E. L. Desautels, and A. F. Frère, "Analysis of asymmetry in mammograms via directional filtering with Gabor wavelets," IEEE Trans. Med. Imag., vol. 20, no. 9, pp. 953–964, Sep. 2001.

[4.]R. J. Ferrari, R. M. Rangayyan, J. E. L. Desautels, R. A. Borges, and A. F. Frère, "Automatic identification of the pectoralmuscle in mammograms," IEEE Trans. Med. Imag., vol. 23,Feb 2004. b.2004IEEE Trans. Med. Imag., vol. 23, no.2, pp. 232–245, Nov. 21–23, 2013.

[5.]S. Banik, R. M. Rangayyan, and J. E. L. Desautels, "Detection of architecturaldistortion in prior mammograms," IEEE Trans. Med. Imag., vol. 30, no. 2, pp. 279–294, Feb. 2011.

[6.]P. Casti, A. Mencattini, M. Salmeri, A. Ancona, F. Mangieri, and R. M.Rangayyan, "Masking proceduresand measures of angular similarity for detection ofbilateralasymmetryin mammograms," in Proc. IEEE e-Health and Bioeng. Conf., Iasi, Romania, Nov. 21–23, 2013.

[7.]P. Casti, A. Mencattini, and M. Salmeri, "Characterization of the breast region for computerassisted Tabár masking of paired mammographic images," in Proc. 25th Int. Symp. IEEE Comput.-Based Med. Syst., 2012,pp.1-6.

[8.]Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Trans.,vol. 13,no.4,pp.600-612, Apr.2004.

[9.]E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, "Shiftable multi-scale transforms," IEEE Trans. Inf. Theory, vol. 38, no. 2, pp. 587–607, Mar. 2010.

[10.]A. Mencattini, M. Salmeri, and P. Casti, "Bilateral asymmetry identification for the early detection of breastcancer," in Proc. IEEE Int. Workshop Med. Meas. Appl.Proc., 2011, pp. 613–618.