Cancer Mass Detection from Mammogram Based on Enhanced Feature Extraction Method

Jasmin M R[1]
PG Student,Dept:of Computer Science
TKM Institute of Technology
Kollam,Kerala

Soorya P[2]
Asst:Professor,Dept of computer Science
TKM Institute of Technology
Kollam,Kerala

***Abstract*—In the current breast cancer screening ,radiologists often miss approximately 10-30% of tumors because of the ambiguous margins of lesions and visual fatigue resulting from long time diagnosis. Thereby a computer-aided detection system have been developed to aid radiologists in detecting mammogram lesions which may indicate the presence of breast cancer. The digital mammogram image undergoes pre-processing, segmentation and feature extraction steps to identify the abnormal region is cancer affected or not. The proposed study uses surf feature extraction method on each suspicious areas. Finally uses linear discriminant analysis to classify abnormal regions by selecting and rating the individual performance of each feature**.

Keywords- Breast cancer, mammogram, mass, SURF feature extraction.

I. INTRODUCTION

Breast cancer is one of the leading causes of cancer related death among women. The death rate can be reduced if the cancer is detected at its early stages. Early diagnosis and treatment increases the chance of survival. A computer-aided detectionsystem have been developed to aid radiologists in detecting mammogram lesions which may indicate the presence of breast cancer. The digital mammogram image undergoes pre-processing, segmentation and feature extraction steps to identify the abnormal region and distinguish whether it is cancer affected or not. But all this methods are having several disadvantages. A malignant tumor that begins in cells of the breast and gets into the surrounding tissues as well. Mammography is the most used screening tool for abnormality detection, because it allows an easy way to identify the cancer. Masses and calcification are the two primary signatures of abnormality in mammograms.

Existing research results show masses are more difficult to recognize because of their abundant appearances and ambiguous margins than calcification. This paper concentrated on the mass detection in the mammograms.Inorder to extract the features of the mass,here used a new method called SURF feature extraction method.

II. RELATED WORKS

A number of methods have been developed for mammographic mass detection. Eltonsy *et al* [5] developed a concentric morphology model for detecting masses in the breast region.the average performance achieved a sensitivity of 92.1% for malignant masses with 5.4 false positives(FP's) per mammogram. X.Gao *et al* [4] introduced morphological component analysis as a preprocessing step which maintains the local contrast of an original mammogram also suppress the structural noises, blood vessels and glandular tissues. Sameti *et al*[3] introduced a method to detect the breast cancer through the complex feature extraction in the last screening mammograms. Several experiments are carried out to distinguish normal and abnormal mass through feature extraction.this system flagged regions which had a 72% chance of developing a malignant mass by the time of the next screening. Shengzhou Xu *et al*[2] proposed a new method to detect the masses through hierarchical template matching. The test result indicates the sensitivity of 96.2% for malignant masses and false positive rate is 5.2 per image.H.Bay *et al*[6] proposed a new feature extraction criteria called Speeded Up Robust Feature,a novel scale and rotation invariant interest point detector and descriptor.

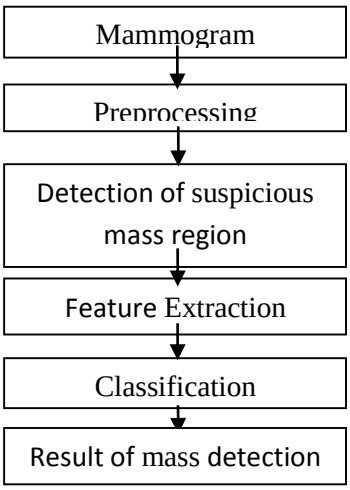

Fig .1. Block diagram of proposed mass detection

The percentage of recognition rate is slightly high when compared with other descriptors.
In this study,hierarchical template matching method is used to segment the suspicious mass from the breast region and an enhanced feature extraction using SURF are proposed.Analysis shows that SURF results better detection compared with complex texture feature extraction methods.The most important advantage of using the Surf is the reduction of false positive rates.

III. METHODS

Detection of masses is difficult because many areas of these mammograms appear to have features that are mass-like but not masses. Moreover, masses have great variability in shape, size, and contrast. Particularly, challenging masses are those that occur near the breast skin line, the chest wall, and generally those appearing in dense breasts. Thus false positives are produced which detract from the effectiveness of the algorithm. The proposed CADe system consists of four stages. They are mammography preprocessing, detection of suspicious region, Feature extraction, and classification. Fig 1 describes the block diagram of the proposed mass detection from the mammograms. This section shows the detailed description about each stages.

i. Mammogram preprocessing

To extract the breast region for the detection of mass, undergoing the preprocessing step. This is used to avoid the pectoral muscles and some of the noises present in the mammogram. Firstly, Otsu's method is used for segmenting the foreground image from the background image. The foreground image will consists of the breast region and pectoral muscles if it is the mediolateral oblique views of the mammogram. Gamma correction is performed to enhance the contrast of the pixels in the image. So that we can easily identify the pectoral muscles like noises from the mammographic views. Then some morphological filtering operations are performed to discard such noises.

Mammographic preprocessing is done to avoid the pectoral muscles and also the other noises like blood vessels and glandular tissues which leads to false positives in the detection of mass regions in the breast region. In this work, morphological operations such as opening and closing filters are used.

ii. Detection of suspicious mass region

Several methods are there to detect the suspicious mass region from the mammograms. In this study ,using a method that is proposed by Xu and Pie for the detection purpose of mass region. Here template matching is carried out. Best template are selected according to the comparison. By taking into account the performances of each template, Sech template had given the best result. It is defined as

$$S(x,y)= \frac{2}{exp(\beta*\sqrt{x2+y2}) + exp(-\beta*\sqrt{x2+y2})}$$

To perform hierarchical template matching, two templates with different size are created. One is local part based template (33 × 33) pixels and complete template(66 × 66) pixels. Here x, y are the coordinates of the template and different templates are obtained by changing the value of β.The obtained breast region after the preprocessing stage are then correlated with the template to check the similarities.It is defined as

$$Cor(T,I) = \frac{E[(I-\mu I)-(T-\mu T)]}{\sigma I \sigma T}$$

T is the template ,I is the mammogram to be detected, μ_I is the average of I ,μ_T is the average of T, and σ_I , σ_T are the standard deviation of I and T respectively. Two correlation maps were obtained from the above correlation equation. Next, find out the average correlation map from the two correlated images to segment the suspicious region. This suspicious regions are called Region of Interest.

iii. Feature extraction

Feature extraction are carried out on the image which is obtained through the hierarchical template matching method. This stage are helpful to extract the characteristics of the ROI. Feature extraction is one of the crucial stage for the detection of cancerous masses. So here used a new extraction criteria called SURF(Speeded Up Robust Features).It is a new scale and rotation invariant detector and descriptor. By using SURF, will get different interest points. Feature detection is carried out on the ROI as well as the detected interest points. This will helps to reduce the false positive rates. Feature detector is based on Hessian matrix[6].

Each suspicious interest points should have a unique descriptor that does not depend on the features scale and rotation called descriptors. The Surf descriptor is based on Haar Wavelet Responses. Feature detection and descriptor are carried out on the integral image. From this feature extraction method, will get a set of descriptors

iv. Classification

The proposed classification module consists of reduced features, which were selected by Linear Discriminant Analysis(LDA).The advantage of this method are that it implicitly accounts for the correlation of features and selects the number of features to be used as inputs. For Linear discriminant Analysis, the mapping is defined as

$$g(\vec{x}) = (\overrightarrow{\mu 1} - \overrightarrow{\mu 2})^{T} S^{-1}_{pooled} \vec{x}$$

where $\vec{x}$ is the vector of input features, $\overrightarrow{\mu 1}$-$\overrightarrow{\mu 2}$ denote means of the abnormal and normal training datasets, and S represents the pooled covariance matrix of the training dataset. The observation $\vec{x}$ is called abnormal if $g(\vec{x}) \geq y_c$ and normal if $g(\vec{x}) < y_c$, where y_c is the decision variable threshold.

So the classifier called Linear Discriminant Analysis (LDA) are used for classifying the ROI is cancerous or not. So that need to create test dataset and trained dataset. Test dataset is obtained through the feature extraction step. Trained datasets are of two types. One is normal trained dataset and another one is abnormal trained dataset. Both of these datasets are created with the help of SURF feature extraction method. The obtained test dataset is then compared with the normal and abnormal trained dataset to find out whether the mass is cancerous or not

IV. PERFORMANCE EVALUATION

Performance testing is carried out on 300 mammogram. The main aim is to reduce the false positive rates, for that the proposed work uses the SURF feature extraction method. Mammogram mass detection based on complex texture feature achieves the detection sensitivity with the FP rate of 4.8%. In the proposed system, it was found out that the FP rate is less than the previous system. Figure(2) shows the performance of the system using SURF feature is better than the ODCM-Optical density feature. SURF

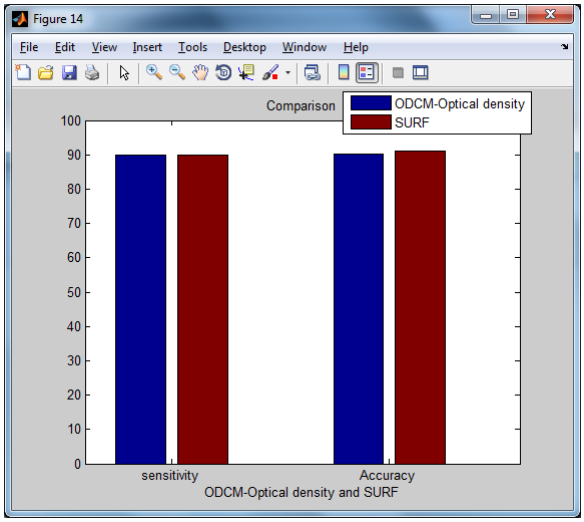

Fig. 2. Performance Comparison of ODCM-Optical density and SURF feature

feature express the better performance in the higher density mammogram

V. CONCLUSION

Breast cancer is one of the most common and deadliest one among the cancer. Early detection can increase the chance of survival. This work focuses on the mass detection on the breast region which include the steps preprocessing, detection of suspicious region, feature extraction and classification. Here SURF feature extraction method is used for finding out whether the mass is cancerous one or not, hence reduced the false positive rates.

REFERENCES

[1] Shen-Chuan Tai,Zih-Siou Chen,and Wei-Ting Tsai, "An Automatic Mass Detection System in MammogramsBased on Complex Texture Features ",IEEE Journal of Biomedical And Health Informatics,Vol.18,No.2,Mrch 2014

[2] .Xu and C.Pei, "Hierarchical matching for automatic detection of masses in mammograms," in Proc. Int. Conf. Electr. Control Eng., Sep. 2011, pp. 4523–4526

[3] M. Sameti, R. Ward, J. Morgan-Parkes, and B. Palcic, "Image feature extraction in the last screening mammograms prior to detection of breast cancer,"IEEEJ.Sel.TopicsSignalProcess.,vol.3,no.1,pp.46–52,Feb. 2009.

[4] X. Gao, Y. Wang, X. Li, and D. Tao, "On combining morphological component analysis and concentric morphology model for mammographic mass detection,"IEEETrans.Inf.Technol.Biomed.,vol.14,no.2,pp.266– 273, Mar. 2010.

[5] N. Eltonsy, G. Tourassi, and A. Elmaghraby, "A concentric morphology model for the detection of masses in mammography," IEEE Trans. Med. Imag., vol. 26, pp. 880–889, Jun. 2007.

[6] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)",Computer Vision and Image Understanding, 110(3):346–359, 2008.

[7] P M Panchal, , S R Panchal, S K Shah." A Comparison of SIFT and SURF", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 1, Issue 2, April 2013.

[8] Edouard Oyallon, Julien Rabin," An analysis and implementation of the SURF method, and its comparison to SIFT", Image Processing On Line, February 25, 2013

[9] Jae Young Choi, Dae Hoe Kim, Konstantinos N Plataniotis and Yong Man," Computer-aided detection (CAD) of breast masses in mammography: combined detection and ensemble classification", Institute of Physics and Engineering in Medicine, 59,2014.

[10] Jian Wu, Zhiming Cui, Victor S. Sheng , Pengpeng Zhao, Dongliang Su, Shengrong Gong," A Comparative Study of SIFT and its Variants", Measurement Science Review, Volume 13, No. 3, 2013