Graph-Based Density Evaluation for Mammographic Image Classification Based On Intensity Values

Dr.S.Julian Savari Antony

Department of Electronics and Communication Engineering, Keshav Memorial Institute of Technology, Hyderabad- 500029, Telangana , India,
Email:savari.sm@gmail.com

ABSTRACT

The classifying mammograms from low scale X-Ray images have studied in different categories of medical fields. The efficiency of classification suffers due to the nature of mammogram images because the x-ray images are low in quality and have taken from low strength rays. The x-rays are low quality different ultrasonic and other images. But the mammogram classification is performed at an initial level of cancer identification. The earlier approaches suffer in identifying microcalcifications. The proposed system has three different steps namely, image enhancement, Mammographic Graph Construction, Density Estimation. A new graph-based density estimation approach will propose to identify the classification of mammograms. The proposed approach enhances the image using histogram equalization technique.The pixel intensity values are adjusted in such a way to increase the contrast of the image that is used to construct the graph.Each pixel is considered as a node and connected with an edge with other neighboring pixel based on the intensity values. It computes the density measures for each distinct graph where the connected nodes have similarly valued pixels. The density measure shows that the region of the image in which has more impact towards cancer cells. The computed density measure is used to classify the region or group of pixels towards mammogram. The proposed method is performed in a single iteration.The graph is constructed to minimum numbers regions according to the number of regions in which has the white and black pixels form regions. This reduces the overall time for segmentation process and reduces the time complexity of mammogram classification.

Keywords : Mammogram, Image Classification, Graph Based Technique, Density Estimation, Pixel.

1 INTRODUCTION

The mammogram is a low energy x-ray taken to identify the presence of breast cancer which is performed at the initial stage. The low energy x-ray is obtained and then it will be monitored by the physician in order to find out the presence of cancer. The mammogram consists of white pixels with more gray values and to form a density region in which it is confirmed. The presence of calcification varies in size and has to be identified efficiently by using different metrics.

Image classification is the process of classifying a single image towards a huge space or a class. A single image can be classified in to a class based on the features of the image. In mammogram image classification, a single image is classified based on the microcalcification identified on the image. The classification algorithm uses many numbers of pre-identified trained images and has extracted their feature vectors to classify the input image.

The graph based approaches are more popular in many image processing works and can adapt them into the problem of mammogram classification.The mammogram image contains white pixels which represents the microcalcifications. Itcan be concluded that image contains cancers or microcalcifications only if they become denser and have grayer values. The general graph based methods represent the pixel as nodes and two nodes can be trained if they have similar or related values. Based on this theory, the density of micro calcification can be computed to identify the breast cancer.

The density of white region can be computed based on the intensity values of the pixels present in the region. For example in any region, the similar white pixels are connected to form a region in the eccentric nodes and their boundary can also be identified. Based on the boundary of the region and number of white pixels present in that, the density can be computed using the white pixels in that particular region to classify the region towards mammograms.

Thereare number of approaches has been deliberated in literature for classifications of mammogramswhile few of them.The classification of mammograms using stochastic neighbor embedding and KNN is proposed in [1], which enumerates the problem of dimensionality reduction handling using stochastic neighbor embedding scheme. The KNN classifier is used to classify the mammogram from the data available and for the evaluation purpose they have used for various mammogram data sets.

Mammographic masses classification: Novel and simple signal analysis method [3] has presented a CAD system for classification of malignant mass in mammogram images. The feature has been extracted using wavelet transform and for the classification they have been used in artificial neural network. The Classification of Microcalcification Using Dual-Tree Complex Wavelet Transform and Support Vector Machine [6]has been discussed for image classification. In this approach dual tree complex wavelet transform technique has been used for feature extraction. The classification is performed using support vector machine.

In Classification of Breast Masses in Mammograms using Support Vector Machine [14-15], describes the multi view CADx System for the mammography images. In this approach, the multi view patterns of mammogram are used and are extracted to perform classification. The view patterns are segmented according to the mass contour and for each view of the image a feature set will be extracted according to the boundary of geometry and they compute the structure of the mammogram in the image. Support vector machine is the classifier used in this approach using the ranked features.

A Classifier with Clustered Sub Classes for the Classification of Suspicious Areas in Digital Mammograms [2] presents a novel methodology for the classification of suspicious areas in digital mammograms. The methodology is based on the fusion of clustered sub classes with various intelligent classifiers. A number of classifiers have been incorporated into the proposed methodology and has been evaluated on the well-known benchmark digital database of screening mammography. The results in the form of overall classification accuracies, TP, TN, FP and FN have been analyzed, compared and presented.

A detection process based on local contrast thresholding and rule-based classification [4] which performed over the pre-processed and segmented mammograms. This approach has produced higher accuracy in detecting the mammogram from small set of images.

A comparison of different Gabor features for mass classification in mammography [5] performs the comparison in the usage of Gabor features in mammogram mass classification. Detection of Cancerous Zones in Mammograms using fractal modeling and Classification by Probabilistic Neural Network [6-7], uses fractal modeling to differentiate the original cancer affected image from others. In Fractional modeling,the original image is first segmented into appropriate fractal boxes followed by identifying the fractal dimension of each windowed section. Then two dimensional box counting is used to place them in appropriate location. They extract eight features of tumor to classify the image.[7] Used histogram equalization method and volumetric values.The classification method was reduced by the false positive result arise in mammogram classification and increased the classification accuracy

Clustered ensemble neural network for breast mass classification in digital mammography [8], proposes a new methodology for mammogram classification using mass of the area affected in the breast image. They have used k-means classifier to create ensamples of neural networks. This technique is designed to improve the classification accuracy of a multi-layer perception style network for mass classification of digital mammograms.

Semi-Supervised K-Means Clustering for Outlier Detection in Mammogram Classification [9] uses k-means classifier for outlier detection. They extract the shape feature and cluster of the image using k-means classifier. The generate association rules using generic association rule miner which generates rules based on the feature set. The generated rules are used to classify the image towards mammograms.

Breast cancer identification in digital mammogram is discussed in [10], which uses curve let approach to extract the feature from the mammogram image. It has two stages namely abnormality detection and classification. At first stage discriminative texture feature are extracted and then for classification ofnearest neighbor classifier the Euclidean distance is used.Classification of Malignant and Benign Micro calcification Using SVM Classifier [11], has used support vector machine for clustering and classifying the mammogram. The SVM classifies the feature according to the hyper plane.

One class classification of mammograms Trace Transform Functional is presented in [12], for the classification of breast cancer images into benign and malignant classes. This replaces the two class classification problem due to the sparse distribution of abnormal mammograms. Trace transform, which is a generalization of the Radon transform, has been used to extract the features. Classifiers such as the linear discriminant classifier, quadratic discriminant classifier, nearest mean classifier, support vector machine, and the Gaussian mixture model (GMM) were used.

Empirical Analysis of Supervised and Unsupervised Filter based Feature Selection Methods for Breast Cancer Classification from Digital Mammograms [13], performs an evaluation and comparative study of various unsupervised and supervised feature selection methods which are available for breast cancer classification from digital mammograms though various classifiers. This study aims towards finding out the better feature selection method and associated classifier which gives better performance.

Classification of mammography images based on cellular automata and Haralick parameters [15], presents a bio-inspired algorithm that helps viewing image classes of mammography. This algorithm uses the concept of cellular automata (CA); the paradigm constructed from simple elements interrelated and forms a regular structure with local neighbur's interactions. This connection generally builds a simple geometry; lattices, the vectors representing the database images located on 2 dimensions, and the local update rule of cells favoring the creation of similar states' groups in neighboring cells. This method helps to understand the underlying structures of all images. A possible solution for the classification of mammographic data by the cellular automata is promising.

All the approaches that are discussed have the problem of false positive results and have low efficiency in identifying breast cancer. A naval approach based on graph techniques and density measures is proposed.

The aim of this paper is to propose a new approach to determine the classification of Mammographic Image Using Graph Based Density estimate. This work is organized as follows. In Section 2, it present the proposed method includes the process of classification. Next, in Section 3, the results are shown. Finally, Section 4 presents some concluding observations.

2 PROPOSED METHOD

The proposed system architecture is shown in Fig.1. The proposed system has three different steps namely, image enhancement, Mammographic Graph Construction, Density Estimation. At the image enhancement stage, the gray scale values of the original images for intensity value. The generated gray scale image is applied with histogram equalization technique where the distributional probability of intensity values throughout the image has been computed. The intensity of pixels is restored with the distributional probability values which increase the contrast of the image and also the image quality. The resultant image is used for the graph generation process. Based on the result of histogram equalization the graph is constructed using compute probability distribution which density estimation is performed.

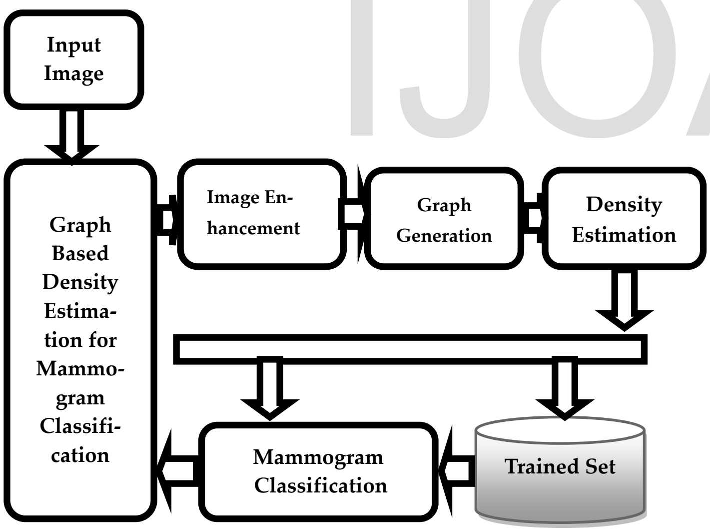

Fig. 1. Proposed system architecture

2.1 Mammogram Graph Construction

At this stage, the enhanced image is used to construct the mammogram graph generation. For each pixel of the image, a node is generated and the node contains the gray scale value of the pixel. Each node of the graph has been connected with neighboring pixels based on mass threshold. The mass threshold shows the lower gray scale value in which a pixel must have to connect with neighboring pixel. A pixel or the node of the pixel will be connected with neighboring pixel only. If it has gray scale values more than the mass threshold the node. Similarly for each of the image pixel, a node is generated and the graph will be constructed using which the density estimation is performed.

The algorithm generates number of graphs according to the histogram values present in the image and the values of neighboring pixels. The number of graph depends on the pixel values and how good they come within the histogram threshold.

2.2 Density Estimation

It calculates the density measure in which shows the white mass value of each region and performed using the pregenerated graph. They have distinct graphs and the pixels are the nodes and they are interconnected with each other. The density measure is computed using the nodes present in the graph and their location in the image which leads to boundary detection. From identified boundary, the area and density of mass region is computed. The area of pixels connected and the unconnected pixels in the region or boundary shows the density of white mass pixels in the region which could be used for the identification of micro calcification.

The algorithm computes the density of white mass value at each region of the image according to the total number of white pixels or the area covered by white mass value and the area not covered by the white mass value.

2.3 Mammographic Classifier

The mammographic classifier computes the density feature for each testing image and trained images. The classifier selects top few density regions and features for each image. The selected feature set is used to compute Euclidean distance between the input and trained set. Based on the distance computed a class label will be assigned to the test image.The mammogram classification algorithm computes the distance between each of the training sample according to the density measure computed in the previous stage of mammogram classification.

3 RESULTS AND DISCUSSION

The proposed approach enhances the image using histogram equalization technique where the pixels intensity values are adjusted in such a way to increase the contrast of the image. The enhanced image is used to construct the graph where each pixel is considered as a node and connected with an edge with other neighboring pixel based on the intensity values. The density measures are computed for each distinct graph where the connected nodes have similar valued pixels. The density measure shows that the region of image which has more impact towards cancer cells. The computed density measure is used to classify the region or group of pixels towards mammogram. The proposed method is performed in a single iteration and the graph is constructed in minimum number according to the number of regions where the white and black pixels form regions. This reduces the overall time for segmentation process and reduces the time complexity of mammogram clas-

ISSN 2278-7763

sification.The proposed methodology has been evaluated with various data sets of mammographic images. Data set shows for the evaluation of the proposed method. The Table I shows the data sets used for the evaluation of the proposed approach. It has used three different data sets like DOD BCRP, MIAS, DDSM which has many numbers of classes with more samples.

The Fig.2 shows the original input image given for mammogram classification and chooses from the data set being used.The input image is applied with the above discussed procedures in mammogram classification. The input image is being pre-processed to remove the noise and enhance the image quality to improve the classification performance. Then the feature of the image is extracted and the measures are computed to perform classification. For the pre-processing the histogram equalization technique is applied, which improves the image quality in efficient manner.

The Table I shows the simulation parameters. The Fig.2shows the original image and the Fig.3 shows the output image produced by histogram equalization to enhance the quality of the image. The image shows the equalized image and it is visible clearly that the input image quality has been improved by applying histogram equalization.

The histogram equalization is performed based on the histogram equalization and values as mentioned below, which is performed for each pixel.Img(i)= 0.2989 * R + 0.5870 * G + 0.1140 * B. //compute the histogram of the value. For each value set, we compute the number of pixels with the same gray scale value and compute the probability of distribution factor value.The Fig. 4 shows the normal image from DDSM data set which has been used for evaluation.

TABLE 1
SIMULATION PARAMETERS

Database	Number of samples	Number of testing images
DoD BCRP	750	26
MIAS	322	16
DDSM	2640	73

The Fig.5 shows the identified interest points and it is clear that the region marked with red color represent the presence of cancer. The Fig.6 shows the comparison of image classification accuracy performed by different methods and it shows that the proposed method has produced more accuracy than other methods. The results are generated using the data set in Table I.The False ratio is the measure computed using the number of false positive results produced and the number of total samples available. The Fig.7 shows the comparison of false classification ratio and it shows clearly that the proposed approach has produced less false classification ratio than others.

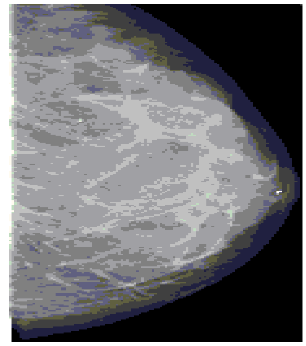

Fig. 2. Original Image

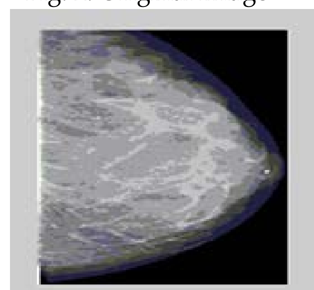

Fig. 3. Histogram equalized image

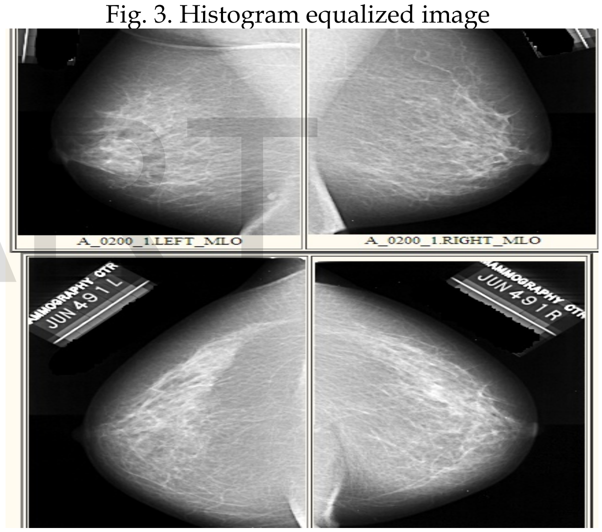

Fig. 4. Sample normal image from DDSM data set

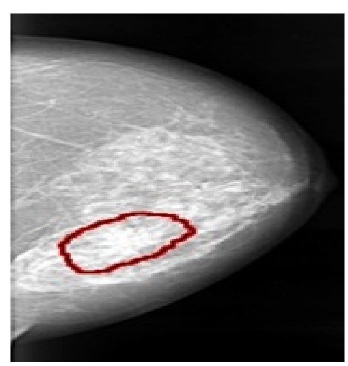

Fig. 5. Identified interest points

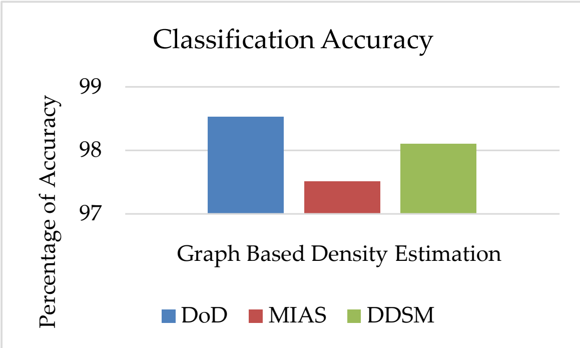

Fig. 6. Comparison of image classification

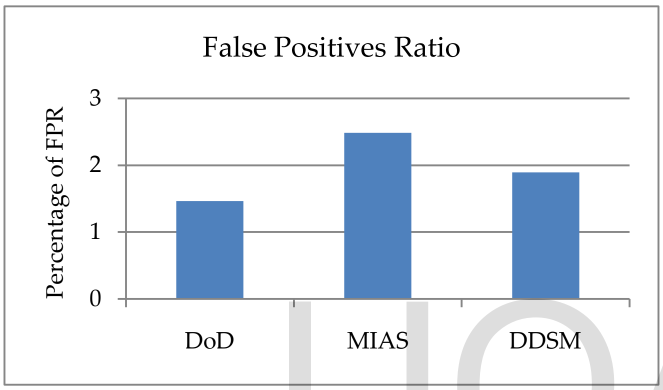

Fig. 7. Comparison of false classification ratio

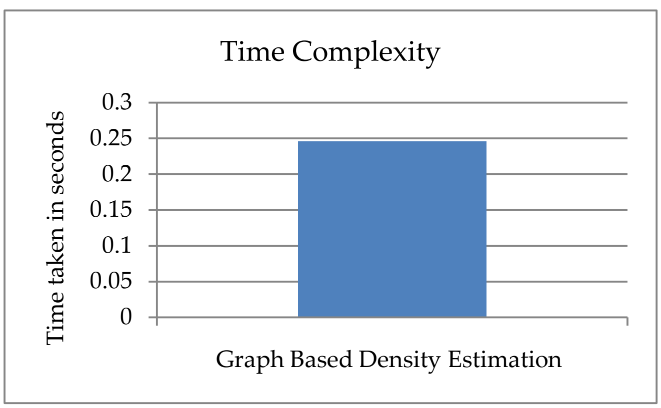

Fig. 8. Comparison of time complexity of different samples

The Fig.8 shows the time taken by different samples for mammogram classification. It shows clearly that the proposed method has taken less time compare to other methods. The efficiency of mammogram classification depends on various other parameters like time taken, complexity. The time complexity of the method is computed using the submission time and result generation time. The duration between these times are used to compute the time complexity and it varies according to the number of samples available in the training set and the number of classes present in the data set.

4 Conclusion

This proposed a new approach to determine the classification of mammographic image using k-means clustering algorithm which uses different features of the image like intensity values, shape and region features and density features to compute the feature vector. They have computed mean values of the intensity values of the pixels in the region extracted to compute the intensity mean value and the density measure is also computed in the similar fashion. The region metric is computed using the extracted region values and it has seven different features hidden. Based on the computed feature vectors the k-means clustering is used to identify the class of the input image. The proposed system produces good results and reduces the space and time complexity. It has produced classification accuracy up to 99 % which is more than other methodologies in this era. In future, it will consider other clustering models for feature extraction in order to improve the classification rate.

References

[1] Erkang Cheng, "Mammographic image classification using histogram intersection," IEEE International Symposium on Biomedical Imaging from Nano to Micro, pp.197-200, 2010.

[2] B.Pereira and K.Mokbel, "Mammary ductoscopy: past, present, and future,"International Journal of Oncology, vol. 10, no.2, pp. 112-116, 2005.

[3] R.M.Haarlick, "Statistical and structural approaches to texture," Proceedings of the IEEE, vol.67, no.5, pp.786-804, 1979.

[4] S.Julian Savari Antony and S.Ravi, "Development of Efficient Image Quarrying Technique for Mammographic Image Classification to Detect Breast Cancer with Supervised Learning Algorithm," Proceedings of the IEEE Xplore on ICACCS, pp.1-7, Dec. 2013.

[5] S.Don, Duckwon Chung, Revathy, K.Eunmi Choi and Dugki Min," A New Approach for Mammogram Image Classification Using Fractal Properties, cybernetics and Information Technologies," Bulgarian academy of sciences, vol.12, no.2, pp.69–83, 2012.

[6] D.RojaRamani, B.BenSujitha and C. Parameswari, "Digital Mammogram Tumor Pre-processing segmentation Feature extraction and classification using K Means Algorithm and Distance Based Skeletonization," International journal of computer application, vol.2, no.2, pp.143-149, 2012.

[7] S. Julian Savari Antony,Dr.S.Ravi ,"A New Approach to Determine the classification of Mammographic Image Using K-Means Clustering Algorithm",International Journal of Advancements in Research & Technology, vol.IV, Issue no.2, pp.40-44, Feb.2015.

[8] L.Zhang, Y.Ren, C.Huang, F.Liu, H.P.Brecht, R.Su and A. Oraevsky, "A novel automatic tumor detection for breast cancer ultrasound Images," Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), vol. 1, pp.401-404, (2011).

[9] B.Liu, H.D.Cheng, J.Huang, J.Tian, X.Tang and J.Liu, "Fully automatic and segmentation-robust classification of breast tumors based on local texture analysis of ultrasound images," Pattern Recognition, vol. 43, no.1, pp.280-298, 2010.

[10] I.K.Maitra, S.Nag and S.K.Bandyopadhyay, "Identification of abnormal masses in digital mammography images," International Journal of Computer Graphics, vol.2, no.1, pp.17-30, 2011.

[11] I.Buciu and Alexandru Gacsadi, "Directional Features for Automatic Tumor Classification of Mammogram Images," Biomedical Signal Processing and Control, vol.6, no.4, pp.370-378, 2011.

[12] M.A.K.Shahid, R. Sabir and M.S.Awan, "A novel approach for mammogram image analysis using the idea of shape features and texture analysis," Peak Journal of Medicine and Medical Science, vol.1, no.2, pp.4-13, 2013.

[13] SubodhSrivastava, Neeraj Sharma and S K Singh,"Empirical Analysis of Supervised and Unsupervised Filter based Feature Selection Methods for Breast Cancer Classification from Digital Mammograms," International Journal of Computer Applications, vol.88, no.8, pp.20-33, February 2014.

[14] G.vairaSuganthi and J.sutha. Article,"Classification of Breast Masses in Mammograms using Support Vector Machine," IJCA Proceedings on International Conference on Recent Advances and Future Trends in Information Technology, pp.1-6, April 2012.

[15] K.Ganesan, "One class classification of mammograms Trace Transform Functionals," IEEE Transaction on instrumentation and measurements, vol.63, no.2, pp.304-311, 2014.