Academic Journal of Cancer Research 9 (4): 70-74, 2016
ISSN 1995-8943

DOI: 10.5829/idosi.ajcr.2016.70.74

Jacobi Moments of Breast Cancer Diagnosis in Mammogram Images Using SVM Classifier

[1]S. Venkata Lakshmi, [2]K. Sujatha and [1]K. Sathyamoorthy

[1]Department of Computer Science, Panimalar Institute of Technology, Chennai, India
[2]Department of Computer Science, Anna University, Chennai, India

Abstract: Feature extraction with mathematical application gives accurate results in the image processing techniques. However there is dimensionality problem arises with multi scale transforms. Hence a simple and fast method was proposed for feature selection using Jacobi Moments with SVM classifier to distinguish between the normal and abnormal breast tissues and to classify tumors as malignant or benign. Experiments were performed on 192 mammograms from the Mammographic Image Analysis Society (mini-MIAS) database using the leave-one-out cross validation. The proposed architecture might be simple but the accuracy is higher compared with other methods in the literature. In summary, Jacobi moments are an efficient and effective way to extract a reduced set of discriminative features for breast cancer diagnosis.

Key words: Mammogram · SVM · DWT · Jacobi

INTRODUCTION

The uncontrolled growth of Breast cells is called as Breast cancer. The abnormal changes in the genes are responsible for the occurrence of cancer. The growth of the genes is affected by the mutation or abnormal change in the cells. The cell growth helps for the growth of new cells by replacing the old ones which die. At the same time mutation also “turns on” and “turns off” the cells. This change keeps growing by dividing in multiple ways resulting in reproducing similar fashion cell which is causes tumor. Mammography is a best screen tool which is used for diagnosing the early detection of Breast cancer.

Breast Cancer is very common disease and leading to death as per the survey in 2012.14.17% of women death is caused due to this breast cancer [1]. The world health organization’s International Agency for Research on Cancer (IARC) estimates that 400, 000 women die each year from this disease [2]. Early detection of breast cancer is essential in reducing life fatalities .The most effectual and viable technique for early detection of breast cancer is to identify the micro calcifications which is commonly known as the calcium deposit that leads to cancer [3, 4].

There are many research works carried out to make the simple and easy and accurate detection of these micro calcification. The proposed method deals with mammogram based on Jacobi Moments and Discrete Wavelet Transform (DWT). There are many applications available employing different orthogonal moments which are utilized to make the pattern features. In this paper the Jacobi moment is applied for extracting the feature by using the Jacobi polynomials with minimum information redundancy between the moments. The purpose of incorporating the moments was to classify the normal and abnormal images that are benign and malignant by computing the Euclidian distance between the vector moments. The approach is carried out with two different steps as first step the features are extracted and the second step is to classify the normal and abnormal cells. Mammographic Image Analysis Society (MIAS) database is used for the experiment. Affected area is separated from the image and the Jacobi moments is applied for 4*4 overlapping window and analysis made to select the range for moment to select the feature. A set of 16 abnormal images 34 normal images are taken for the test.

Image feature extraction is a significant process in the image processing. Extraction of feature can be done directly from the spatial data or by spatial domain transformation either with wave let or curvelet or riglet [5-9].

Corresponding Author: K. Sathyamoorthy, Department of Computer Science, Panimalar Institute of Technology, Chennai, India.

The World Health Organization estimates that more than 385000 women worldwide die of the disease each year [10]. Considering the severity of the disease double reading is not practiced because of the huge amount of mammogram to be analyzed by the radiologist it is necessary to develop a Computer Aided Design (CAD) system which can detect the cancer cells at early stages. Hence there are different researcher who work on this Approach [11]. Thus, computer-aided diagnosis (CAD) systems can be used to reduce the expense and to assist radiologists in mammogram interpretation [12]. Typically, a CAD system is composed of three major phases involved in them like preprocessing as the images may have some details which have to be removed, Next phase is the feature extraction and the final phase will be classification .Implementation of this Cad system reduces the time complexity of diagnosis and the accuracy is improved [13].

A new approach has been proposed with Random Forest Decision Classifier (RFDC) for classifying mammograms. A wavelet based feature extraction method is applied for the classification of mammogram. A matrix is constructed by with row and column value of the image and with the threshold value and the Euclidian distance the features are selected [14]. Nezha *et al*. [15] presented a high accuracy computer-aided diagnosis scheme. The goal of the developed system is to classify benign and malignant micro calcifications on mammograms. It is mainly based on a combination of wavelet decomposition, feature extraction and classification methodology using Fisher's linear discriminant. The contribution of wavelet decomposition is to de noise and to enhance regions of interests (ROI) containing abnormalities. Feature extraction is performed using Spatial Grey Level Dependence (SGLD) matrices. The purpose of classification is to assign an object to a certain class. Many classification methods have been described. Fisher's linear discriminant is used in this method and it is particularly useful for discriminating between two classes in a multidimensional space. Since it is based only on the first and second moments of each distribution, it is not a computationally intensive method.

There are number of works concentrated on extraction of features with wavelet and curvelet transform [16, 17] the limitation in the curvelet is that it may end up with dimensionality problem which leads to increase cost. Hence there is always a challenge for the researchers to minimize the features using curvelet.

A multi scale statistical approach to texture analysis was proposed by Kramer and Aghdasi *et al.* [18]. These techniques are used to classify micro calcifications in digitized mammograms as benign or malignant. The multi scale statistical texture signatures, based on the co-occurrence matrix, as well as wavelet based texture signatures from the regions of interest containing the micro calcifications is described. The discriminatory ability of these texture signatures is demonstrated by their ability to successfully distinguish between benign and malignant cases. Classification is performed by means of a k-nearest neighbor classifier.

After a brief study of literature it is learned that there is a need for best method for feature selection and minimization of the features and the best classification method. Hence this paper contributed a Jacobi moments for selection of features with SVM classifier and the method was compared with the other existing method.

MATERIALS AND METHODS

The main objective of this proposed work was to extract the feature based on moments. The proposed method focused on the classification of mammogram image based on Jacobi moments. Here the moments were used for feature extraction .The method has two phases like feature extraction and classification. The block diagram is as shown in the figure 1.As shown in the figure 1 the images from Mammographic Image Analysis Society (MIAS) database was taken and the Jacobi moments were calculated and the features were extracted and stored in the feature library. In case of classification the extracted feature was classified using SVM classifier.

Feature Extraction Using Jacobi Moments: Generally the while using the machine learning and pattern recognition system the preprocessing step becomes an essential. The feature extraction consists two major steps like feature selection or feature construction. Selection of features should have enough information about the normal and abnormal regions as well they should have the property to differentiate them.

Jacobi Moments: The Jacobi moments consists of Jacobi polynomial of n^{th} order with the parameter α? and β which is defined as,

$$P_n(x,\alpha,\beta) = ((\alpha+1)n/n!xx_2F_1\left(\begin{matrix}-n,n+\alpha+\beta+1\\ \alpha+1\end{matrix}\middle|\frac{1-x}{2}\right) \qquad (1)$$

where $x \in [-1, 1]$. The generalized hyper geometric function, *KrFs*, is defined.

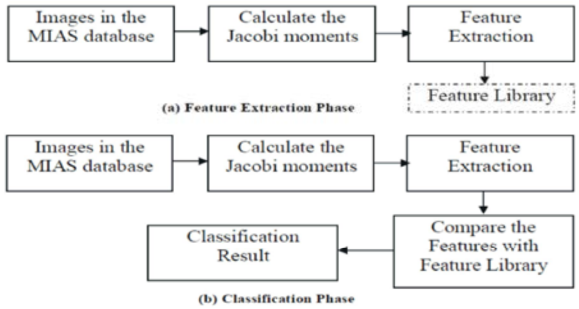

Fig. 1: Block diagram of feature Extraction and classification phase

$$K_r F_s\left(\begin{matrix} a_1,\ldots,a_r \\ b_1,\ldots,b_r \end{matrix} \middle| Z\right) = \sum_{k=0}^{\infty} \frac{(a_1,\ldots,a_r)k}{(b_1,\ldots,b_r)k} \frac{z^k}{k!} \tag{2}$$

where (a1, ar) k = (a1) k… $(ar)k$ and the Pochhammer-symbol

$$(a)k = a(a+1)\,(a+2)\,\ldots\,(a + k - 1) \tag{3}$$

With k= 1, 2, 3, … and (a)0 = 1 and thus have the explicit expression

$$P_n(x,\alpha,\beta) = \frac{\Gamma(\alpha+n+1)}{n!\Gamma(\alpha+\beta+n+1)} \sum_{m=0}^{n} \binom{n}{m} \frac{\Gamma(\alpha+\beta+n+m+m1)}{\Gamma(\alpha+m+1)} \left(\frac{z-1}{2}\right)^m \tag{4}$$

The Jacobi moment of order (p+q) of an image $f(x, y)$ with M N pixels is defined as:

$$J_{mn} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y)\overline{p_m}(x)\overline{P_n}(x) \tag{5}$$

where Pm (x) =Pm (x; 1, 1) and Pn(y) =Pn(y; 2, 2) (6)

and

$$\text{so } \overline{P_n}(x;\alpha,\beta) = P_n(x;\alpha,\beta)\sqrt{\frac{\omega(x;\alpha,\beta)}{\rho(n,\alpha,\beta)}} \tag{7}$$

$$\sum_{x=0}^{N-1} \overline{P_m}(x;\alpha,\beta)\overline{P_n}(x;\alpha,\beta) = \delta_{mn} \tag{8}$$

The term P_n (x; α, β) is the weighted Jacobi polynomials.

This section gave the brief idea about the Jacobi moments and this could be applied for the selection of feature. The input image was fed and the ROI was selected and the Jacobi moment was calculated using the equation (8) as shown in the following section, input image was of size 256*256. Though there were a number of features available the essential feature of 10 features was selected using the subset. All these important features were stored in the feature library and in the classification phase the extracted features were classified using the classifier.

Classification Using SVM: SVM classifier is trained with the feature vector being framed in the previous phase. SVM is one of the machines learning method which is very effective in classifying the binary classes using the class boundaries. The training data sample which helps to separate the data is called support vectors and the margin is the distance between the support vectors and the class boundary hyper planes.

In SVM is the supervised algorithm which intends to partition the entities with respect to the degree of margin. Consider a group of training samples with W feature vectors, which are needed to be classified into two classes C_i namely (+ve, -ve). This work considers +ve as malignant and –ve as normal. In order to classify between these two classes, a hyperplane is necessary. The hyper plane separates the entities into two different classes and is given below.

$$\psi.ji + b \geq +ve \text{ for } C_i = +ve \tag{9}$$

$$\psi.\, j_i + b \leq \text{ for } C_i = -ve \tag{10}$$

The classification accuracy is determined by the distance between the hyper planes. The distance between two different hyperplane is computed by $\frac{2}{\|\psi\|}$. The classification results are better, when the $\|\psi\|$ is minimized. The optimal hyperplane is obtained by applying Lagrange's function and is provided in the following equation.

$$f(x) = \sum_{i=1}^{Q} \alpha_i \psi_i (j_i . j) + m \tag{11}$$

In the above equation, α_i is the lagrange multiplier which tends to separate the hyperplane ψ_i (j_i. j) and the threshold to separate hyperplane is denoted by m. This makes sense that when the value of $f(x)$ is greater than 0, then the entity to be classified is +ve else –ve.

RESULTS AND DISCUSSION

From MIAS database input images are taken for both training and testing. The over lapping window size of 4*4 was considered. Then the boundary was used as a characteristic vector to classify the unknown image as normal or abnormal as well as benign or malignant.

There are 194 normal images and 24 micro calcification images available in the MIAS. The 256*256 size of image is considered .The performance is measured using classification accuracy and sensitivity and specificity. The results with SVM classifier is as shown in the Table 1.

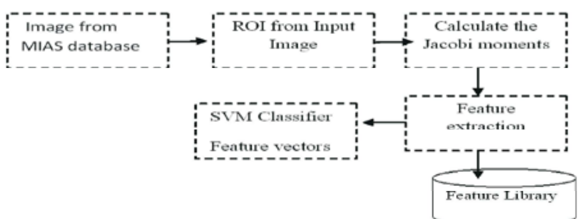

Fig. 2: Block diagram of proposed method

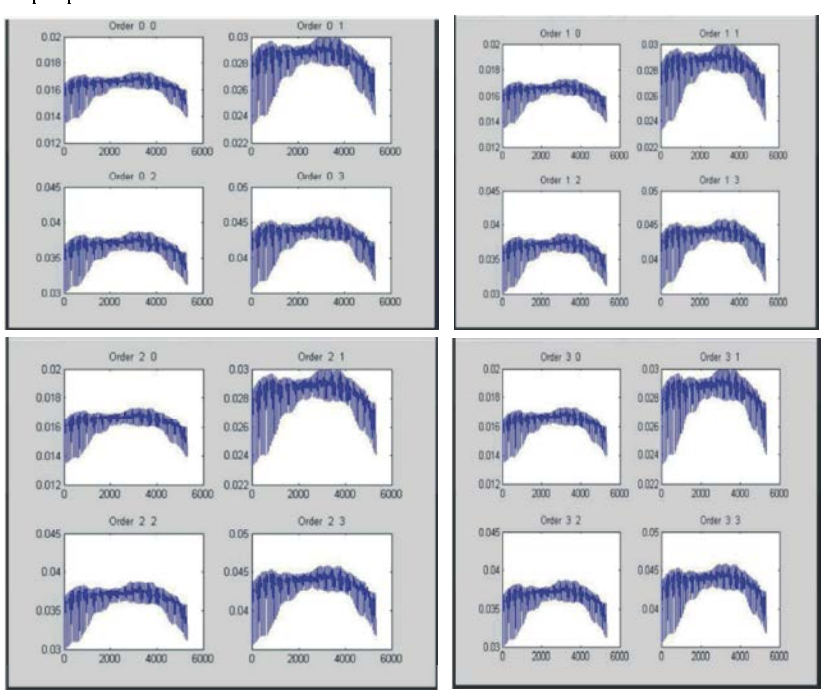

Fig. 3: Jacobi moments for 4 orders

Table 1: Performance metrics

Input images	Training Image	Testing image	Classification Accuracy
Normal	130	64	96.02%
Abnormal	14	10	92.56%
Benign	7	5	93..45%
Malignant	8	5	93.3%

CONCLUSION

In this paper the Jacobi moment was used to extract the feature and the SM classifier is used to classify and the accuracy of the results was improved. But as the Jacobi has the limitation of window size the feature reduction method was very essential to improve the accuracy. These SVM classifiers are used to test 4 kernel sizes with reduced features and might be very efficient and hence the reduction of feature could be considered in the future work. In future some of the multi resolution technique like curvelet has to be attempted to improve the accuracy.

REFERENCES

1. Pal, N.R., B. Bhowmick, S.K. Patel, S. Pal and J. Das, 2008. A multistage neural network aided system for detection of micro calcifications in digitized mammograms, Neuro Computing, 71: 2625-2634.
2. Fu, J.C., S.K. Lee, S.T. Wong, J.Y. Yeh, A.H. Wang and H.K. Wu, 2005. Image segmentation feature selection and pattern classification for mammographic micro calcifications. Computerized Medical Imaging and Graphics, 29(6): 419-429.
3. Cheng, H.D., X. Cai, X. Chen, L. Hu and X. Lou, 2003. Computer-aided detection and classification of microcalcifications in mammograms: a survey. Pattern Recognition, 36(12): 2967-2991.
4. Rashed, E.A., I.A. Ismail and S.I. Zaki, 2007. Multiresolution mammogram analysis in multilevel decomposition. Pattern Recognition Letters, 28(2): 286-292.
5. Hofvind, S., G. Ursin, S. Tretli, S. SebuodegÅrd and B. Moller, 2013. Breast cancer mortality in participants of the Norwegian breast cancer screening program, Cancer, 119: 3106-3112.
6. Dinnes, J., S. Moss, J. Melia, R. Blanks, F. Song and J. Kleijnen, 2001. Effectiveness and costeffectiveness of double reading of mammograms in breast cancer screening: findings of a systematic review, The Breast, 10: 455-463.
7. Astley, S. and F. Gilbert, 2004. Computer-aided detection in mammography, Clin. Radiol., 59(5): 390-399.
8. Moura, D. and M. Guevara López, 2013. An evaluation of image descriptors combined with clinical data for breast cancer diagnosis, Int. J. Comp. Ass. Rad. Surg., 8(4): 561-574.
9. Vibha, L. and G.M. Harshavardhan, 2006. "Classification of Mammograms Using Decision Trees", IEEE 10th International Database Engineering and Applications Symposium, December, pp: 263-266.
10. Dinnes, J., S. Moss, J. Melia, R. Blanks, F. Song and J. Kleijnen, 2001. Effectiveness and costeffectiveness of double reading of mammograms in breast cancer screening: findings of a systematic review, The Breast, 10(6): 455-463.
11. Astley, S. and F. Gilbert, 2004. Computer-aided detection in mammography, Clin. Radiol., 59(5): 390-399.
12. Moura, D. and M. Guevara López, 2013. An evaluation of image descriptors combined with clinical data for breast cancer diagnosis, Int. J. Comp. Ass. Rad. Surg., 8(4): 561-574.
13. Ibrahima Faye, BrahimBelhaouari Samir and M.M. Eltoukhy, 2009. "Digital Mammograms Classification Using a Wavelet Based Feature Extraction Method", IEEE Second International Conference on Computer and Electrical Engineering, pp: 318-322.
14. Nezha Hamdi, Khalid Auhmani and Moha M'rabet Hassani, 2008. "Design of a High-Accuracy Classifier Based on Fisher DiscriminantAnalysis: Application to Computer-Aided Diagnosis ofMicrocalcifications", International Conference on Computational Sciences and Its Applications ICCSA, pp: 267-273.
15. Gardezi, S., I. Faye and M. Eltoukhy, 2014. Analysis of mammogram images based on texture features of curveletsubbands, in: Fifth International Conference on Graph and Image Proceedings.
16. Narváez, F., G. Díaz, F. Gómez and E. Romero, 2012. A content-based retrieval of mammographic masses using the curvelet descriptor, in: Proceedings of SPIE, Medical Imaging 8315, pp: 1-7.
17. Dani Kramer and Farzin Aghdasi, 1998. "Classification of Microcalcifications in Digitized Mammograms Using MultiscaleStatistical Texture Analysis", IEEE Proceeding of the South African Symposium.
18. Hasan Abdel Qader and Syed Al-Haddad, 2007. "Fingerprint Recognition using Zernike moments", The International Arab Journal of Information Technology, 4: 372-376.