Current Medical Imaging, 2020, *16*, 703-710 703

RESEARCH ARTICLE

Classification of Benign and Malignant Breast Masses on Mammograms for Large Datasets using Core Vector Machines

J. Jeba Sonia[1] and J. Dheeba[2,*]

[1]*Amal Jyothi College of Engineering, Kerala 686518, India;* [2]*School of Computer Science and Engineering, Vellore Institute of Technology, Vellore Campus, Katpadi, India*

ARTICLE HISTORY

Received: August 18, 2018
Revised: March 25, 2019
Accepted: April 09, 2019

DOI:
10.2174/1573405615666190801121506

Abstract: ***Background*:** Breast cancer is one of the most leading causes of cancer deaths among women. Early detection of cancer increases the survival rate of the affected women. Machine learning approaches that are used for classification of breast cancer usually takes a lot of processing time during the training process. This paper attempts to propose a Machine Learning approach for breast cancer detection in mammograms, which does not depend on the number of training samples.

***Objectives*:** The paper aims to develop a core vector machine-based diagnosis system for breast cancer detection using the date from MIAS. The main motivation behind using this system is to reduce the computational and memory requirement for large training data and to improve the classification accuracy.

***Methods*:** The proposed method has four stages: 1) Pre-processing is done to extract the breast region using global thresholding and enhancement using histogram equalization; 2) identification of potential mass using Otsu thresholding; 3) feature extraction using Laws Texture energy measures; and 4) mass detection is done using Core vector machine (CVM) classifier.

***Results*:** Comparative analysis was done with different existing algorithms: Artificial Neural Network (ANN), Support Vector Machine (SVM), and Fuzzy Support Vector Machines (FSVM). The results illustrate that the proposed Core Vector Machine (CVM) classifier produced a promising result in terms of sensitivity (96.9%), misclassification rate (0.0443) and accuracy (95.89%). The time taken for training process is 0.0443, which is less when compared with other machine learning algorithms.

***Conclusion*:** Performance analysis shows that CVM classifier is superior to other classifiers like ANN, SVM and FSVM. The computational time of the CVM classifier during the training process was also analysed and found to be better than other discussed algorithms. The results achieved show that CVM classifier is the best algorithm for breast mass detection in mammograms.

Keywords: Breast cancer, computer aided diagnosis, core vector machine, laws, mammograms, ANN.

1. INTRODUCTION

Breast cancer (BC) diagnosis at an early stage is essential for reducing the mortality rate among women. As reported by the World Health Organization (WHO), BC is the leading cause of cancer deaths among women in developing countries [1]. This is because of the reason that the majority of the cases are diagnosed at a later stage. Presently, in India, more number of young women in the age group of 30 - 40 years are diagnosed with breast cancer. Data reports from various National Cancer registries show that the mortality rate is 12.7 per 100,000 women [2]. Out of all cancer deaths among females, 21.5% of women die because of breast cancer in India (statistics from WHO). The major reason for the mortality rate is due to the lack of diagnosis of disease at the early stage [3].

Digital mammography is one of the most effective imaging modalities in analysing and visualizing breast abnormalities. Mammography is capable of detecting breast lumps of 3 mm size by utilizing low dose X-ray projections [3]. As suggested by the American Cancer Society (ACS), mammogram screening has to be done once in two years after the age of 40. It is a challenging task to read a large volume of images and analyse the data accurately and consistently by radiologists. The presence of dense mass or variation in the texture is the key indicator of breast abnormality. Mass detection is challenging, due to variability in appearance, irregular margins and it is overlapped with a dense breast tissue. Breast abnormalities can be benign (non-cancerous) or malignant (cancerous). Mass with regular shape is usually benign and

*Address correspondence to this author at the School of Computer Science and Engineering, Vellore Institute of Technology, Vellore Campus, Katpadi, India; E-mail: dheeba.jacob@gmail.com

1573-4056/20 $65.00+.00

those with irregular shapes are malignant. They are characterized by shape and margin [4]. About 80% of the tumors biopsied are found to be benign. So, there is a need for a non-invasive method that discriminates the benign and malignant tumors. A Computer-Aided Detection (CAD) tool cannot completely replace human intelligence, but these tools can help the radiologist in finding information which is not visible to the human vision. The CAD system can act as the second reader in identifying the potential areas of abnormality and in classifying them as benign or malignant.

Several CAD approaches have been proposed and reviewed for classifying the abnormalities in breast [5-7]. Kooi *et al*. [5] have shown a deep learning model in the form of convolution neural network trained on a large set of data, the model learns from the data and does not rely on the domain experts. Yang *et al*. [6] developed a microcalcification detection system using modified simplified pulse coupled neural network (SPCNN) in mammograms. Hu *et al*. [8] proposed a CAD model based on Hidden Markov Model for diagnosis of microcalcifications in mammograms. Peng *et al*. [9] used 222 images from mammographic image analysis society (mini-MIAS) database and 100 random images from the independent Banco Web database to implement a confirmatory system for analysis of mammograms. They used a seeded region growing algorithm for segmentation and an artificial neural network for classification, achieving an accuracy of 96% for 100 test cases from MIAS. Li *et al*. [10] developed a mass detection algorithm based on two-concentric masks with discriminating texton, achieving an accuracy of 86.92%.

De Bruijne [11] discussed the scientific and practical challenges that need to be addressed for training strong models on limited data. Swiderski *et al*. [12] presented an automatic system for mammogram analysis to recognize normal, malignant and benign data using principal component analysis. Dhungel *et al*. [13] described a complete minimal user intervention CAD system for detection, segmentation and classification of masses in mammograms; detection algorithm uses a cascade of deep learning and random forest for generating the mass candidate by achieving a sensitivity of 98% and specificity of 70%. Sami Dhahbi *et al*. [14], used Curvelet transform to extract texture features of 252 images from MIAS and 11553 from Digital Database for Screening Mammography (DDSM); k-nearest neighbor is used to distinguish between malignant and benign masses achieving an accuracy 91.27%.

A CAD system can act as a second reader in clinical setting only if it is robust to false positives and false negatives. The effectiveness of the system lies in the features extracted and the learning algorithm employed to make the decision. Features are extracted based on the knowledge of the radiologists such as the contrast details of the benign and malignant lesions, speculated patterns as in case of architectural distortions and the border regularity. Hu *et al*. [15] dealt with the different characteristics of the masses. Space occupying lesions can be of different classes [16]; Spiculated masses (SPIC), Circumscribed (CIRC) masses, Asymmetry (ASYM), Architectural Distortions (ARCH) and ill-defined masses (MISC). SPIC and ARCH have lines originating from the centre and moving towards the margin showing a textured pattern. CIRC masses can have a definite shape or can be irregular; round and oval masses are benign and irregular masses suggest a greater likelihood of malignancy. Ill-defined masses and ASYM are characterized by gray level features.

In this paper, we employed a segmentation algorithm to segment the mass region using region growing algorithm and classify the masses based on shape, grey level and texture features with Core Vector Machine (CVM) Classifier. CVM is a kernel-based algorithm suited for large data classification applications [17]. As mammographic images taken in screening centres are large in number, there should be an efficient classification approach that can handle a large volume of data. CVM has the advantage in dealing with data in high dimensional space. Chang [18] has dealt with the geometric relationship between the CVM and SVM. The performance of the SVM and CVM in terms of classification is more or less the same. The advantage of using CVM in large data set is the space complexity and time complexity. CVMs do not depend on the number of data samples, it is independent of the dataset and hence the algorithm converges at $O(\frac{1}{\varepsilon})$ iterations, $\varepsilon \in (0,1)$. The running time of SVM greatly depends on the training samples used and has a computational complexity of $O(n^3)$, where n is the training data samples. A computerized decision support system using CVM helps in analysing medical imaging data and aids the physicians in diagnostic management.

The rest of the paper is organized as follows;

2. MATERIALS AND METHODS

The paper aims to develop a core vector machine-based diagnosis system for breast cancer detection using the date from MIAS. The main motivation for using this system is to reduce the computational and memory requirement for large training data and to improve the classification accuracy. Here, we reformulated the SVM's quadratic problem (QP) as a minimum enclosing ball problem to obtain the solution. The proposed idea can help the CAD system to increase the accuracy at a reduced computational time. Fig. (**1**) shows the proposed framework for a CAD system using the CVM classifier. As shown in the figure, the first step was to obtain the mammogram images. Here, we used mammograms available in the standard database mini-MIAS. In the second phase, the breast region is segmented out and the region is enhanced. The potential areas of abnormalities are detected using the Otsu thresholding method. The features are extracted using the Laws texture analysis. Based on the features extracted, the CVM classifier is trained and the abnormalities in mammograms are detected.

2.1. Database Used

Experiments were performed on 294 images from MIAS database [16] comprising CIRC, SPIC, ARCH, ASYM, MISC, and NORM. The MIAS database has 322 mammograms; from these, we used 294 images for our experimental purpose. The images were digitized to 200-micron pixel edge and resized to 1024 x 1024 pixel. The database is pro-

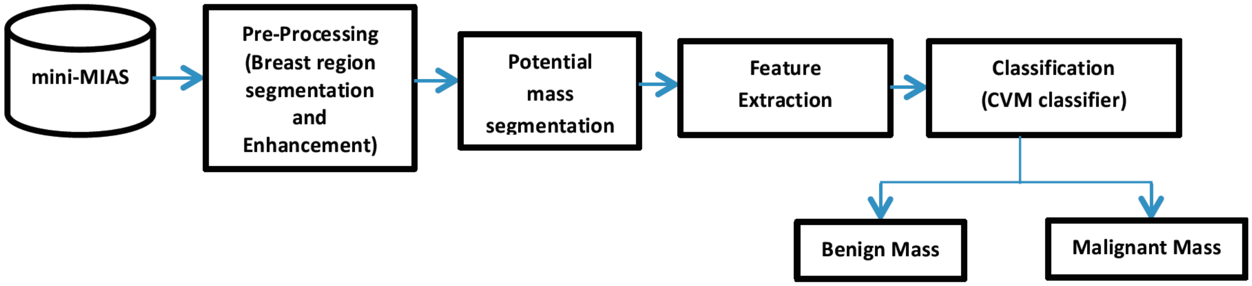


Fig. (1). Framework of a CAD system using CVM classifier.

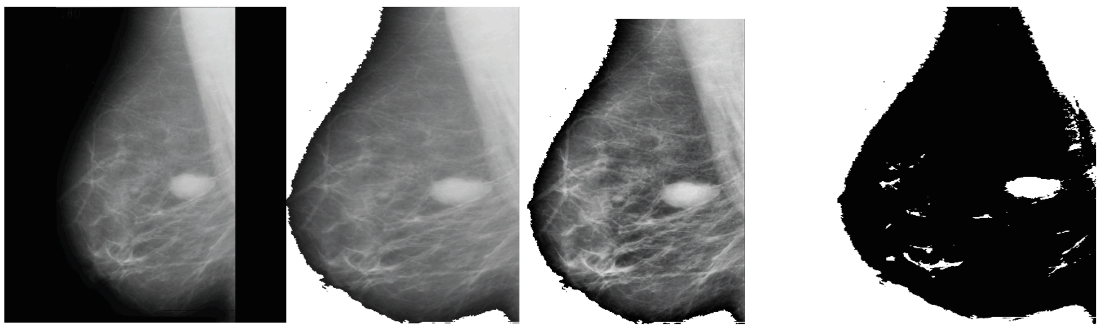


Fig. (2). (**a**) Original mammogram image. (**b**) Segmented Breast Region. (**c**) Contrast-Enhanced Image (CLAHE). (**d**) Segmentation of potential regions using Otsu thresholding.

vided with the ground truth by showing the x and y – coordinate values.

2.2. Pre-processing

Pre-processing is done in the mammogram to separate the breast region from the background. The background contains markers and artifacts, which need to be removed to avoid misclassification. We developed a global thresholding mechanism to segment out the background image and separate the breast region. Global thresholding can be applied for segmenting the background as the background of the mammogram is mostly of 0 pixel value. Considering a mammogram image $I(x, y)$, the breast region from the mammogram was removed by setting a threshold value. The objects in the mammogram image whose values are greater than the threshold T are considered as the breast region. The value of T is chosen to be 30, based on the knowledge from the expert radiologist. ACS recommends MRI for higher-risk women, but it is expensive and many women do not have access to this technology. As an alternative way, we need to look for a cost-effective and economical mechanism and sustainable methods for healthcare. Mammography is a low-cost imaging technology which is made available in rural areas also. To facilitate with high performance and accuracy, contrast enhancement is integrated in the X-ray mammograms, which act as a powerful tool like magnetic resonance imaging (MRI). As a pre-processing stage, Contrast limited adaptive histogram equalization (CLAHE) [19] is used to improve the local detail of the mammogram. The algorithm divides the image into non-overlapping tiles (8 x 8). Contrast enhancement limit is applied to each tile; the value should be between [0, 1]. As higher limits will lead to high contrast and even noises in the images will be enhanced so an optimal value of 0.005 is taken as clip limit. Now, histogram equalization is applied in each tile with bilinear interpolation. Fig. (**2a**) shows the detected breast region using global thresholding and Fig. (**2b, c**) shows the contrast-enhanced image.

2.3. Potential Mass Segmentation

Abera *et al*. [20] evaluated some of the thresholding mechanisms in segmenting a two-phase porous media. Medina-Carnicer *et al*. [21] used thresholding mechanism for edge detection in images and found that Otsu is the best segmentation technique. In this paper, we employed an Otsu thresholding mechanism to segment the potential regions in the mammograms. Otsu method depends on the data intensity and not on the shape and geometric characteristics; this helps in segmenting the potential regions and highlights the pixels which have to be analysed further. The thresholding method chosen is simple to compute, robust and adaptable to many computer vision applications [22]. Fig. (**2**) shows the segmentation results using Otsu thresholding. In Fig. (**2d**) the pectoral muscle region is removed to avoid misclassification.

2.4. Feature Extraction

Texture features play a major role in distinguishing malignant cells from benign cells in images. The changes in the image texture are difficult to diagnose by humans. An automated CAD system will help them in extracting the texture properties from the mammogram and provide information for the classification process. In this paper, Laws texture energy measure (LTEM) for extracting the local texture features was utilized [23]. A set of $1-D$ convolution kernels was applied to get the $2-D$ convolution kernel masks. The $1-D$ kernels are given as, $5\ (Level) = [1\ 4\ 6\ 4\ 1]$,

1	4	6	4	1
4	16	24	16	4
6	24	36	24	6
4	16	24	16	4
1	4	6	4	1

L5L5

-1	-4	-6	-4	-1
-2	-8	-12	-8	-2
0	0	0	0	0
2	8	12	8	2
1	4	6	4	1

E5L5

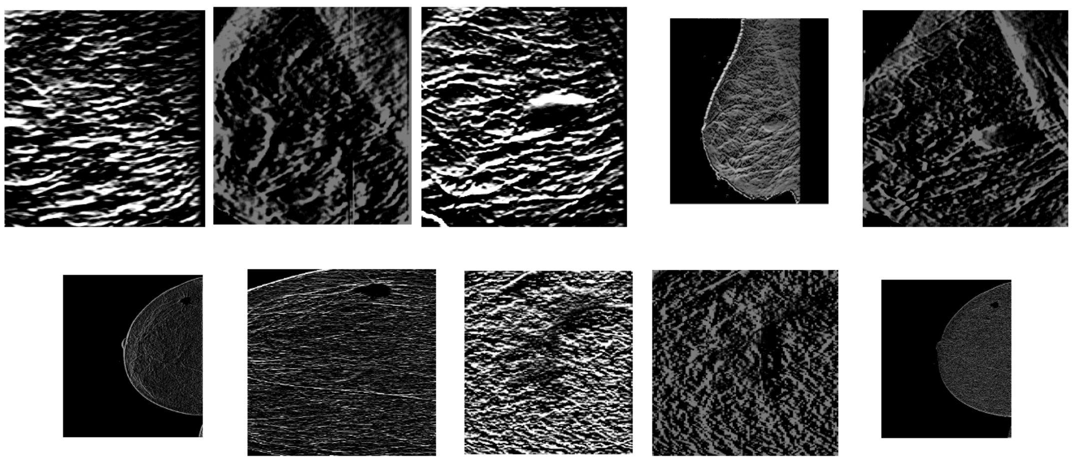


Fig. (3). TEM filtered mammogram images.

$E5\ (Edge) = [-1\ -2\ 0\ 2\ 1]$, $S5\ (Spot) = [-1\ 0\ 2\ 0\ -1]$, $R5\ (Ripple) = [1\ -4\ 6\ -4\ 1]$, $W5\ (Wave) = [-1\ 2\ 0\ -2\ -1]$. Convolution of the $1 - D$ kernel gave a result of 25 two dimensional masks.

The mask obtained was convoluted with the original image to produce a number of images. The local mask was applied onto a $15\ x\ 15$ window for extracting the texture features. The 25, $5\ x\ 5$ masks was applied to get filtered images; these images are called texture energy measure. They estimate the energy within the passband of the associated filters. The distinct pairs are taken out and their mean value is calculated. The distinct pairs are replaced with the mean. Fig. (**3**) shows the TEM filtered images of the abnormal images. Based on the kernel mask, these images are obtained by convoluting the mask with the image pixel values.

2.5. Classification

CVM is a kernel-based algorithm best suited for large data classification applications. As the images in the screening centers are large in number, there should be an efficient classification approach that can handle a large volume of data. CVM has the advantage in dealing with data in high dimensional space [24]. CVM converts Quadratic programming problem of SVM into a Minimum Enclosing Ball (MEB) problem. The approximate optimal solution of CVM is obtained by utilising an approximate MEB approach. The classification results of the CVM are more or less the same as that of SVM. As MEB problems are independent of the number of data samples used and the computational complexity, while in terms of space and speed, it is much faster than SVM. The basis of the MEB problem and approximation is given in a study [24]. The basis of CVM is discussed in this section.

For a given set of points (features) $S = \{p_1, p_2, \dots, p_n\}$, where $p_i \in R$, we find the smallest circle with radius R^* and center C^* that encloses the points in S. Points which have similar properties are enclosed together. All the points are mapped to a high dimensional feature map with function φ. The mapped points (features) are represented as $S_\varphi = \{\varphi(p_1), \varphi(p_2), \dots, \varphi(p_n)\}$. The smallest enclosing circle that encloses all the mapped points S_φ is given as a minimum function. For a given point p and $\varepsilon > 0$, a sequence of centers is generated c_i, where $i = 1, 2, \dots, \frac{1}{\varepsilon}$.

Algorithm 1

CVM

Initialize S_0, c_0, r_0

$\forall$ training points p

do

Find p farthest away from c_t

$S_{t+1} = S_t \cup \{p\}$

$c_{t+1} = cMEB(S_{t+1})$

$r_{t+1} = rMEB(S_{t+1})$

$t = t + 1$

Until there are no training points

The idea is to increase the ball $B(c_t, r_t)$ at the current iteration by finding a point outside the ball $B(c_t, (1+\varepsilon)r_t)$. The iteration is repeated until all the points in S are covered. For $\varepsilon > 0$, a subset of $X \subset S$, is called a ε – core set. The core set is defined as a subset of input points from the original training samples and when training is applied on these points, it gives an optimized result.

$$(c^*, R^*) = {}^{min}_{c,R} R^2$$

$$\|c - \varphi(x_i)\|^2 \leq R^2 \; \forall i$$

Its dual problem is given as:

$${}^{max}_{\alpha} -\alpha^T K \alpha$$

Subject to, $\alpha^T 1 = 1, \alpha \geq 0$

$$K = [k(x_i, x_j)]_{m \times m} = [\varphi(x_i)^T \varphi(x_j)]_{m \times m}$$

$$m{\times}m \rightarrow kernel\ matrix$$

Any QP problem of the form ${}^{max}_{\alpha} -\alpha^T K \alpha$ is considered as an MEB, if kernel satisfies $k(x, x) = \rho$, where ρ is constant. The core set computed depends only on ε and not on the size of the training samples. $(1+\varepsilon)$ approximation is used for computing MEB, so that it becomes efficient for very small positive values in large data set. The time complexity of CVM is formulated as:

$O(\frac{1}{\varepsilon})$ iterations

$T = \sum_{t=1}^{\tau} O(tm + t^3)$

$= O(\tau^2 m + \tau^4)$

$= O\left(\frac{m}{\varepsilon^2} + \frac{1}{\varepsilon^4}\right)$

The training converges at $O\left(\frac{1}{\varepsilon}\right)$ iterations. The classification results applied to mammograms using SVM and CVM are more or less the same. But, the computational complexity was highin terms of space and speed was much faster than SVM, as the minimum enclosing ball problems are independent of the number of data samples used. For problems with the dimension of samples greater than 30, solutions were not good. CVM is best suited for a large dataset as it does not depend on the training samples used.

3. RESULTS AND DISCUSSION

The proposed CVM classifier was evaluated using database 53 benign mammograms and 38 malignant mammograms from the mini-MIAS database. The classifier was trained with 70% of the images from mini-MIAS and the remaining 30% of the images were taken for testing. Linear, polynomial, and Gaussian function kernels are the popular kernels used in support vector machine. Here for the CVM classifier, radial basis function kernel function was used. CVM is trained with 25 texture energy measures extracted from the segmented breast region. For the Law's texture energy measures, the mean of the Law's energy and standard deviation of the Law's energy were computed. Texture analysis performed on mammograms shows that TEM has the capability to identify micro patterns in images and it is used for detecting edges, spots, ripples, levels, waves and spots at chosen vectors in both the horizontal and vertical direction. The set of features, which contains real numbers, is mapped in the mammograms which characterise the abnormal mammogram. For the purpose of discriminating between the benign and malignant masses, classification was done using the CVM classifier. Fig. (**4**) shows the minimum enclosing ball obtained from the feature vector. The red stars show the data points (25 features) obtained from the Laws feature extraction. The blue circle shows the minimum enclosing ball that can enclose the data points (features). The three points on the

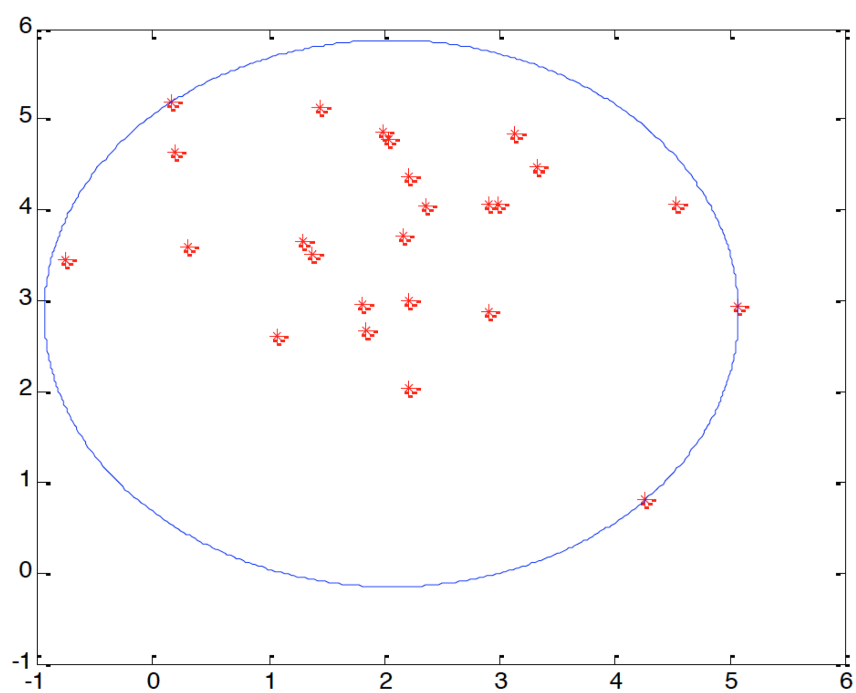
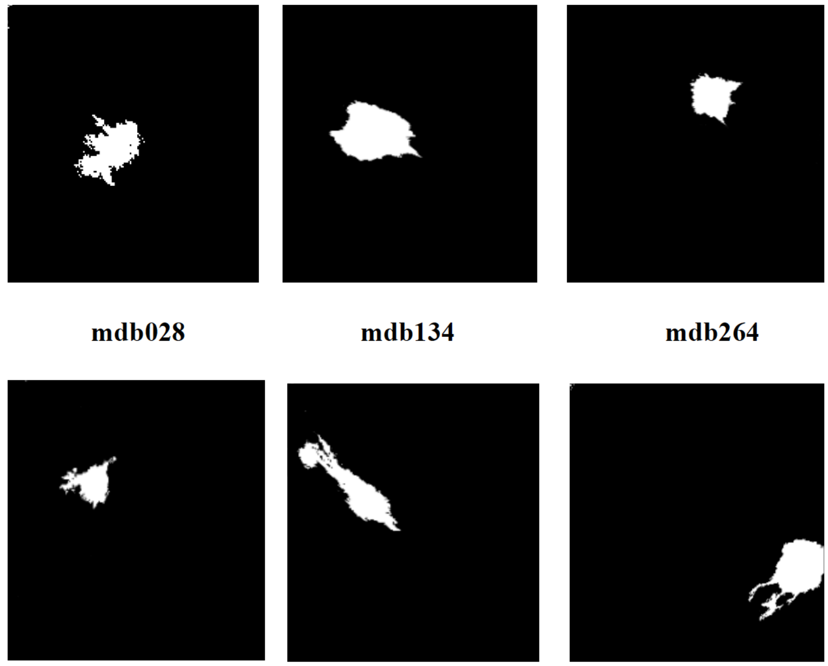


Fig. (4). MEB for 25 TEM features. *(A higher resolution / colour version of this figure is available in the electronic copy of the article).*



mdb028 **mdb134** **mdb264**

mdb155 **mdb274** **mdb271**

Fig. (5). Classification of Breast Masses using CVM classifier.

ball are the core vectors of the MEB. The classifier was trained and tested using $\varepsilon = 0.25$ and an upper bound of $C = 10$ was used for abnormal samples. $C = 100$ was given for normal samples. Fig. (**5**) shows the classification results, the images were taken from the mini-MIAS database and the mammograms shown have malignant masses. The accuracy of the classifier was assessed by comparing the results obtained from the proposed CAD system with the ground truth available in the database description. The evaluation metrics of the proposed detection methodology is based on classification accuracy (A), Sensitivity (Se), Specificity (Sp), Youden index (J) [25], Area under the ROC curve (AUC), misclassification rate (MR) and FN - False Negative. Sensitivity is measured as the ratio of the accurately recognized abnormal cases - True Positive (TP), to all the representatives of this class $(TP + FN)$, FN is False Negative. Similarly, specificity is the ratio of negative samples correctly classified by the classifier as true negative True Negative (TN), to all the representatives of this class $(TN + FP)$, FP which is False positive. Accuracy refers to the ratio of the correctly diagnosed cases $(TP + TN)$ to all the examined cases $(TP + FP + TN + FN)$. The receiver operating curve is the plot of the $sensitivity$ versus $1 - specificity$. The classifier is considered to perform well if its ROC curve ascends towards the upper left corner of the graph. The area under the ROC curve (AUC) is analysed if the value is 1 or nearer to 1 then, the classifier is thought to perform well.

The discriminating ability of the proposed classifier was evaluated and the performance of the proposed classifier was tested by comparing with different classification algorithms (ANN, SVM, FSVM). ANN was built using 25 input neurons, 120 hidden neurons and 2 output neurons. For training the images, backpropagation neural network was utilized. One main drawback of using ANN is, determining the structure of the neural network. ANN achieves an appropriate network structure through trial and error. Moreover, the system has to be fine-tuned to achieve better performance. Hence the training time increases, as ANN takes 32.741 sec for training. SVM can solve complex problems with the help of kernels. A radial basis kernel function is used for our training purpose. SVMs can handle high dimensional data, but it takes a long training time in large datasets. The training time taken by SVM is 20.94. FSVM helps to handle outlier data using fuzzy membership function. Each data point is assigned with a fuzzy value. For this study, we used a Gaussian membership function. The training time is higher when compared with the SVM, because of the fuzzy value allocation to all the data points. The existing methods described in the literature have a drawback of handling a large dataset. The CVM classifier helps to solve the drawbacks of the existing method. It is capable oftraining the data in a very minimum time 8.75sec, which is less when compared with the other algorithms in the literature.

Table **1** shows the performance obtained using the different classifier. The sensitivity of discriminating abnormal masses from the normal ones using CVM classifier is higher by a value of 96.9% at a specificity rate of 92.9%. The increase is reflected in the ROC curve, which has an AUC value of 0.9784. The main motivation in using the CVM classifier is to reduce the time complexity during the training process. Table **2** shows the comparison of the time taken for training the samples using ANN, SVM, FSVM and CVM. The elapsed time for training the feature set using different classifier was compared. The time taken for training the features using CVM is the least, as the classifier does not depend on the dimension/size of the training data. The increase

Table 1. Performance measures in detection of abnormal masses in mammograms.

Classifier/Metrics	ANN	SVM	FSVM	CVM
Se (%)	87.705	88.90	93.99	96.9
Sp (%)	77.77	84.84	92.308	92.9
A (%)	85.443	89.342	93.671	95.89
AUC	0.861	0.9153	0.9595	0.9784
Y	0.65483	0.7284	0.8625	0.8978
MR	0.1455	0.126	0.0633	0.0443

Table 2. Elapsed time for training the mammograms.

Classifier	ANN	SVM	FSVM	CVM
Elapsed time (Sec)	32.741	20.94	28.20	8.75

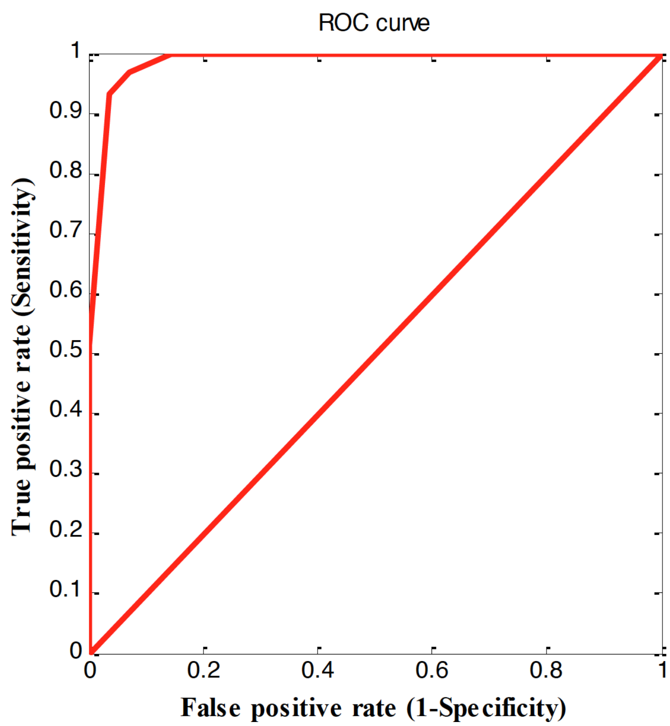


Fig. (6). ROC curve for CVM classifier. *(A higher resolution / colour version of this figure is available in the electronic copy of the article).*

in the elapsed time of the FSVM classifier when compared with SVM is because of the time taken to calculate the membership function (Fig. **6**).

CONCLUSION

This paper focuses on developing a CAD system for detection of breast masses in mammograms using a classifier which can work well in large datasets. Laws texture energy measure is used for extracting texture features from the extracted potential mass regions. The features extracted are used to train the CVM classifier. Performance analysis shows that CVM classifier is superior to other classifiers like ANN, SVM and FSVM. The computational time of the CVM classifier during the training process was also analysed and found to be better than other discussed algorithms. The results achieved shows that CVM classifier is the best algorithm for breast mass detection in mammograms. In this paper, we considered the breast abnormality detection as a two-class problem. In the future, classification based on the BI-RADS category can be implemented to classify the pixels based on the levels of severity.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No animals/humans were used for studies that are base of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

Not applicable.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

[1] Balleyguier C, Boyer B, Athanasiou A, Vanel D, Sigal R. Understanding CAD (computer-aided diagnosis) in mammography. J Radiol 2005; 86(1): 29-35. http://dx.doi.org/10.1016/S0221-0363(05)81319-8 PMID: 15785414

[2] Malvia S, Bagadi SA, Dubey US, Saxena S. Epidemiology of breast cancer in Indian women. Asia-Pacific J Clin OncolPubl 2017; 2017: 9.

[3] Ng KH, Muttarak M. Advances in mammography have improved early detection of breast cancer. J-Hong Kong Coll Radiol 2003; 6: 126-31.

[4] Timp S, Karssemeijer N. A new 2D segmentation method based on dynamic programming applied to computer aided detection in mammography. Med Phys 2004; 31(5): 958-71. http://dx.doi.org/10.1118/1.1688039 PMID: 15191279

[5] Kooi T, Litjens G, van Ginneken B, *et al.* Large scale deep learning for computer aided detection of mammographic lesions. Med Image Anal 2017; 35: 303-12. http://dx.doi.org/10.1016/j.media.2016.07.007 PMID: 27497072

[6] Yang Z, Dong M, Guo Y, *et al.* A new method of microcalcifications detection in digitized mammograms based on improved simplified PCNN. Neurocomputing 2016; 218: 79-90. http://dx.doi.org/10.1016/j.neucom.2016.08.068

[7] Ferlay J, Shin H-R, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer 2010; 127(12): 2893-917. http://dx.doi.org/10.1002/ijc.25516 PMID: 21351269

[8] Hu K, Yang W, Gao X. Microcalcification diagnosis in digital mammography using extreme learning machine based on hidden Markov tree model of dual-tree complex wavelet transform. Expert Syst Appl 2017; 86: 135-44. http://dx.doi.org/10.1016/j.eswa.2017.05.062

[9] Peng W, Mayorga RV, Hussein EMA. An automated confirmatory system for analysis of mammograms. Comput Methods Programs Biomed 2016; 125: 134-44. http://dx.doi.org/10.1016/j.cmpb.2015.09.019 PMID: 26742491

[10] Li Y, Chen H, Wei X, Peng Y, Cheng L. Mass classification in mammograms based on two-concentric masks and discriminating texton. Pattern Recognit 2016; 60: 648-56. http://dx.doi.org/10.1016/j.patcog.2016.06.021

[11] de Bruijne M. Machine learning approaches in medical image analysis: from detection to diagnosis. Med Image Anal 2016; 33: 94-7. http://dx.doi.org/10.1016/j.media.2016.06.032 PMID: 27481324

[12] Swiderski B, Osowski S, Kurek J, *et al.* Novel methods of image description and ensemble of classifiers in application to mammogram analysis. Expert Syst Appl 2017; 81: 67-78. http://dx.doi.org/10.1016/j.eswa.2017.03.031

[13] Dhungel N, Carneiro G, Bradley AP. A deep learning approach for the analysis of masses in mammograms with minimal user intervention. Med Image Anal 2017; 37: 114-28. http://dx.doi.org/10.1016/j.media.2017.01.009 PMID: 28171807

[14] Dhahbi S, Barhoumi W, Zagrouba E. Breast cancer diagnosis in digitized mammograms using curvelet moments. Comput Biol Med 2015; 64: 79-90. http://dx.doi.org/10.1016/j.compbiomed.2015.06.012 PMID: 26151831

[15] Hu K, Gao X, Li F. Detection of suspicious lesions by adaptive thresholding based on multiresolution analysis in mammograms. IEEE Trans Instrum Meas 2011; 60(2): 462-72. http://dx.doi.org/10.1109/TIM.2010.2051060

[16] Suckling J, Parker J, Dance D, *et al.* The mammographic image analysis society digital mammogram database. Exerpta Medica Inter Congress Series 1994; 1069: 375-8.

[17] Wang D, Zhang B, Zhang P, Qiao H. An online core vector machine with adaptive MEB adjustment. Pattern Recognit 2010; 43(10): 3468-82. http://dx.doi.org/10.1016/j.patcog.2010.05.020

[18] Chang L. The geometric relationship between core vector machine and support vector machine. In: 7th World Congress on Intelligent Control and Automation; Chongqing, China; IEEE; 2008; pp. 4439-43. http://dx.doi.org/10.1109/WCICA.2008.4593638

[19] Zuiderveld K. Contrast limited adaptive histogram equalization. Graph Gems 1994; 1994: 474-85. http://dx.doi.org/10.1016/B978-0-12-336156-1.50061-6

[20] Abera KA, Manahiloh KN, Nejad MM. The effectiveness of global thresholding techniques in segmenting two-phase porous media. Constr Build Mater 2017; 142: 256-67. http://dx.doi.org/10.1016/j.conbuildmat.2017.03.046

[21] Medina-Carnicer R, Madrid-Cuevas FJ, Fernández-Garcia NL, Carmona-Poyato A. Evaluation of global thresholding techniques in non-contextual edge detection. Patt Recognit Lett 2005; 26(10): 1423-34.

[22] Otsu N. A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern 1979; 9(1): 62-6. http://dx.doi.org/10.1109/TSMC.1979.4310076

[23] Laws KI. Texture energy measures. In: Proceedings of the Image Understanding Workshop 1979; pp. 47-51.

[24] Tsang IW, Kwok JT, Cheung P-M. Core vector machines: Fast SVM training on very large data sets. J Mach Learn Res 2005; 6(3): 363-92.

[25] Youden WJ. Index for rating diagnostic tests. Cancer 1950; 3(1): 32-5. http://dx.doi.org/10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3 PMID: 15405679