Websites:
http://www.sciencepub.net
http://www.sciencepub.net/researcher

Emails:
editor@sciencepub.net
marslandresearcher@gmail.com

Researcher

MARSLAND PRESS
Multidisciplinary Academic Journal Publisher

A Study On Mammography Computer Aided Diagnosis System Using Machine Learning Methods

Dr.K.Dinakaran[1] ,V.Sivakrithika[2]

[1.] Department of Computer Science and Engineering, RMD Engineering College, Tamilnadu, India
[2.] Department of Computer Science and Engineering ,TRP Engineering College, Tamilnadu , India
[1.]dina_72@yahoo.com, [2.]sivakrihtikav@gmail.com

Abstract: Early diagnosis is an important aspect of successful treatment for breast cancer. Mammogram is the most reliable imaging technique available. It is a challenging task for radiologists to detect the abnormalities in the mammograms. Computing helps the radiologists in diagnosing the abnormalities in the mammogram. Computer Aided Diagnosis System involves computerized biomedical image analysis to classify the mammography into benign or malign. In a decade of research work number of algorithms had been proposed to classify the image that employ data mining techniques, image processing methods, machine learning methods and pattern recognition. In this paper such algorithms in previous research work is studied and their performance is discussed.
[K. Dinakaran, V. Sivakrithika. **A Study On Mammography Computer Aided Diagnosis System Using Machine Learning Methods.** *Researcher* 2021;13(4):54-65]. ISSN 1553-9865 (print); ISSN 2163-8950 (online). http://www.sciencepub.net/researcher 8. doi:10.7537/marsrsj130421.08.

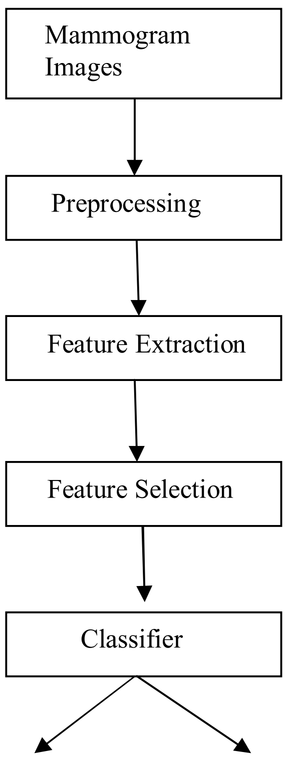
Keywords: Medical Image Analysis, Data Mining, Cancer Diagnosis, Mammogram Classification, Computer Aided Diagnosis ,Feature Extraction, Feature selection, Classifiers, Mammogram Abnormalities.

1. Introduction

Breast cancer is one of the most leading fatal diseases among women. Early detection in breast cancer helps to a successful treatment. Regular screening could find the breast cancer earlier. Screening involves the techniques like ultrasound, MRI, X-rays and Mammograms. Among all these techniques mammograms has been very successful in detecting cancerous tissues in earlier stage and has less side effects. But detection accuracy depends upon the radiologist. Radiologists need to interpret thousands of mammograms, due to over load and over looked they might miss some diagnosis. And there is a need that the radiologist should have strong knowledge and a specialist in this area. In the wide range of computer applications, the computer processing aids in medical image interpretation also. It could give a second opinion for radiologist and eliminate the unnecessary biopsies. And the system is called as computer aided diagnosis (CAD). CAD helps in reducing the cost and improving the effectiveness in double reading [1] of mammograms. Mortality rate has been highly reduced due to usage of CAD system. CAD has wide spread use in diagnosis of diseases. Over a decade, research work had been carried out in detecting the abnormalities in medical images and classifying it as benign or malign in case of cancer diagnosis. CAD can be of two variations CADe and CADx [2]. CADe is used to detect the abnormalities and CADx used to diagnose the abnormalities. Radiologists also make use of CAD system to analyze the mammograms in diagnosing breast cancer. CAD system is used as second readers in diagnosis by the radiologists. In a decade research it has been seen that use of CAD system improved the detection rate when compared with radiologists.

CAD system involves image processing methods, image enhancement methods, stochastic modeling methods, multiscale decomposition methods and machine learning methods. Earlier CAD system made use of statistical methods, wavelets and linear classifiers for classifying the abnormalities. But making use of data mining techniques and evolutionary algorithms in machine learning method made CAD system to be very successful. Machine learning is the computer algorithm that interprets pattern from training data and make decisions. These computational techniques along with soft computing methods produce precise, accurate and fast solutions. In the previous research, number of algorithms had been proposed. Each method had its advantage, such as high accuracy, speed or even be a cost effective method. But all the work involves in detecting the suspicious area that require further investigation and to classify such area as Malign (cancerous) or Benign (non cancerous).In this paper a study is made on various methods employed in CAD system for analyzing mammograms. The typical steps in mammography CAD system is shown in the figure1.

The rest of the paper is organized as follows. Section 2 depicts the breast image captured using mammograms and describes the various sign of abnormalities in mammogram. Preprocessing is discussed in Section 3. Feature Extraction and Features Selection is presented in Section 4 and Section 5 respectively. In Section 6 the various existing classifier is discussed. Section 7 depicts how the classifier performance is evaluated.



2. Mammograms

Mammograms are similar to X-rays. Mammogram is the most reliable technique in diagnosing breast cancer, since it has fewer side effects compared to other methods. There are two types of mammogram [3]: film mammography and digital mammography. Digital mammography is more effective than film mammography. Even in digital mammography the visual clues of abnormalities would be subtle. And therefore radiologist finds it difficult to interpret mammography. It is a challenging task even for specialist. Computer processing helps the radiologist to interpret the images. There are two views in mammograms mediolateral oblique (MLO) and Cranio Caucial(CC) views. In mammography, breast abnormalities [4] could be detected with earlier signs such as

1. Mass
2. Calcification
3. Architectural Distortion
4. Bilateral Asymmetry

Mass can be defined as space occupying lesion. Speculated mass in the image indicates high probability of cancer. Discriminating lesion and mass is challenging task. Calcifications are deposits of calcium and appear as a dark spot with high intensity in mammograms. Micro cluster of calcification is the most common sign found in malign images. The third most common sign is architectural distortion. Distortion of architecture along with mass or calcification could have more possibility for malign. Bilateral Asymmetry is used to detect primary cancer.

Number of research papers have been published in detecting the masses and micro calcification in mammograms. But only limited number of research work has been done in identifying the architectural distortion though it is the most associated early sign of breast cancer.

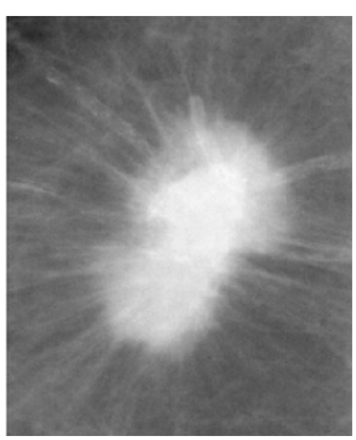


Figure 1. Speculated Mass

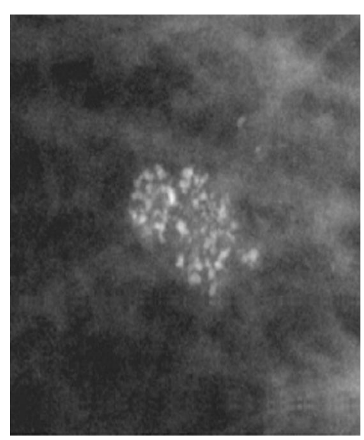


Figure 2. Micro clusters

Figure 3. Architectural Distortion

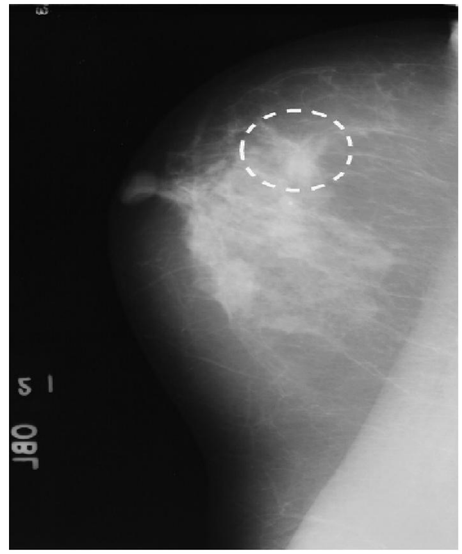


Figure 4 .Bilateral Asymmetry

3. Preprocessing

First step in the CAD system is preprocessing. The mammogram obtained may contain noise and artifact. Preprocessing is the process in which denoising and enhancement of image is done before the image is analyzed. Removal of these could help in interpreting the mammogram easily and accurately. Even though digital mammography has more advantages it has low contrast. So the image is also enhanced with respect to the contrast. And background is removed using median filter method .Using the segmentation methods such as water shed method [5], region growing method [6] the suspicious area could be separated.

From this area meaningful features could be extracted for classification. Number of image processing methods had been proposed in the literature such as median filter, contrast enhancement threshold methods and fixed intensity range for preprocessing the mammograms. Adaptive image enhancement [7] outperforms the classical image enhancement techniques. Adaptive image enhancement considers the local context of the mammograms while detecting micro calcifications. In case of detecting the mass the unenhanced images would give better results because sometimes enhancement techniques could change the anatomical information. Interpreting digital mammography does not require enhancing methods often.

Features	Description
Histogram Features	smoothness, uniformity, third moment , entropy are extracted from the histogram of the mammogram
Intensity Features	Mean median, mode, standard deviation , variance calculated from the intensity values of the pixel
Co-occurrence Features	Features obtained from spatial gray level co-occurrence matrix(GLCM)
Wavelet Features	Entropy, Energy, Sum-average Sum-variance, Cluster tendency calculated from wavelet transform image
Ridgelet Feature	Texture features extracted from ridgelet transform image.
Gabor Filter Bank method	Features are computed from image applied with gabor filter method
Fractal Dimension	Features extracted from fractal model of image
Cluster Feature	Number of micro calcification in an area, cluster area are obtained in the presence of micro calcification
Morphological Feature	Speculation, shape, size are obtained from the image for identifying architectural distortion and bilateral asymmetry
Physical Features	Patient age, surgery undergone, symptoms may be combined with the above feature to improve the accuracy.

4. Feature Extraction

After preprocessing the mammogram visual clues should be identified by processing the pixel intensity and density values in the image. Features are the high level information about the digital image that helps in analyzing the image. So for researchers had made use of many types of features such as

1. Region based feature
2. Shape based feature
3. Image structure based feature
4. Texture feature

The list of features may differ with respect to the type of medical images. Similarly for mammograms, the type of feature may differ in detecting the different sign of abnormalities. Texture feature gives the spatial information as well as it could distinguish the density of the masses. Texture feature helps in detecting masses and micro calcification. In detecting micro calcification wavelet transform provide meaningful information. Texture features defined by Haralick [8] had been widely used in the literature. Texture features are extracted from co-occurrence matrix, wavelet transform, ridgelet transform. Among these methods wavelet transform outperformed according to R.P.Ramos et.al [9]. In case of architectural distortion and bilateral asymmetry morphological features such as shape,

Table 4.1 Features extracted from mammograms

size, speculation, margin and contrast is used. The various features used in the literature is shown in the Table4. 1.

5. Feature Selection

Number of features could be extracted from the image as discussed above. But if the feature vector is too large the computational time would be longer. The feature set affects the accuracy of the CAD system. Therefore the features that discriminate the classes of the images can be selected. The features are significant if they have the following four properties such as discrimination, reliability, independence and optimality. Li et.al [10] had discussed the guidelines in selecting the optimal features. In the literature [11]-[13], number of approaches are available for selecting optimal feature vector are hill climbing, best first search, feature weighting technique, heuristic search, class separation distance method, genetic algorithm. Among these methods genetic algorithm outperformed with great optimization.

Feature extracted from the image is a continuous data. Handling of continuous data is a difficult task. Therefore extracted feature is discretized. In the literature there are several methods employed in discretizing the continuous values. Among them Fuzzy Discretization[14] proved to be very fast method. Fuzzy partition method overcomes the limitation such as uncertainty, loss of information at the boundaries. N number of features can be extracted from the image. But all the features cannot be used since some of them have low discrimination power and the data base could become very large. So feature selection process is done which reduces the feature dimensionality. If N features are extracted from the image then 2N possible subset features could be produced. These combinations could be formed using genetic algorithm efficiently. In the literature the features such as entropy, contrast, correlation , mean, deviation, area, shape , Fourier transform and texture features proved to give best results.

6. Classifiers for detecting mass and micro calcifications

With the features extracted from the mammograms, the CAD system is trained to classify the new test samples accurately. Over a decade researchers had proposed number of classification algorithms. In this paper classification algorithms that employ data mining techniques are discussed. Machine learning methods are computer algorithms that infer a hypothesis from the diagnosed samples and make decisions for a new sample. Classifiers used for detecting and classifying abnormal signs such as mass and micro clusters are discussed.

6.1 Neural Network

Neural network is used to achieve artificial intelligence it looks similar to biological neural network. It is a multilayer network. The first layer is the input layer and the last one is output layer. In between layer are known as hidden layer that contain computational nodes. Back propagation algorithm is used to train network in which the weight and biases in the hidden layer is calculated using activation function. The weights are adjusted by propagating backward from the output layer to hidden layers. The architecture of neural network consists of three layers. The number of nodes in the input layer represents the number of features in feature vector. The nodes in the output layer represent the classes. In classifying mammograms we have two classes malign and benign and therefore the output layer contain one node. The weights are adjusted in the learning process to minimize the mean squared error between the networks prediction and layer value. It is not sure that the learning process may terminate automatically. The training period also takes long time.

Kupinski [15] minimized over training given to the neural network. Early stopping technique is used in which training is stopped at early stage when the performance of the network is maximum for an independent data set. David.B.Foge [16] added patient age as a feature and used ANN as classifier. The proposed method used evolutionary algorithm, a population based stochastic search algorithm to train up the network .Results were best when mass is considered Az =0.91 but when micro calcifications were included Az=0.84 which was poor. Dhawan et.al [17] used Back propagation network along with 10 SGLD features and tested the method with 85 mammograms. They obtained accuracy of 74%. Verma & Zakos [13] in their research work they detected the suspicious micro calcification area in the image using fuzzy logic. Then features are extracted from the suspicious area and applied feed forward BPN for classification. The method achieved 88.9% of accuracy. Panchal Rinku [18] explored the significance of 14 gray scale features and patient age in detecting the mass type abnormalities in the mammograms and using neural classifier they achieved 91% of classification rate. Brijesh verma and Ping zhang [19] used double genetic algorithm for training the neural network and selecting the best feature subset. They obtained classification rate of 85%. But the training process using genetic algorithm is also quite slow. It is also seen that number of hidden layers, threshold, feature subset and learning rate of the network affects the classification rate in neural network.

6.2 Linear Discriminant Analysis

LDA classifier minimize the within class distance and maximize the between class distance there by it gives good discrimination to classify the images. Chang et.al [20] used LDA in two step process. In the first step the texture features are extracted from the SGLD matrix and the features with discriminative power is selected. As a second step discrimination function of the classifier is optimized by adjusting the weights of the features in the function so that the function could classify the classes with great accuracy. The classifier obtained an area of .82 under ROC. The major advantage of LDA classifier is that it uses only second order moments. At the same time they may be inaccurate and local information is missed by LDA classifier.

Bruce et.al [21] used multi resolution mass shape features and LDA as classifier. They obtained classification rate of 83% using leave one out test method which outperformed compared to uniresolution features. Datong wei [22] used multi resolution local and global features extracted from the wavelet decomposition and used LDA method to classify the mass and obtained area Az of 0.92 under ROC curve. Berkman sahiner [23] proposed a method that makes use of mass segmentation as the first step. The features are extracted from the segmented region and used LDA classifier area of 0.89 is obtained under ROC. They also devised one more technique that involves contour segmentation and morphological and texture features were extracted to interpret the speculation in segmented mass region. They used LDA classifier and obtained Az of 0.91 under ROC. Mohammed et.al [50] combined LDA classifier with SVM classifier and achieved 90% specificity compared to LDA and SVM classifiers. Nicholas Petrick [24] used density weighted contrast method for segmenting the suspicious mass regions. Multi resolution texture features are extracted from the segmented region. The test images are classified using LDA and obtained.

6.3 Support Vector Machines

Support Vector Machine (SVM) is used to classify linear and non linear data. SVM embeds kernel learning which results as a best classifier. It uses nonlinear mapping to transform the original data to high dimensional data and uses support vector to find the margin. Using SVM optimal solution could be obtained. SVM is a pattern classifier that uses hyper plane to classify the data. The training samples along the hyper plane and near the class boundary are called support vectors. Kernel learning is used to decide the hyper plane that separates the test data. SVM transforms the non linear feature into linearly separable data it helps in classifying the mammograms, since mammograms have non linear feature space.

Campanini et al.[25] used two support vector machine classifiers to detect mass in the mammograms and obtained 80% true positive rate. Moyaedi et al. [26] proposed support vector based fuzzy neural network classifier. The classifier combines the efficiency of SVM in high dimensional space data and uncertainty is handled by fuzzy logic. This combination produce good classification rate. Timp et al. [27] used SVM classifier to analyze the temporal changes in mammogram. And ROC of 0.77 is obtained using temporal features. In [28] Abdall et.al used textural features with support vector machine and achieved 82.5% accuracy. Massotti [29] used SVM and achieved accuracy of 90%.J.S.Leena Jasmine et.al [30] used contourlet coefficients to extract the features, then the extracted features were given to SVM classifier the proposed method obtained an accuracy rate of 81%. Y.Ireaneus Anna Rejani [31] in their proposed method segmented the tumor region using thresholding technique and then morphological features are extracted from the segmented region. Finally using SVM classifier they achieved an accuracy rate of 88.75%. Leonardo de Oliveira Martins et.al [32] segmented the mass region from mammograms using K means algorithm .And from the segmented region co-occurrence based features are extracted and fed to SVM classifier. The accuracy obtained was 93.1%.

RVM [42] is used in detecting the micro calcification of clusters in digital mammograms. Relevance vector machine is based on Bayesian estimation theory. RVM makes use of sparse decision function i.e minimum number of vectors is considered for detecting the clusters. Detection accuracy is similar to SVM. In SVM the vectors are taken near the boundary but in RVM the most representative vectors away from the boundary is considered. But the disadvantage is computational efficiency.

6.4 Bayesian Belief Network

Bayesian belief network is probabilistic classifiers. It is a graphical model of relationships on which learning can be performed. Network consists of directed acyclic graph and set of conditional probability. The nodes represent the feature and any of the nodes in the network is identified as output (i.e) class label. The network determines the probability of each class. The weights are assigned to random probability values and they are updated until it converges to the optimal solution. Connections between the nodes indicate the probabilistic dependence of two features. Bayesian belief network outperformed neural network .But in many cases both the classifiers had fluctuation.

Ramirez et.al [33], used Bayesian classifier in detecting the breast lesion and achieved accuracy rate of 93%. Bayesian network [34] resulted in covering area of 0.873 under ROC curve and achieved 0.936 for small set of data base. Wang X.H. [35] integrated image and non image based feature and obtained area of 0.89 under ROC .Non image based features such as patient clinical history and physical examination were considered. Bin Zheng [36] selected 16 local and global features from 150 features using genetic search and used BBN and obtained ROC curve of 0.873.Probabilistic learning concept in BBN performed well compared to other classifiers. J.L.Vilton [37] used speculation and boundary as features and made use of Bayesian classifier to detect the mass. B.zheng [38] compared ANN and Bayesian belief performance of both classifier converged at same level with the use of GA optimization.

6.5 K Nearest Algorithm

K-Nearest algorithm is very simple and easiest algorithm. But it is an efficient classifier .It is based on pattern similarity approach, distance is computed from sample pattern to every other nearest pattern. The algorithm gives best results when detecting micro calcifications than other signs. The test pattern is assigned to the class with maximum micro calcifications. Woods et.al [39] modified the algorithm by selecting the unknown test sample and assigning to the class that has maximum k micro calcifications.

6.6 Decision Tree

A decision tree is a tree structured network, where internal node represents a test on a feature, each branch indicates outcome of the test and leaf node represents the class label. A path is traced from root to leaf node for the class prediction for the test image .The learning and classification steps of decision tree are so simple and fast. The decision tree [40] is constructed with starting feature as root node and descending downwards by the values of the feature until it reaches the terminal node i.e. predicted class label. Decision tree require less computational effort than neural network and Bayesian belief network. There is no probability distribution of features which is an overhead. When decision tree is combined with fuzzy membership functions it outperformed neural network and Bayesian belief network.

Table 1. Confusion matrix for binary classification

Class	Benign	Malign
Benign	TN	FP
Malign	FN	TP

Kegelmeyer et al.[41] used binary decision tree to classify the speculated mass and obtained 100 %.As an outset fuzzy based decision tree obtained 96% of sensitivity. But if the tree is large it would be difficult to interpret. And this method suffers from sub tree repetition and replication. C4.5 is one of the decision tree algorithms that extract rule along with pruning procedure. It obtained high classification rate than other decision tree algorithms. Decision tree is a non linear classifier which performs better when combined with fuzzy logic. They are robust with non linear feature vector space.

6.7 Rule Mining

Association rules are mined using apriori algorithm from the database. The association rules are extracted from the database which contains feature vector extracted from the mammography and the class the image belongs to, benign or malign. The antecedent part of the rule contains the features and the consequent of the rule represents the classes. Association rule mining is so fast in building up a classifier and computational over head is less. Four statistical parameters were used as feature set and to avoid large set of rules pruning is done. Pruning eliminates specific rules and chooses general rules with high confidence.

M.L.Antonie et.al [43] used apriori algorithm. The rules are mined using low confidence which mislead the classifier result and make the process slow. The classifier counts the number of rules that matches the rules in trained set and chooses the class. Accuracy of the system is 80.33% .X.Wang et.al [44] used association rules for classifying the mammograms. But they discretized the features in equal sized interval which lead to loss of significant information. Marcela et.al [45] used association rule mining and achieved 96.7% of accuracy. But the classifier works good only when the rules in the different classes are balanced. The rules are extracted from the database with minimum support and minimum confidence. And moreover minimum confidence affects the quality of the rules mined. Marcela [46] used IDEA (Image Diagnosis Enhancement through Association rules) method in 2009 that employs apriori method for mining rules. To overcome the minimum confidence drawback, it is set to 97%. The proposed system used a special classifier ACE classifier which returned multiple classes. The classifier finds the class by calculating the weights of the fully matched rules, partially matched rules, unmatched rules. The higher value of weight gives high confidence. Ali Keles et.al [47] developed a system using neuro-fuzzy rules for diagnosing cancer and obtained 96% of positive prediction rate. Sumeet Dua et.al [48] made use of weighted rules which outperformed well than existing association rule mining algorithm.

7. Evaluation of Classifiers

Performance of a classifier is measured by analyzing how it recognizes the different classes. Confusion matrix [49] is used to measure the performance of the classifier.

In the diagnosis of mammography images the malign type of images are identified as positive and the benign images as negative. Classifying the malign image as malign is known as true positive and malign images as benign are false positive. Similarly detection of benign images as benign is true negative and benign image as malign is false negative. Two measures sensitivity, specificity reveals the misdiagnosis rate. Sensitivity measures true positives and specificity measures the negatives. The formulae are stated in table. One more factor to measure the performance of the system is the accuracy of the system built.

Table 2. Performance measures

Performance measures	Formula
Sensitivity	$\frac{\text{Number of true positives}}{\text{Number of true positives} + \text{Number of true negatives}}$
Specificity	$\frac{\text{Number of true negatives}}{\text{Number of true positives} + \text{Number of true negatives}}$
Accuracy	$\text{sensitivity} * \frac{\text{pos}}{\text{pos} + \text{neg}} + \text{specificity} * \frac{\text{neg}}{\text{pos} + \text{neg}}$

To measure the accuracy of the model ROC curve is used. Receiver operating characteristic curves [50] shows the tradeoff between true positive and false positive. The vertical axis of ROC curve represents true positive and horizontal axis of ROC curve represents false positive. For each test sample if it is classified as true positive, move up in the ROC curve and plot the point. If it is false positive move towards right and plot the point. To measure the accuracy of the classifier, the area under the curve is measured. If the area is close to 0.5 then the system is less accurate and if the area is 1.0, then the system is more accurate.

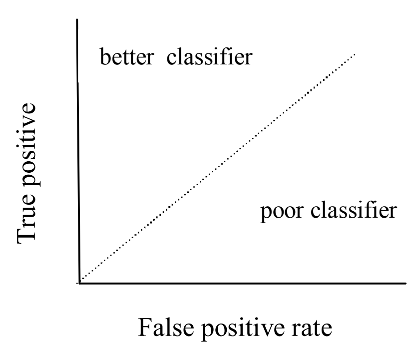


Figure 1. Receiver operating curve

8. Comparison of classifier performance

Features extracted from the mammograms play a vital role in any classification method. Relevant feature vector helps in improving the classification rate. So feature selection is an important step in any CAD system. There are number of feature selection algorithm and classification algorithm in data mining gadgets. Table 8.1 depicts the performance statistics of the classifier. When we consider neural network classifier two limitations are there one is the nonlinearity decision functions of the network makes them more complex, and the second is the network may prematurely converges at suboptimal solutions. Bayesian belief networks outperform neural network by providing the conditional probabilities and handling outliers. SVM chooses the vectors nearer to the boundary and detect the classes which may not be more accurate .RVM is similar to SVM but chooses the more representative vectors. Kernel learning algorithm out performs other classifiers but computational complexity is high. Rule mining classifiers incorporating fuzzy rule mining, weighted rules outperform association rule mining using apriori algorithms. Figure 8.1 shows the performance of different classifier. Among the five different classifier rule mining and kernel learning classifier out performed. But still the imbalanced dataset affects the accuracy. False positive rate is still needed to be reduced even though the classification rate is good .Rule mining method that involves evolutionary algorithms reaches high accuracy.

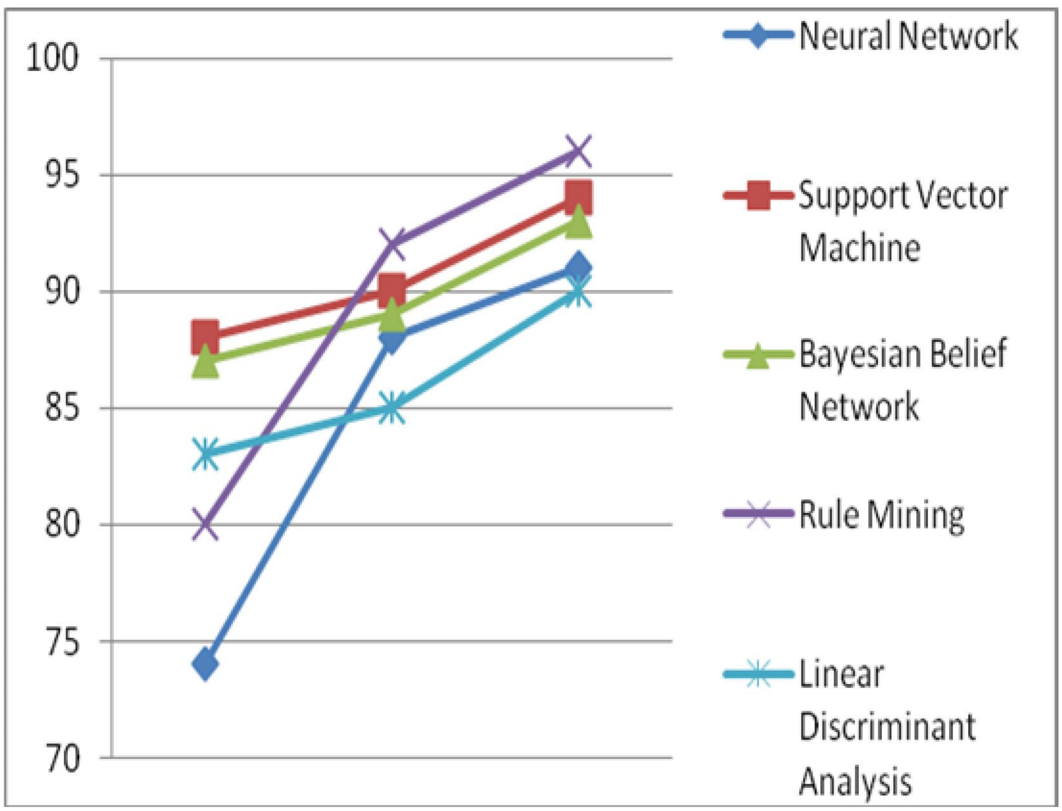


Figure 2. Comparision of classifier performance

Table 3. Performance statistics of selected classifiers

Classifier	Method	Abnormality Sign	Result Az : area under ROC % of accuracy
Neural Network	Feature vector includes patient age and 12 radiographic feature, Stochastic search algortihm to train the network, dataset : 111 malignant ,105 benign. 10 SGLD features, backpropagation algorithm to train up the network, dataset : 85 mamamograms. Combination of entrophy, standard deviation and number of pixels,fuzzy technique to detect the microcalcification, 14 gray leveled based features,four BIRADS feature and patient age feature, dataset: DDSM benchmark database.	Mass Micro calcification Mass Micro calcification Mass	Az 0.91 Az 0.84 74% 88.9 91%
LDA	Texture features using SGLD matrix from ROI, LDA in texture feature space, data set :digital mammograms. Mass shape feature extracted using wavelet transform modulus-maxima method to classify mass as round, nodular or stellate ,dataset : digitized mammograms. Texture features extracted from wavelet transformed images, LDA classifier, dataset : 168 digitized mammograms Three stage segmentation process (clustering, active contour, speculation detection stages), texture and morphological features with LDA classifier. Feature selection using forward stepwise linear regression method, Combined LDA and SVM classifier	Mass Mass Mass Mass Mass	Az 0.82 83% Az 0.92 Az 0.91 90%
Bayesian belief network	Bayesian network classifier, dataset : real world database of breast lesion. Image and non image features (patient clinical history ,physical examinations),Network incorporates both the features, data set : 419 cases. Adaptive region growing segmentation method, 12 local and 4 global features were selected using GA, data set :digitized mammograms two different data sets were used.	Mass(lesion) Mass Mass	93.04 Az 0.89 Az 0.87
Support Vector Machine	Statistical textural features are extracted , compared with linear ,non linear classifier SVM outperformed ,dataset :120 digital mammograms from digital database. Gray scale invariant ranklet texture feature along with SVM classifier reduced false positive rate. Image enhancement methods (filtering,top hat operation, DWT), suspicious area is segmented using threshold method, features are extracted from segmented image ,data set : 75 mammographic images. K-means clustering algorithm is used to segment the mammograms and SVM classifier is used to classify whether the segmented region is mass or not texture features from co -occurrence matrix and shape features are used.	Mass Mass Tumors Mass	82.5% 90% 88.75% 93.1%
Rule Mining	Four statistical parameters ,association rule based classification by category classifier, dataset : MIAS database (322 images)	Malignant, Benign	80.3%

9. Conclusion

Diagnosis is the most important aspect in the treatment of any fatal disease. Mammography Computer Aided Diagnosis system helps the physician in diagnosing abnormalities in early stage. CAD system employs different classification methods. In the literature we could see that no single algorithm is used to detect all the abnormal signs in mammograms.

Combining the results of the different classifier and feature selection could improve the classifier results. The performance of the classifier depends upon the learning principle and robustness against the imbalanced training samples, insignificant features. Fuzzy logic could be used in mammogram analysis which could deal with outliers and noisy data. Data extracted from the mammograms are non linear so classifier should be able to handle such type of data. There is only limited number of commercial systems available including R2 system, the iCAD system and Kodak's system. But still none of the algorithm proved good commercially. This is due to inconsistent results obtained from the classification algorithms and non iterative classification systems could be fast. There is a wild search for a methodology that could obtain above 0.95 areas under ROC curve and meet the requirements of clinics and screening centers. Research area focuses on reducing the false positive rate and increasing the sensitivity. Making use of soft computing techniques in the classifier could handle the non linear data and improve the accuracy rate. In future we have planned to focus on the effectiveness of soft computing techniques in classifier.

References

[1] .S. Ciatto, M.R.D. Turco, P. Burke, C. Visioli, E. Paci, M. Zappa, Comparison of standard and double reading and computer-aided detection (CAD) of interval cancers at prior negative screening mammograms: blind review, Br. J. Cancer 89 (2003) 1645–1649.

[2].Jelena Bozek, Kresimir Delac, Mislav Grgic, Computer-Aided Detection and Diagnosis of Breast Abnormalities in Digital Mammography, 50th International Symposium ELMAR-2008, 10-12 September 2008, Zadar, Croatia

[3].J.M. Lewin, C.J. D'Orsi, R.E. Hendrick, L.J. Moss, P.K. Isaacs,A. Karellas, G.R. Cutter, Clinical comparison of full-field digital mammography and screen-film mammography for detection of breast cancer, Am. J. Roentgenol. 179 (2002) 671–677.

[4] M. Kallergi, K. Woods, L.P. Clarke, W. Qian, R.A. Clark, Image segmentation in digital mammography: comparison of local thresholding and region growing algorithms, Comput.Med. Imag. Graph. 16 (1992) 323–331.

[5] Marcela X. Ribeiro, An Association Rule-Based Method to Support Medical Image Diagnosis With Efficiency, IEEE Transactions on Multimedia, Vol. 10, no. 2, February 2008 .

[6] W.M. Morrow, R.B. Paranjape, R.M. Rangayyan, J.E.L. Desautels, Region-based contrast enhancement of mammograms, IEEE Trans. Med. Imag. 11 (3) (1992) 392–406

[7] R. Sivaramakrishna, N.A. Obuchowski, W.A. Chilcote, G. Cardenosa, K.A. Powell, Comparing the performance of mammographic enhancement algorithms—a preference study, Am. J. Roentgenol. (2000) 45–51.

[8] R.M. Haralick, K. Shanmugam, and I. Distein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3,pp. 610–621, 1973

[9] R.P.Ramos,Marcelo Zanchetta, Texture extraction: An evaluation of ridgelet, wavelet and co-occurrence based methods applied to mammograms, Expert Systems with Applications 39 (2012) 11036–11047.

[10] H. D. Li, M.~Kallergi, L. P. Clarke, V. K. Jain, and R. A. Clark, "Markov random field for tumor detection in digital mammography," IEEE Transactions on Medical Imaging, vol. 14, pp. 565-576, September 1995.

[11]R. Kohavi, D. Sommerfield, Feature subset selection using the wrapper method: overfitting and dynamic search space topology, in: The First International Conference on Knowledge Discovery and Data Mining, 1995.

[12]C. Guerra-Salcedo, D. Whitley, Genetic Approach to Feature Selection for Ensemble Creation, GECCO-99, 1999.

[13] B. Verma, J. Zakos, A computer-aided diagnosis system for digital mammograms based on fuzzy-neural and feature extraction techniques, IEEE Trans. Inform. Technol. Biomed. 5 (1) (2001) 46–54.

[14] V.Sivakrithika, B.Shanthi, Comparative Study on Cancer Image Diagnosis using Soft Computing Techniques, International Journal of Computer Applications, No.5-Article 4,2011.

[15] Kupinski, M.A,"Investigation of regularized neural networks for the computerized detection of mass lesions in digital mammograms ", Engineering in Medicine and Biology Society, Proceedings of the 19th Annual International conference of the IEEE,1336-1339,vol.3,Nov.1997.

[16] David B. Fogel , Eugene C. Wasson, et..al Evolving artificial neural networks for screening features from mammograms,Artificial Intelligence in Medicine 14 (1998) 317–326

[17]A.P.Dhawan, Y.Chitre, Artificial Network based Classification of mammographic micro calcifications using image structure features,Proc SPIE 1905(1993)820-831.

[18] Panchal, R & Verma, B, 2006, 'Neural Classification of Mass Abnormalities with Different types of Features in Digital Mammography', International Journal of Computational Intelligence and Applications, vol. 6, no. 1, 61-75.

[19]Brijesh Verma , Ping Zhang ,A novel neural-genetic algorithm to find the most significant combination of features in digital mammograms, Applied Soft Computing 7 (2007) 612–625

[20] Chan HP, Wei D, Helvie MA, Sahiner B, Adler DD, Goodsitt MM, Petrick N.,Computer-aided

classification of mammographic masses and normal tissue: linear discriminant analysis in texture feature spac,. Phys Med Biol. 1995 May;40(5):857-76.
[21] Bruce LM, Adhami RR.,Classifying mammographic mass shapes using the wavelet transform modulus-maxima method, IEEE Trans Med Imaging. 1999 Dec;(12):1170-7.
[22] Datong Wei, Heang-Ping Chan, Nicholas Petrick,, False-positive reduction technique for detection of masses on digital mammograms: Global and local multiresolution texture analysis, Med. Phys. 24, 903 (1997)
[23] Berkman Sahiner, Heang-Ping Chan, Nicholas Petrick, Mark A. Helvie, and Lubomir M. Hadjiiski ,Improvement of mammographic mass characterization using spiculation measures and morphological features , Med. Phys. 28, 1455 (2001); http://dx.doi.org/10.1118/1.1381548.
[24] Nicholas Petrick, Heang‐Ping Chan, Datong Wei, Berkman Sahiner, Mark A. Helvie, and Dorit D. Adler ,Automated detection of breast masses on mammograms using adaptive contrast enhancement and texture classification , Med. Phys. 23, 1685 (1996); http://dx.doi.org/10.1118/1.597756.
[25] R. Campanini, D. Dongiovanni, E. Iampieri, N. Lanconelli, M. Masotti, G. Palermo, A. Riccardi, M. Roffilli, A novel featureless approach to mass detection in digital mammograms based on support vector machines, Phisics in Medicine and Biology 49 (2004) 961–975.
[26] Moayedi, F.; Boostani, R.; Azimifar, Z.; Katebi, S., "A Support Vector Based Fuzzy Neural Network Approach for Mass Classification in Mammography," Digital Signal Processing, 2007 15th International Conference on , vol., no., pp.240,243, 1-4 July 2007
[27] Timp S, Varela C, Karssemeijer N,Temporal change analysis for characterization of mass lesions in mammography, IEEE Trans Med Imaging. 2007 Jul;26(7):945-53.
[28] Al Mutaz M.Abdalla,; Deris, Safaai; Zaki, Nazar; Ghoneim, Doaa M., "Breast Cancer Detection Based on Statistical Textural Features Classification," Innovations in Information Technology, 2007. IIT '07. 4th International Conference on , vol., no., pp.728,730, 18-20 Nov. 2007.
[29]Matteo Masotti, Nico Lanconelli, and Renato Campanini, Computer-aided mass detection in mammography: False positive reduction via gray-scale invariant ranklet texture features, Med. Phys. 36, 311 (2009).
[30] J.S.Leena Jasmine, Dr.S.Baskaran and Dr.A.Govardhan. Article:An Automated Mass Classification System in Digital Mammograms using Contourlet Transform and Support Vector Machine. International Journal of Computer Applications31(9):54-61, October 2011.
[31] Y.Ireaneus Anna Rejani,, Dr.S.Thamarai Selvi, Early detection of breast cancer using svm classifier technique, International Journal on Computer Science and Engineering Vol.1(3), 2009, 127-130.
[32] Leonardo de Oliveira Martins, Geraldo Braz Junior, Detection of Masses in Digital Mammograms using K-means and Support Vector Machine, Electronic Letters on Computer Vision and Image Analysis 8(2):39-50, 2009.
[33] Cruz-Ramirez, N., H. Acosta-Mesaa, H. Carrillo-Calvet,L. Nava-Fernandezc and R. Barrientos-Martineza,2007. Diagnosis of breast cancer using Bayesian networks: A case study. Comput. Biol. Med.,37: 1553-1564.
[35] Wang XH, Zheng B, Good WF, King JL, Chang YH., Computer-assisted diagnosis of breast cancer using a data-driven Bayesian belief network, International Journal of Medical Informatics, Volume 54, Issue 2, May 19 99, Pages 115–126.
[36] Bin Zheng ; Yuan-Hsiang Chang ; Xiao Hui Wang ; Walter F. Good ; David Gur, Application of a Bayesian belief network in a computer-assisted diagnosis scheme for mass detection, Proc. SPIE 3661, Medical Imaging 1999: Image Processing, 1553 (May 21, 1999)
[37]J.L. Viton, M. Rasigni, G. Rasigni, A. Liebaria, Method for characterizing masses in digital mammograms, Opt. Eng. 35 (12)(1996) 3453–3459.
[38] Bin Zheng; Yuan-Hsiang Chang; Xiao-Hui Wang; Good, W.F., "Comparison of artificial neural network and Bayesian belief network in a computer-assisted diagnosis scheme for mammography," Neural Networks, 1999. IJCNN '99. International Joint Conference on , vol.6, no., pp.4181,4185 vol.6, Jul 1999.
[39] Kevin S.Woods et al ,Comparative evaluation of pattern recognition techniques for detection of microcalcifications in mammography, Int.J.Patt. Recogn. Artif. Intell. 07, 1417 (1993).
[40] L. Zhen; A.K. Chan,"An artificial intelligent algorithm fortumor detection in screening mammogram", Medical Imaging,IEEE Transactions on , vol.20, no.7, pp.559-567, July 2001
[41] Kegelmeyer W P, Pruneda I M. Bourland P D, Hillis A. Riggs M W and Nipper M L 1994 Computer-aided mammographic screening for spiculated lesions Rodiolo~ 191 331-7
[42] Liyang Wei, Yongyi Yang, Relevance Vector Machine for Automatic Detection of Clustered Microcalcifications, IEEE Transactions on medical imaging, vol. 24, no. 10, october 2005
[43]M.-L. Antonie, O.R. Zaane, A. Coman, Associative classifiers for medical images, in: LNAI 2797, MMCD, Springer-Verlag, 2003, pp. 68–83.
[44] X. Wang, M. Smith, R. Rangayyan, Mammographic information analysis through

association-rule mining, in: IEEE CCGEI, 2004, pp. 1495–1498.
[45] Marcela X.Riberio “An Association Rule-Based Method to Support Medical Image Diagnosis With efficiency”,IEEE Transactions On Multimedia,feb 2008
[46] [46]Marcela X. Ribeiro , Pedro H. Bugatti , Caetano Traina Jr. , Supporting content-based image retrieval and computer-aided diagnosis systems with association rule-based techniques, Data & Knowledge Engineering 68 (2009) 1370–1382.
[47]Ali Keles, Ayturk Keles,Expert system based on neuro-fuzzy rules for diagnosis breast cancer,Expert System with Applications 38 (2011),5719-5726.
[48]Sumeet Dua , Harpreet Singh, Associative classification of mammograms using weighted rules, Expert Systems with Applications 36 (2009) 9250–9259.
[49]Jiawei Han ,Micheline Kamber ,Data Mining concepts and techniques, Second Edition, Elsevier 2006
[50] Alolfe, M.A.; Mohamed, W.A.; Youssef, A.-B.M.; Mohamed, A.S.; Kadah, Y.M., "Computer aided diagnosis in digital mammography using combined support vector machine and linear discriminant analaysis classification," Image Processing (ICIP), 2009 16th IEEE International Conference on, vol., no., pp.2609,2612,7-10Nov. 2009.

1/5/2021