International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 11 (2015) pp. 30045-30054

http://www.ripublication.com

A Review of Medical Image Classification and Evaluation Methodology For Breast Cancer Diagnosis With Computer Aided Mammography

Venmathi A R
Research Scholar, Anna University, Chennai, India.
Assistant Professor, Department of Electronics & Communication Engineering,
Kings Engineering College, Irungattukottai, Chennai, India-602117
E.mail: arvenmathi@yahoo.co.in
Dr. E N Ganesh
Professor & Dean, Department of Electronics & Communication,
Rajlakshmi Institute of Technology, Kuthambakkam Post, Chennai, India-600124.
E.mail: enganesh50@yahoo.co.in
Dr. N. Kumaratharan
Professor, Department of Information & Technology,
Sri Venkateswara College of Engineering, Pennalur, Sriperumbudur Taluk,
Chennai, India-602117. E. mail: kumaratharan@svce.ac.in

Abstract

Mammography considered as most potent technique for the screening of breast cancer and abnormalities detection. Screening is one of the key factors to reduce the death rates. Due to strong reciprocity between disorders and unnoticeable cancer tissues shows that radiologists may get benefit by the CAD (Computer aided Diagnosis) system with capabilities of automatic classification of breast tissue. The aim of this paper is to conduct a review and present a survey of existing mammograms enhancement and segmentation methods with the distinct classification and analysis of each technique.

Keywords: Mammography, Microcalcification, Feature Extraction, Classification, Enhancement and segmentation

Introduction

Breast cancer is more chronic and widespread cancer caused from breast tissue Breast cancer occurs majority in women. Breast cancer is a kind of malignant tumor that starts in the breast cells initially. It causes more death in women each year [1]. As per cancer fact sheets of World Health Organization, every year more than 1,50,000 women worldwide die because of breast cancer. Approximately 8-13% women develop breast cancer during their life span. The rate of survival depends upon the stage when it was detected. Initial Detection of clusters of microcalcification in digital mammograms is the foremost sign of breast cancer and is the best indicators for malignancy. In the detection process signs of breast cancer can be interpreted with high resolution quality of images. Nowadays X-ray mammography is the most efficient diagnostic method for early breast cancer detection. A typical X-ray mammogram provides various information, which represents, vessels, tissues, ducts, breast edge, chest skin, the film, and the X-ray characteristics. The computer aided systems for mammograms are used in two ways: as computer aided detection system (CADe) and as computer aided diagnosis system (CAD). CADe is able to identify the Regions of Suspicion (ROS), but CAD can make a decision whether a ROS is benign or malignant. Since micro calcifications are small and subtle abnormalities, they may be overlooked by an examining radiologist. Image enhancement and filtering is always the root process in many medical image processing applications

Pre Stages In Classification

In digital mammogram classification of micro calcifications clusters can be done in three stages Preprocessing stage, Feature Extraction stage and Classification stage.

A) Preprocessing

Preprocessing of the mammogram images commonly involves removing reflections, low-frequency image noise, normalizing the intensity of the individual particles images, and masking some portions of images. Image preprocessing is the method of enhancing data images prior to computational processing. During image preprocessing image data is improved and ready for feature extraction stage.

B) Image Enhancement

In mammogram images, the image enhancement technique means to improve the quality of a digitally stored image by manipulating the image to achieve denoising and contrast enhancement. The advanced image enhancement techniques also support many filters for altering images in various ways. Specialized programs for image enhancement are sometimes called image editors. The enhancement method which performs well on a particular type of images may not be satisfactory on the other images. Thus there is no general image enhancement method that can enhance all the images. Depending upon the application and the image source, best method can be selected. The existing image enhancement technique can be classified into two heads: Spatial based domain image enhancement and Frequency based domain image enhancement.

Spatial based domain image enhancement employs directly on pixels. The real advantage of spatial based domain technique is that they are conceptually simple technique which can be used for real time implementations but these techniques normally lacks to provide adequate robustness and imperceptibility requirements. In Frequency based domain image enhancement the analysis of mathematical functions or signals with respect to frequency and operates directly on the transform coefficients of the image. For example, Fourier Transform, Discrete Wavelet Transform (DWT), and Discrete Cosine Transform (DCT), many enhancement techniques are tested on digital mammograms present in mini-MIAS database.

C) Segmentation

Mammogram segmentation usually involves classifying mammograms into several distinct regions, including the breast border [5], the nipple [6] and the pectoral muscle. The principal feature on a mammogram is the breast border, otherwise known as the skin-air interface, or breast boundary. The breast contour can be obtained by partitioning the mammogram into breast and non-breast regions. The extracted breast contour should adequately model the soft-tissue, air interface and preserve the nipple in profile.

There are a number of problems associated with the accurate segmentation of the breast region. Firstly, owing to the nature of x-ray each pixel in a mammogram represents two or more tissues; indeed all pixels contain a component due to attenuation by skin. Superimposition of different tissue types makes it difficult to differentiate between different regions. As a result of the mammogram acquisition process, there is a region of decreasing contrast near the breast contour where the breast tapers off. This region constitutes the uncompressed region of the breast commonly referred to as the “breast edge”, and is caused by a lack of uniform compression of the breast tissue. This tapering effect causes a lack of visibility along the peripheral region of the mammogram, making it difficult to perceive the breast contour and identify the nipple position. The process of digitization may further decrease this visibility through the addition and accentuation of noise. There have been various approaches proposed to the task of segmenting the breast profile region in mammograms. Some of these have focused on using threshold [7][8], gradients [9], modeling of the non-breast region of a mammograms.

D) Feature Extraction

Feature extraction has a vital role in pattern recognition, it starts from an initial set of measured data and builds derived values (features) intended to be more informative, non redundant, facilitating the subsequent learning and generalization steps, in some cases extracting spatial data such as texture, size, etc. It improves the performance by restructuring the data and removes superfluous image data or very less value data in the process. This method helps in pattern recognition

Classification

Classification is the most important analysis in image processing, which is used to analyze the numerical properties of various image features which is extracted through the feature extraction stage and identifies data extracted from the target data. In mammogram images, initially the mammograms are classified into normal and abnormal. Abnormal mammogram is again classified into malignant or benign. In digital mammogram several methods like wavelet transform, wavelet based contourlet transform, etc. are used for classification of microcalcifications.

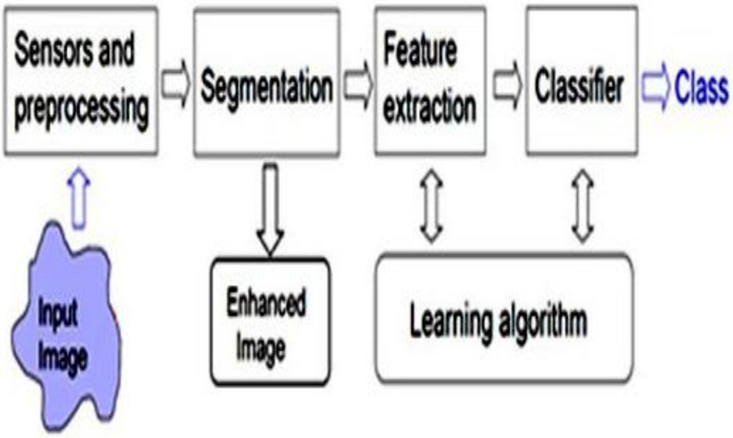

Figure 1: Image classification system

Analysis of Existing Classification Methods

In this digitized mammogram classification the Netherlands breast cancer screening program is taken as reference. In this program mammography initially views of each breast are taken for screening. The two views are Medio-Lateral Oblique (MLO) and a Cranio Caudal (CC) view. This study was carried out for several numbers of cases and observed that some were benign and some are of malignant. These images were further classified with CAD scheme. Year wise screening outcomes for subsequent examinations, women aged 50–69 years are shown in fig. 2

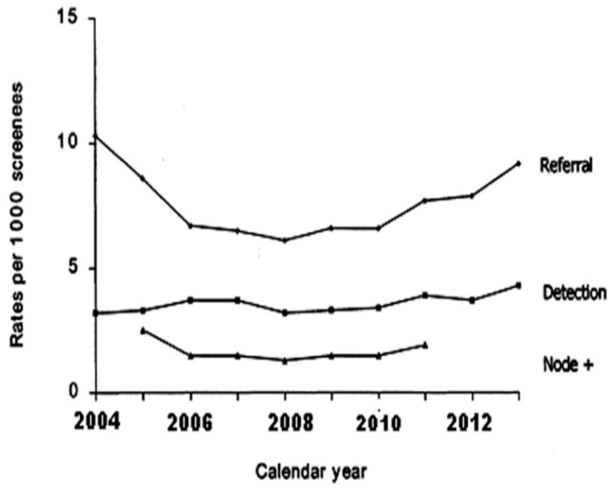

Figure 2: Trends In Screening Outcomes For Subsequent Examinations, Women Aged 50–69 Years

A) CAD scheme

In CAD classification method generally following Steps followed

1. Likelihood detection: In this step CAD scheme initially segments an image into breast tissue and background, using a skin line detection algorithm and it is done in the edge of the pectorals muscle. For MLO image view thickness equalization algorithm is applied to improve the periphery of the breast. A similar algorithm is applied; background equalization is used in the pectorals muscle, to stop the detection of masses located on or near the pectoral boundary. With these features a neural network classifies each pixel and assigns a level of suspiciousness to it. The neural network is trained using pixels sampled inside and outside of a representative series of malignant masses. Finally, the result of this method is an image whose pixel values represent the likelihood at a malignant mass or architectural distortion is present.
2. Initial step of detection results likelihood image and a number of suspect image locations.
3. Region segmentation detection is to provide the possibility suspicious locations as seed points to characterize and detect them by parameter analysis (shape, position, size).
4. Regions Classification as true abnormalities and false positives.

B) Naïve Bayesian Classifier

A Naive Bayes classifier is a simple probabilistic classifier based on limited topology applicable to learning tasks. It is constructing classifiers: models that assign class labels to problem instances, represented as vectors of feature values, where the class labels are drawn from some finite set. It is not a single algorithm for training such classifiers, but a family of algorithms based on a common principle: All naive Bayes classifiers assume that the value of a particular feature is independent of the value of any other feature, given the class variable where each instance is described by a conjunction of feature values and a class value. This method is also used to estimate the unknown data item using probabilistic statistics model. Challenges in the Bayesian classification is to determine the class of data sample which have some number of attributes.

Naive Bayes classifiers can handle an arbitrary number of independent variables whether continuous or categorical. Given a set of variables, X = {x1,x2,x3...,xd}, we want to construct the posterior probability for the event Cj among a set of possible outcomes C = {c1,c2,c3...,cd}. In a more familiar language, X is the predictors and C is the set of categorical levels present in the dependent variable. Using Bayes rule:

$$p\left(C_j \mid x_1, x_2, \ldots, x_d\right) \propto p\left(x_1, x_2, \ldots, x_d \mid C_j\right) p\left(C_j\right)$$

where p(Cj | x1,x2,x...,xd) is the posterior probability of class membership, i.e., the probability that X belongs to Cj. Since Naive Bayes assumes that the conditional probabilities of the independent variables are statistically independent we can decompose the likelihood to a product of terms:

$$p(X \mid C_j) \propto \prod_{k=1}^{d} p(x_k \mid C_j)$$

and rewrite the posterior as:

$$p(C_j \mid X) \propto p(C_j) \prod_{k=1}^{d} p(x_k \mid C_j)$$

Using Bayes rule, we label a new case X with a class level Cj that achieves the highest posterior probability Although the assumption that the predictor (independent) variables are independent is not always accurate, it does simplify the classification task dramatically, since it allows the class conditional densities p(xk | Cj) to be calculated separately for each variable, i.e., it reduces a multidimensional task to a number of one-dimensional ones. In effect, Naive Bayes reduces a high-dimensional density estimation task to one-dimensional kernel density estimation. Furthermore, the assumption does not seem to greatly affect the posterior probabilities, especially in regions near decision boundaries, thus, leaving the classification task unaffected. Naive Bayes can be modeled in several different ways including normal, lognormal, gamma and Poisson density functions:

$$p(x_k \mid C_j) = \begin{cases} \dfrac{1}{\sigma_{kj}\sqrt{2\pi}} \exp\left(\dfrac{-(x-\mu_{kj})^2}{2\sigma_{kj}}\right), \quad -\infty < x < \infty, -\infty < \mu_{kj} <, \sigma_{kj} > 0 & \text{Normal} \\ \mu_{kj}: \text{mean}, \ \sigma_{kj}: \text{standard deviation} & \\ \dfrac{1}{x\sigma_{kj}(2\pi)^{1/2}} \exp\left\{\dfrac{-\left[\log(x/m_{kj})\right]^2}{2\sigma_{kj}^{\ 2}}\right\}, \quad 0 < x < \infty, m_{kj} > 0, \sigma_{kj} > 0 & \text{Lognormal} \\ m_{kj}: \text{scale parameter}, \ \sigma_{kj}: \text{shape parameter} & \\ \dfrac{\left(\dfrac{x}{b_{kj}}\right)^{c_{kj}-1}}{b_{kj}\Gamma(c_{kj})} \exp\left(\dfrac{-x}{b_{kj}}\right), \quad 0 \le x < \infty, b_{kj} > 0, c_{kj} > 0 & \text{Gamma} \\ b_{kj}: \text{scale parameter}, \ c_{kj}: \text{shape parameter} & \\ \dfrac{\lambda_{kj} \exp(-\lambda_{kj})}{x!}, \quad 0 \le x < \infty, \lambda_{kj} > 0, x = 0,1,2,\ldots & \text{Poisson} \\ \lambda_{kj}: \text{mean} & \end{cases}$$

Poisson variables are regarded here as continuous since they are ordinal rather than truly categorical. For categorical variables, a discrete probability is used with values of the categorical level being proportional to their conditional frequency in the training data. In this method of classification dataset are divided into two sets, training and testing respectively. Training dataset is considered as prior information and model is constructed on the basis of this training dataset.

C) Support Vector Machines

SVM classification method is used for the diagnosis of breast tumors by evaluating the potentiality image recognition and image classification tasks. SVM is a deeply inspired and one of the radiance techniques among the many machine learning algorithms in the last decades. The general framework to measure the accuracy of a SVM on a given database is composed of the following stages and as mentioned in figure 3.

1. Preprocessing of the images in the database
2. Separation of the database into training and test sets
3. Choice of the representation of the input data
4. Choice of the way of training, which includes :
    a) Method of multi-class training
    b) Value of the penalty term C
    c) Choice of the kernel

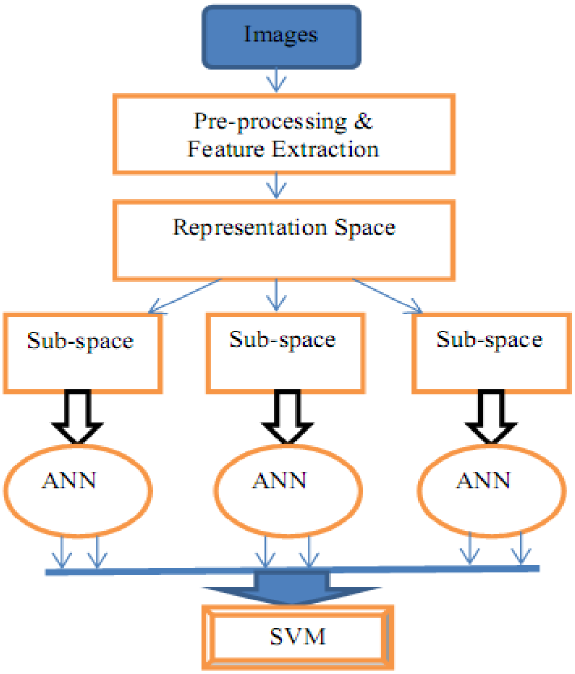

Figure 3: Image classification stages used in SVM

The benefit of the SVM scheme depends on algorithm which is being presented by CORTES AND VAPNIK. This is used for solving classification tasks and has been successfully applied in various areas of research. The basic idea of SVM is that it projects image data points from a given two-class training set in a higher dimensional space and finds an optimal hyper plane. The optimal one is the one that separates the data with the maximal margin.

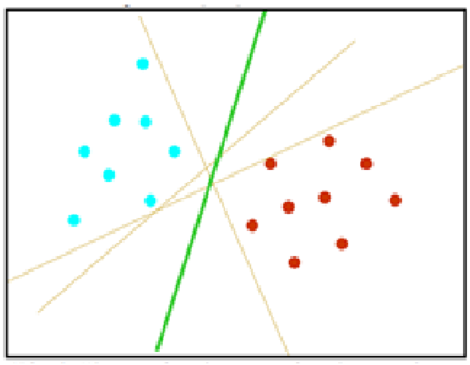

Figure 4: The optimal separating hyperplane

SVM identify the image data points near the optimal separating hyper plane which are called support vectors. The distance between the separating hyper plane and the nearest of the positive and negative data points is called the margin of the SVM classifier. The separating hyper plane is defined as

$$D(x) = (w \cdot x) + b$$

Where, x is a vector of the dataset mapped to a high dimensional space and w and b are parameters of the hyper plane that the SVM will estimate. The theoretical advantage of SVMs is that by choosing a specific hyperplane among many that can separate the data in the feature space, the problem of over fitting the training data is reduced. They are often able to characterize a range training set with a small subset of the training points. Also, SVMs can work on features with arbitrary distributions, without the need to make any independence assumptions.

Evaluation Methods

In CAD research, it is often necessary to assess the diagnostic performance of CAD systems and evaluate the differences between them. Also, it is most important to test the performance of CAD systems against practically proven results, obtained by experienced radiologists [5]. The performance of a diagnostic system is usually measured by its sensitivity and specificity. Sensitivity is the measure of how reliable a system is at making positive recognition, otherwise where it is correctly identifying that which is inspected as being correctly that which is desired. A super sensitive system will recognize what it is looking for most of the time, and rarely produces a false negative. On the other hand, Specificity is a measure of how well a system can make a negative picking out, or pick out when something inspected is not what is being desired, but something else. A classification system with greater specificity will rarely make the mistake of identifying what is being inspected as what is being desired.

There are other benchmark that includes both Sensitivity and Specificity, which measures the global performance, accuracy of the algorithm about the correct decisions. In the evaluation process of sensitivity and specificity, the results are normally defined in terms of Receiver Operating Characteristic (ROC) curve, Figure 5, which corresponds to the trade-off between the true-positive rate and the false-positive rate inherent in selecting specific thresholds on which predictions might be based [5].

The receiver operating characteristic analysis is a statistical method which carries out binary classification tasks, by analyzing, visualizing and comparing. The performance evaluation of a CADx system is a task which carries out classification for a suspicious region into benign or malignant, ROC analysis is typically used. The Receiver Operating Characteristic curve expresses sensitivity as a function of 1− specificity, and utilizes the Az metric, where Az is the area under the ROC curve. The system with the dashed ROC curve performs better than the system with the solid ROC curve.

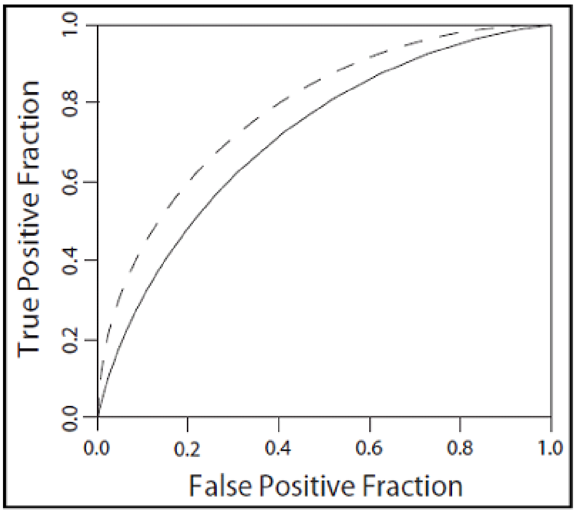

Figure 5: Receiver Operating Characteristic (ROC) Curve

Conclusion and Result

The conventional CAD scheme was employed in this study including: (i) image preprocessing to enhance the texture characterization of the image (ii) feature extraction; (iii) classification of benign and malignant clusters; and (iv) performance evaluation using ROC based on visual inspections. The receiver operating characteristic (ROC) analysis is a statistical method which carries out binary classification tasks, by analyzing, visualizing and comparing. This evaluation of a CADx system is a task which carries out classification for a suspicious region into benign or malignant. The preliminary results that tested on a small dataset demonstrated that the presented simple method was capable of detecting microcalcifications by visual inspection of digital mammograms. However, the performance of the presented CAD algorithms need to be further improved to achieve higher sensitivity with lower false positive rate based on a large testing dataset.

Future Works

In the future research work, the techniques which are discussed in this paper will be tested with larger database. These techniques can be improved by using an unsupervised multiple stage classification based on nearest neighbor classifier with hybrid wavelet feature design will result in more accurate CAD scheme classification technique.

References

[1] Y. Kopsinis and S. Mclaughlin, Development of EMD-Based Denoising Methods Inspired by Wavelet Thresholding, IEEE

[2] Cheng, H.D., et al., “Computer-aided Detection and Classification of Microcalcifications in Mammograms: A Survey”, Pattern Recognition, 2003, 36(12): p. 2967-2991.

[3] Bellotti, R., “A Completely Automated CAD S Mass Detection in a Large Mammographic D Medical Physics, 2006. 33(8): p. 3066-75.

[4] Qian, W., L. Li, and L. Clarke, “Image Feature Extraction for Mass Detection in Digital Mammography: Influence of Wavelet Analysis”,Med Phys., 1999, 26(3): p. 402-408.

[5] Laine, J. Fan and W. Yang, Wavelets for contrast enhancement of digital mammography, IEEE Engineering in Medicine and Biology, Sep./Oct. 1995, pp. 536-550.

[6] H.Wang, Y. Chen, T. Fang, J. Tyan and N. Ahuja, Gradient Adaptive Image Restoration and Enhancement, IEEE Proc. Int. Conf. Image Processing, Oct. 2006, pp. 2893–2896.

[7] R. C. Gonzalez and R. E. Woods, Digital Image Processing, 2nd edition Pearson Education, 2002. Trans. Signal Processing, Apr. 2009, vol.57, no.4, pp.1351- 1362.

[8] S. Anand and R. Aynesh Vijaya Rathna, Detection of Architectural Distortion in Mammogram Images using Contourlet Transform, IEEE International Conference on Emerging Trends in Computing, Communication and Nanotechnology, 2013.

[9] Rekha Lakshmanan and Vinu Thomas, Enhancement of microcalcification features using morphology and contourlet Transform, IEEE International Conference on Advances in Computing and Communications, 2012.

[10] Sharanya Padmanabhan and Raji Sundararajan, Enhanced Accuracy of Breast Cancer Detection in Digital Mammograms using Wavelet Analysis, IEEE, 2012.