

ANALYSIS OF MAMMOGRAPHIC IMAGES FOR EARLY DETECTION OF BREAST CANCER USING MACHINE LEARNING TECHNIQUES

*A Thesis submitted in partial fulfilment of the requirements for the degree of
Master of Technology*

In

Biomedical Engineering

By

RAM SHANKAR SAHU

213BM1016

Under The Supervision of

DR. SUBHANKAR PAUL



**Department of Biotechnology & Medical Engineering
National Institute of Technology
Rourkela-769008, Orissa, India
2015**

Dedicated To

My loved Parents



National Institute of Technology, Rourkela

CERTIFICATE

This is to certify that the thesis entitled “**ANALYSIS OF MAMMOGRAPHIC IMAGES FOR EARLY DETECTION OF BREAST CANCER USING VARIOUS MACHINE LEARNING TECHNIQUES**” by **RAM SHANKAR SAHU (213BM1016)** submitted to the National Institute of Technology, Rourkela for the award of Master of Technology in Biomedical Engineering during the session 2013-2015 is a record of bonafide research work carried out by him in the Department of Biotechnology and Medical Engineering under my supervision and guidance.

To the best of my knowledge, the matter embodied in the thesis has not been submitted to any Other University / Institute for the award of any Degree or Diploma.

Place: NIT Rourkela

Date:

Dr. Subhankar Paul
(Associate Professor)
Department of Biotechnology & Medical Engineering
National Institute of Technology
Rourkela-769008

ACKNOWLEDGEMENT

This work was carried out at the Department of Biotechnology and Medical Engineering, National Institute of Technology, Rourkela, Orissa, under the supervision of Prof. Subhankar Paul.

I would like to express my sincere gratitude to Prof. Subhankar Paul for his guidance and support throughout my work. Without him I will never be able to complete my work with this ease. He was very patient to hear my problems that I am facing during the project work and finding the solutions. I am very much thankful to him for giving his valuable time for me.

I would like to express my sincere thanks to the PhD scholars Mr. Deependra Kumar Ban and Mr. Sailendra Kumar Mahanta and my friends Prashant Kumar, Krishna Murthy, Usha Pandey, Gudia Jha , Jyoti Jha for their precious suggestions and encouragement to perform my project work.

I thank God for giving me the chance to pursue my academic goals and a wonderful, complete life full of amazing people.

Finally, hearty thanks to my family members for their sacrifice that encouraged me to proceed with the work.

Date:
Place: NIT Rourkela

Ram Shankar Sahu
Roll No.-213BM1016

Abstract

Breast cancer is the main reason for death among women. Radiographic images obtained from mammography equipment are one of the most frequently used techniques for helping in early detection of breast cancer. The motivation behind this study is to focus the tumour types of breast cancer images. It is methodology to anticipate a sickness in view of the visual conclusion of breast disease tumour types with precision, particularly when numerous features are related. Breast Cancer (BC) is one such sample where the phenomenon is very complex furthermore numerous features of tumour types are included. In the present investigation, various pattern recognition techniques were used for the classification of breast cancer using mammograms image processing techniques. The pattern recognition techniques for tumour image enhancements, segmentation, texture based image feature extraction and subsequent classification of breast cancer mammogram image was successfully performed. When two machine learning techniques such as Artificial Neural Network (ANN), Support Vector Machine (SVM) were used to classify 120 images, it was observed from the results that Artificial Neural Network classifiers demonstrated the highest classification rate 91.31% and the SVM with both Radial Basis Function (RBF) and linear kernel classifiers demonstrated the highest classification rate of 92.11% and RBF classification rate is 92.85%.

Keywords: Breast Cancer, Mammographic images, Image enhancement, Image Segmentation, feature extraction, Principal Component Analysis, ANN, SVM.

CONTENTS

Acknowledgements	iv
Abstract	v
Contents	vi
List of Figures	viii
List of Tables	ix
Notations and Abbreviations	x
1.INTRODUCTION	2
1.1 Objective of The Thesis	6
2. LITERATURE REVIEW	8
3. METHODOLOGY	13
3.1 Experimental Documented Data collection	15
3.2 Experiment methods	15
3.2.1 Image Enhancement	15
3.2.1.1 Gaussian Filter	16
3.2.1.2 Gabor Filter	17
3.2.2 Image segmentation	17

3.2.3 Feature Extraction	18
3.2.3.1 Chip histogram based Texture Feature	18
3.3 Machine Learning	20
3.3.1 Dimensionality reduction using PCA	21
3.3.2 ANN Classifier	22
3.3.2.1 Multilayer Feed forward Neural Network classification	23
3.3.2.2 Testing and Training images	24
3.3.3 Introduction to SVM	24
3.4 Performance Evaluation Criteria for Classifier	27
4. RESULTS AND DISCUSSION	29
4.1 ANN Classifier performance	33
4.2 SVM Classifier performance	37
4.3 Comparisons of the Classifiers Performance	38
5. CONCLUSION	40
Future Work	40
References	41

LIST OF FIGURES

Fig no.	Figure Caption	Page No.
1	Flow chart shows the pattern recognition methods and classified mammogram images.	14
2	Shows the original image and filtered image.	16
3	Block diagram of Multilayer feed forward neural network.	23
4	Working of SVM for the nonlinear distinguishable information where dark circle symbol used for cancerous data and circle for noncancerous data.	25
5	To take the images and find out different filters results.	30
6	Original image than after find segmented threshold image.	30
7	Morphology operation using Dilation and Erosion.	31
8	Image shows the ROI by selected region of Tumour.	32
9	ANN tool used to find different parameter for classifier.	34
10	Variation of training MSE with respect to number of iterations.	35
11	Confusion Matrix for positive and negative results.	36

LIST OF Tables

Table no.	Tables Caption	Page No.
1	General procedure to solve by confusion matrix for binary classification.	27
2	In ANN to get optimize training performance these given various parameter.	33
3	General analysis by confusion matrix.	36
4	Given the SVM parameters and its values for RBF.	37
5	Given the SVM parameters and its values for linear kernel.	37
6	Confusion Matrix for RBF kernel.	37
7	Confusion Matrix for Linear kernel.	38
8	Comparisons of classification rate, sensitivity, specificity and ROC.	38

NOTATION/ABBREVIATION

CC	Craniocaudal
MLO	Medio Lateral Oblique
DDSM	Digital Databased For Screen Mammography
PCA	Principal Component Analysis
SVM	Support Vector Machine
ANN	Artificial Neural Network
PSNR	Peak Signal to Noise Ratio
RBF	Radial Basis Function
CR	Classification Rate
ROC	Receiver Operating Characteristics

CHAPTER 1

INTRODUCTION

&

OBJECTIVE

Introduction

Breast cancer is a major health problem in Western countries as well as our country. In USA, A study developed by the American Cancer Society estimate that, in between 8 of one and 12 of one women will develop breast cancer during their life (ACS, 2003). Breast cancer is 2nd most commonly diagnosed cancer world-wide (Parkin et al. 2001). There are main two types of breast cancer namely cranio-caudal (CC) and medio-lateral oblique (MLO). Breast cancer appeared by usually in the ducts, tubes that convey milk to the nipple, and lobules, organs that create milk. In both men and women, it occurs, although male breast cancer is an uncommon disease. A few works have reported that the development rate of a tumour is corresponding to its temperature (H.Usuki et al.1990). Screening test searches for cancer disease before an individual exhibits any symptom. It can help to discover cancer tumour at an early stage. In this examination, four images are gotten, two relating to right side breast and two to left side breast of the projections cranio-caudal (CC) and medio- lateral oblique (MLO). The utilization of CC and MLO images enhances visualization of breast tissue and expands the chance of detecting the presence of non-palpable breast cancer. During the examination, the radiologist consolidates data from these two perspectives to expand the possibilities of deciding from the earlier locales with irregularities characterized as genuine true positive (TP) and diminish the quantity of districts without anomalies, i.e., decrease false positive (FP) areas (J. Tang et al. 2009). Numerous masses are benign that implies the anomalous growth is mainly restricted to a circumscribed, single and growing mass of cells (Gokhale, 2009). A few tumours are dangerous that implies the anomalous growing attacks the encompassing tissues and that may metastasize or spread to remote zones of the body (Rangayyan, 2004).

The masses have the potential to lead to complications whereas the malignant tumours are extremely serious cancer. The larger part of breast tumours will have metastasized before attaining a tangible size. In this way, a different quantities of imaging techniques are found for breast tumour recognition in tissue level. These are arranged as Infrared Imaging and Mammography. In early stage, Infrared imaging or Thermography is one of the best routines in detecting breast cancer, furthermore helpful for recording advanced stages of breast cancers.

Mammography is one of the most efficient techniques for the detection of early stage breast cancer. Mammograms are for the most part investigated by radiologists to distinguish the early stage of breast cancer (Rojas Domínguez et al. 2009). The essential indications of anomalies and disease acquired from mammography are Calcifications, Localized increment in thickness, Masses, symmetry variation between the left and right breast images and Architectural distortion (Moradmand et al. 2011). A mass can be malignant or benign. These masses can shape as an after effect of diverse interior techniques that influence the breast in distinctive ways. Samples of kind breast masses incorporate fibroadenomas, fibrocystic malady, atypical hyperplasia of the breast, phyllodes tumor, periductal mastitis and papillomas (Kopans, 1998). Malignant breast masses can either be kept to the ducts where they are framed or they can be invasive spreading through the channels to lymph hubs and other inaccessible locales. For diseases limited to ducts, the most widely recognized illustrations incorporate lobular carcinoma insitu also, intraductal carcinoma. Obtrusive (invasive) breast malignancies can be ductal, lobular, medullary, comedo carcinoma, papillary, scirrhous or tubular. Mammographically, benign masses are all around encircled contrasted with dangerous masses anyway it is 4 hard to sum up. When all is said in done, all masses identified need further investigation unless they are traditional representation of no doubt understood sorts.

There are a few routes for recognizing and diagnosing breast variations from the norm, for example, self-examination furthermore, clinical breast exams, mammography, and open surgery (biopsy). Clinicians prescribe mammography on the grounds that it is thought to be sheltered, less hurtful contrasted with biopsy, and more exact than self-examination where the tumour can be identified before it can be felt. Mammography is considered as the best strategy for right on time recognition of breast cancer, and the rate of patients that can be cured at ahead of schedule stages is generally high (Tucker, 1993; Egan, 1998).

The most important thing about estimation of mammography is that it can identify breast variations in the early stages of cancer even before the development any physical symptoms. The American Cancer Society's rules for early breast malignancy location stress mammography and physical examinations. Clearly there are numerous different methods and techniques that are utilized for breast screening and every strategy accomplishes an alternate level of clarity in showing breast images. On the other hand, mammography is the main procedure that has been ended up being successful for breast tumour screening. One of the primary favourable circumstances of utilizing mammography is its cheap expense of usage for a huge populace of subjects. Since overall radiologists screen more than hundreds of films each day, keeping up consistency and exactness in analysis is not simple. This implies that computer supported analytic systems have the best seek after enhancing breast disease identification and reducing morbidity from the disease.

The X-ray (Mammogram) are checked by an advanced scanner whose optical attributes are straight forwardly identified with the nature of computerized images delivered. Unfortunately, it is not at present conceivable to get computerized images specifically which would take out a portion of the issues that we have with simple to advanced transformation. Improvement (enhancement) can be performed in either the spatial or spectral domain. Poor quality of

compression and/or low resolution of scanning mechanism can leads to poor signal to noise ratio in final images.

In this paper organised following section. In section 1 took the Mammogram images which is involved different noise and a lot of information and identifies cancer and non-cancer.

In section 2 present the standard histogram comparison strategy for image enhancement alongside our proposed fuzzy enhancement method. Enhancement techniques are using different filter for improve image. These result showed various information such as peak signal to noise ratio (PSNR). This parameter is a close estimation to human impression of reproduction quality. Higher the PSNR better is the picture quality. PSNR is the assessment standard of the recreated picture quality, and is critical estimation highlight. So Gabor channel used to picture. In section 3 present segmentation, using region based and edge detection based segmentation techniques which are used in feature extraction. In the present experimental study, we have utilized the computerized mammogram images which were given from online mammogram database (MIAS database 2007). Artificial neural Network (ANN) is the most prominently used machine learning algorithm. It is based on the biological learning systems but it does not reflect it fully. The well propagation (BP) algorithm is well known to be robust especially to problems with errors in the training set. Further, support Vector Machines (SVM) is identified as a new and efficient learning algorithm. It can also be used to learn target functions. One can distinguish SVM and ANN is in the way they optimize the error. The objective of learning in ANN, is to generate a set of weight values. These values minimize the training error. On the other hand, in SVM the training error is set to minimum. Correlation of the performance of different classifiers utilized under this pattern recognition methods regarding Classification rate (CR), affectability (sensitivity), specificity and area under receiver operating characteristics (ROC) curve, to focus the best classifier for the breast cancer images classification.

1.1 Objective of the Thesis

The main goal of this examination includes the improvement of new computerized pattern recognition technique for the classification of cancer and non-cancer breast images from mammogram images. This technique includes:

- (i) Collection of breast cancer images as well as non-cancerous images from the open source online data base.
- (ii) Developing the various pattern recognition techniques to classify breast cancer and noncancerous image from these mammogram images.
- (iii) Correlation of the performance of different classifiers utilized under this pattern recognition methods regarding Classification rate (CR), affectability (sensitivity), specificity and area under receiver operating characteristics (ROC) curve, to focus the best classifier for the breast cancer images classification.

CHAPTER 2

LITERATURE

REVIEW

2. Literature Review:

During the past year, Breast cancer is a standout amongst the most widely recognized diseases among women of the creating nation on the planets, and it has also converted a foremost source of death. Different commitments have been made in literature regarding to the utilization of example pattern recognition techniques for breast tumour conclusion in tissue level.

Hala al-shamlan and his group proposed to extraction the feature values for analysis breast cancer mammogram images to classify the breast tumour (hala al-shamlan et al. 2010). They are resolved situated essential and discriminating breast cancer feature extraction. They are used two process. First process which is improved the contrast of image. Another process segmentation which finds the locale for mass identification. They found out good results. The results were the reach qualities expected in every feature extraction. Martin and his group proposed the method for recognition of mass on digitized mammograms (Martins et al. 2009). They utilized K-means bunching calculation for picture division and dim level co-event grid to portray and break down the surface of divided structures in the picture. The grouping of these structures was attained to through Support Vector Machines (SVM), which is divide them into 2 gatherings; utilizing the shape and composition descriptors, masses and non-masses. The arrangement exactness acquired from that system was 85%.

Vishnukumar K. Patel and his group to proposed a mammographic images of breast cancer detection by using image enhancement techniques (V.K. Patel et al. 2012). They are used different techniques in enhancement algorithm such as frequency domain and spatial domain. A recurrence space smoothing-sharpening system is proposed and its effect is surveyed to supportively enhance mammography image. They are used contrast enhancement algorithm and using gabor filter. Gabor filter is used for find out good image. They took the mammogram image and find out different filter values than after used the PSNR algorithm and found out

better results in numerical values. The result pictures turned up new areas which may be of expressive redirection. Combining Gabor figuring with quick Fourier change and overlay cover division showed to be a greatly powerful techniques for wiping out clamor and updating edges, along these lines improving the sign to-commotion extent. Superimposition of pictures changed using diverse techniques into single picture inside indicated to profitably upgrade the perceivability and facilitate the ID of profitable data to the human eye contrast with different routines for separating. Athanasiadis I. Emmanouil and his gathering proposed mammographic picture improvement utilizing wavelet based handling and histogram leveling (Athanasiadis I.E et al. 2004). The purposed of this research was investigate the effectiveness of wavelet transform by processing the discrete wavelet transform (DWT) detail coefficients with a sigmoid function and also used Histogram Equalization Mapping Functions (HEMF) They took the 6 parameters and enhancing the images and found out the result 91% over all images. It was watched that the sigmoid wavelet-based channel fulfilled better perception of the bosom skin, the thoracic muscle, vessels, veins, and channels, while it improved the separation of the picture representation of average greasy bosoms in an alluring level.

Wener Borges Sampaio and his group proposed to Detection of masses in mammogram images using CNN, geostatistic functions and SVM (Wener et al. 2011). This work displays a computational system that helps powers recognize bosom masses in mammogram pictures. The primary period of the method intends to improve the mammogram picture. This stage contains in emptying dissents outside the bosom, diminishing commotion and highlighting the inside structures of the bosom (Sampaio et al. 2011). Next, cell neural frameworks are used to part the regions that may contain masses. These areas have their shapes explored through shape descriptors (uncommonness, circularity, thickness, indirect divergence and round thickness) and their arrangements separated through geostatistic capacities. Bolster vector machines are

used to gathering the candidate regions as masses or non-masses, with affectability of 80%, rates of 0.84 false positives each picture and 0.2 false negatives each photo, and an area under the ROC bend of 0.87. M. Sundaram and his group worked on Histogram Modified Local Contrast Enhancement for mammogram images (Sundaram et al. 2010). They are used different enhancement techniques and find different results and modify the local histogram. Contrast change for mammogram pictures is presented. The proposed work gives a superior differentiation upgrade and information protection of data mammogram picture. The test outcomes are similarly more effective without trading off differentiation furthermore interesting information. As it uses the information histogram, which does not change basically, the proposed work does not exhibit relics in the yield. The proposed system is more suitable for an extensive variety of mammogram pictures of greasy, greasy glandular and thick glandular mammogram pictures and its execution is surveyed for each of the 22 amounts of Mias mammogram pictures with microcalcification. The subjective and target measures are furthermore consoling. This work may be connected with test its execution for mammogram pictures in vicinity of clamor and this procedure can bolster for distinguishing of microcalcification in mammogram. It is clear that the proposed framework has extended the perceptibility of microcalcifications present in the mammogram pictures.

Indra Kanta Maitra and his group prosed to analysis image using Technique for preprocessing of digital mammogram (Indra et al. 2011). They are utilized create calculations for PC helped location (CAD). This estimation is used to reduction clamor, edge-shadowing impact, absolutely perceive pectoral muscle and smother the pectoral muscle adequately without losing any information from the straggling leftovers of the mammogram (Maitra et al. 2011). The resultant mammogram can be used further for the computerized variations from the norm discovery of human bosom like calcification, encompassed masses, spiculated masses and

other not very much portrayed masses, outlined injuries, asymmetry examination and so on. Further work may be coordinated to make and smoothening the pectoral muscle segmentation.

Mohamed Meselhy Eltoukhy and his group proposed a statistical based feature extraction method for breast cancer diagnosis in digital mammogram using multiresolution representation (Mohamed et al. 2011). They are used multi-resolution representations, wavelet or curvelet transform. A highlight extraction technique is made in light of the quantifiable t-test strategy. The framework is situating the elements (segments) as demonstrated by its ability to discrete the classes. At that point, a dynamic utmost is joined with improve the amount of highlights, which can perform the best characterization accuracy (exactness) rate. The game plan order rate is made headway. Bolster vector machine (SVM) is utilized to orchestrate between the typical and anomalous tissues and to perceive kind kindhearted and dangerous tumors. Using curvelet coefficients, the gathering arrangement rate go to 95.98% to gathering amidst ordinary and anomalous and 97.30% to make sense of if the tumor is obliging benign or malignant. Nawazish Naveed and his group proposed to Quantum and impulse noise filtering from breast mammogram images used two modules, noise detection and noise filtering. For discovery reason, neural network is utilized which successfully distinguish the noise from exceptionally corrupted images (Naveed et al. 2012). The proposed strategy has been attempted on salt & pepper and quantum commotion present in mammogram pictures. Crest sign to clamor proportion (PSNR) and basic comparability list measure (SSIM) are used for examination of proposed framework with distinctive existing systems. Examinations shows that proposed framework convey better results as stand out from existing routines.

CHAPTER 3

METHODOLOGY

3. Methods:

The strategy Pattern recognition fundamentally begins from the requirement for automated machine recognition of signals, pictures and objects or any choice based methodology, on the premise of the arrangement of feature. The objective of Pattern recognition is to predict the right level comparing to given list of capabilities feature taking into account a superior learning got through training. The example of pattern recognition can surely knew by considering a case: the person can undoubtedly recognize the gender taking into account the face while machine can't, so the point is to prepare the machine by considering different feature as: outward appearance, facial bone structure, hair length and others. In the wake of preparing, the machine can undoubtedly recognize the obliged class the new test article has a place. From this above case, it is very clear that the heart of example recognition pattern is feature extraction and characterization.

The main goal of this research study is managing the classification of the cancerous and non-cancerous images from mammogram breast images, to used image processing and machine learning strategies. In our research the pattern recognition techniques to classification of feature and identify the cancerous and non-cancerous image given by the flow chart, as show in figure 1. Here, the proposed Pattern recognition framework is created with the assistance of 120 quantities of mammogram breast images; out of these 60 quantities of image are have cancerous and 60 has belong to non-cancerous. The software programming to execute this framework is created in MATLAB programming. Recognition of tumours in mammogram is partitioned into three principle stages. The principal step includes an enhancement strategy, image enhancement strategies are utilized to enhance an image, where to expand the SNR and to make certain feature less challenging to get by altering the modifying colours or brightness. After that the brightness conformity is a greatness of image values to another reach. The test

archived the images collection from mammogram and image processing parts are discussed about in this section where the machine learning part is examined in section 3. Under image processing part, image segmentation i.e. to segment image background and feature extraction; i.e. to extract of different significant features from every segmented image.

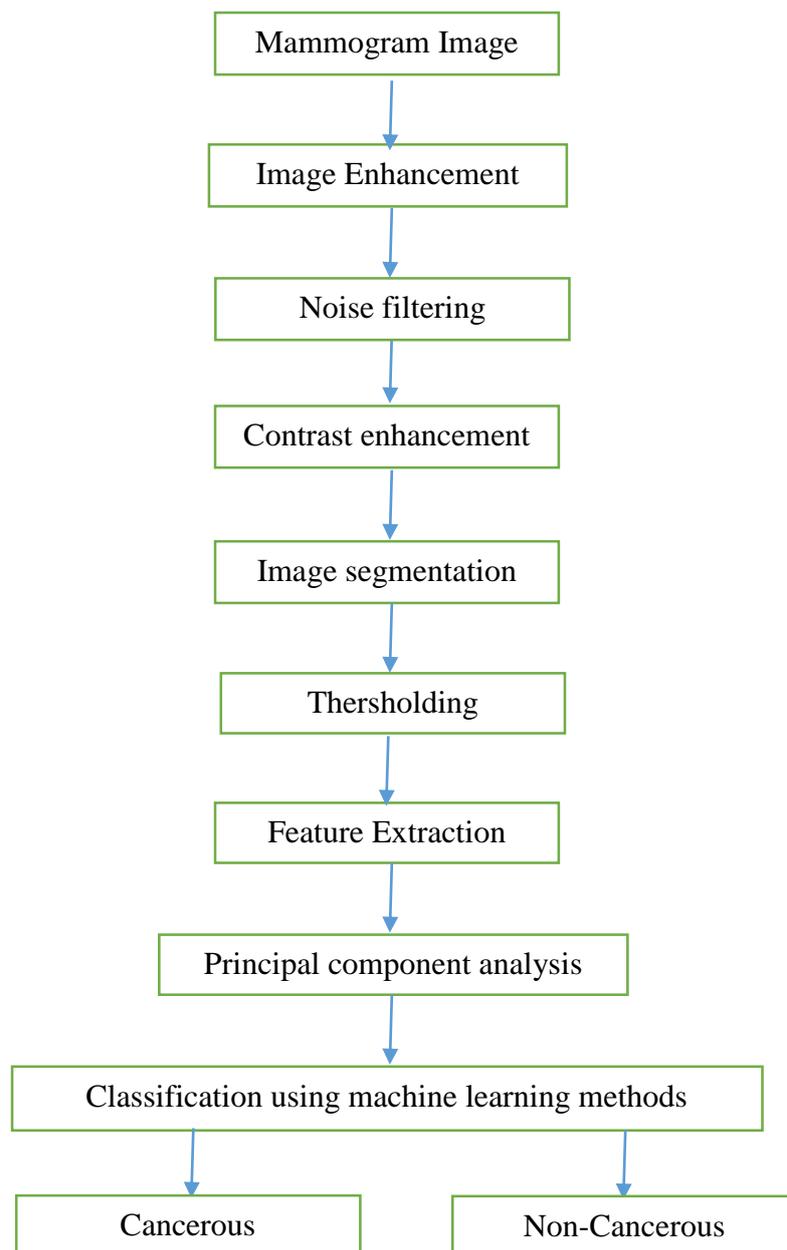


Figure 1. Flow chart shows the pattern recognition methods and classified mammogram images.

3.1 Experimental Documented data collection:

In my research work to collected the data for online mammogram database which name is Mammography image analysis society (MIAS). MIAS is open source database used to research work and researchers to be able directly compare the results. The Mammographic Image Analysis Society (MIAS) is an association of UK researcher groups inspired by the comprehension of mammograms and has produced a database of computerized mammograms. The Digital Database for Screening Mammography (DDSM) is another asset for conceivable use by the mammogram images investigation investigation groups. To take the mammogram images than applied different methods to improve the good quality of image.

3.2 Experiment methods:

Digital image processing is a part of signal processing which is input signal as form an image such as a video or photograph frame. The output find out such as image or parameters characterizations which is linked to the image. Maximum image transforming processes include regarding the image as a 2-D (two-dimensional) signal and applying standard signal processing techniques to it.

In my research I used different techniques such as;

- (1) Image Enhancement,
- (2) Image Segmentation,
- (3) Feature Extraction.

3.2.1 Image Enhancement:

In my research work to take image than after enhancement the image. Enhancement is the technique which is used the improve quality of image such as improving contrast, blurring, noise reduces. Enhancement means improvement, in image processing different reason to

defect the image such as instrumentation noise which is called electronic noise and other one is called photon noise. So that need the image enhancement. Following parameter are used to enhanced the image and find out best parameter results used the enhancement image for segmentation such as.in image enhancement to used different filters and find out different PSNR (peak signal to noise ratio). If the peak signal to noise ratio is high than image quality is better shown in fig.2

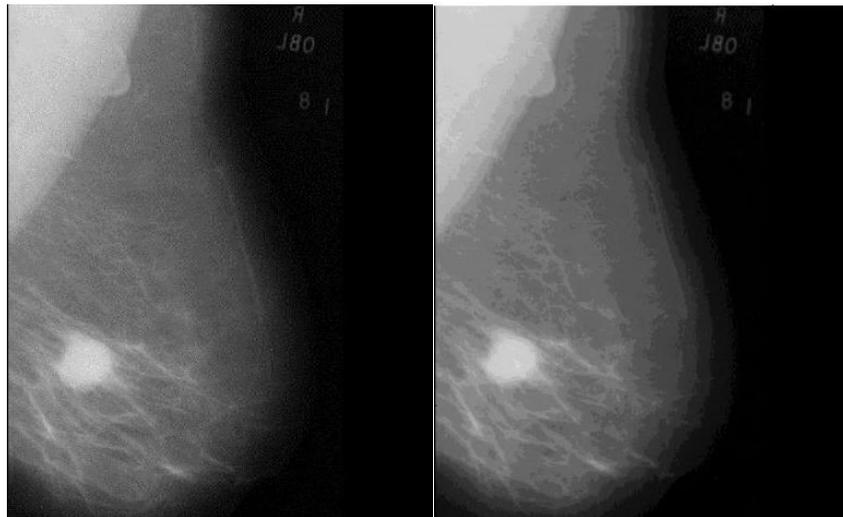


Figure 2: Shows the original image and filtered image.

Gaussian filter and Gabor filter are used to image enhancement.

3.2.1.1 Gaussian Filter: Gaussian filter is a most important parameter of image enhancement. The Gaussian pressing function is a two dimation convolution parameter that is applied to `blur' pictures and to clear feature and noise. In a gaussian filter standard deviation is play very important roll in its property.gaussian fuction is used no. of research ares such as,

- It characterizes a probility distribution for data or noise
- Smoothing operator
- Mathematics.

Given formula below,

$$p(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \quad (3.1)$$

Where μ is denoted mean and σ is denoted standard deviation.

3.2.1.2 Gabor filter: Gabor filter is work like a band pass filter which is used unidimensional signal. In gabor filter ,impulse response created by multiplying of gaussian envelope fuction with a multipule oscillation. Basically we can say that it is multiple of gaussian filter to used feature exctraton and texture analysis. Mathematical equation is given below,

$$f(x, y, \omega, \theta, \sigma_x, \sigma_y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[\frac{-1}{2}\left(\left(\frac{x}{\sigma_x}\right)^2 + \left(\frac{y}{\sigma_y}\right)^2\right) + j\omega(x \cos \theta + y \sin \theta)\right] \quad (3.2)$$

Where ω is frequency, σ is special spread and Θ is orientation.

3.2.2 Image Segmentation:

Image Segmentation is a basically image division or partition. It is procedure of partition of image into various region that can be connected with the properties of one or more criterion. Image segmentation is commonly used to find objects and boundary (based, edge etc.) in pictures. All the more unequivocally, image segmentation is the procedure of appointing a label to each pixel in an image such that pixels with the same label share certain characteristics. Here I am using texture based image segmentation which give statistical value. I am using step by step image segmentation, the least difficult case is having just two region, an object region and a background region. In segmentation procedure is segment an image into areas without knowing what these region present (Niblack, 1986).The main prompts to comprehension areas of interest are based of further investigation on their composition and shape. Most ways to deal with Segmentation fall under one of following classifications.

- Thresholding.
- Edge Detection.
- Region Detection.

3.2.3 Feature Extraction:

Extricating feature from mammographic images utilizing workstation is the most ideal approach towards distinguish breast cancer. That implies, in the image processing techniques will advantage the medicinal process in identifying breast cancer and it will help to enhancement medicinal advancement. These are various classifications for feature in image processing. There are many categories of feature extraction such chip histogram feature, wavelet based feature extraction.

3.2.3.1 Chip histogram based feature extraction:

To take the image which is converted in gray level scale. In mage which present no. of gray levels vector which varies from 1 to L. where L is gray level of image. Various chip histogram is extracted from each cancer and no cancerous segmented image .it explained in detailed below, that functions are to get of the chip histogram based structures such as:

- Mean.
- Variance.
- Skewness.
- Kurtosis.
- Energy.
- Entropy.

- (a) **Mean:** In a mathematics, mean is defined the sum of all pixel values of image divided by counts of pixels. Formula given below:

$$\text{Mean} = \frac{\sum_i \sum_j I(i,j)}{n} \quad (3.3)$$

- (b) **Variance:** Variance is measured by taking the differences between each number in the set and the mean, squaring the refinements (to make them positive) and differentiating the total of the squares by the amount of characteristics in the set. A variation of zero shows that all the abilities are identically.

$$\text{Simple variance} = S^2 = \frac{\sum (X - \bar{X})^2}{n-1} \quad (3.4)$$

- (c) **Skewness:** Skewness measures the asymmetry of the likelihood probability distribution of original value irregular variable. A dispersion has positive skew (right-skewed) if the benefit (higher worth) tail is longer or fatter and negative skew (left skewed) if the left (lower quality) tail is longer or fatter. The Skewness of the regular scattering (or any flawlessly symmetric circulation) is zero.

$$\text{Skewness}(S) = \frac{\sum_{r_k=1}^L P(r_k) \times (r_k - \mu)^3}{\sigma^3} \quad (3.5)$$

- (d) **Kurtosis:** Kurtosis is a calculated of whether the information are flat or peaked with respect to a probability dispersal. That is, information data with maximum kurtosis have an inclination to have a particular peak close to the mean, decrease rather quickly, and have heavy tails. Information data sets with low kurtosis have a tendency to have a flat highest close to the mean as opposed to a piercing peak. Kurtosis calculated by given formula below,

$$\text{Kurtosis}(Ku) = \frac{\sum_{r_k=1}^L P(r_k) \times (r_k - \mu)^4}{\sigma^4} \quad (3.6)$$

(e) **Entropy:** Entropy is a calculated of the instability connected with irregular variable.

Entropy in a data sense is calculate of unpredictability. It is a statistical measure of arbitrariness that can be utilized to describe the texture of the image. To calculated by given formula below,

$$\text{Entropy} = - \sum_{r_k=1}^L P(r_k) \times \log P(r_k) \quad (3.7)$$

(f) **Energy:** Energy of an image gives the ideas about measure of the information. Here and there the energy can be a negative calculate to be lessened and in some cases it is a positive calculate to be highest. Energy is utilized to depict a measure of data when detailing an operation below a probability system, for example, maximum a priori (MAP) estimation in conjunction with Markov random Fields (Lambrou, 2011). To calculated by given formula below,

$$\text{Energy} = \sum_{r_k=1}^L (P(r_k))^2 \quad (3.8)$$

3.3 Machine Learning:

Machine learning is a co-field of software engineering that developed from the investigation of example acknowledgment and computational learning theory in artificial intelligence. Machine Learning is a top notch ticket to the most energizing professions in information to data analysis. As information sources multiply alongside the computing power to process them, setting off straight to the information is a standout amongst the most direct approaches to rapidly pick up bits of knowledge and make predictions. Machine learning is a subfield of software engineering that developed from the investigation of example acknowledgment and computational learning theory in artificial intelligence. Machine learning investigates the development and investigation of calculations that can gain from and make expectations on information. Such calculations work by building a model from sample inputs to make

information driven expectations or choices instead of taking after static program instructions. The term "Machine Learning" originates from the artificial intelligence group yet now a day it for the most part the centering zone for some branches of engineering and science (Alpaydin 2004). Learning is refer to learning from data set or feature set. There are many methods of learning for staticstical data analysis. There are many types of methods such as supervised, unsupervised and support learning.

These objectives of supervised learning is typically sum up that's data input output relationship and facilitating the expectation of output connected with previously unseen inputs. The essential supervised learning assignments are the arrangement and regression. Additionally, the objective of the unsupervised learning is ordinarily not identified with the future perceptions. Rather, one looks to comprehend structure in information test itself, or to deduce a few qualities of the probability distribution. The fundamental unsupervised learning assignments are dimensionality lessening, the clustering and density estimation. Dimensionality reduction can likewise be taken as the class of supervised learning however there ought to be the input output relationship. In the event of reinforcement learning in the input pattern are observed successively and after every observed pattern the learner must make a move. The fundamental objective of the learner in supervised learning is to focus an arrangement to maximum the long term reward. This kind of learning is chiefly helpful in financial matters, robotics (way arranging and route) and different areas (Hastie et al. 2005). In this study, the PCA calculation, which goes under the unsupervised learning undertaking, is utilized for the dimensionality lessening (Kambhatla et al.1997).

3.3.1 Principal component Analysis Using for Dimensionality Reduction:

In machine learning techniques most important technique is Principal component analysis which is a measurable method that uses an orthogonal change to change over an arrangement

of observations of perhaps associated variables into an arrangement of estimates of straight uncorrelated variables called Principal component. The principal components is not precisely or equal to the amount of interesting variables. This change is described in such a way, to the point that the first foremost parts has the greatest possible distinction (that is, records for however a significant part of the variability in the data as could sensibly be normal), and each succeeding part along these lines has the most raised variance possible under the impediment that it is orthogonal to (i.e., uncorrelated with) the previous parts. The essential parts are orthogonal in light of the fact that they are the Eigen vectors of the covariance grid, which is symmetric. PCA is delicate to the relative scaling of the first variables.

Mathematical principal component analysis is describe consider \mathbf{X} a matrix and \mathbf{p} is dimensional vector of weight where

$\mathbf{w}_{(k)} = (w_1, \dots, w_p)_{(k)}$ that map individually row vector $\mathbf{X}_{(i)}$ of \mathbf{X} to a new vector of principal component *scores* $\mathbf{t}_{(i)} = (t_1, \dots, t_p)_{(i)}$, given by

$$t_{k(i)} = \mathbf{X}_{(i)} \cdot \mathbf{w}_{(k)} \quad (3.9)$$

t considered over the information set progressively get the most extreme conceivable change from \mathbf{x} , with every stacking vector \mathbf{w} obliged to be a unit vector. Calculated the covariance matrix

$\mathbf{S} = \frac{1}{m} \mathbf{x}^T \mathbf{x}$ where m is row. Than after obtained Eigen value and Eigen vector of matrix \mathbf{S} from each row which denoted by ' \mathbf{pi} ' matrix obtained \mathbf{P} .

Than after reducing feature using PCA methods, the classifiers (ANN, SVM and LSSVM) which go under the supervised learning undertaking of machine learning are utilized to group this arrangement of features into cancerous and noncancerous classes. In machine learning and measurements the order is an undertaking of recognizing, to which class the new perception has a place. A calculation which actualizes the arrangement is essentially called as the

classifier. For the most part order is partitioned into two classes i.e. the binary classification arrangement and multi class characterization.

3.3.2 Artificial Neural Network (ANN) as classification:

Artificial Neural Networks are by and large rough electronic models in view of the neural structure of the Brain. The brain fundamentally gains as a matter of fact or we can say that learning from experiences. This brain model likewise guarantees a less specialized approach to create machine arrangements. This better approach to manage computer moreover gives a more smooth corruption system over-weight than its more conventional partners. It comprises of a large number of highly interconnected handling components (called neurons), which are working on simultaneously at the same time to tackle a particular assignment in machine learning. These assignments are essentially the order and capacity estimation. Like person the ANN is additionally gaining from the occurrences or cases (Mohanty et al. 2010). To plan an ANN based framework, a quantities of programming ideas and numerical models are helpful.

3.3.2.1 Multilayer feed forward Neural Network (MFNN) Classification

In ANN using three layer which is call input layer than after hidden layer than after output layer. Input layer is a basically applied input parameters for examples chip histogram parameters are applied as an input layer. In input layer I am using PCA data reduction set as an input. Hidden layer is most important layer which is present in middle of input layer which present input features parameter and output layer. Output layer is defined to found out the desire output, it assigned as class 1 for cancer and class 0 for noncancer. The three layers are showing in Fig 3

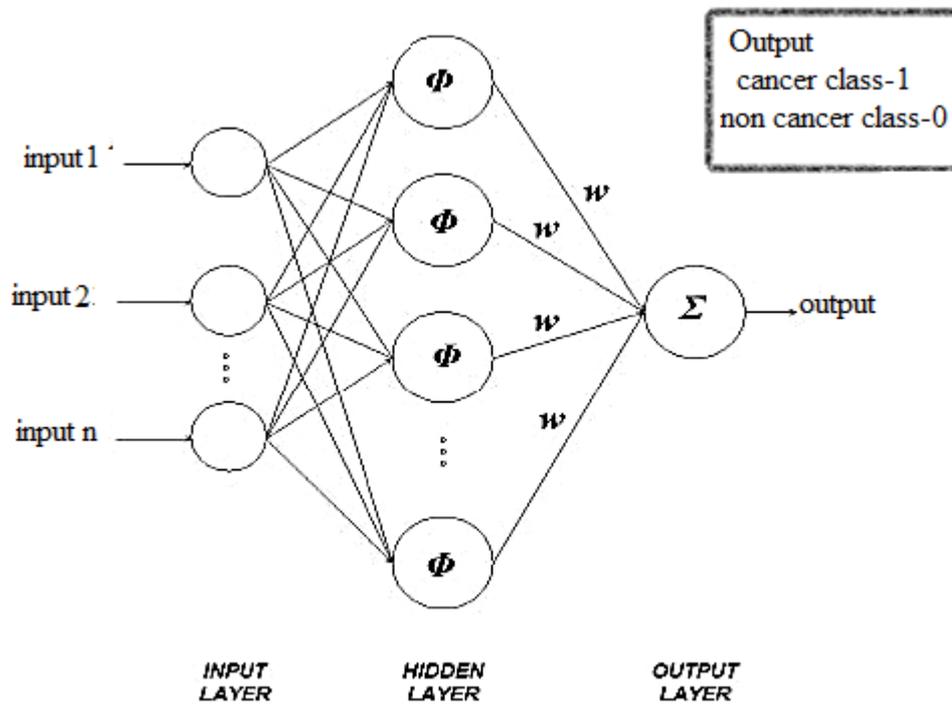


Figure 3: Block diagram of Multilayer feed forward neural network.

3.3.2.2 Testing and training images for ANN:

In neural network there are number of preparing calculation are utilized to prepare the neural system, while in ANN back spread calculation is a straightforward utilized as a part of neural system which is easy to understanding and executed. It is divided four stages, one stage is Initialization of weight values, second stage is feed forward Propagation, 3rd step is Back propagation of errors, and final stage is updating the weight values (Sivanandam et al. 2006).

Considered training vector matrix is $x = (x_1, x_2, x_3, \dots \dots x_{n1})$ as input matrix .now I am using PCA method to reduce the data set feature so obtained the T output unit vector which is binary value (T=0 for noncancer, and T=1 for cancer).

3.3.3 Support Vector Machine:

SVM is the classification of supervised learning strategies, which is fundamentally utilized for illuminating grouping and relapse assignments in machine learning. The reduced features acquired from PCA stage are straight forwardly to both SVM and LS-SVM Classifiers. At the output, the classifier predicts, the features comparing to image have a place with either cancerous or non-cancerous classes. The SVM and LS-SVM are goes about as non-probabilistic parallel direct classifiers and manufactures a Hyperplane which separates the two classes. The most imperative idea of SVM and LS-SVM is 33 dividing Hyperplane which can know about SVM by these three point (Cristianini et al. 2000; Suykens et al. 1999).

- The most extreme edge Hyperplane (This relates to the best Hyperplane is the particular case that divides the two classes and lies at a maximal separation from any examples).
- The delicate edge which is called soft margin (This allows some level of the misclassification).
- The kernel function (This capacity maps the info reduce feature to a higher measurement highlight space, with the objective of discovering a partition between the information).

The margin implies the maximal width of the piece parallel to the Hyperplane that has no inside data centres (Campbell et al., 2011).

In the training data set which is used $\{x_n, y_n\}_{n=1}^k$ with the input data set $x_n \in R^m$ and output data set $y_n \in R$. now decision function is associated with Hyperplane to recognize the binary classes.

$$Y(x) = \text{sgn}(w^T \varphi(x) + b) \quad (3.10)$$

where w for weight vector, b for bias value. These weight vectors and bias values for ideal Hyperplane SVM were gotten by augmenting the separation between the bias and Hyperplane.

Then achieved by maximizing the margin which defined by $M = \frac{2}{|w|}$ and minimization of

$$\frac{\|w^2\|}{2} = \frac{1}{2} w^T w \tag{3.11}$$

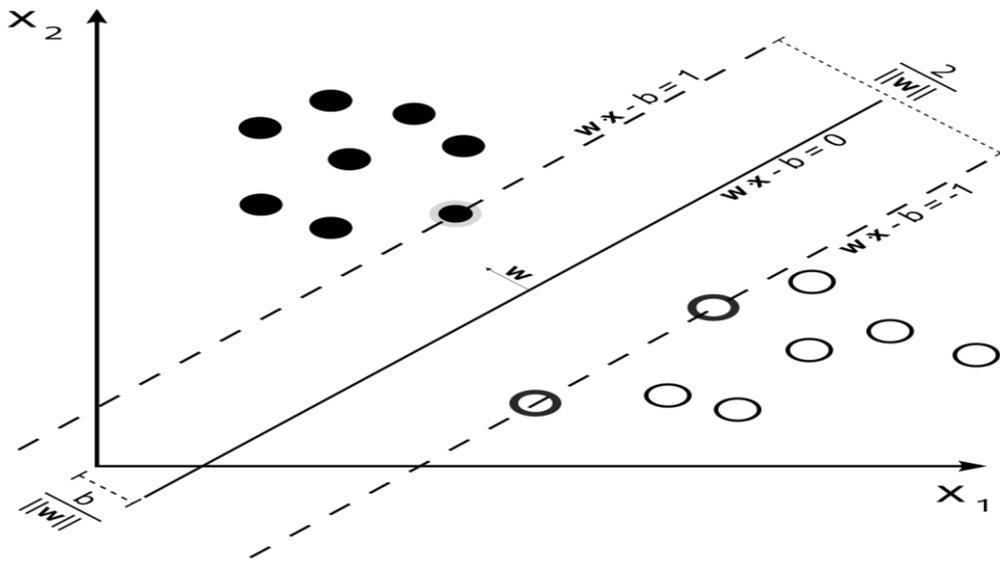


Figure 4: Working of SVM for the nonlinear distinguishable information where dark circle symbol used for cancerous data and circle for noncancerous data.

The output feature are separated by decision as corresponding to output class

(y = +1) if

$$w^T \phi(x_n) + b \geq +1 \tag{3.12}$$

Similarly (y = -1) than

$$w^T \phi(x_n) + b \leq -1 \tag{3.13}$$

There are two set which is inequality so combine these set than find out it from eqtn 4.4 and eqtn 3.12 than

$$y(w^T \phi(x_n) + b) \geq 1, \tag{3.14}$$

Where n = 1 2 3...k

Now eqn 3.13 and eqn 3.16 is combine than

$$\min_{w,b,\xi} J_{\text{Primal}}(w, \xi) = \frac{1}{2} w^T w + c \sum_{n=1}^K \xi_n, \text{ such that, } y(w^T \phi(x_n) + b) \geq 1 - \xi_n \quad (3.15)$$

Where $\xi_n \geq 0$ and $n=1,2,\dots,K$.

Now solve the problem and lagraingian is present below

$$L(w, b, \xi, \alpha, v) = J_p(w, \xi) - \sum_{n=1}^K \alpha_n (y_n [w^T \phi(x_n) + b] - 1 + \xi_n) - \sum_{n=1}^K v_n \xi_n \quad (3.16)$$

Where α_n and v_n are lagraingian multiplier $n = 1, 2, 3 \dots k$

The eqn 3.16 is solve by differentiation with reverence to w, b and ξ_n to assign the initial value is zero than

$$\frac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{n=1}^K \alpha_n y_n \phi(x_n) \quad (3.17)$$

$$\frac{\partial L}{\partial b} = 0 \rightarrow \sum_{n=1}^K \alpha_n y_n = 0 \quad (3.18)$$

$$\frac{\partial L}{\partial \xi_n} = 0 \rightarrow 0 \leq \alpha_n \leq c, n = 1, \dots, K \quad (3.19)$$

To solve quadratic optimization problem which is called dual problem with respsect to the

primal problem is given below:

$$\max_{\alpha} J_{\text{Dual}}(\alpha) = -\frac{1}{2} \sum_{n,l=1}^K y_n y_l K(x_n, x_l) \alpha_n \alpha_l + \sum_{n=1}^K \alpha_n \text{ Such that } \sum_{n=1}^K \alpha_n y_n = 0 \quad (3.20)$$

This quadratic optimization used in Kernel fuction K and find out the mathematical model for

input features of Kernel fuction given below:

$$K(x_n, x_i) = \phi^T(x_n) \phi^T(x_i) \quad (3.21)$$

Now find out new data for using kernel task

$$y(x) = \sum \alpha_i y_i k(x_n, x_i) + b \quad (3.22)$$

3.4 Performance Evaluation Criteria for classification:

The performance evaluation depends on various factors such as classification rate (CR), sensitivity, specificity and area under Receiver operating characteristics (ROC) curve (Lu et al. 2004).

Table 1. General procedure to solve by confusion matrix for binary classification.

TP	FN
FP	TN

Where TP for that no. of true positive, TF for that no. of true false, FN for that no. of false true and FP for that no. of false positive.

$$CR = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (3.23)$$

Where CR is classification Rate.

Sensitivity: It is used for identification of true positive cases

$$\text{Sensitivity} = \frac{TP}{(TP+FN)} \quad (3.24)$$

Specificity: It is used statistical analysis and also find binary classifier for true negative class.

$$\text{Specificity} = \frac{TN}{(TN+FP)} \quad (3.25)$$

If specificity is one for alarm rate the calculate false negative class

$$\text{False alarm rate} = 1 - \text{Specificity} = \frac{FP}{(TN+FP)} \quad (3.26)$$

CHAPTER 4

RESULT

&

DISCUSSION

Initially, I have used enhancement techniques. In enhancement techniques using different filter and find out better result by help of PSNR show the results of different images

To take the images and find out different filters results, where

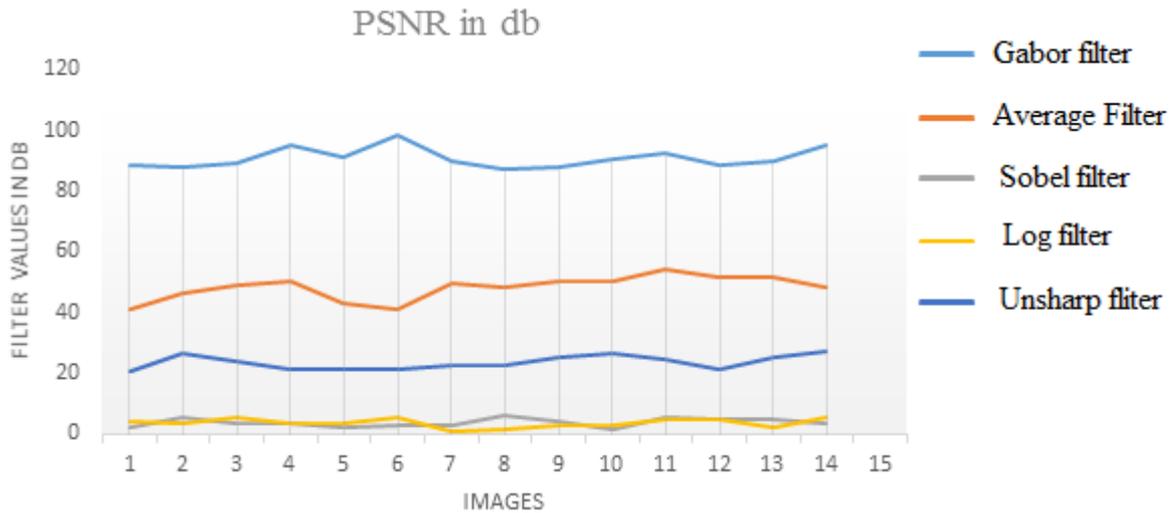


Figure 5: To take the images and find out different filters results.

In PSNR property that if high psnr any image than that image give more information (Vishnukumar et al. 2011). Now to take the enhancement image for segmentation techniques and find out boundary and edge detection shown

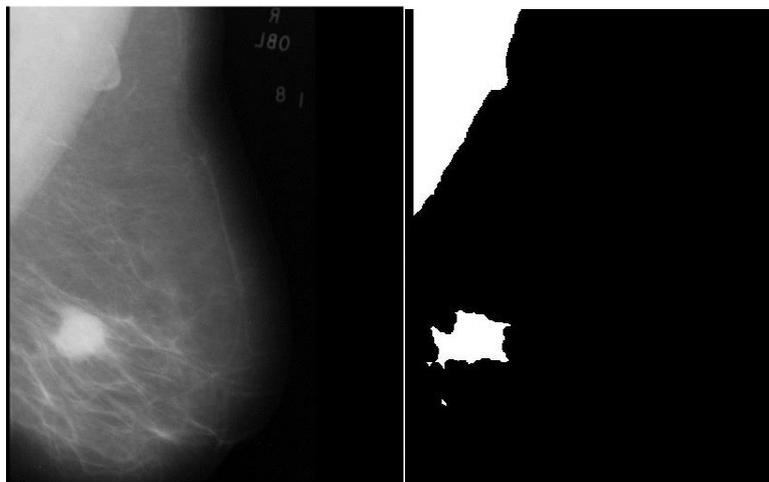


Figure 6: Original image than after find segmented threshold image.

In segmented image using dilation and erosion which give information maximum and minimum pixel values show

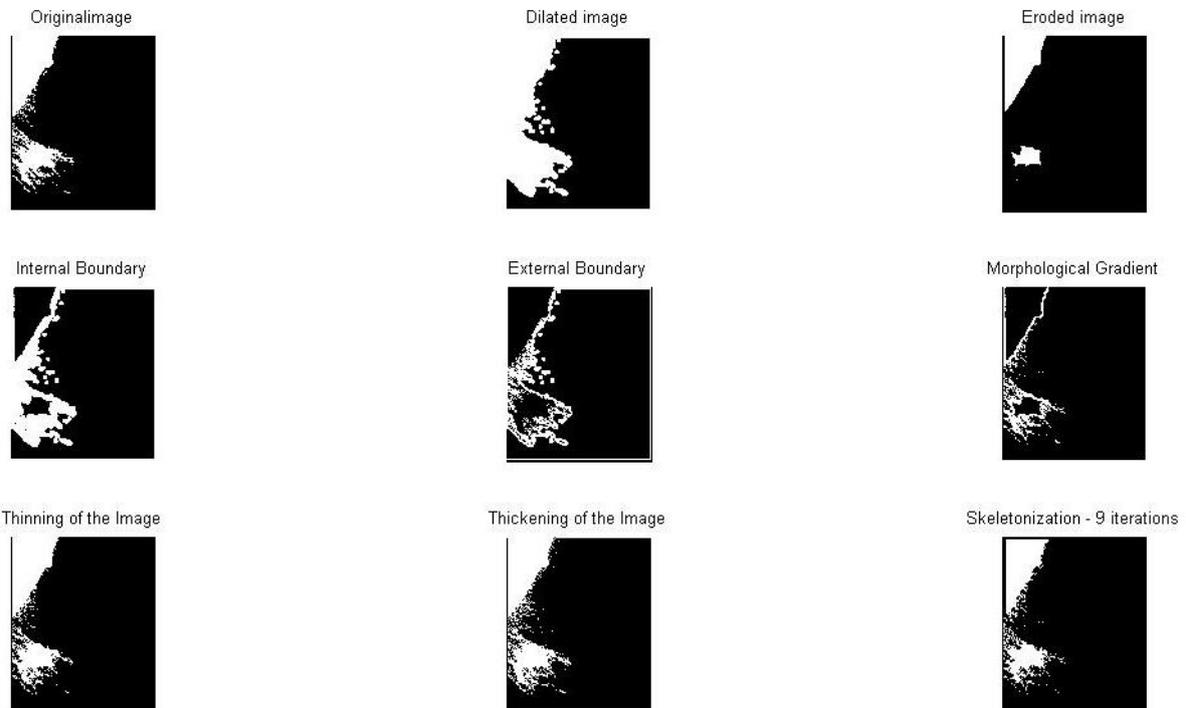


Figure 7: Morphology operation using Dilation and Erosion.

After segmented image find out the region of interest (ROI) shown in figure below;

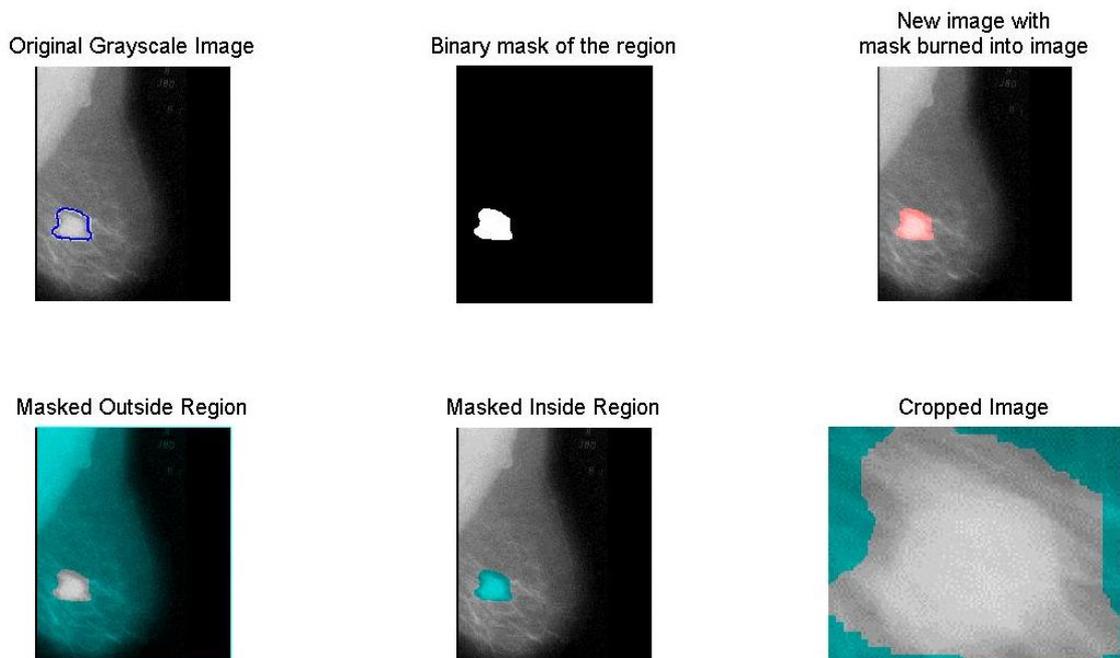


Figure 8: Image shows the ROI by selected region of tumour.

After segmented image then required texture feature extraction such as chip histogram and extract feature of each segmented image. I take 120 image and extracted 90 feature of single image and make the feature data set and analysis. The size of data set 120×90 so I am using PCA method to reduce the data set and converted into 120×15 which is easy to classification purpose and denoted class 1 for cancer and class 0 for noncancer.

The features reduce by principal component analysis method for using as an input for different techniques for example ANN and Support Vector Machine (SVM) classification for training and testing purpose. Now the output used a binary value 0 for non cancer and 1 for cancer image. I am taking 50 % data for training and remaining data for testing analysis. After tested data and trained data set are evaluated and find out predict output and actual output to compare for classifier performance.

4.1 Artificial Neural Network Classification:

In artificial neural network classifier the different parameters are used to find out no. of hidden neurons, learning rate, moment factor, numbers of iterations and training parameter goal. Show in table 2 below;

Table 2: In ANN to get optimize training performance these given various parameter.

ANN parameter	Values
Number of hidden neurons (Layer)	15
Learning rate	0.06
Number of iterations	1000
Training Parameter Goal	10^{-9}

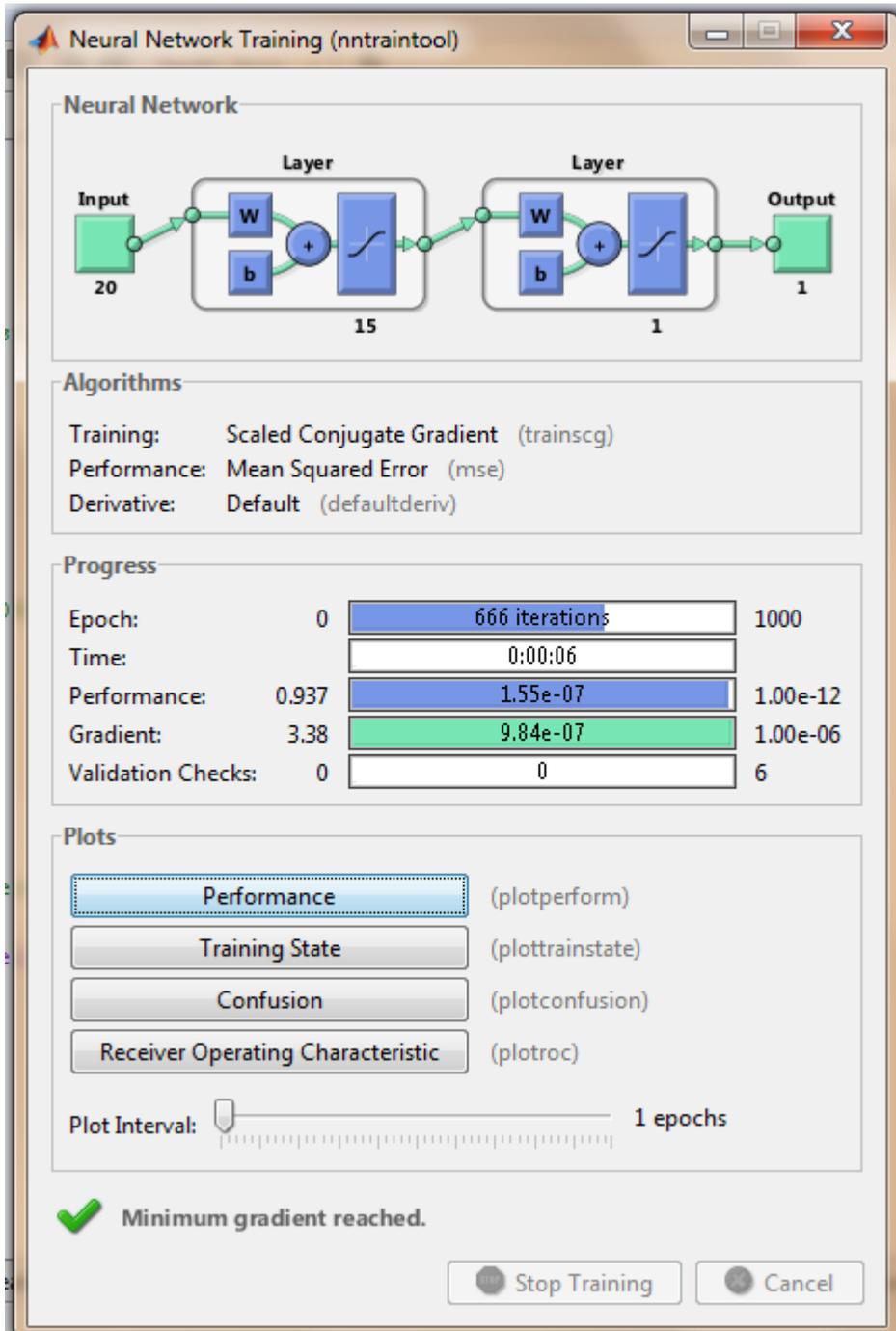


Figure 9: ANN tool used to find different parameters for classifier.

In ANN, the variation of mean square error (MSE) with respect to epoch is shown in the performance graph in Figure 10. After 666 cycles, the optimal training goals were achieved, and further test information was assessed by passing it to the advanced trained ANN network. The data set was used to determine the classification rate (CR), sensitivity, and specificity.

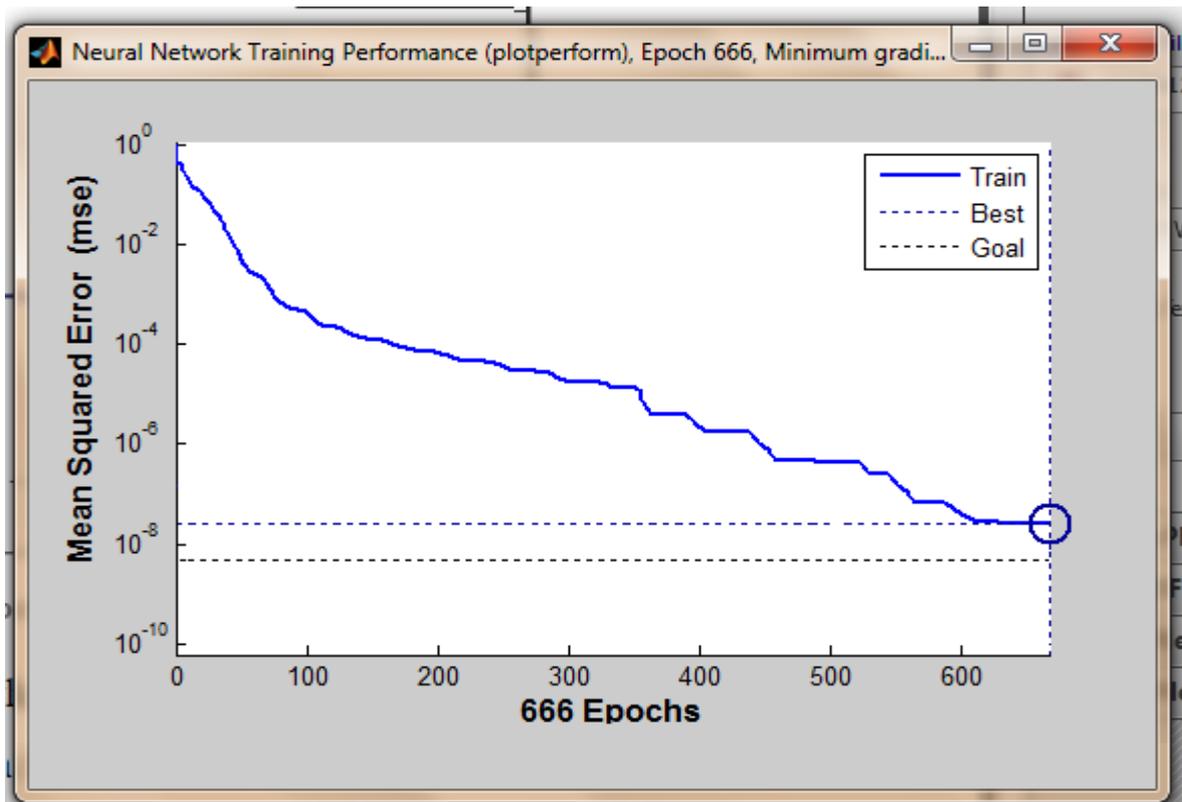


Figure 10: Variation of training MSE with respect to number of iterations.

In that window demonstrates, the information has been isolated utilizing the separator and capacity. In ANN, the preparation technique is utilizing with the methods square lapse execution capacity. In ANN review that these all default setting for food forward system. In this execution chart which is indicating MSE and ages. Through the preparation, the advancement is frequently upgraded in the preparation capacity. Of most hobby are the execution, the size of the inclination of execution and the quantity of approval checks. The size of the angle and the quantity of approval checks are utilized to end the preparation. The MSE is just as it says: The mean (average) magnitude of the squares of the error: i.e., the distance between the model's estimate of your test values and the actual test value. (Squaring just converts things to an absolute value rather than fiddling with under or overshooting). The gradients will turn out to be little as the preparation achieves at least the execution. A very

much prepared ANN ought to have a low MSE toward the end of the preparation stage. The importance of MSE being little (near to zero) is that the sought yields and the ANN's yields for the preparation set have turn out to be near to one another. Now find out the confusion matrix, which give the information about predictive analytics. Which confusion matrix is find out true positive, true negative and false positive, false negative results than found out the accuracy given in tables below;

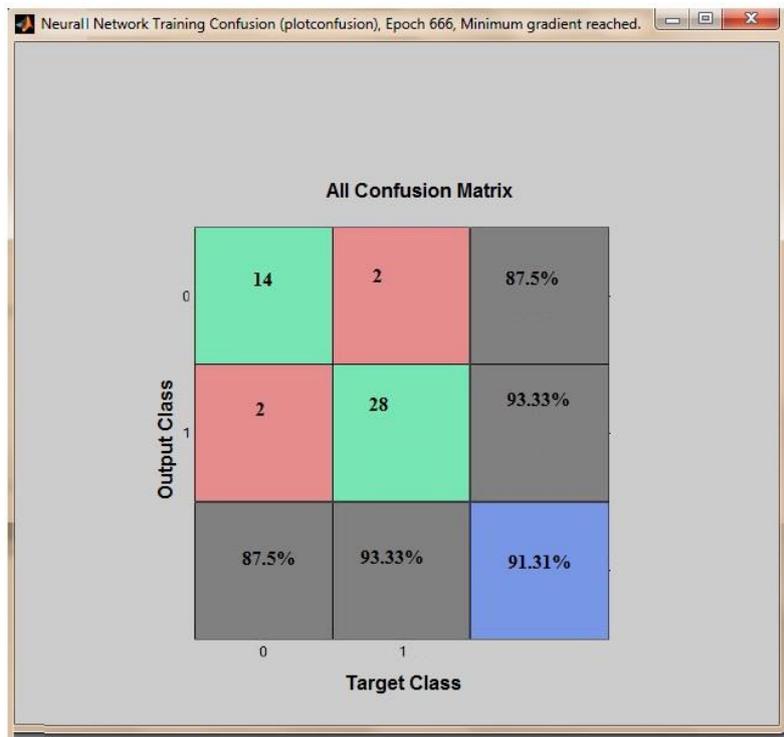


Figure 11: Confusion Matrix for positive and negative results.

Table 3: General analysis by confusion matrix.

Prediction \ True	Cancerous	Noncancerous
Cancerous	14	2
Noncancerous	2	28

Now find out of the parameters by given confusion matrix.

$$CR = \frac{14+28}{14+2+2+28} \times 100 = 91.31\%, \quad \text{Sensitivity} = \frac{14}{14+2} \times 100 = 87.5\%,$$

$$\text{Specificity} = \frac{28}{28+2} \times 100 = 93.33\%.$$

4.2 SVM classifier Performance:

To obtain the classification performance by SVM technique with both RBF kernel and linear kernel function. Here I used number of training parameters in SVM technique and find out classifier performance with number of support vector and number of iterations.

In given table to take support vector and iteration below;

Table 4: Given the SVM parameters and its values for RBF.

SVM Parameter	Values
Support Vector	40
Number of Iteration	1000

Table 5: Given the SVM parameters and its values for linear kernel.

SVM Parameter	Values
Support Vector	27
Number of Iteration	1000

The data set are found out by using it to the optimized trained SVM with RBF and linear kernel. The performance parameters of the evaluation of data set are given by confusion matrix and ROC plot. Show in table 6

Table 6: Confusion Matrix for RBF kernel.

Prediction \ True	Cancerous	Noncancerous
Cancerous	7	1
Noncancerous	1	16

$$CR = \frac{7+16}{16+1+1+7} \times 100 = 92.11\%, \quad \text{Sensitivity} = \frac{7}{7+1} \times 100 = 87.51\%,$$

$$\text{Specificity} = \frac{16}{16+1} \times 100 = 94.11\%.$$

Table 7: Confusion Matrix for Linear kernel.

Prediction \ True	Cancerous	Noncancerous
Cancerous	18	0
Noncancerous	2	8

$$CR = \frac{18+8}{18+0+2+8} \times 100 = 92.85\%, \quad \text{Sensitivity} = \frac{18}{18+0} = 1,$$

$$\text{Specificity} = \frac{8}{8+2} \times 100 = 80.00\%.$$

4.3 Comparisons of classifier performance:

After calculation of classification rate (CR), Sensitivity, specificity and under area of receiver operating characteristics (ROC) curve for various classification results shown in table below,

Table 8: Comparisons of classification rate, sensitivity, specificity and ROC.

Classification models	CR (%)	sensitivity	specificity	ROC
ANN	91.31	0.875	0.933	NA
SVM for Linear kernel	92.11	0.855	0.9411	0.60
SVM for RBF	92.85	1.00	0.811	0.90

CHAPTER 5

CONCLUSION

&

FUTURE WORKS

Initial stage of detection of breast cancer has the ability to reduce morbidity and mortality. There are many diagnostic technology and tests to diagnose breast BC. Conventional determination strategies for BC detection is Screening tests (such as yearly mammograms), Diagnostic tests (such as biopsy) etc. The selected ANN and the SVM with RBF kernel model to find out better performance classification rate compare with other models. It was gotten by producing SVM with RBF model accuracy (CR) 92.85%, other than models. In ANN model it is very fast and accuracy is so good, but SVM machine learning is given accurate result to predicted training optimized model. The performances of this propose for evaluated in terms of CR, sensitivity, Specificity and ROC, which are compares other models. It was calculated 120 images which present benign and malignant masses with different size, shape and contrast. The process works properly in all cases and the projected structure was evaluated with and without initial de-noising steps. In both cases results are found to be comparable. Using the proposed ANN classifier, 91.31% sensitivity and 85.7% specificity is achieved 93.33%. In SVM model which is used in two categories with RBF and linear kernel and classification results better than ANN.

Future work

The future advancement of this application may be the followings:

- i. To implement same pattern recognition techniques used in lungs cancer and other cancers.
- ii. To implemented the tumours detection techniques by using AutoCAD and find out the tumours size and patterns.
- iii. To extract feature using LS-SVM and fuzzy-SVM techniques to find out classification rate.

REFERENCES

- Alpaydin E (2004) Introduction to machine learning, The MIT Press.
- American Cancer Society (2003). Cancer prevention and early detection Facts and Figures. Atlanta GA.
- Arora N, Martins D, Ruggerio D, Tousimis E, Swistel AJ, Osborne MP, Simmons RM (2008) Effectiveness of a noninvasive digital infrared thermal imaging system in the detection of breast cancer. American journal of surgery 196(4):523-6.
- Athanasiadis E, Piliouras G (2004) mammographic image enhancement using wavelet-based processing and histogram equalization, ic-scce vol 1.
- Campbell C, Ying Y (2011) Learning with support vector machines. Synthesis Lectures on Artificial Intelligence and Machine Learning 5(1):1-95.
- Cristianini N, Shawe J (2000). An introduction to support vector machines and other kernelbased learning methods. Cambridge university press.
- Gokhale S (2009) Ultrasound characterization of breast masses. The Indian journal of radiology & imaging 19(3):242-7.
- Husuki U, Onoda Y, Kawasaki T, Misumi T, Murakami, Komatsubara H, Teramoto S(2009) Relationship between thermo graphic observations of breast tumours and the DNA indices obtained by flow cytometry, Biomedical Thermology 10 (4) 282–285.
- Hala S, Ali E (2013) Feature Extraction Values for Breast Cancer Mammography Images, IEEE 335-340.
- Hastie T, Tibshirani R, Friedman J, Franklin J (2005) The elements of statistical learning: data mining, inference and prediction. The Mathematical Intelligencer 27(2):83-85.
- Kambhatla N, Leen TK (1997) Dimension reduction by local principal component analysis. Neural Computation 9(7):1493-1516.
- Kingsbury N (1998). Wavelet transforms in image processing. Signal Analysis and Prediction: 27.
- Kopans, D B (1998), Breast Imaging, Lippincott-Raven Publishers, Philadelphia, Tucker.
- Lambrou, Papadopoulos H and Gammerman A (2011), “Evolutionary Conformal Prediction for Breast Cancer Diagnosis”, IEEE 9th International Conference on Information Technology and applications in Biomedicine: 1-4.
- MAIS database (2007), <http://peipa.essex.ac.uk/info/mias.html>.

Mohanty S, Ghosh S (2010) Artificial neural networks modelling of breakdown voltage of solid insulating materials in the presence of void. *Science, Measurement & technology, IET* 4(5):278-288.

Moradmand H, Setayeshi S, Targhi HK (2011) Comparing Methods for segmentation of Microcalcification Clusters in Digitized Mammograms, *International Journal of Computer Science Issues*: 6.

Nawazish N, Ayyaz H, Arfan M, Tae-Sun C (2012). Quantum and Impulse noise filtering from breast mammogram images. *Computer Methods and Programs in Biomedicine*. 108(3):1062–1069.

Niblack, W (1996). *An introduction to digital image processing*, Prentice-Hall International.

Parkin DM, Bray F, Ferlay J, Pisani P (2001) Estimating the world cancer burden: Globocan 2000. *International journal of cancer* 94(2):153-156.

Rangayyan R, Faramawy N, Desautels JEL, Alim OA (1997) Measures of acutance and shape for classification of breast tumors. *Medical Imaging, IEEE Transactions on* 16(6):799-810.

Rojas Domínguez A, Nandi AK (2009) toward breast cancer diagnosis based on automated segmentation of masses in mammograms. *Pattern Recognition* 42(6):1138-1148.

Sivanandam S, Deepa S (2006) *Introduction to neural networks using Matlab 6.0*. Tata McGraw-Hill Education.

Tang J, Rangayyan R M, Xu J, Yang Y (2009) Computer-aided detection and diagnosis of breast cancer with mammography: recent advances, *IEEE Transactions on Information Technology in Biomedicine* 13 (2) 236–251.

Tucker, A.K. (2001). *Textbook of mammography*, Churchill Livingstone.

Unser M, Aldroubi A (1996) A review of wavelets in biomedical applications. *Proceedings of the IEEE* 84(4):626-638 doi:10.1109/5.488704.

Vishnukumar K. Patel, Prof. Syed Uvaid, Prof. A. C. Suthar (2012), Mammogram of breast cancer detection based using image enhancement algorithm, *International Journal of Emerging Technology and Advanced Engineering*, 8(2): 143-147.

Wener S , Edgar Moraes D , Anselmo Cardoso de Paiva, Marcelo Gattass (2011) Detection of masses in mammogram images using CNN, geostatistic functions and SVM, *Computers in Biology and Medicine* 41(1):653–664.