INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X

(Volume2, Issue4)

Available online at: www.Ijariit.com

Classification through Artificial Neural Network and SVM of breast Masses Mammograms

Kamaldeep Kaur
Student
(Computer Science and Engineering)
Patiala institute of engineering and technology
kamalpatiala@yahoo.com

Er. Pooja
Assistant professor
(Computer Graphics)
Patiala institute of engineering and technology
p.poojagarg22@gmail.com

Abstract— Breast Cancer is one of the most common types of cancer among women. Breast cancer occurs inside the breast cells due to excessive amount increase in production of cells. Most often this can cause death if not cure at a right time. There are many techniques to detect breast cancer and various abnormalities which are described in this report. But, in this research mammography technique is used to deal with the abnormality type: breast masses. These mammograms (X-ray images) of breast masses are stored in the standard mini-MIAS/DDSM databases. To finding the region of interest there are two methods are applied on it these are: segmentation and noise removal by using neural segmentation and thresholding respectively. After the extraction of abnormal part or region of interest, feature extraction is done through using three features: GLCM, GLDM and geometrical feature on which feature selection is applied to get higher accuracy. After calculating the value of each and every feature the classification is done through using method ANN (Artificial neural network) in which 40 mammograms are used to evaluate the terms named as True Positive, True Negative, False Positive, and False Negative with the help of confusion matrix. By using these confusion matrices, the system can understand the stage of each case. Performance evaluation explains that how much effective and beneficial the new research is. Hence, ANN are used to evaluate the performance through defining Accurateness (precision), Sensitivity and Specificity and also compare the results with existing SVM classification technique.

Keywords—Artificial neural network, Feature extraction, Mammograms, Segmentation, Support vector machine.

1. INTRODUCTION

Cancer can affect each and every part of the body. Cancer begins with in the breast area is known for breast cancer. Breast cancer is the second- leading cause of cancer deaths among women in the 40-45 age groups. Lung cancer is first with 67,600 deaths expected in 2000. Biggest risk factors are being female and aging.

Approximately 182,000 new cases of breast cancer are diagnosed and 46,000 women die of breast cancer each year in the United States. Up to now, screening technique mammography is an effective way for get identify the presence of breast cancer or breast masses abnormality. Mammograms are the best breast cancer detection technique at this time. But the limitation of mammography is that the mammogram misses many cancers in dense-breasted women. To identify the presence of cancer, the detection of cancer can be done by retrieving the mammographic images available in the standard database such as DDSM or MI-AS database which helps in analysing the type of cancer whether it is benign or malignant and according to the stage of that cancer, there are many different ways to treat mass lesions. This paper deals with the retrieval of images and then applies methods to locate cancer if present. All the mammograms are publicly available, which are used in this research. The MIAS (Mammography Image Analysis society) database have images with PGM format. Portable gray map is the abbreviation for PGM format. It contains the mammographic images which are classify into benign, malignant and normal types of cancer. The second database is DDSM and it stands for database for screening mammography images. The basic idea behind this research shown with the help of following general flow chart of proposed method.

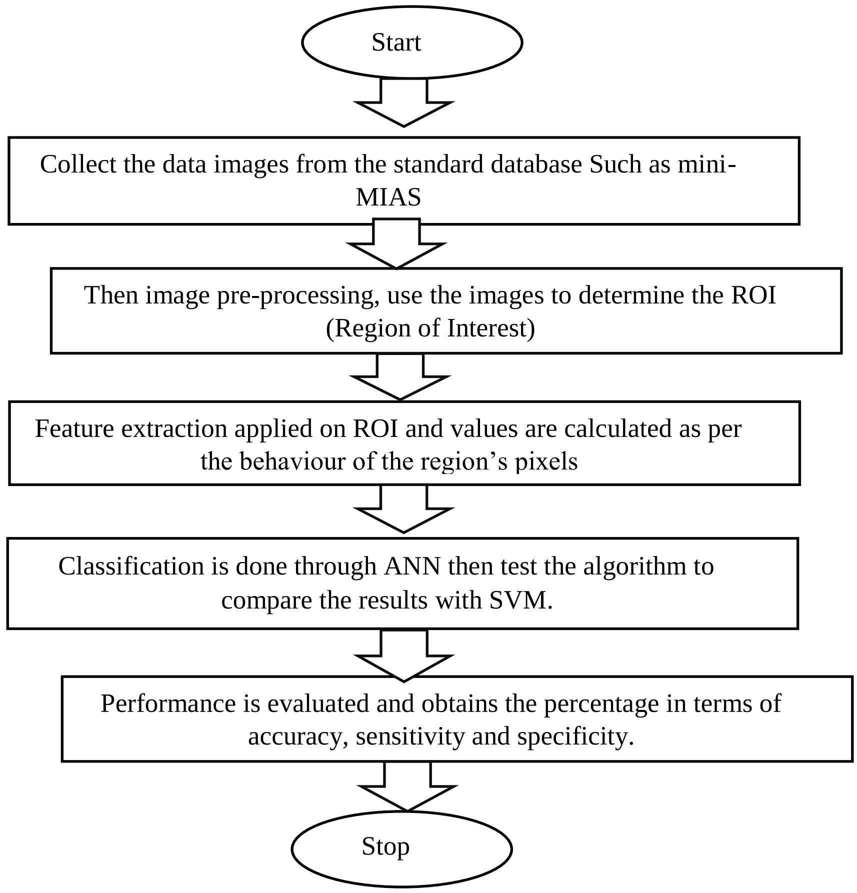

Figure 1. General flow chart of proposed method.

1.1. Breast Cancer Detection Techniques

To detect breast cancer there are several screening techniques exist such as magnetic resonance Imaging, ultrasound, mammography, Thermography, and Positron emission tomography (PET). But in this research mammograms are used basically mammogram is an X-ray image and extracted with the help of mammography technique. Description of various techniques:

Mammography: This technique retrieving images through X-rays called mammograms or mammographic images. Combination performance of mammography and MRI is 94.6%.Helpful for 68.4% of people. But, mammography misses many cancers in dense -breasted women as this technique is only used to detect the presence of cancer because it is not harmful for patients.

Ultrasound: This technique uses high-frequency sound waves to detect 87% of cancer. In more than 50 % of studies with dense breast cases While, Automated whole breast ultrasound (AWBU) is improved form as compare to mammography technique, but this technique is only applied when there is the confirmation for the presence of cancer.

Thermography: Compare temperature of cancerous and normal cells and able to diagnose at least in 10-year advance. It helps in accurate detection and widely used technique.

Magnetic Resonance Imagining (MRI): It Uses hydrogen nucleus for imagining and used for identifying Tumour but costly approach. Also 87.5% accuracy level. Sensitive than mammography and Useful in higher risk. Can be modified to detect at early stages.

Positron Emission tomography (PET): Produce contrast b/w normal & cancerous cells. It detects 55.5% of cancer by using high-resolution camera used to evaluate at high risk.

1.2. Breast Cancer Abnormalities

There are different kinds of abnormalities are examined under breasts which can lead to the death. Such abnormalities can be categorized into different types and can be cured by different methods or techniques. Such abnormalities are Breast masses, Macro-calcification and Micro-calcification, Architectural Distortion, Asymmetry Breast Tissue etc. But this research is based on abnormality type breast masses.

Masses: Masses abnormality defined as it is with a smooth, well- defined border, this is often benign (non-cancerous). It is same as a tumor or as a lump. It is an area that look abnormal contain fluid, is called a cyst. It can be either in Round shape OR Oval in shape.

Architectural distortion: Basically it is a disruption of the normal ‘random’ pattern of linear radiopaque structures seen with the help of X-rays. There is no visible mass but distortion often appears. This kind of abnormality is difficult to detect and remove.

Asymmetry Breast Tissue: Asymmetry breast tissue is an observation made with respect to the same area on the other breast. It is fairly way to find the no focal mass, no distortion, no density and no classifications. Around 3% of mammogram screening shows asymmetry breast tissue. To diagnose a small percentage of women sent for biopsy.

Calcification: In breast tissues it is a small mineral deposits. In mammography it look like a tiny white spot. There are two types of calcification first is macro-calcification and second is micro-calcification.

2. IMAGE PRE-PROCESSING

The image pre-processing stage are divided into two operational branches image segmentation and noise removal respectively. The goal of segmentation is represent an image in something that is easier to analyse. The main purpose of these branches is to eliminate background and annotation. The first step of image pre-processing is segmentation. There are various method of segmentation like edge detection, region growing etc. but this research is done through nucleus segmentation method and applying segmentation technique over retrieved images in which the boundary of the object is from the image by eliminating the calculating the eccentricity between the inside space of breast and the outside background part and whichever is closest through that eccentricity is obtained and the region is calculated. After segmentation technique is applied on mammogram next operation is noise removal. Noise removal process is completed by thresholding method. With the help of noise removal process eliminate the extra space from background and also remove the noise from the segmented images. The outputs for the image pre-processing stage are shown in following figure 2.

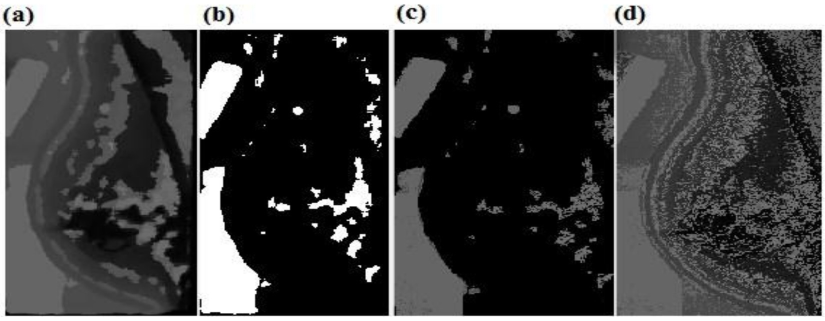

Figure 2. (a)Segmentation (b) Extraction of segmented part (c) Noise removal result (d) fully enhance mammogram.

3. FEATURE EXTRACTION

The selected region is then used for extracting features in teams of shape descriptor and texture descriptor. In feature extraction there are numbers of features but their function is divided into mass feature vectors, boundary feature vectors and background feature vectors. In this research, the total features to be calculated are 33 out of them 7 features are geometric features which represent more efficiency in calculating the nature and shape of the presence of cancer. Those features are Area, Eccentricity, Euler number, equivalent diameter, perimeter etc. there are two texture features are used in this implementation these are sub divided into 26 sub features. 22 sub features for GLCM and out of them 7 are GLDM. GLCM stands for gray level co-occurrence matrix. Following table-1 shows the equations for some GLCM and geometric features.

TABLE 1. Expressions for shape and texture features.

S.NO	FEATURES	EXPRESSION
1.	Area	Total Pixels in mass.
2.	Perimeter	Total pixels in edge border of mass.
3.	Entropy	$\sum_{m=0}^{G-1}\sum_{n=0}^{G-1} p(m,n)\log p(m,n)$
4.	Equivdiameter	$\sqrt{4*\frac{Area}{\pi}}$
5.	Compactness	$\left(\frac{2*\sqrt{Area*\pi}}{Perimeter}\right)$

6.	Euler number	Number of connected components, i.e. no. of holes.
7.	Standard derivation of edges	$\sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i-\overline{y})^2}$
8.	Energy	$\sum_{m=0}^{G-1}\sum_{n=0}^{G-1} P(m,n)^2$
9.	Contrast	$\frac{1}{(G-1)^2}\sum_{m=0}^{G-1}\sum_{n=0}^{G-1}(m-n)^2\, p(m,n)$
10.	Correlation	$\frac{\sum_{m=0}^{G-1}\sum_{n=0}^{G-1} mnp(m,n) - \mu_x\mu_y}{\sigma x \sigma y}$
11.	Homogeneity	$\sum_{m=0}^{G-1}\sum_{n=0}^{G-1}\frac{p(m,n)}{(1+\|m-n\|}$

Gray level difference matrix is abbreviation for GLDM. GLDM feature are used to define the probability density functions and therefore, it includes 4 features and they are pdf-1, pdf-2, pdf-3, pdfj-4 are used to define the changes occurs in the image density. This feature can be seen through plotting its density graph.

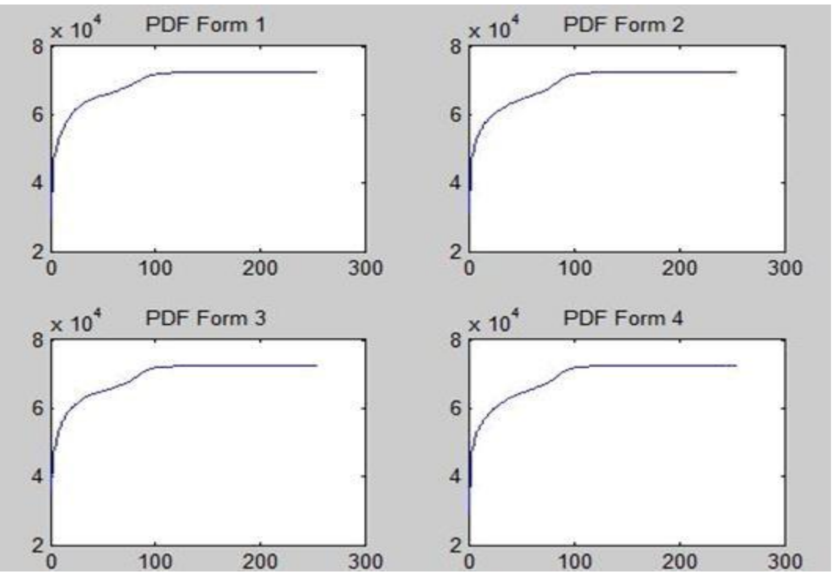

Figure 3. GLDM feature density graph.

4. ARTIFICIAL NURAL NETWORK AND SUPPORT VECTOR MACHINE

After calculating or extracting all the feature values in team of both shape as well as texture features system divide the classification into two parts three classes (benign, malignant and normal) and second one with two classes (benign and malignant). The output of these classes is shown in teams of True positive, True negative, false positive and false negative with the confusion matrix. These outputs are four outcomes, performed to detect the diseases type or stage. For outcomes explained with the help of figure 4.

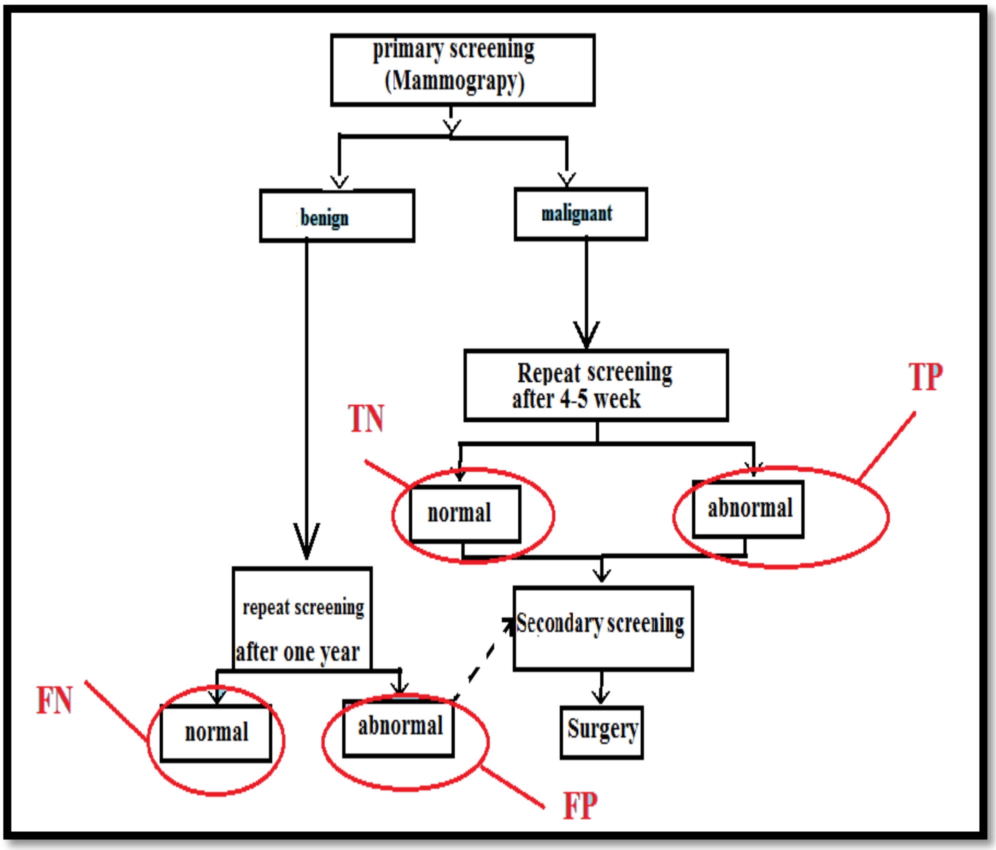

Figure 4. Expected outcomes in terms of TP, TN, FP and FN.

After the detection of disease type in team of outcomes, ANN and SVM classifiers are used to evaluate the performance for both classes i.e. two classes and another with three classes in terms of Accuracy, sensitivity and specificity.

Accuracy: This means as many times the different samples or images are tested with the same algorithm and the machine or system provides results how much Accurate or correct. This is called the accuracy of the system.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

Sensitivity: Sensitivity means that how accurately a test identifies people or cases who do have cancer or disease.

$$\text{Sensitivity} = \frac{TP}{TP+FN}$$

Specificity: Specificity means that how accurately a test identifies people or cases who do not have cancer or disease.

$$\text{Specificity} = \frac{TN}{TN+FP}$$

5. EXPERIMENTAL RESULTS

PERFORMANCE EVALUATION

5.1. ANN Performance Table

TABLE 2. ANN Performance

	ACCURACY	SENSITIVITY	SPECIFICITY
ANN with two classes	83.35%	75%	75%
ANN with three classes	63.4%	50%	63.2%

5.2. SVM Performance Table

TABLE 3. SVM Performance

	ACCURACY	SENSITIVITY	SPECIFICITY
SVM with two classes	55 %	52%	58%
SVM with three classes	73%	72%	60%

5.3. Comparison Between ANN AND SVM

TABLE 4. Comparison

	ACCURACY	SENSITIVITY	SPECIFICITY
ANN with two classes	83.35%	75%	75%
SVM with two classes	55 %	50%	58%
ANN with three classes	63.4%	50%	63.2%
SVM with three classes	73%	72%	60%

5.4. Graphical Representation of Comparison between BOTH ANN and SVM.

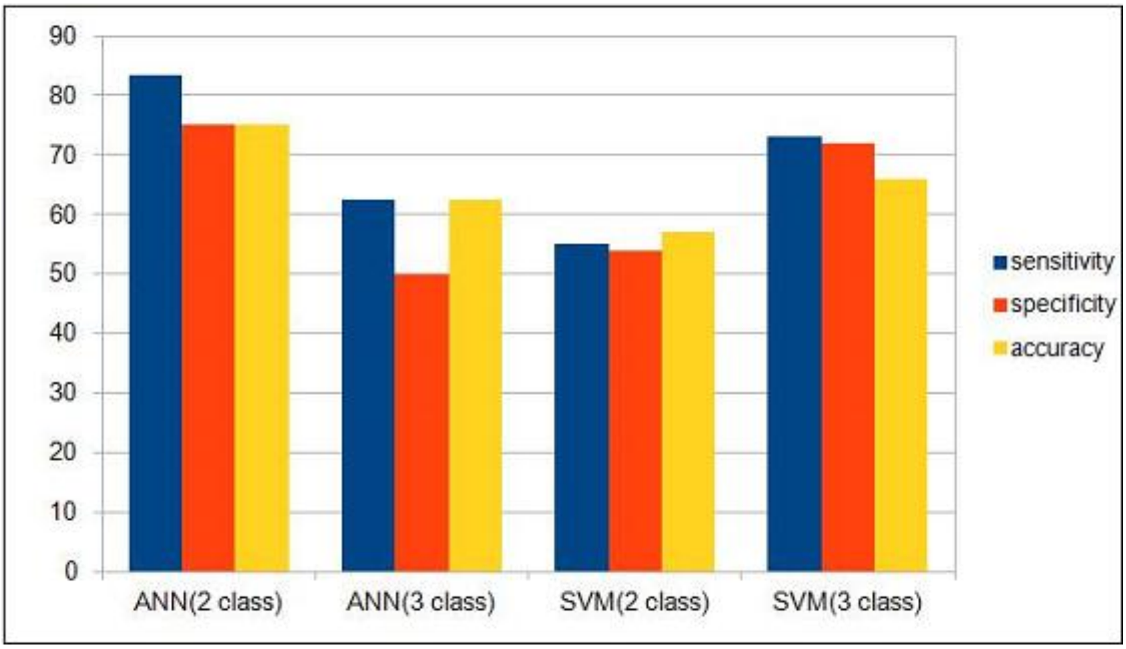

Figure 5. Performance graph.

6. CONCLUSION

This paper presents the system for mammographic masses classification based on artificial neural network. From the above results, the problem on which the research has been carried out has been solved using different measures like: the selection of mammographic images, use of more than one feature to analyse the presence of cancer and at the end calculating the percentage of getting the correct results from a large set of database. In earlier cases, the detection of cancer is processed through calculating the one of the different features descriptor, but there is a small approach is made i.e. instead of using single feature descriptors, the use of two different features in this case they are: shape and texture. Here 7 shape descriptors are used and 22 GLCM and 4 GLDM texture feature descriptors are used. Hence, by combining the features provides the better chances of analyzing the cancer type and character. Therefore, the conclusion of this research is that using more features builds the chances of spotting the cancer more effectively.

REFERENCES

[1] B.Surendiran and A.Vadivel, Classifying mammographic masses into BI-RADS shape categories using various Geometric Shape and margin features, International journal of biomedical signal processing,2(1), 2011, PP, 43-47.

[2] Htet Thazin Tike Thein and Khin Momo tun, An approach for breast cancer Diagnosis classification using neural network, Advanced computing: An international journal (ACIJ), Vol.6, No.1, January 2015.

[3] Sandeep chaurasia et al, An application of classification techniques on breast cancer prognosis, International journal of computer applications, volume 59-No 3, December 2012.

[4] W.Xie, et al, breast mass classification in digital mammography based on extreme learning machine, nerocomputering,http://dx.doi.org/10.10.1016/k/neucom.2015.08.048

[5] Sailesh GC, chulwoo pack, Sung Shin and Hyung D. choi, oct 9-12, 2015, Breast cancer classification of mammographic Masses using Improved shape features.

[6] Texture Features based breast cancer detection: A survey by Aunps, maria sabu, D.narain ponraj, Dr.poongodi, Vol.3, No.9.sep. 2012, journal of emerging Trends in computing and Information sciences.

[7] Aswini kumar Mohanty, Swapnasikta Beberta, Saroj Kumar Lenka/Internaitional Journal of Engineering Research and Applications (IJERA), Vol. 1, Issue 3, pp.687-693.

[8] R.NITHYA, B.SANTHI, Comparative study on feature extraction method for breast cancer classification, Jourmal of Theoretical and Applied Information Technology, 30 nov 2011. Vol. 33, No. 2.

[9] International journal of computer & Organization Trends- Volume 3 Issue8- sep 2013, breast cancer detection in mammogram Images using Region-Growing and contour-Based Segmentation Techniques by D.Sujitha Priya, Dr. B.Sarojini.

[10] International Journal of Advanced research in computer science and software engineering, classification of breast masses in digital mammograms using support vector Machines.,by Nasseer M. Basheer, Mustafa H. Mohammed , vol3, Issue 10, October 2013.

[11] Vinitha Sree S et al. Breast imaging: A survey, April 10, 2011, volume 2, Issue 4.