Breast Cancer: Classification of Breast Masses Mammograms Using Artificial Neural Network and Support Vector Machine

Kamaldeep Kaur
Patiala Institute of Engineering and Technology for Women
Patiala Punjab India

Er. Pooja
Patiala Institute of Engineering and Technology for women
Patiala Punjab India

Abstract— this paper presents the diagnosis of breast cancer by using ANN and SVM.To deal with the different kinds of abnormalities causing Cancer, this report consists of all the modalities which help in detecting cancer and well as different methods of feature extraction. Such modalities can be named as: Mammography, Ultrasound, MRI etc [1]. Currently, Electrical impedance and nuclear medicine are used widely for diagnosis. These modalities Based on the image processing i.e. identification of abnormality is done through the reading and retrieving information from images. But this research is based on mammogram images. Before retrieving information one should know about all kinds of abnormalities like: micro classification, masses, architectural distortion, asymmetry, and breast density etc. [2]. And after the process of extracting the abnormal part or can say that ROI (Region of Interest) on which the treatment is applied. To extract ROI various methods are used like region growing, edge detection, segmentation etc. [3][4]. Then, feature extraction is done from which a lot of features are extracted on which feature selection is applied to get higher accuracy. After going through all researches done till now here I have got the conclusion that for determining the presence of cancer researcher uses different features but till now only few researcher used two features named shape and texture which needs good classification technique[1]. Then, classify into classes of normal and abnormal classes. From the statistical study it has been found that the trend in increasing cancer every year, thus, the best and most effective way to cure cancer is the removal of cancerous part.

Keywords— Support Vector Machine, Mamogram, Breast Cancer, Machine Learning.

Introduction

Breast Cancer in women's is 2nd most common type of diseases causing deaths in women because of the production of out of proportion cancerous cells which replaces neighboring tissues or cover all over the body. Most often this can lead to death if not cure at right time [4]. According to the National Cancer Foundation estimation there are 200000 new breast cancer cases are noticed and from these 40000 leads to the death. It is found that, if the cancer is detected at earlier stages; there occur better chances to get the proper treatment. Therefore, women at the age of 40 are asked to attain the regular mammographic screening, as mammography is most reliable method for cancer detection at early stages. Mammogram is the most common method used for breast imaging. It helps in examine the presence of cancer at early stages and help in reducing the mortality rate by 25-30% in screened women. There occur many different types of breast cancer such as: mass, micro classification, architectural distortion, asymmetry breast tissue [2]. But this paper carries the masses problem and deals with its shape and texture feature for classification [1].

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 and MI-AS database which helps in analyzing the type of cancer and according to the stage of that cancer, there are many different ways to treat mass lesions[4][8]. This paper deals with the retrieval of images and then applies methods to locate cancer if available and also provide the idea of how to classify that cancer if it is present.

1.1 Table 1
Types of cancer detection modalities.

S. No.	Technique	Definition
1.	**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.
2.	**Mammography**	This technique retrieving images through X-rays called mammograms or mammographic images. Combination performance of mammography and MR 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
3.	**Positron Emission tomography (PET)**	Produce contrast b/w normal & cancerous cells. It detects 55.5% of cancerby using high-resolution camera used to evaluate at high risk.
4.	**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.
5.	**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.

TECHNIQUE USED
There are various technique are used for detection of the cancerous part like Mammography, Ultrasound, MRI, PET (positron Emission Tomography), optical Imaging etc. In this research, we are going to use Mammographic images. Mammograms are breast x-rays. In which breast being squashed between two plates so that the radiography can clearly see the tissue. To retrieving the mammographic images available in the standard database such as DDSM or MI-AS database which helps in analyzing the type of cancer and according to stage of that cancer, there are many different ways to treat mass lesions [5].

Sumeet Dua et al present a unique technique for the categorization of mammograms with the use of a novel weighted involvement rule on the basis of classifier. Images are preprocessed to disclose regions of concern. Texture elements are elicitate from divided parts of the image and discretized for rule detection. Involvement rules are consequently among several texture elements elicitate from parts of images, and employ for categorization on the basis of their intra- and inter-class dependencies. These rules are subsequently utilized for the categorization of a generally used mammography dataset, and precise experimentation is achieved to estimate the rules' effectiveness under distinct categorization scenarios. The experimental outcome proves that this technique works well for type of datasets, acquire efficiency as high as 89%, which outperforms the accuracy rates of other rule on the basis of categorization methods [9].

1.2 Types of Breast Abnormalities

Around all over world, 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:-
Mass (lump) :- with a smooth, well-defined border -- this is often benign (non-cancerous). <u>Ultrasound</u> is needed to characterize the inside of a mass (for example, if the mass contains fluid, it is called a cyst). It can be either in Round shape OR Oval in shape

Round   **Oval**

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.

But, this report highlights the treatment methods and the ways to extract features and classify them for the detection or present of Masses in breasts.

 The main aim of this paper is:

To retrieve the image from database and then pre-process by segmenting it and then extracting mass features which describes the availability of mass lesions in breast cells.

The extraction of features are collected in more than one form such as two different features are classified named as shape and texture features and based on result the identified region is being classified as cancerous cell or normal unaffected one[7].

After extracting the region of interest, different features extraction methods are applied to get better accuracy [6].

And the result is shown on the behalf of comparison between various features in the form of accuracy, specificity, and sensitivity.

2. Feature extraction.

Feature extraction refers to represents the images in its compact and reduced form in order to facilitate and speed up the decision-making process example classifications and clustering. In this research, we will discussing various features.

Various features.
The selected region is then used for extracting various features like some shape features, texture features, intensity histograms, GLCM etc. there are different types of shape features like perimeter, area, compactness, eccentricity, circularity, Zernike moments, convexity, density etc. [6]. Texture features are divided into statistics: First order statistics and second order statistic [7]. First include extraction of features on the basis of image histogram, shows the probability and this includes the various sub-features like mean, variance, Kurtosis, Skewness, Entropy. Second order statistics deal with calculating the probability of a relationship between two pixels. The first stage of 2nd order is forming the matrix co-occurrence and second is specifying the characteristics as a function of the matrix. This also include various sub features like correlation, Angular second Moment, Inverse difference moment, Entropy, Contrast. GLCM stands for Grey level Co-occurrence Matrix. GLCM is a 2-D array which takes into account the specific position of a pixel relative to other pixels.

Methods

After the Calculation of all feature values of various features, there classification is done through using methods SVM and ANN in which the number of samples are used to train the algorithm and evaluate the result named True Positive (TP), False positive (FP), True Negative (TN), False Negative (FN).Support vector machine is learning algorithm used to analyse data. We use SVM in our research because of its better performance generation even in the case of multi-dimensional data. By using SVM and ANN, the proposed idea will be implemented in MATLAB which is broadly used in all areas of research works.

Four Expected Outcomes
1. The outcome which is expected from the use of more than one feature to detect and classify cancer inside breast cells is that how much accurate the algorithm works and gain the increase in accuracy from previous researchers. Four expected outcomes are: True Positive (TP), False positive (FP), True Negative (TN), and False Negative (FN) [11].

   2. The expected performance can be calculated through:

   1. Sensitivity: The Sensitivity of a system means how much accurate is the identification of cancerous part.

   2. Specificity: The Specificity is the ability to represent how much able the system is to identify the image or data which are not affected from disease.

3. Accuracy: Accuracy deals with the precision which means as many times the new different images are tested with the algorithm the machine provides the accurate or correct result.

Four Possible Outcomes.

	Patient devlops cancer within 1 yr	Patient does not devlops cancer within 1 yr
Abnormal mammogram	True Positive	False Positive
Normal mammogram	False Negative	True Negative

Conclusion

As we know that one of the most common and dangerous types of cancer is breast cancer. Every year more than 12,000 women are diagnosed with breast cancer in a single country and more than 2,600 women lose their life every year. To identify different abnormal features of Breast cancer, this paper discuss various Modalities or techniques which helps to detect cancer. To identify the presence of cancer, the detection of cancer can be done through retrieving mammographic images available in the standard database such as DDSM or MI-AS database which helps in analyzing the type of cancer and according to the stage of that cancer. Basically, this research is based on mammograms of breast masses. With the help of mammograms, various feature extraction methods are used to extract ROI. After the Calculation of all features values of various features their classification is done through using methods SVM and ANN in which the number of mammograms are used to evaluate the result. This Performance evaluation explains that how much effective the new research by evaluating performance in terms of accuracy, sensitivity, and specificity.

Acknowledgment

References:

[1] S. V. S. R. U. A. O. F. Eddie Yin-Kwee Ng, "Breast Imaging:A Survey," World Jounral of Clinical Oncology, vol. 2, no. 4, p. 8, 2011.

[2] G. R.Bhanumathi, "Latest Advances in Computer-Aided Detection of Breast Cancer by mammography," International journal in IT and Engineering, vol. 1, no. 6, p. 13, 2013.

[3] R. ShilpaKamdi, "Image Segmentation and Region Growing Algorithm," IJCTEE (Internatioanl Journal of Computer Technology and Electronices Engineering), vol. 2, no. 1, p. 5.

[4] R. M. Hilary Alto, "Content-Based Retrieval and Analysis of Mammographic Masses," Jounral of Electronic Enginering, vol. 14, no. 2, p. 17, 2005.

[5] M. C. I. F. G. Qaisar, "Breast Mass Segmentation Using Region-Based and Edge-Based Methods in a 4-Srage Multiscale System," ELSEVIER, p. 11, 2013.

[6] S. Y. Chia-Hung, "Mammogram Retrieval on Similar Mass Lesions," ELSEVIER, p. 15, 2012.

[7] A. H. Sri Hartati, "Selection Mammogram Texture Descriptors Based on Statistics Properties Backpropogation Structure," internation journal of computer science and information seurity, vol. 11, no. 5, p. 5, 2013.

[8]   L. j. G. d. P. Martins, "Detection of Masses in Digital Mammograms Using K-means and Support Vector Machine," vol. 8, no. 2, p. 12, 2011.

[9] Dua S, Singh H, Thompson HW (2009) Associative classification of mammograms using weighted rules. Expert Syst Appl 36(5):9250–9259. doi:10.1016/j.eswa.2008.12.050

[10] G. R.Bhanumathi, "Latest Advances in Computer-Aided Detection of Breast Cancer by mammography," International journal in IT and Engineering, vol. 1, no. 6, p. 13, 2013.

[11] M. C. I. F. G. Qaisar, "Breast Mass Segmentation Using Region-Based and Edge-Based Methods in a 4-Srage Multiscale System," ELSEVIER, p. 11, 2013.

Books:-

Image Databases Search and Retrieval Imagery by Wiley.

Fundamentals of Digital Image Processing by Chris Solomon and Toby Breckon.