International Journal of Engineering Sciences &Research Technology

(A Peer Reviewed Online Journal)
Impact Factor: 5.164

IJESRT

RESEARCHERID
THOMSON REUTERS

ISSN: 2277-9655
[Aict 2019] Impact Factor: 5.164
IC™ Value: 3.00 CODEN: IJESS7

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

REVIEW ON MAMMOGRAPHIC CLASSIFICATION TECHNIQUES

Revathy M[*1], Sabeena M[*2]
PG Scholar, Assistant Professor
[*] Dept. of Electronics and Communication, College of Engineering Karunagappally, Kollam, Kerala, India

DOI: 10.5281/zenodo.2629268

ABSTRACT

Breast cancer is cancer that forms due to the uncontrolled growth of cells in the breast region. Breast cancer cause in both men and women, but mostly affected by women. It is the second leading cause of women death in today.The early prediction of breast cancer is the key to reduce women mortality. The major symptoms of breast cancer is the occurrence of microcalcification clusters or small amount of calcium accumulated in the breast region. For the efficient detection of breast cancer manymedical techniques are used to diagnosis the breast cancer like biopsy , clinical examination, ultrasonography, MRI, and mammography etc. Mammography is the process of using low energy x-rays to analyze the human breast for diagnosis and screening of tumor in the breast region. In this paper discussing about the various classification techniques are used in mammographic images.

KEYWORDS: Micro calcification, MRI , Ultrasonography, Mammography

1. INTRODUCTION

Cancer is the leading cause of death in world wide. It is accounted to 17 million deaths in 2017.The detection of a cancer at its early stage is necessary for reducing mortality rate. Breast cancer starts when cells in the breast region begin to grow out of control. These cells usually form a tumor that can often be diagnosisan x-ray or other cancer detection techniques. Breast cancer can occur in different areas of the breast . The main parts of the breast region are lobules, ducts, and connective tissue. The ducts are the tubes that carry milk to the nipple. The connective tissue which is made up of fibrous and fatty tissues which surrounds and holds everything together. Most breast cancers begin in ducts or lobules. It can spread outside the breast through blood vessels and lymph vessels. When breast cancer spreads to the other parts of the body that is called metastasized .Breast cancer occurs almost entirely in women, but men can get breast cancer,too.

It is the most frequent cancer in women, in every year 2.1 million women impacting breast cancer which also causes the largest number of cancer death among women. In 2018 ,it is estimated that 627,000 women died due to breast cancer. While breast cancer rates are higher among women in more developed regions compared to other cancers. In order to overcome breast cancer outcomes and survival ,early detection of cancer is very critical. Early diagnosis and screening are the two early detection strategies for breast cancer. Limited resource settings with weak health systems are the reasons ofthe majority of women are diagnosed in late stages. It should prioritize early diagnosis programmes based on awareness of early signs and symptoms and prompt referral to diagnosis and treatment.

The major signs of breast cancers are

- The lymph nodes swollen under the arm or around the collarbone.
- Swelling of all or part of the breast region.
- Skin irritation or dimpling
- Breast or nipple pain
- Nipple retraction
- Redness, roughness or thickening of the nipple or breast skin.
- Nipple discharge

A woman's breasts contain specialized glands that can produce milk. The breast structure consists of 15-20 lobes. Each lobe is made up of many smaller lobules which have groups of tiny glands that can produce milk. The milk travels through a network of tiny tubes (ducts) to a reservoir that lies just below the nipple. The dark round area of skin surrounding the nipple is called the areola. The breast also contains blood & lymph vessels and lymph nodes.

The Lymph System of the Breast

One of the main reason of breast cancer spreads through the lymphatic system.Our bodies have a network of lymph vessels and lymph nodes Lymph vessels carry a clear fluid called lymph which drains into lymph nodes. Lymph nodes are small bean-shaped structures which contain cells that compact infections (immune system cells).

Lymph vessels from the breast drain into the axillary lymph nodes and supraclavicular lymph nodes of the breast.

Lymph nodes are important in cancer care because any cancer cells that have broken away from a cancerous tumor that can be carried by the tissue fluid to the nearest lymph nodes. In some cases lymph nodes becomes enlarged for other reasons which are not related to the cancer. Breast tissue extends under the armpit region that is made up of many lymph nodes, also known as lymph glands.

a. BREAST CANCER DETECTION TECHNOLOGIES

Several medical techniques have been developed for breast cancer detection. These systems include some that use radioactive compounds that concentrate in cancerous tissue to image breast cancer. Some of which are used in combination until a diagnosis is proven.

- **Clinical examination**

A clinical breast examination involves through a physical examination of the whole breast area, including both breasts, nipples, armpits and collarbone. The doctor will also ask about the women's personal and family history of breast cancer and whether she has any symptom of breast cancer.

- **Magnetic resonance imaging(MRI)**

An MRI produces an image of the body using magnetic fields. Women under 50 years of age who are at high risk of breast cancer are eligible routine screening with MRI undermedicare. To access the service younger women must be referred by there specialist.

- **Ultrasound**

An ultrasound uses sound waves to outline a part of the body. A breast ultrasound is used to see whether a breast lump filled with fluid. An ultrasound doesnot replace the need for a mammogram, but it is often used to check abnormal result from a mammogram.

- **Biopsy**

A Biopsy is the removal of small sample of tissue from the breast lump or lymph nodes. With examines under the microscope. Analysis by a pathologist will help to diagonise both the presence of breast cancer and its type. This will confirm diagnosis and help to determine the appropriate treatment plan.

- **Mammography**

Mammography is also called mastography. Mammography is the process of using low energy x-rays to examine the human breast for diagnosis and screening of cancer. An x-ray picture of the breast is called mammogram. It is used to check cancer in women who have no signs or symptoms of the cancer. It also be used if you have a lump or other sign of breast cancer.Screening mammography is the type of mammogram that checks when you have no symptoms of disease. It can also help to reduce the number of deaths from breast cancer among women ages 40 to 70. But it can also have some drawbacks. Mammograms can sometimes find something that looks any abnormality in the breast but isn't cancer. This leads to further testing and can cause you anxiety for patient. Sometimes mammograms can miss cancer when it is present in the breast. It also exposes radiation usually around 30kvp.Mammograms are also recommended for younger women who have symptoms of breast cancer or who have a high risk of the disease.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

2. METHODS

In this paper discussing about the various classification techniques are used for large data set of mammograms.

- **SVM classifier**

Support Vector Machine (SVM) isa classifier that formally defined by a separating hyper plane. In other wise it is defined for a given labeled training data (*supervised learning*), the algorithm outputs an optimal hyper plane. In two dimensional space of this hyper plane is a line dividing a plane in two parts where in each class lay in either side.

The SVM scheme is based on algorithm. which is used for solving classification tasks and have been successfully applied in various areas of research. The basic idea of SVM is that it projects data points from a given two-class training set in a higher dimensional space and finds an optimal hyper plane. The optimal one is the one that separates the data with the maximal margin. SVMs classifier identify the data points near the optimal separating hyper plane which are called support vectors. The distance between the separating hyper plane and the nearest of the positive and negative data points is called the margin of the SVM classifier.

The separating hyper plane is defined as

$D(x) = (w.x) + b(4)$

Where, x is a vector of the dataset mapped to a high dimensional space, and w and b are parameters of the hyper plane that the SVM will estimate.

- **CAD scheme**

The CAD scheme is described briefly by following steps:

a) Likelihood detection- In segmentation first step is of our CAD scheme is segmentation of an image into breast tissue and background, using a skin line detection algorithm. Additionally, it in the edge of the pectorals muscle. If the image is a MLO view. Then, a thickness equalization algorithm is applied to improve the periphery of the breast. Background equalization is done using a similar algorithm in the pectorals muscle, in order to avoid the problem with detection of masses located on or near the pectoral boundary. A neural network classifies each pixel using these features and assigns a level of suspiciousness to it. The neural network is trained using pixels sampled inside and outside of a representative series of malignant masses.

Finally, the result of this method is an image whose pixel values represent the likelihood at a malignant mass or architectural distortion is present.

b) Initial detection step resulting in a likelihood image and a number of suspect image locations.
c) Region segmentation, any dynamic programming, using the suspicious locations as seed points.
d) Finally, classifying the regions as true abnormalities and false positives

- **Naive Bayesian Classifier**

The naive Bayesian classifier is a Bayesian network with a limited topology applicable to learning tasks where each instance is described by a conjunction of feature values and a class value. This method is also used to estimates the unknown data item using probabilistic statistics model. Challenge in the Bayesian classification is to determine the class of data sample which have some number of attributes. Let Z be the dataset with 'u' objects such that X1, X2,…,Xu. Each object have n attributes i.e. A1, A2…… An. Let there are m classes C1,C2…Cm. Naive Bayesian classifier predicted the unknown data sample X which is without the label to the class Ci if and only

$$\mathbf{P(ci \mid x) > P(cj \mid x)\ for i<=j}$$

Where <=m, j<>i

Then X is assigned to Ci. This is called Bayes decision rule. In this method of classification dataset are divided into two sets, training and testing respectively. Training dataset is considered as prior information and model is constructed on the basis o this training dataset.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

- **K-Nearest Neighbor algorithm** (KNN)

KNN is supervised learning algorithm used for statistical estimation and pattern recognition. It classifies objects into separate classes based on closest training examples in feature vector space. It is used to estimate the class of an unknown instance from its neighboring returns the most common value. All training samples are stored in n dimensional pattern space. The class of an unknown sample is decided by a majority vote of its neighboring samples in the training pattern space. There are many distance measures that can be used in KNN classification such as Euclidian distance, Manhattan distance, Minkowski distance, Mahalanobis distance etc. However Euclidean distance is the most commonly used distance instances. It stores all possible cases and classifies new cases based on distance metric .

The KNN classifier can perform efficiently, if the optimal value of 'k' is used in the classifier design. In the present work, after experimentation, the optimal value of 'k' is decided as 5.
Let k denotes desired number of nearest neighbors and F: Pattern Space: set of training samples { F1 ,F2, F3,….FN,}: The set of feature vector of training mammograms in the form of Fi = {Xi, Cj} Where Xi: feature vector of the point Fi and Ci is the class that belongs to. F' = {Xi',Ci'} .

Test mammogram feature vectorFor a test mammogram, following steps are followed to make prediction:
(i) Compute the Euclidian distance d (X', Xi) between test mammogram and all training mammograms.
(ii) Sort all points Fi according to distance
(iii) Select the first K points from the sorted list. These are k closest samples to F.
(iv) Choose the closest label.
(v) Return Majority label and assign class to test mammogram F' based on majority vote

3. LITERATURE REVIEW

Al-Najdaw et al. have investigated combining several images of mammogram applying enhancement algorithms to improve the performance of breast-region segmentation. The masses that show in mammogram images are further analyze and classified into 4 categories It consists of benign, probable benign and possible malignant, probable malignant and possible benign, and malignant. The main concept of this work is to reveal the optimal combination of different enhancement techniques and to segment breast region. This will helps to get better visual interpretation, analysis, and classification of mammogram masses to help radiologists in making correct and better decisions. The dataset consists of more than 1300 mammogram images from both the King Hussein Cancer Center and Jordan Hospital. Results achieved cancer classification accuracy values of 90.7%. Moreover, the results shows a sensitivity of 96.2% and a specificity of 94.4% for mass classifying algorithm .

Liu et al. Designs thepossibilistic fuzzy c-means (PFCM) clustering algorithm and weighted support vector machine (WSVM) for the detection of micro calcification clusters formed in full-field digital mammograms (FFDM). For each case of image, the suspicious micro calcification regions are extracted with region growing and active contour segmentation technique. Then geometry and texture features are extracted for each suspicious micro calcification, a mutual information-based supervised criterion is used to select important features, and PFCM is applied to the cluster samples into two clusters. Weights of the samples are calculated based on the possibilities and typicality values from the PFCM, and the ground truth labels. A weighted nonlinear support vector machine (SVM) is trained. During the test process, when an unknown image is presented, segmentation step is used for locating suspicious regions, selected features are extracted, and the suspicious micro calcification regions are classified as containing micro calcification. Finally, the micro calcification regions are analyzed with spatial information to locate micro calcification clusters. The proposed method is evaluated using a MIAS dataset and compared with the standard un weighted support vector machine (SVM) classifier. The detection performance is evaluated by the response receiver operating (ROC) curves and the free-response receiver operating characteristic (FROC) curves. The proposed method obtained an area under ROC curve of 0.8676, while the standard support vector machine (SVM) obtained an area of 0.8268 for micro calcification detection. In the case of cluster detection, the proposed method obtained a high sensitivity of 92% by a false micro calcification -positive rate of 2.3 clusters image, and it is better than standard support vector machine (SVM)with 4.7 false-positive clusters image at the same sensitivity

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

.Aswini Kumar Mohanty et al. proposes a method for classification of breast masses using correlated rule mining. The images are preprocessed by resizing and finding out the regions of interest. The Statistical and texture features are formed using the gray level co occurrence matrix method (GLCM).The classification is done through the Correlated association rule mining which takes association framed using the features extracted. The classification is done using the CorClass algorithm. The classification is done based on the decision list and the weighted combination of the association rules based on which the best rule is selected for classification.

Sharma et al. proposed an experiment using CAD system to classify malignant and nonmalignant mammogram patches. Suitable methods were used for preprocessing, which is helped to removes the artifacts, unwanted components, and extracts the breast region from its background along with extraction of abnormal regions from mammograms. This paper proposes a method to avoid the processing of full mammogram, methods to extract the fix-sized ROI patches . The Zernike moments are based on texture feature extraction technique and it is used to recognize the pattern of malignancy or non malignancy in the mammogram patches.

The variations in the outcome are observed by the experimenting with the low-order and high-order Zernike moments. Experiments are performed with the other well known texture descriptors gray level co-occurrence matrix (GLCM) and discrete cosine transform and (DCT), and it is observed that the proposed CAD system works well with Zernike moments. It has been observed that SVM with RBF kernel attains the highest sensitivity and specificity values at lower orders of Zernike moments. CAD system has improved the accuracy of diagnosis and promises to perform a second reader role .

Antony et al. proposed a new approach to determine the classification of mammographic image byusing k-means clustering algorithm. It can be used to extract different features of the image like shape, intensity values and density features and region features to compute the feature vector. They computed the mean values of intensity values of the pixels in the region extracted to compute the intensity mean value. The density measure is also computed in the similar fashion. The region metric is computed with the extracted region values and it has seven different features hidden. k-means clustering is used based on the computed feature vectors to identify the class of the input image. The proposed system reduces the space and time complexity and produces good results. It has produced classification accuracy up to 99% which is more than other methodologies in this era .

Digambar A Kulkarni et al proposed a K means clustering algorithm and used SVM for classifier construction. K means is simple algorithm and give good accuracy of classifier . The disadvantage is that it is difficult to analyze K value and there dissimilarity in final clusters due to early different partitions.

4. CONCLUSION

In this paper we discussed about various classification techniques of breast cancer.An unsupervised Kohonen neural network which can extract features from extracted areas then feed those features to another feed forward neural networks. Support Vector Machine is a machine learning technique [12] used to analyze data, recognize patterns and classify data into different classes by constructing a hyper plane in an infinite dimensional space.KNN is supervised learning algorithm used for statistical estimation and pattern recognition. It classifies objects into separate classes based on closest training examples in feature vector space. It is used to estimate the class of an unknown instance from its neighboring instances.The limitation of KNN classifier is false classification of test image when majority of the nearest neighbors have closely matched features.In this review we have conclude that classification of benign and malignant cancer is done using SVM classifier having more accuracy than KNN classifier because SVM works better than KNN classifier. Using KNN classifier it produces classification accuracy 99% which is best and reduces time and space complexity

REFERENCES

[1] E. Burnside, D. Rubin, and R. Schachter, “A Bayesian network for mammography,” in Proceedings of AMIA Annual Symposium, pp. 106–110, 2000.
[2] X. Wang, B. Zheng, W. Good, J. King, and Y. Chang, “Computer assisted diagnosis of breast cancer using a data-driven Bayesian belief network,” International Journal of Medical Informatics 54, pp. 115–126, May 1999.
[3] C. Cortes and V. Vapnik, “Support vector networks,” Machine Learning 20(3), pp. 273–297, 1995.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

[4] S. Timp, Analysis of Temporal Mammogram Pairs to Detect and Characterise Mass Lesions. PhD in medical sciences, Radboud University Nijmegen, 2006. ISBN 9090205500.
[5] T. Nattkemper, B. Arnrich, O. Lichte, W. Timm, A. Degenhard, L. Pointon, C. Hayes, and M. Leach, "Evaluation of radiological features for breast tumour classification in clinical screening with machine learning methods," Artificial Intelligence Medical 34, pp. 129–139, June 2004.
[6] M. Mavroforakis, H. Georgiou, N. Dimitropoulos, D. Cavouras, and S. Theodoridis, "Significance analysis of qualitative mammographic features, using linear classifiers, neural networks and support vector machines," European Journal of Radiology 54, pp. 80–89, April 2005.
[7] S. Li, T. Fevens, and A. Krzyzak, "Automatic clinical image segmentation using pathological modeling, PCA and SVM," Engineering Applications of Artificial Intelligence 19, pp. 403–410, June 2006.
[8] G. te Brake, N. Karssemeijer, and J. Hendriks, "An automatic method to discriminate malignant masses from normal tissue in digital mammograms1," Physics in Medicine and Biology 45(10), pp. 2843–2857, 2000.
[9] N. Karssemeijer, "Automated classification of parenchymal patterns in mammograms," Physics in Medicine and Biology 43(2), pp. 365–378, 1998.
[10] P. Snoeren and N. Karssemeijer, "Thickness correction of mammographic images by anisotropic filtering and interpolation of dense tissue," in Medical Imaging 2005: Image Processing. Edited by Fitzpatrick, J. Michael; Reinhardt, Joseph M. Proceedings of the SPIE, Volume 5747, pp. 1521-1527 (2005)., J. M. Fitzpatrick and J. M. Reinhardt, eds., pp. 1521–1527, April. 11
[11] J. S. Drukteinis, B. P. Mooney, C. I. Flowers, and R. A. Gatenby, "Beyond mammography: new frontiers in breast cancer screening.," Am. J. Med., vol. 126, no. 6, pp. 472–9, 2013.
[12] X. Liu, M. Mei, J. Liu, and W. Hu, "Microcalcification detection in full-field digital mammograms with PFCM clustering and weighted SVM-based method," EURASIP J. Adv. Signal Process., vol. 73, no. 1, pp. 1–13, 2015.
[13] N. Al-Najdawi, M. Biltawi, and S. Tedmori, "Mammogram image visual enhancement, mass segmentation and classification," Appl. Soft Comput., vol. 35, pp. 175–185, 2015.
[14] S. J. S. Antony and S. Ravi, "A New Approach to Determine the Classification of Mammographic Image Using K-Means Clustering Algorithm," Int. J. Adv. Res. Technol., vol. 4, no. 2, pp. 40–44, 2015.
[15] A. H. Abbas, A. A. Kareem, and M. Y. Kamil, "Breast Cancer Image Segmentation using Morphological Operations," no. MAY, pp. 2–9, 2015.
[16] Digambar A Kulkarni, Vijaylaxmi K Kochari,"Detection of Breast Cancer using K- Means Algorithm" in IJETAE, Volume 6, Issue 4, April 2016. Rejani YIA, Selvi ST., "Early Detection of Breast Cancer using SVM Classifier Technique", International Journal of Computer Science Engineering 2009; 1: 127-130.
[17] M. ArfanJaffar, "Dct Features based Malignancy and abnormality type detection method for mammograms "International Journal of Innovative Computing, Information and Control ICIC ISSN 1349-4198Volume 7, Number 9, September 2011
[18] Deepa Deshpande, ArchanaRajuskar, "Texture based associative classifier-An application of data mining for mammogram classification", Computational Intelligence in Data Mining 1,Springer
[19] Kanchanamani M, VaralakshmiPerumal, "Performance evaluation and comparative analysis of various machine learning techniques for diagnosis of breast cancer", proceedings of biomedical research 2016.
[20] Aswini Kumar Mohanty, Sukanta Kumar Swain, Pratap Kumar Champati, Saroj Kumar Lenka, "Image Mining for Mammogram Classification by Association Rule Using Statistical and GLCM features", International Journal of Computer Science Issue, 2011.