

Early Detection of Breast Cancer using Soft Computing

Amjad Khan¹ Zahid Ahmed Ansari²

¹Department of Electronics and Communication Engineering ²Department of Computer Science & Engineering

^{1,2}P.A College of Engineering Mangaluru-574153, India

Abstract— Breast cancer is the most harmful of all types of cancers that leads to the death of women. The initial screening test for breast cancer is made using Mammography. Masses, Calcification and Architectural distortion are the three major signs notable as causes for cancer recognized using mammogram images. Soft Computing methods provide solutions to biologically inspired problem of medical domain like breast cancer. Neural Networks, Fuzzy Logic and Genetic Algorithms contribute novel algorithms to deal with breast cancer. Breast cancer can be diagnosed using soft computing methods. The effective diagnosis of breast cancer can be achieved by using feature reduction and classification methods. The feature reduction method applied is Principal Component Analysis (PCA) and the classification method includes Support Vector Machines (SVM). The objective of this study is to improve the breast cancer early detection with the application of soft-computing techniques.

Key words: Breast Cancer, Mammography, Soft Computing, Fuzzy Sets, Feature reduction, Classification

I. INTRODUCTION

Breast cancer is the most harmful of all types of cancers that leads to the death of women. The initial screening test for breast cancer is made using Mammography. It is recognized that breast cancer ranks to 25th position worldwide as a severe root cause of death [7]. Approximately 4 lakh women are affected by this cancer out of which 50 thousand are from India. Masses, Calcification and Architectural distortion are the three major signs notable as causes for cancer recognized using mammogram images. So, early detection and diagnosis of this cancer becomes more critical. The early detection of breast cancer is very essential to provide the timely treatment so that the probability of curing the disease increases [3]. Sensitivity of breast Magnetic resonance images (MRI) shows an increased rate in detecting small cancer tumors in women at high risk. The sensitivity of 97% is impressive in terms of better diagnosing the breast tumors [2]. The pain and toxicity can be minimized if the disease is found in initial stages. Screening is the most popular method to detect the disease in the early stages before the symptoms are shown [29]. In order to analyze the Breast cancer images, two of the most popular data mining techniques are “clustering” and “classification”. Clustering is an unsupervised classification technique used for grouping the images into various disease classes. The selection of appropriate method for clustering is an important task in medical diagnosis and the choice of efficient clustering method should be done carefully to suits the desired task [13],[22]. The fundamental objective of carrying out image classification or clustering in medical image mining is to analyze the image contents and classify them into appropriate disease categories [14], [23]. Some of the most popular medical data used are CT scan, PET scan

and MRI scan images. Analyzing this large amount of data is a challenging task therefore it has become necessary to use data mining techniques to help in efficient and accurate prediction and diagnosis of cancer. The main goal of medical image mining is to extract clinically relevant information from the medical images to assist the physicians in accurate diagnosis and detection of diseases [20], [21].

Soft Computing played an important role in early detection and diagnosis of Breast cancer diseases analytically with improved effectiveness and suitable accuracy with the help of methods and proper attributes reference value. Soft computing is a consortium of methodologies that provides flexible information processing capability [14], [19]. Its aim is to exploit the tolerance for imprecision, uncertainty, approximate reasoning, and partial truth in order to achieve tractability, robustness, and low-cost solutions [11]. The outcome of this would help doctors, scientists, pharmacists in understanding the characteristic and association of attributes which is responsible for these diseases and provide proper diagnosis method and in discovering new drugs. The early detection of breast cancer can be helpful in curing the disease carefully. So the requirement of techniques to detect the occurrence of breast cancer nodule in early stage is increasing [7].

In the following, section II presents a review on various breast cancers based papers and in Table-1 gives about soft computing methods applied to breast cancer and along with performance. Next in section III, prediction of Breast cancer processes along with the diagram were explained. Section IV, discusses about the various imaging techniques and tools which were utilized for the breast cancer detection. Section V, discusses about soft computing and in section V, final conclusions were discussed.

II. REVIEW ON BREAST CANCER

Here are the reviews on various soft computing based breast cancer disease research papers are described following briefly.

Breast cancer is known as one of the most common cancers to afflict the female population. Computer assisted diagnosis can be helpful for doctors in detection and diagnosing of potential abnormalities. Several techniques can be useful for accomplishing this task. In this regard the following papers were reviewed to find the solutions for the breast cancer which assist to cure the diseases.

Ashwini et al.[1] proposed a hybrid approach of feature selection. It uses a new decision tree approach for classification to classify and detect the cancerous tissue. The approach can reduce the computation cost of mammogram image analysis. The algorithm uses simple statistical techniques in collaboration to develop a novel feature selection technique for medical image analysis.

Shruthi et al.[2] proposed a hybrid of Fuzzy c-means algorithm and Self organizing map algorithm to

segment the input image and then categorize tumor affected breast images and normal breast images. Breast MR Image segmentation presents many challenges related to shape and feature extraction in various segmentation techniques.

Kowsalya et al [3] presents the analysis of various detection and diagnosis techniques which is used to identify three major signs using mammogram images. The signs of breast cancer include Masses, Calcification and Architectural distortion which can be identified in the early stages. The mammogram image is processed in different stages to detect whether a person is suffering with benign or malignant breast cancer.

Anita et.al [23] aim to get the more accurate results by using various enhancement and segmentation techniques. In the segmentation stage the Watershed and Thresholding Segmentation is used and comparison has been made.

K.vijaya sri et al [4] expressed the net effect in the diagnosis of breast cancer using feature reduction and classification. Breast cancer can be well diagnosed using soft computing techniques. The net-effect of the classification before and after feature reduction process is stated.

The feature reduction method applied is Principal Component Analysis (PCA) and the classification method includes Support Vector Machines (SVM). The result of the proposed method produced better outcome when applied on Wisconsin Breast Cancer Data Set (WBCD) [4]

Sushma S.J et al [5] presents about the advancement of the medical image processing for the detection of breast cancer. Also discusses on imaging techniques such as ultrasound, CT scan, PET scan, MRI etc. it is proposed that a model for enhancing the contrast of mammogram for superior resolution will lead to better precision of detection of cancerous portion.

Soumadip et al [6] proposed the detection of breast cancer using a neuro- fuzzy based classification and then compare its performance Multilayer Perceptron and Support Vector Machine classification method. The method utilizes the primary benefits of Artificial Neural Networks such as immense parallelism, adaptivity, robustness and imprecision.

R Delshi Howsalya et al [7] proposed a prediction of breast cancer using decision tree classifier. The proposed approach consists of three steps. The first step groups the data in to number of clusters using Farthest First clustering algorithm, outliers are detected from breast cancer dataset using ODA (Outlier Detection Algorithm) and identifies whether the cancer is benign or malignant from the pre-processed data set using J48 classification algorithm. These

findings will help the doctors to diagnose the breast cancer and also helping the patients in recovery.

SL Aarthy et al [8] made a computerized approach on breast cancer detection using classification. The classifier is used to find out the different stages in the cancerous patients. When comparing with all the image modalities digital mammogram is inspected to improve the specificity and sensitivity of the breast cancer diagnosis.

Dansheng et.al [25] provides centrosomal numeral and morphological abnormalities. Here Regions of interest were selected based on segmentation, feature abstraction, and optimization using linear discriminant analysis (LDA) and support vector machines (SVM).

Ibrahim et al [9] presented an effective hybrid system for breast cancer classification. The proposed system combines K-means clustering algorithm, fuzzy rough feature selection (FRFS), and discernibility nearest neighbor (D-KNN) classifier. The proposed model performs with accuracy up to 98.9%.

Samy S et al [10] presented the design of the proposed Rule Based System which uses SL5 object language, symptoms of the breast cancer disease and possible ways to prevent it.

Fatima et al.[29] proposed outlines an approach for recognizing breast cancer diagnosis using neuro-fuzzy inference technique namely ANFIS (Adaptative Neuro-Fuzzy Inference System). According to the results obtained it is noted that the used method is very promising approach for recognition.

III. DETECTION OF BREAST CANCER

Early Detection Breast cancer by analyzing the related medical images is the proposed concept as shown in the figure-1. It is recognized that breast cancer ranks to 25th position worldwide as a severe root cause of death. Approximately 4 lakh women are affected by this cancer out of which 50 thousand are from India [7]. The Soft Computing methods provide solutions to biologically inspired problem of medical domain like breast cancer. The cancer image patterns are designed and these patterns are compared with the sample image data to find out the affected cancer patterns by applying the Soft Computing Techniques [19], [22].

The Soft Computing Techniques applied to medical images are used to predict the breast cancer. Also their comparisons are listed in the Table-1. By applying Suitable Clustering and classification techniques and their analysis gives the prediction of breast cancer.

Soft computing Methods used	Accuracy	Sensitivity	Image	Citation	Year
Genetic Algorithm, J48 classifier	95%	97%	mammogram	[1]	2010
FCM,neural networks	87.5%	91%	MR images	[2]	2015
k-means clustering, BBN	92.5%	83%	mammogram	[3]	2015
PCA-SVM	97.12%	96%	WBCD set	[4]	2016
Fuzzy logic based optimization	93.%	96.6%	mammogram	[5]	2016
Neuro-fuzzy based Classification,NFS, MLP,SVM	97.8% 86.3% 87.6%	95.8% 84.2% 82.5%	WBCD set	[6]	2016
Decision tree classifier,J48	99.6%	95%	WBCD set	[7]	2016
SVM ,linear classifier	93%	94%	MRI	[8]	2016

Hybrid classifier, SVM+ k-means	97.38%	---	USI data set	[9]	2016
Rule based system	95.6%	94.7%	mammogram	[10]	2016

Table 1: Comparison of soft computing methods applied to Breast cancer images

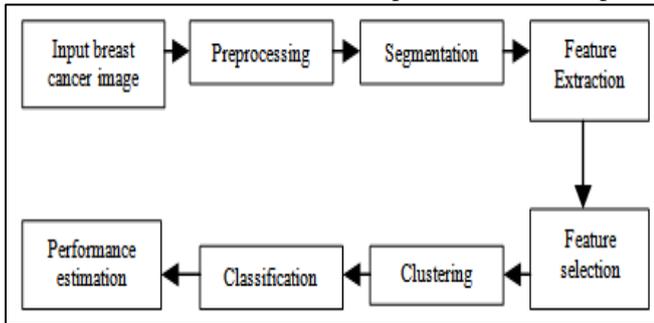


Fig. 1: Early detection of Breast Cancer

IV. MEDICAL IMAGE MINING

Medical image mining refers to the application in data mining techniques on medical image data. The main goal of medical image mining is to extract clinically relevant information from the medical images to assist the physicians in accurate diagnosis and detection of diseases. This section discusses about the various imaging techniques and tools which were utilized for the breast cancer detection [13],[20].

Some of the important breast cancer imaging techniques is discussed over here.

- 1) Mammography: Mammography is an X-ray technique that was introduced specifically to examine the breast lesions.
- 2) Computed Tomography (CT): There are so few benefits are associated with the Computed Tomography as high cost and high exposure of radiation is another demerit of this technique.
- 3) Position Emission Tomography Screening (PET): In this technique a radioactive chemical is applied to the patient before performing scan in order to see the hotspot areas under the scan. PET is used to diagnose the benign from malignant lesions of breast cancers [26].
- 4) Magnetic Resonance Imaging (MRI): Some few types of MRI such as Dynamic Contrast Enhancement MRI's can be useful to detect the malignancy of the non-palpable lesions of breast cancer.

V. SOFT -COMPUTING TECHNIQUES

Soft computing is a consortium of methodologies that provides flexible information processing capability. Its aim is to exploit the tolerance for imprecision, uncertainty, approximate reasoning, and partial truth in order to achieve tractability, robustness, and low-cost solutions [14]. Principal Soft computing techniques includes fuzzy sets, neural networks, genetic algorithms, and rough sets are most widely applied in the medical image mining. Generally fuzzy sets are suitable for handling the issues related to understandability of patterns, incomplete/noisy data, and mixed media information which provide approximate solutions faster [21].

Neural networks are nonparametric, robust, and exhibit good learning and generalization capabilities in data-

rich environments. Genetic algorithms provide efficient search algorithms to select a model from mixed media data, based on some objective function. Rough sets are suitable for handling different types of uncertainty in medical image data Clustering is a common technique for statistical data analysis, which is used in many fields, including machine learning, data mining, pattern recognition, image analysis and bioinformatics [15].

Among the fuzzy clustering methods, fuzzy c-means (FCM) algorithm is the method used in image segmentation because it has robust characteristics for ambiguity and it has a good performance in a large class of images [18]. The k-Means clustering algorithm is one of the most commonly used methods for partitioning the data. A fuzzy clustering: This method assigns degrees of membership in several clusters to each input pattern. A fuzzy clustering can be converted to a hard clustering by assigning each pattern to the cluster with the largest measure of membership [19]. Artificial neural networks (ANNs) for clustering: the features of ANNs are inherently parallel and distributed processing architectures.

VI. CONCLUSION

Breast cancer needs to be providing accuracy of diagnosis at the early stages. Mammography is one of the best methods in breast cancer detection Hence it is seen that Soft Computing methods provide solutions to biologically inspired problem of medical domain like breast cancer. Neural Networks, Fuzzy Logic and Genetic Algorithms contribute novel algorithms to deal with breast cancer. Breast cancer can be diagnosed using soft computing methods.

The effective diagnosis of breast cancer can be achieved by using feature reduction and classification methods. Applying soft computing techniques for breast cancer will improve the performance and fetch better results. The sensitivity of 97% is impressive in terms of better diagnosing the breast cancer. Soft computing methods based on image mining can help to make the prediction of breast cancer diseases more reliable, more effective and efficient. Hence help to prepare some methods for diagnosis, prognosis and decision making most accurate and efficient.

REFERENCES

- [1] Ashwini Kumar, Saroj Kumar, "Efficient Image Mining Technique for Classification of Mammograms to Detect Breast Cancer" ICCT-2010, 3rd-5th December 2010, pages 99-106.
- [2] Shruthi Anand, Viji Vinod and anand rampure, "Application of Fuzzy c-means and Neural networks to categorize tumor affected breast MR Images" International Journal of Applied Engineering Research, ISSN 0973-4562 Vol. 10 No.64 (2015),pages,274-281
- [3] Kowsalya, shanmuga priyaa, "A Survey on Diagnosis Methods of Breast Cancer using Mammography" International Journal of Engineering Technology

- Science and Research, ISSN 2394 – 3386 Vol. 02, Issue 4(April 2015),pages,100-107
- [4] K. Vijaya Sri1, K. Usha Rani, “net-effect in the diagnosis of breast cancer using feature reduction and classification” IJRET: International Journal of Research in Engineering and Technology, eISSN: 2319-1163, Volume: 05 Special Issue: 05, ICIAC-2016 | May-2016,pages, 62-65.
- [5] Sushma S J, S C Prasanna Kumar “Advancement in Research Techniques on Medical Imaging Processing for Breast Cancer Detection” International Journal of Electrical and Computer Engineering (IJECE), ISSN: 2088-8708, Vol. 6, No. 2, April 2016, pp. 717~724.
- [6] Soumadip Ghosh, Sushanta Biswas and Debasree Chanda, “Breast Cancer Detection using a Neuro-fuzzy based Classification Method” Indian Journal of Science and Technology, ISSN (Print): 0974-6846, Vol 9(14), April 2016, pages, 1-15
- [7] Delshi Howsalya Devi et al.,“ outlier detection algorithm combined with decision tree classifier for early diagnosis of breast cancer” International Journal of Advanced Engineering Technology ,E-ISSN 0976-3945, Vol. VII/Issue II/April-June.,2016/93-98.
- [8] SL. Aarthy, S. Prabu., “a computerized approach on breast cancer detection and classification” IIOAB Journal, ISSN: 0976-3104, Vol. 7/Issue 5/2016/157-169.
- [9] Ibrahim el-hasnony and Hazem M El-Bakry., “Classification of Breast Cancer Using Softcomputing Techniques” International Journal of Electronics and Information Engineering, Vol.4, No.1, PP.45-54, Mar. 2016.
- [10] Samy S. Abu Naser, Bashar G. Bastami, “A Proposed Rule Based System for Breasts Cancer Diagnosis” World Wide Journal of Multidisciplinary Research and Development, e-ISSN: 2454-6615, Vol.2, Issue.5, PP.27-33, 2016.
- [11] Hsu, Wynne, Mong Li Lee, and Ji Zhang. "Image mining: Trends and developments." Journal of Intelligent Information Systems 19, no. 1 (2002): 7-23.
- [12] Hema, a., and e. annasaro. "a survey in need of image mining techniques." international journal of advanced research in computer and communication engineering (ijarcece) issn (print): 2319-5940.
- [13] S.Zulaikha Beevi , M.Mohammed Sathik , K.SenthamaraiKannan, “A Robust Fuzzy Clustering Technique with Spatial Neighborhood Information for Effective Medical Image Segmentation” (IJCSIS) International Journal of Computer Science and Information Security, Vol. 7, No. 3, March 2010.
- [14] P.K.Vaishali , A.Vinayababu , “Application of Data mining and Soft Computing in Bioinformatics” International Journal of Engineering Research and Applications (IJERA), ISSN: 2248-9622, Vol. 1, Issue 3, pp.758-771
- [15] Ramadass Sudhir , “ A Survey on Image Mining Techniques: Theory and Applications” Computer Engineering and Intelligent Systems, ISSN 2222-1719, Vol 2, No.6, 2011.
- [16] J. Priya, “A Survey on Image Mining Techniques for Image Retrieval” International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), ISSN: 2278 – 1323, Volume 2, Issue 7, July 2013.
- [17] Yushu Xiong, “A Clustering Algorithm Based on Rough Set and Genetic Algorithm” TELKOMNIKA, Vol. 11, No. 10, October 2013, pp. 5782 ~ 5788, e-ISSN: 2087-278X.
- [18] Mignotte, M. “A de-texturing and spatially constrained K-means approach for image segmentation” Pattern Recognition Lett. (2010), doi:10.1016/j.patrec.2010.09.016. Elsevier.
- [19] Piero P. Bonissone, “Fuzzy Logic and Soft Computing: Technology Development and Applications” Schenectady NY 12309, USA. July 10, 1997.
- [20] K.J. and G.W. Moore Cios, "Uniqueness of medical data mining," Artificial Intelligence in Medicine, pp. 1-24, 2002.
- [21] R., Zhang, Y., Katta, "Medical Data Mining," Data Mining and Knowledge Discovery, pp. 305-308, 2002.
- [22] N.S.Nithya, Dr.K.Duraiswamy, P.Gomathy, “A Survey on Clustering Techniques in Medical Diagnosis” International Journal of Computer Science Trends and Technology (IJCSST) Vol. 1 Issue2, Nov-Dec 2013
- [23] E. Barati, M. Saraee, A. Mohammadi, N. Adibi and M. R. Ahamadzadeh, “ A Survey on Utilization of Data Mining Approaches for Dermatological (Skin) Diseases Prediction”, Journal of Selected Areas in Health Informatics (JSHI): March Edition, page.1-11,2011.
- [24] A.K. jain, M.N. murty, P.J. flynn, “Image Segmentation Using Clustering” 1996 IEEE Computer Society Press.
- [25] Dansheng Song, Tatyana A. Zhukov, Olga Markov, Wei Qian3, Melvyn S. Tockman, “Prognosis of cancer patients through quantitative analysis of centrosomal features”, iee, pp: 1607-1610, 3012.
- [26] Aparna Kanakatte, Nallasamy Mani, Bala Srinivasan, Jayavardhana Gubbi, “Pulmonary Tumor Volume Detection from Positron Emission Tomography Images”, International Conference on Biomedical Engineering and Informatics, pp: 213-217, 2008.
- [27] Yang Liu, Jinzhu Yang, Dazhe Zhao, Jiren Liu, “A Method of Pulmonary Nodule Detection utilizing multiple Support Vector Machines”, International Conference on Computer Application and System Modeling (ICASM 2010), pp: 118-121,2010
- [28] JIA Tong, ZHAO Da-Zhe, YANG Jin-Zhu,WANG Xu, “Automated Detection of Pulmonary Nodules in HRCT Images”, IEEE, 2007.
- [29] Bekaddour Fatima and Chikh Mohammed Amine, “ A Neuro-fuzzy inference model for breast cancer recognition” International Journal of Computer Science & Information Technology (IJCSIT) Vol 4, No 5, October 2012