

# **REVIEW ON DIAGNOSIS CANCEROUS TISSUE USING IMAGE DATA MINING**

Jaynesh H. Desai<sup>1</sup> and Eshakumari H. Desai<sup>2</sup>

**Abstract-** In today's scenario leading cause of death in the world is cancer, Early and errorless detection of cancer is critical to the wellbeing of patients. Breast cancer is the most common types of cancer in women. Diagnosis of Breast Cancer for early detection of cancer's masses from Images (Mammogram Images, Ultrasound Images (UI), Magnetic Resonance Imaging (MRI), Microscopy Images) through Image Segmentation by method of Image Processing and Data Analysis using different Data Mining Techniques. This diagnosis is less sensitive in younger women especially cancers in dense breasts become difficult to detect.

**Keywords –** Image mining, Domain specific application, Data Mining, Image Segmentation.

## **I. INTRODUCTION**

A disease caused by an uncontrolled division of abnormal cells in a part of the body called as cancer. Breast cancer has been determined to be the quite high cause of cancer death in women in India, and the most common type of cancer in women. Statistics from Globocan 2008 data shows that there is rapid growth in death of women suffering from this disease i.e. one death for every two cases detected [27]. So, early detection of the stage of cancer allows treatment which could lead to high survival rate. Early detection of cancerous region, Mammography is considered the most reliable method [3] [4], However it suffers from relatively high missed- and false-detection rates and involves uncomfortable compression of the breast. X-rays are also ionizing and therefore not generally suited to frequent screening [5].

Mining has been done in accordance with the integrated collections of images and its related data [1]. Image mining is the concept used to extract implicit and useful data from images stored in the large databases and further mining on that data, so it is an extension of Data Mining to Image Processing. An advance in image acquisition and storage technology has been led to tremendous growth in very large and detailed image databases. A vast amount of image data is generated in our daily life and in each field, such as medical image (CT images, ECT images & MR, mammograms images, etc.). These images involve a great number of useful and implicit information that is difficult for user to discover. Image mining can automatically discover these

---

<sup>1</sup> *B.M.College of Computer Application, Surat, Gujarat, India*

<sup>2</sup> *B.M.College of Computer Application, Surat, Gujarat, India*

implicit information and patterns from the high volume of images and is rapidly gaining attention in the field of data mining [2].

Data mining of medical images is used to collect effective models, relations, rules, abnormalities and patterns from large volume of data. This procedure can accelerate the diagnosis process and decision-making. Different methods of data mining have been used to detect and classify anomalies in mammogram images such as wavelets[6], statistical methods and most of them used feature extracted using image processing techniques. Some other methods are based on fuzzy theory and neural networks.

## II. REVIEW WORK

Data mining of medical images is used to collect effective models, relations, rules, abnormalities and patterns from large volume of medical data. This procedure can speed up the diagnosis process and decision-making. Different methods have been used to classify and detect variation in medical images, such as wavelets, Statistical methods and most of them used features extracted using image-processing techniques. In addition, some other methods were presented in the literature based on fuzzy set theory and neural networks.

A number of Association Rule (AR) mining algorithms apply on medical database has been available. The standard/traditional way for mining association rules is the Apriori algorithm [7]. But multiple times scanning of transactional database and generation of candidate item set is the main issue. So improvement to Apriori is essential concern.

D. Deshpande et al. [8] present mammogram classification using texture based association rule mining but result of classification obtained in malignant type stage is only 84%.

Kom et al. [9] present an algorithm for suspicious masses identification in mammogram but, result of mass detection obtain sensitivity is 95.91%

Eltonsy et al. [10] present a technique for the automatic detection of malignant masses in mammogram while screening but it only examine on malignant type stage.

L. F. A. Campos et al. [11], present A discrimination and classification method for mammographic image with benign, malignant and normal tissues implementing independent component analysis and multilayer neural networks.

C. Velayutham, et al. [12], presents A novel feature extraction method using spectral shape in digital mammogram image. Researcher compare texture description methods GLCM, GLDM, SRDM, NGLCOM, NGLDM, GLRLM with Spectral Shape but analysis accuracy value is 93.85 for Spectral Shape.

R. Nilavalan et al. [13], present Numerical analysis of microwave detection of breast tumours using synthetic focussing techniques.

Chi-Shih Yang et al. [14], present Parametric data mining and diagnostic rules for digital thermographs in breast cancer. Experimental results indicated that a total of 1750 abnormal regions (703 positive and 1047 negative) were detected. 61 positive abnormal regions ( $61/703=8.6\%$ ) from 42 cancer patients ( $42/71=59.2\%$ ) can be found. So, conclusion is less result gain from this technique.

AK Mohanty, et al. [15], presents An improved data mining technique for classification and detection of breast cancer from mammograms. The proposed system consists of three major steps: (1) ROI extraction of 256 X 256 pixels size. (2) The feature extraction, where a set of 26 features are used and these features are capable of differentiating between normal and cancerous breast tissues are found. (3) The classification process; the technique of the association rule miner is used to classify between normal and cancerous tissues.

T Goudas et al. [16], present Characterization of the cells and the quantification of breast cancer in Microscopy Images.

Gokhan Zorluoglu et al. [17], present Diagnosis of Breast Cancer Using Ensemble of Decision Trees (DT), Support Vector Machines (SVM), Artificial Neural Network (ANN).

A. Mohanty et al. [18], present Diagnosis of Breast Cancer by Association Rule Using Statistical and GLCM features.

Vrushali Gaikar et al. [19], present Clustering of Breast Cancer Tumor using Third order GLCM feature. For benign Tumor gave 81.0% result and 93% accuracy.

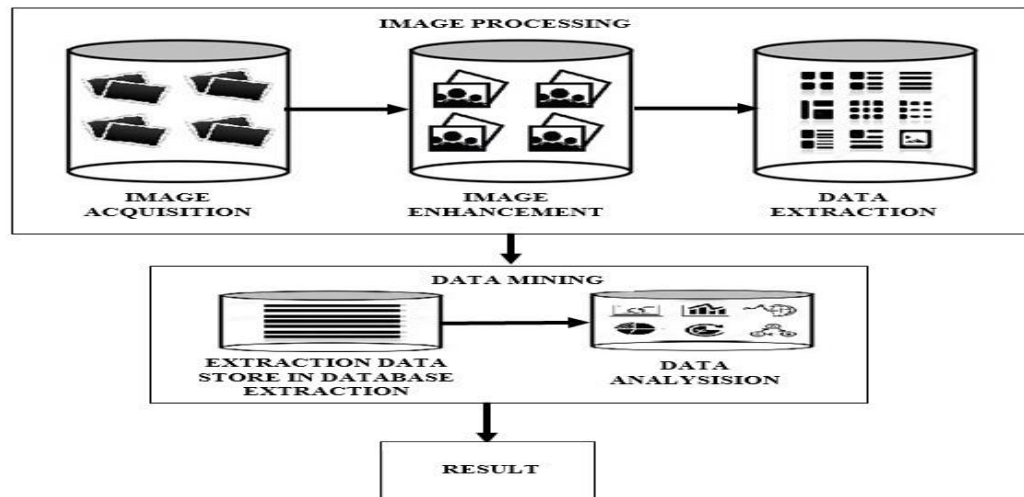
### **III. REVIEW FINDING**

According to review work in Breast cancer diagnosis majority work is done with the use of 2D-Mammogram image dataset. Here scope of research will be using 3D-Mammogram dataset image for quickly discover diagnosis on cancerous tissue.

Mining the large archives of digitized medical images of hospitals, healthcare institutions, and physicians' offices provides unique opportunity for advances in medicine and health care. MRI, PET scans, mammograms, ultrasound images, protein crystallography – all these image categories are excellent candidates for image mining. Through image-content database query techniques, data mining can discover common and indicative patterns leading to early and automated discovery of abnormal organ Tissues, lesions, cancer and others. Moreover, it is also possible to use image-content data mining as a clinical tool for new cancer screenings as well as to provide clues to deeper understanding of the nature of different sicknesses and their correlation with environmental and genetic factors as derived from patient records.

### **IV. METHODOLOGY**

As per review work, in below figure I illustrate the methodology for image mining process for Cancerous Tissue images.



## V. CHALLENGES

- The challenge of using mammography is how to locate the area, which is indeed a solitary geographic abnormality.
- To identify whether image belonging in “Malignant” stage or “Benign” Stage with improved clarity.
- To classify the cancerous tissue size, lymph node in cancerous tissue cell.

## VI. CONCLUSION

As given in the review work, there are various algorithms and methods available for mining Cancerous images of breast which can be quickly discover frequent item-sets and can be applied to mining Cancerous image dataset.

## REFERENCES

- [1] “Survey on Image Mining Techniques: Theory and Applications”, Computer Engineering and Intelligent Systems ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online) Vol 2, No.6, 2011 A Ramadass Sudhir, Research Scholar of Manonmaniam Sundaranar University.
- [2] “Image Mining: Trends and Developments”, I Zhang Wynne Hsu Mong Li Lee, School of Computing National University of Singapore, Singapore 117543, 2002.
- [3] “Automatic Detection and Classification of Cancerous Masses in Mammogram”, S.Pitchumani Ngayarkanni, Dr.Nadira Banu Kamal, Dr. V Thavavel, IEEE Trans, 2012.
- [4] “Application of Data Mining Techniques for Medical Image Classification”, Maria-Luiza Antonie, Osmar R. Zaiane, Alexandru Coman, 2001.
- [5] “Numerical analysis of microwave detection of breast tumours using synthetic focussing techniques”, R. Nilavalan; J. Leendertz; I. J. Craddock; A. Preece; R. Benjamin Antennas and Propagation Society International Symposium, 2004. IEEE
- [6] “Efficient Image Mining Technique for Classification of Mammograms to Detect Breast Cancer”, Aswini Kumar Mohanty, Saroj Kumar Lenka, 2010.
- [7] “Fast algorithms for mining association rules”, R. Agrawal and R. Shrikant, Proc. 20th Int’l. Conf. Very Large Databases, pp.487-499, 1994
- [8] “Medical image analysis an attempt for mammogram classification using texture based association rule mining”, Deepa S. Deshpande; Archana M. Rajurkar; Ramchandra M. Manthalkar, IEEE- 2013
- [9] “Automated detection of masses in mammograms by local adaptive thresholding”, G. Kom, A. Tiedeu, M. Kom, Comput. BioI. Med.37 (I) (2007) 37-48.

- 
- [10] "A concentric morphology model for the detection of masses in mammography", N. Eltonsy, G. Tourassi, A. Elmaghraby, Medical Imaging, IEEE Transactions. (2007)
- [11] "Diagnosis of breast cancer in digital mammograms using independent component analysis and neural networks", L. F. A. Campos, A. C. Silva, A. K. Barros, X Iberoamerican Conference on Pattern Recognition, Havana. (2005)
- [12] "A novel feature extraction method using spectral shape in digital mammogram image", C. Velayutham; K. Thangavel. IEEE-2011
- [13] "Numerical analysis of microwave detection of breast tumours using synthetic focussing techniques", R. Nilavalan; J. Leendertz; I. J. Craddock; A. Preece; R. Benjamin, Antennas and Propagation Society International Symposium, 2004. IEEE
- [14] "Parametric data mining and diagnostic rules for digital thermographs in breast cancer", Chi-Shih Yang, Ming-Yih Lee, Engineering in Medicine and Biology Society, IEEE 2008.
- [15] "An improved data mining technique for classification and detection of breast cancer from mammograms", AK Mohanty, MR Senapati, SK Lenka - 2013 – Springer
- [16] "An Advanced Image Analysis Tool for the Quantification and Characterization of Breast Cancer in Microscopy Images", T Goudas, I Maglogiannis - Journal of medical systems, 2015 – Springer
- [17] "Diagnosis of Breast Cancer Using Ensemble of Data Mining Classification Methods", Gokhan Zorluoglu, Mustafa Agaoglu , AIS – 2015
- [18] "Image Mining for Mammogram Classification by Association Rule Using Statistical and GLCM features", Aswini kumar mohanty, Sukanta kumar swain ,Pratap kumar champati ,Saroj kumar lenka, IJCSI-2014
- [19] "Clustering of Breast Cancer Tumor using Third order GLCM feature", Vrushali Gaike, Rahul Mhaske, IEEE-2015
- [20] "A Survey of Image Processing Algorithms in Digital Memmography" Jelena Bozek, Mario Mustra, Kresimir Delac, and Mislav Grgic IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 5, No 3, September 2009
- [21] "Breast Tissue Classification Using Statistical Feature Extraction of Mammograms", Holalu Seenappa Sheshadri and Arumugam Kandaswamy, 2006.
- [22] "Application of Fuzzy Cluster analysis for Medical Image Data Mining", Shuyan Wang, Mingquan Zhou and Guohua Geng, Proceedings of the IEEE International Conference on Mechatronics & Automation Niagara Falls, Canada,pp. 36 – 41,July 2005.
- [23] "Image Mining using Association rules derived from Feature Matrix", L.Jaba Sheela , V. Shanthi , D.Jeba Singh (2009) In proceedings of the 2009 International Conference on Advances in Computing, Communication and Control (ICAC3'09), Mumbai. World of Computer Science and Information Technology Journal (WCSIT) ISSN: 2221-0741 Vol. 1, No. 7, 311-316, 2011
- [24] "Using Image Mining to Discover Association Rules between Image Objects", Prof. Dr.Hilal M. Yousif\* Dr.Abdul- Rahman Al-Hussaini Mohammad A. Al-Hamami, 2003.
- [25] "Enhancement of classification accuracy of our Adaptive Classifier using image processing techniques in the field of Medical Data Mining" By Chandra, S. Sch. of Computer Sci. & Eng., Lovely Prof. Univ., Phagwara, India [2015]
- [26] "Analysis of fractal image compression technique in different medical imaging modalities" Research By Sumathi poobal [2008]
- [27] <http://globocan.iarc.fr/>