ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 8, Issue 9 , September 2021

# Preprocessing and Segmentation of Digital Mammogram Images for Early Detection of Breast Cancer

**O.R.Yusupov, Kh.S.Abdiyeva, A. Primov**

PhD, Department of Software Engineering, Samarkand State University, Uzbekistan
Assistant teacher, Department of Software Engineering, Samarkand State University, Uzbekistan
Student, Department of Software Engineering, Samarkand State University, Uzbekistan

**ABSTRACT**:The impact of digital images on modern life is tremendous, and image processing is now an important component in science and technology. The rapid development in computerized medical image reconstruction, and the associated progresses in analysis methods and computer-aided diagnosis, has propelled medical imaging into one of the most important sub-fields in scientific imaging. Nowadays, some cancer types can be detected by using medical images such as mammograms, CT images and others. A mammography has become the most effective technique for early breast cancer detection. Digital mammogram takes an electronic image of the breast and stores it directly in a computer. Imaging techniques play a significant role in helping performance of digital mammogram, especially of abnormal areas that cannot be felt by radiologist but can be seen on a conventional mammogram. Before any image-processing algorithm of mammogram pre-processing steps are very important in order to limit the search for abnormalities without undue influence from background of the mammogram.

**KEY WORDS**: Mammogram image, breast cancer, image enhancement, image segmentation.

## I.INTRODUCTION

Cancer is a leading cause of death and decreases life expectancy in every country of the world. Breast cancer arises in the lining cells (epithelium) of the ducts (85%) or lobules in the glandular tissue of the breast [1]. At the beginning, the growth of cancerous cells is confined to the duct where it generally causes no symptoms and has minimal potential for spread (metastasis). Over time, in the first stage of the disease cancers may progress and invade the surrounding breast tissue (invasive breast cancer) then spread to the nearby lymph nodes (called “regional metastasis”) or to other organs in the human body (called “distant metastasis”). Because of widespread metastasis, a woman dies from breast cancer. A treatment of breast cancer can be highly effective, especially when the disease is identified in early stages. Treatment of this illness often consists of a combination of surgical removal, radiation therapy and medication ( hormonal therapy, chemotherapy and targeted biological therapy) to treat the microscopic cancer that has spread from the breast tumor through the blood. In 2020, there were 2.3 million women diagnosed with breast cancer and 685 000 deaths globally. As of the end of 2020, there were 7.8 million women alive who were diagnosed with breast cancer in the past 5 years. Breast cancer occurs in every country of the world in women at any age after puberty but with increasing rates in later life. Breast cancer mortality changed little from the 1930s through to the 1970s. Improvements in survival began in the 1980s in countries with early detection programs combined with different modes of treatment to eradicate invasive disease.[1]

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 8, Issue 9 , September 2021

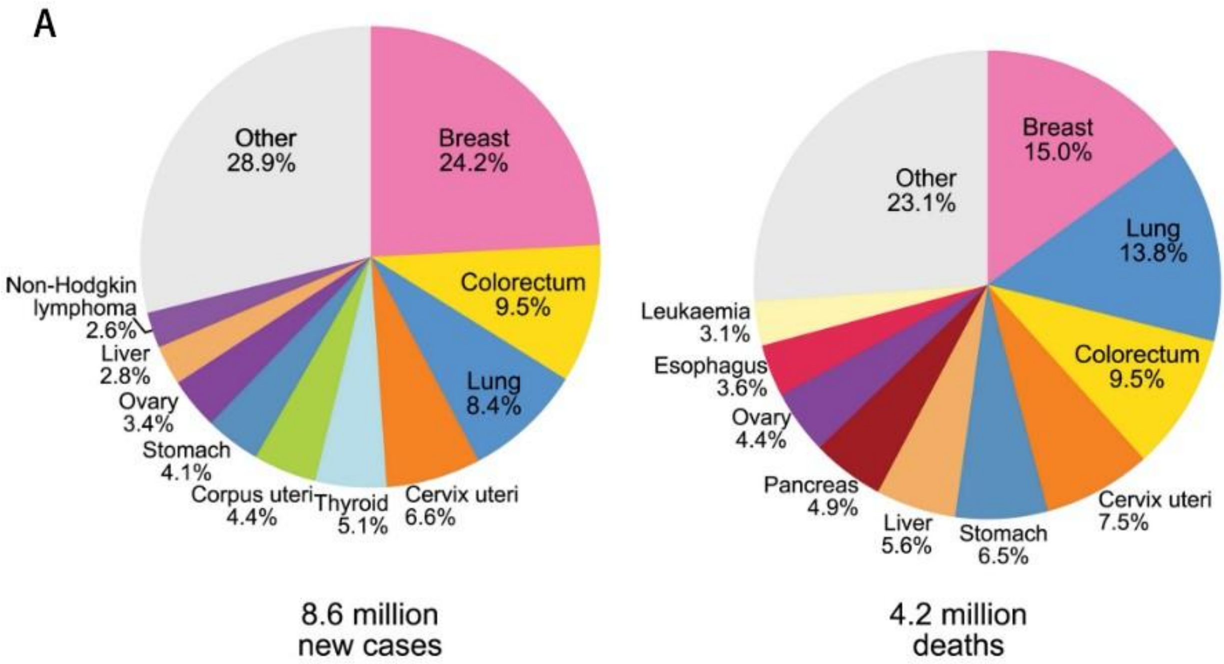



Fig1: Cancer distribution worldwide 2020

Currently, three methods are used for breast cancer diagnosis: mammography, surgical biopsy and fine needle aspirate. Mammography has a reported malignant sensitivity which varies between 68 and 79% [2]. Fine needle aspirate depends on extracting fluids from a breast lump and inspecting it under the microscope. This method has a reported sensitivity varying from 65 to 98% [2]. Surgical biopsy is more evasive and costly but it is the only test that can confirm malignancy. Efficient machine learning algorithms can enhance the performance of mammogram analysis and provide an equivalent performance in terms of robustness and accuracy for surgical biopsy without its evasiveness and cost. Mammographic screening allows early detection of non-palpable, non-invasive and early invasive tumors. Hence, it can reduce the mortality from breast cancer by 20-30% [3]. There is an increasing need for automatic and accurate detection of cancer cells. However, the low contrast between the breast cancer cells and normal cells increases the difficulty of early detection. Most of the work in mammography aims at detecting one or more of the three abnormal structures in mammograms [4]: micro-calcifications [5], circumscribed masses [6] and speculated lesions [7]. Other methods depend on classifying the breast lesions as benign or malignant [8]. There are problems with the subjective analysis of mammographic images by radiologist [7]. Subjective analysis depends mainly of the experience of the human operator, but it is also affected by fatigue and other human-related factors. In addition, the interpretation is a repetitive task that requires lot of attention to minute details. Hence, it requires lot of staff time and effort, which results in slowing the diagnosis time. On the other hand, the objective analysis of mammograms, which is carried out by automated systems, provides consistent performance but its accuracy is usually lower. Due to the sensitivity of this problem, we believe that radiologists should be involved and computers should not replace them completely. However, computer systems can help them perform better by enhancing the quality of images, highlighting the suspicious regions and providing better analysis tools.

## II. SIGNIFICANCE OF THE SYSTEM

The paper mainly focuses on pre-processing and segmentation of digital mammogram images for early detection of breast cancer. The study of literature survey is presented in section III, Methodology is explained in section IV, section V covers the experimental results of the study, and section VI discusses the future study and Conclusion.

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 8, Issue 9 , September 2021

## III. LITERATURE SURVEY

Mammograms are medical images that are difficult to interpret, thus a pre-processing phase is needed in order to improve the image quality and make the segmentation results more accurate [7]. In [7], an automated system was developed for assisting the analysis of digital mammograms. Computer image processing techniques were applied to enhance images and this was followed by segmentation of the region of interest (ROI). The textural features were extracted from the ROI. The texture features were used to classify the ROIs as either masses or non-masses. JelenaBozek and others [8] gave a survey of image processing algorithms that have been developed for detection of masses and calcifications. An overview of algorithms in each step (segmentation step, feature extraction step, feature selection step, classification step) of the mass detection algorithms was given. Wavelet detection methods and other recently proposed methods for calcification detection were presented. An overview of contrast enhancement and noise equalization methods was given as well as an overview of calcification classification algorithms. Ayman A. and others [9] introduced a preprocessing technique for reducing the size and enhancing the quality of USF and MIAS mammogram images . The algorithm analyzed the mammogram image to determine if 16-bit to 8-bit conversion process is required. Enhancement was applied later followed by a scaling process to reduce the mammogram size. The performances of the algorithms were evaluated objectively and subjectively. On average 87% reduction in size was obtained with no loss of data at the breast region.
To achieve the segmentation Sara Dehghani and others[10] proposed a "three-phase" based method:
Stage A: omit the excessive parts of the image
Excessive background parts which do not cover breast region, are the right and left part of the image.
Stage B: diagnose the image direction and put it to one direction
Stage C: omit labels and breast region from background.
Shruthishree S.H, and others [11] sketched some of the fundamental concepts of medical image processing.

## IV. METHODOLOGY

The first step of the pre-processing of image is to change the image from rgb form to grey level form. Then according to the grey level of threshold limit we make the original image binary. There is a sample of binary.

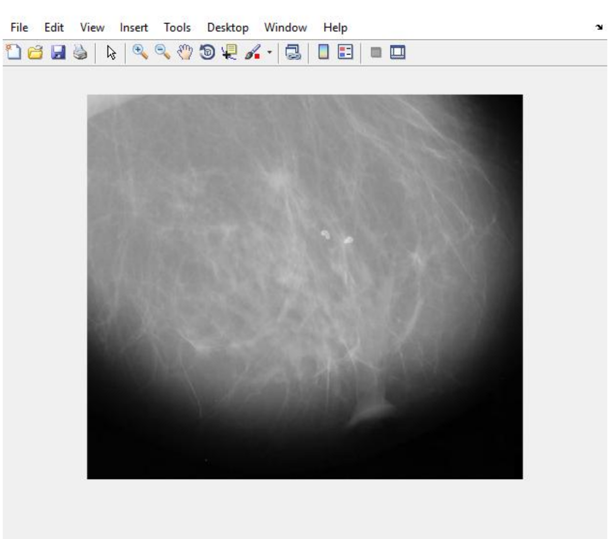



Fig2: A sample of grey level image

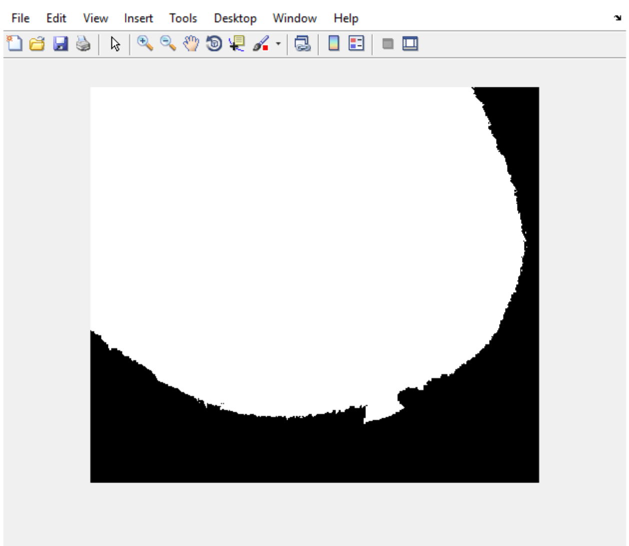



Fig3: Binary form of mammogram

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 8, Issue 9 , September 2021

Threshold limit or the amount of pixels's brightness of binary picture is equal to 1. Then with high attention the breast region is separated from the background and the other parts (the grey black surface) become zero. With this action in addition to breast tissue segmentation, notes have been omitted simultaneously. Figure 4 shows the breast region. The same as you can see the labels in the upper part of the image which is considered as a noise is omitted in this stage.

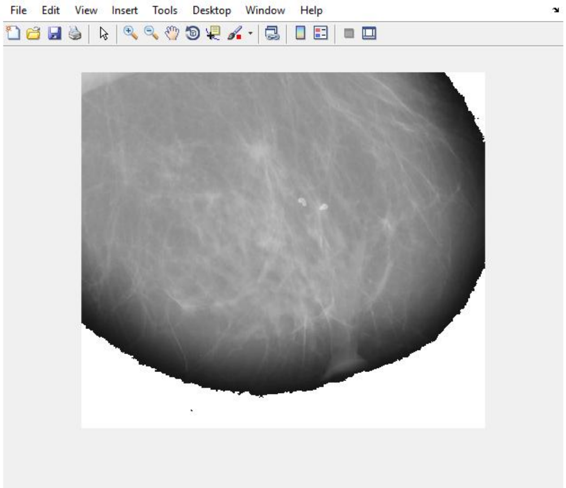
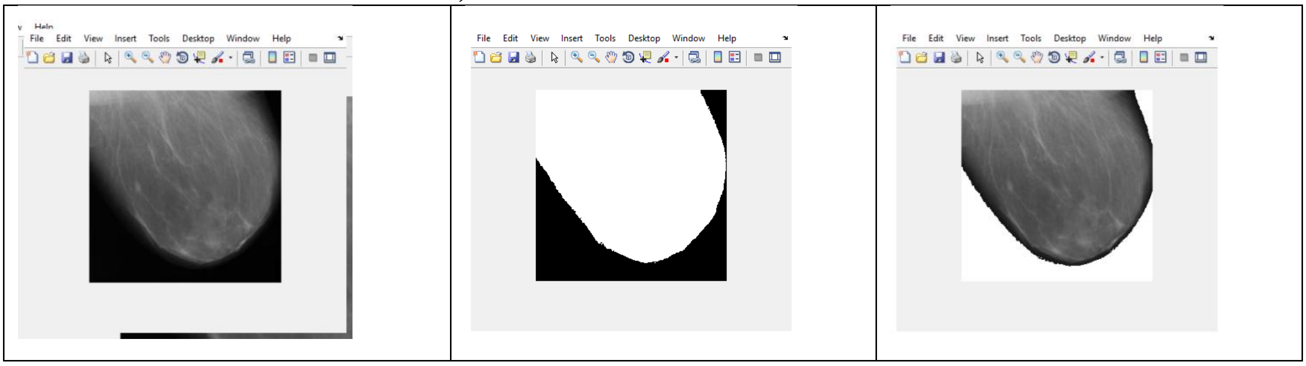


Fig4: Breast region

## V. EXPERIMENTAL RESULTS

We used the public database Mini MIAS[25] to test our method. It is a reduced version of the original MIAS Database (digitized at 50 micron pixel edge) that has been reduced to 200 micron pixel edge and clipped or padded so that every image is 1024x1024 pixels. Figure 5 shows three representative results. We have tested over 70 images, and we have obtained a 90% of "near accurate" results, which include the "accurate" results.



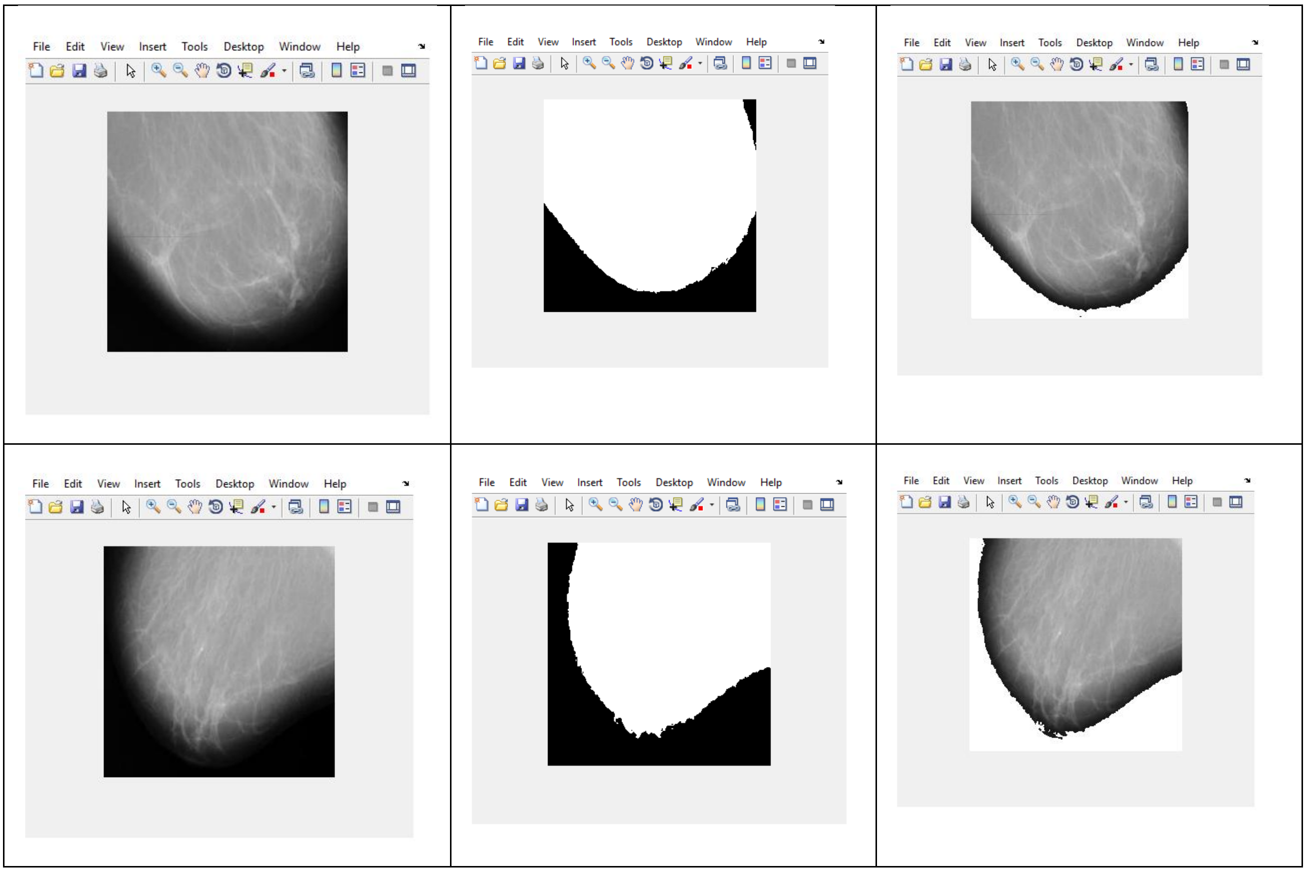



Fig5: An example of the performance of the presented approach of the profile of three different breasts

## VI.CONCLUSION AND FUTURE WORK

The literature survey will be a useful resource for others researching in this area.To summarize, the results obtained by the method show that it is a robust approach but it can be improved in terms of accuracy. Even so, we accept this method because it provides useful regions (there is no meaningful loss of information).

## REFERENCES

[1] World health organization. www.who.int
[2] Mangasarian, O.L., 1995. Breast cancer diagnosis and prognosis via linear programming. Oper. Res., 43: 570-577.
[3] Coakley, K., F. Quintarelli, T. van Doorn and C. Hirst, 1994. Classification of equivocal mammograms through digital analysis. The Breast, 3: 222-226.
[4] Comer, M., S. Liu and E.J. Delp, 1996. Statistical segmentation of mammograms. Proc. 3rd Intl. Workshop on Digital Mammography, Chicago, pp: 475-478.
[5] Strickland, R.N. and H.I. Hahn, 1996. Wavelet Transforms for detecting microcalcifications in mammograms. IEEE Trans. Med. Imaging, 15: 218-229.
[6] Giger, M.L., F.-F. Yin, K. Doi, C.E. Metz, R.A. Schmidt and C.J. Vyborny, 1990. Investigation of methods for the computerized detection and analysis of mammographic masses. Proc. SPIE, Washington, pp: 183-184.
[7] Samir Kumar Bandyopadhyay ,2010. Preprocessing of mammogram images. International Journal of Engineering Science and Technology Vol. 2(11), 6753-6758
[8] JelenaBozek, Mario Mustra, KresimirDelac, and MislavGrgic. 2009. A Survey of Image Processing Algorithms in Digital Mammography. , SCI 231, pp. 631–657
[9] Ayman A. AbuBaker, R.S.Qahwaji, Musbah J. Aqel, Hussam Al-Osta and Mohmmad H. Saleh.2006. Efficient Pre-processing of USF and MIAS Mammogram Images Journal of Computer Science 3 (2): 67-75
[10] Sara Dehghani .,MashallahAbbasiDezfooli. A Method For Improve Preprocessing Images Mammography International Journal of Information and Education Technology, Vol. 1, No. 1, April 2011

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 8, Issue 9 , September 2021

[11] Shruthishree S.H, and Harshvardhan Tiwari. (2017). “A REVIEW PAPER ON MEDICAL IMAGE PROCESSING.” International Journal of Research - Granthaalayah, 5(4) RACSIT, 21-29. https://doi.org/10.29121/granthaalayah.v5.i4RACSIT.2017.3344.
[12] Alpaydin E (2004) Introduction to machine learning, The MIT Press.
[13] Geoff Dougherty. 2009. Digital Image Processing for Medical Applications.
[14] Rafael C. Gonzalez., Richard E. Woods. Digital Image Processing.Third edition.
[15] S. Sharma, M. Kharbanda, and G. Kaushal, “Brain tumor and breast cancer detection using medical images,” International Journal of Engineering Technology Science and Research, vol. 2, 2015.
[16] Samir M. Badawy1, Alaa A. Hefnawy2, “Breast Cancer Detection with Mammogram Segmentation: A Qualitative Study”, (IJACSA) International Journal of Advanced Computer Science and Applications, vol. 8, No. 10, 2017.
[17] “Breast cancer facts & figures 2017-2018.” [Online]. Available: https://www.cancer.org/content/dam/cancerorg/ research/cancer-facts-and-statistics/breast-cancerfacts-and-figures/ breast-cancer-facts-and-figures-2017- 2018.pdf.
[18]Uday, Kumbhar, Vishal, Patil, and Shekhar, Rudrakshi. “Enhancement of Medical Images Using Image Processing In MATLAB.” International Journal of Engineering Research & Technology, vol. 2, No. 4, pp. 2359-2364, 2013
[19] A.S., Bhide, Priyanka, Patil, and Shraddha, Dhande “Brain Segmentation using Fuzzy C means clustering to detect tumor Region.” International Journal of Advanced Research in Computer Science and Electronics Engineering, vol.1, No. 2, pp. 85-90, 2012.
[20]Pratibha, Sharma, Manoj, Diwakar, and Sangam, Choudhary. “Application of Edge Detection for Brain Tumor Detection.” International Journal of Computer Applications, vol. 58, No. 16, pp. 21-25, 2012.