

Mammogram Pre Processing and Data Manipulation using Novel Threshold Based Pre Reduction Technique

S. Thilagavathi, S. Ahamed Johnsha Ali, S. Lashmidevi, K. Brindha

Abstract — Mammography is the best existing assessment for the discovery of early indications of breast cancer, for example, masses, calcifications, respective asymmetry and building twisting. Furthermore, mammograms are hard to interpret, particularly of cancer at their beginning times. The presence of relics and commotions can exasperate the location of breast cancer and decrease the pace of accuracy in the computer aided analysis (CAD). For this reason, pre-processing of mammogram images is significant during the time spent breast cancer analysis since it could decrease the quantity of false positive. Then again, it could assist radiologists with making an examination between mammograms. This paper proposed to Novel Threshold Based Pre Reduction Technique and Pre Processing research strategy of mammography breast cancer image processing.

Keywords: Threshold, Mammogram, Pre Processing, Effectiveness.

I. INTRODUCTION

Breast cancer is the most incessant cancer in ladies around the world. [1] The sickness is reparable whenever distinguished early enough. Screening is done based on mammograms, which use x-beam images to uncover protuberances in the breast. Calcium stores can likewise show the presence of a tumor. In any case, the stores are often just a couple of tenths of a millimeter in size thus profoundly installed in thick tissue that they are almost imperceptible in the images.

[3] Mammogram tests with stamped harmful tumor. Digital mammography is demonstrated as effective apparatus to recognize breast cancer before clinical indications show up. Digital mammography is as of now considered as standard method for breast cancer diagnosis, different man-made brainpower techniques are utilized for arrangement issues in the territory of medicinal diagnosis. [4] Feature extraction of image is significant advance in mammogram order. These features are extricated utilizing image processing techniques. A few sorts of feature extraction from digital mammograms including position feature, shape feature and surface feature and so forth. Surfaces are one of the significant features

Revised Manuscript Received on February 22, 2020.

* Correspondence Author

S. Thilagavathi*, Department of Computer science, Sri Krishna Adithya College of Arts and Science, Coimbatore, Tamilnadu, India Email: thilaga.hicet@gmail.com

S. Ahamed Johnsha Ali, Department of Computer Application, Sri Krishna Adithya College of Arts and Science, Coimbatore, Tamilnadu, India, ahamed.doit@gmail.com.

S. Lakshmi Devi, Department of Computer science, Sri Krishna Adithya College of Arts and Science, Coimbatore, Tamilnadu, India Email: lakshmi@skacas.ac.in

K. Brindha, Department of Computer Application, Sri Krishna Adithya College of Arts and Science, Coimbatore, Tamilnadu, India, brindhak@kacas.ac.in.

utilized for some applications. Texture features have been broadly utilized in mammogram characterization. The surface features are capacity to recognize unusual and ordinary cases. [2] Surface can be described as the spa circulation of dim levels in an area. Surface feature have been demonstrated to be helpful in separating typical and irregular example. Extricated surface features give information about textural attributes of the image. Various classifiers are utilized for restorative imaging application including man-made consciousness, wavelet and so forth. Surface measures are two sorts, first request and second request. In the principal request, surface measure are insights determined from an individual pixel and don't consider pixel neighbor connections. [5] Power feature are first request surface computation. In the subsequent request, measures consider the connection between neighbor pixels GLCM is a second request surface estimation. Texture features has been separated and utilized as parameter to improve the arrangement result.

Mammography (additionally called mastography) is the way toward utilizing low-energy X-beams (more often than not around 30 kVp) to look at the human breast for diagnosis and screening. The objective of mammography is the early discovery of breast cancer, commonly through recognition of trademark masses or miniaturized scale calcifications. [8] Like all X-beams, mammograms use dosages of ionizing radiation to make images. These images are then dissected for any unusual discoveries. It isn't unexpected to utilize lower-energy X-beams, commonly Mo (K-shell x-beam energies of 17.5 and 19.6 keV) and Rh (20.2 and 22.7 keV) than those utilized for radiography of bones. Ultrasound, ductography, positron emission mammography (PEM), and magnetic resonance imaging (MRI) are assistants to mammography. Ultrasound is commonly utilized for further assessment of masses found on mammography or substantial masses not seen on mammograms. Conduit grams are as yet utilized in certain foundations for assessment of bleeding areola release when the mammogram is non-analytic. [10] X-ray can be helpful for further assessment of faulty discoveries just as for screening pre-careful assessment in patients with known breast cancer to distinguish any extra sores that may change the careful methodology, for example from breast-saving lumpectomy to mastectomy. Different strategies being explored incorporate tom amalgamation.

II. LITERATURE SURVEY

Milos Radovic, Marina Djokovic, Aleksandar Peulic, Nenad Filipovic presents the CAD (computer aided diagnosis) framework for the identification of ordinary and anomalous example in the breast. The proposed framework comprises of four significant advances: the image preprocessing, the feature extraction, the feature choice and the characterization procedure that orders mammogram into typical (without tumor) and anomalous (with tumor) design. In the wake of expelling clamor from mammogram utilizing the Discrete Wavelet Transformation (DWT), first is chosen the region of interest (ROI). By recognizing the limit of the breast, it is conceivable to expel any antiquity present outside the breast zone, for example, quiet markings. This paper demonstrates that best in class techniques of image processing and the majority discovery are valuable in computer aided diagnosis. Many missed radiologist conclusions can be ascribed to human factors, for example, emotional or shifting choice criteria, interruption by other image features or enormous number of images to be investigated. Therefore, computer-aided diagnosis is a significant research zone. The techniques like one presented in this paper could help the radiologist and improve the accuracy of identification. Identification is done based on textural descriptors acquired from features extraction process. Results demonstrate that not many of proposed information mining calculations can manage the issue of mammogram characterization. This methodology has potential for further advancement as a result of its effortlessness that will persuade continuous breast cancer diagnosis in giving a second sentiment to radiologists. S.N. Deepa and B. Aruna Devi *et.al* proposed a Computer-Aided Diagnosis for the medicinal anticipation. Fake Neural Network forms the base of the canny frameworks. There are various examples any place computerized reasoning is use for the diagnosis of the chest cancer. The wise registering techniques can be utilized for analytic sciences in biomedical image order. Faye *et al* proposed a characterization technique for images is based on preselecting features based on their abilities of separating classes utilizing a T test. Irregular subsets accomplishing a predefined accuracy rate are then used to produce a last arrangement of features. The technique was utilized in this work with wavelet transform with LDA and KNN classifiers. In spite of the fact that the last accuracy rate got in the analyses are generally low, the improvement when consolidating classifiers is exceptionally reassuring Pereira D. C. *et al* presents a lot of computational devices to help division and discovery of mammograms that contained mass or masses in CC and MLO sees. A curio evacuation calculation is first actualized pursued by an image de-noising and dark level improvement strategy based on wavelet transform and Wiener channel. At long last, a strategy for identification and division of masses utilizing numerous thresholding, wavelet transform and hereditary calculation is utilized in mammograms which were haphazardly chosen from the Digital Database for Screening Mammography (DDSM).

III. RESEARCH METHODOLOGY

In this method to select the digital mammogram since they are the best technique for early recognition of cancer tumor. The

mammographic images used in this preliminary are constant images assembled from Scan reports. The images are coordinated in sets of mammograms where each match represents the left and right cancer of a patient. Those images can be classes as the going with dataset, which are altruistic, perilous.

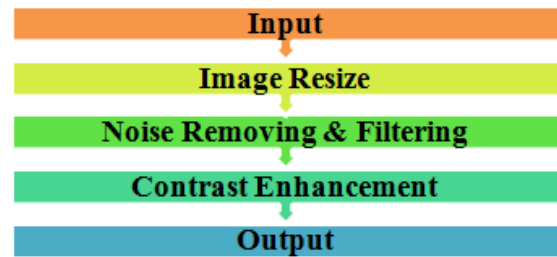


Figure 1: Proposed Pre-Processing Methodology

Image enhancement technique is characterized as a strategy in image preparing with the end goal that the outcome is substantially more reasonable than the first image utilized for a particular application. This change of the digitally put away image is prepared with the assistance of the MATLAB software. The means in the proposed for preprocessing technique (Simultaneous Noise Removal, sifting and Contrast Enhancement)

1. The ongoing mammographic image is gained from the output focus
2. The Color image is changed over to Gray scale image.
3. To resize in various size in given image
4. Noise from the gray image is expelled by using the separating technique
5. Differentiation of the image is improved by applying the Adaptive Histogram Equalization to the stage 4 image
6. They got image is the improved yield image

IV. PROPOSED WORK

4.1 Proposed preprocessing technique – Novel Threshold Based Pre reduction technique (NTBPT)

The Novel preprocessing technique gives the blend of three particular information mining techniques dim change, Noise evacuation and for resizing. So as to recognize the suspicious discovery, the influenced district of the cancer is first fragmented. Recognizing the frontal area of the concerned instate thresholding technique was connected to the image which contains a cancer locale and a pectoral muscle district in most Medio Lateral-Oblique (MLO) perspectives of mammograms. After this preprocessing thresholding more data has been assembled. The more splendid districts are represented with pectoral muscle. This situation impacts the identification results. Translating process with a Gamma remedy is connected to ensure the luminance and the chrominance of a image. The Gama rectification improves the pectoral muscle. This can be killed, disintegration and widening channel. This preprocessing stage decreases the consequences for image clamors the granular tissues and the veins which frequently prompt false positive outcomes in suspicious reason, consequently the preprocessing stage is substantially more accommodating for this exploration objective.



The preprocessing utilizes the blend of 3 calculations for dark transformation, Noise evacuation and for resizing. Two calculations are proposed to diminish the capacity necessities for mammogram images. The info image experiences a contracting procedure that believes the 16-bit images to 8-bits by utilizing pixel-profundity change calculation pursued by upgrade process. The execution of the calculations is assessed equitably and emotionally. [11] A half reduction in size is gotten with no loss of critical information at the cancer district. Two calculations for mammogram measure reduction are proposed. These calculations can effectively diminish the extent of the mammogram images by half. For instance, a image that has a unique size of 21,217,280 bytes ends up 10,616,832 bytes with least processing time which is 12 seconds on 1.8GHz CPU processor and 512MRAM. The contracting calculation that is utilized as a pre-reduction process is produced and executed. The primary issue is to locate a reasonable coefficient that can be appropriate for all the pixel profundities at the image. [12] This coefficient ought to be effective to change over the 16-bit pixel profundity image to 8-bit. The point of this calculation is to locate a reasonable and effective coefficient that can change over the image from 16 to 8 bits while keeping up the vital restorative subtle elements. The algorithm can be described as follows

```

unsigned int max_level = max(Shrunk_Hist[]);
Long double divider = 0.0;
While (true)
{
divider += 0.01
if ((max_level/divider) <= 255)
break;
}
for (y=0; y<myimage_height; y++)
for (x=0; x<myimage_width; x++)
{
myimage[y][x] = myimage[y][x]/divider}
    
```

The initial phase in the calculation is to locate the most extreme level of utilized dim size of the contracted histogram. More often than not, this number for a 16-bit mammogram image is under 65536. In this way, the most extreme level is resolved for each image. At that point, the most extreme level is re-computed to be in the range from 250 to 255 dark levels. In any case, the genuine test is to discover the coefficient that would empower this. [4] Clamor Removing Image commotion is characterized as the arbitrary variety distinguished in brilliance or shading data in images created by therapeutic gadgets or scanners. Image commotion is for the most part seen as an unwanted result amid image procurement. Clamor is regularly characterized as the vulnerability in the flag because of arbitrary vacillations in signs. There are numerous foundations for these vacillations. Every single restorative image contain some visual clamor. The presence of commotion gives a image a mottled, grainy, finished, or frigid appearance. The pectoral muscle represents a noteworthy thickness district in most medio-sidelong diagonal (MLO) perspective of mammograms and can shape the result of image processing strategies. Power based techniques, can present poor execution when connected to separate thick structures, for example, the fibro-glandular circle or little far fetched masses,

since the pectoral muscle shows up as pretty much indistinguishable thickness from the thick tissues of enthusiasm for the image. To build the consistency of limit coordinating, the pectoral muscle can be expelled from the cancer district. Another critical need to recognize the pectoral muscle lies in the likelihood that the nearby data of its edge, alongside an inside investigation of its locale, might be utilized to distinguish the participation of anomalous axillaries lymph hubs, which might be the main sign of mysterious cancer carcinoma. A changed following calculation technique is connected to isolate the pectoral muscle district. The worldwide ideal in the histogram is select as the threshold esteem (around 90% of the histogram bend value). This is utilized to modify the info dark image to the twofold image. At that point by utilizing this paired image as veiling image unique is removed. Pixel estimation of first line read from left corner and point where pixel distinction is more prominent than 50 power esteems is noted. This procedure is proceeded till in which line there is no such sudden variety. Noted focuses are associated and pictorial muscle is chosen. In this way pectoral muscle is extricated and middle separating is connected.

V. EXPERIMENTAL RESULTS

Process of User Identification

Existing 1	Existing 2	Existing 3	Novel Threshold Based Pre Reduction Technique
13	22	7	35
18	27	15	39
21	35	20	44
24	44	22	56
30	51	26	60

Table 1: Comparison table of process user identification
Correlation table of process client recognizable proof depicts three existing strategies() and one proposed strategy. Contrasted with existing techniques the proposed strategies esteems are high. Proposed Novel Threshold Based Pre Reduction Technique strategy esteems begins from 35 to 60.

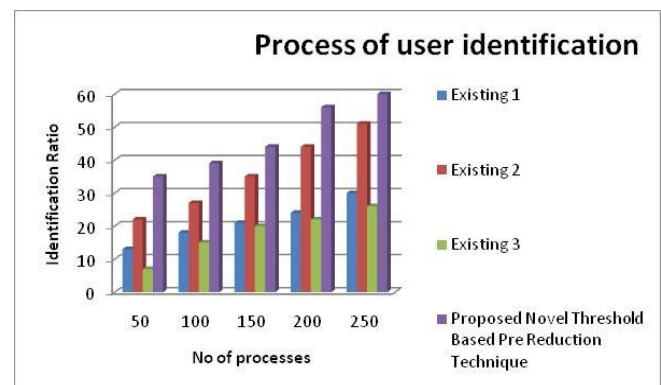


Figure 2: Comparison graph of process user identification

The Comparison outline of process client distinguishing proof demonstrates the diverse benefits of existing techniques and proposed strategy. No of records in x hub and succession level in Y hub.

At the point when analyzed existing strategy and proposed technique the proposed strategy esteems are higher than other. Proposed Novel Threshold Based Pre Reduction Technique esteems begin from 35 to 60.

Absolute Effectiveness

Table 2: Comparison table of absolute effectiveness

Existing 1	Existing 2	Existing 3	Novel Threshold Based Pre Reduction Technique
0.02	0.09	0.04	0.13
0.05	0.14	0.08	0.2
0.09	0.19	0.13	0.28
0.14	0.25	0.19	0.39
0.19	0.3	0.22	0.45

Correlation table of outright adequacy portrays three existing strategies() and one proposed strategy. Contrasted with existing techniques the proposed strategies esteems are high. Proposed Novel Threshold Based Pre Reduction Technique strategy esteems begins from 0.13 to 0.45.

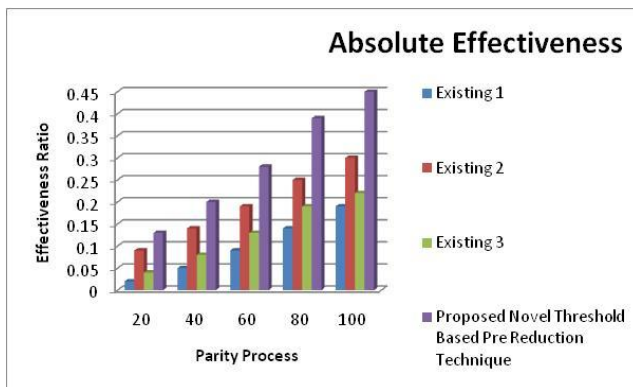


Figure 3: Comparison graph of absolute effectiveness

The Comparison chart of absolute effectiveness shows the different values of existing methods and proposed method. No of records in x axis and sequence level in Y axis. When compared existing method and proposed method the proposed method values are higher than other. Proposed method Novel Threshold Based Pre Reduction Technique values starts from 0.13 to 0.45.

Pre-Reduction Time

Table 3: Comparison table of pre-reduction time

Existin g 1	Existin g 2	Existin g 3	Novel Threshold Based Pre Reduction Technique
2.2	0.09	1.3	0.02
2.7	0.14	1.8	0.05
3.5	0.19	2.1	0.09
4.4	0.25	2.4	0.14
5.1	0.3	3	0.19

Examination table of pre-reduction time depicts three existing techniques() and one proposed strategy. Contrasted with existing techniques the proposed strategies esteems are low. Proposed Novel Threshold Based Pre Reduction Technique strategy esteems begins from 0.02 to 0.19.

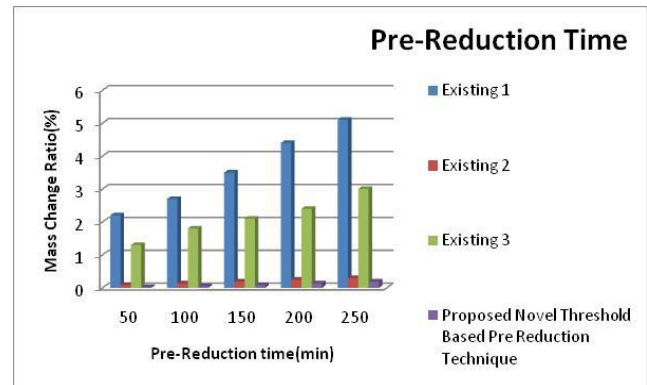


Figure 4: Comparison graph of pre-reduction time

The Comparison diagram of pre-reduction time demonstrates the distinctive benefits of existing techniques and proposed strategy. No of records in x pivot and succession level in Y hub. At the point when thought about existing strategy and proposed technique the proposed Novel Threshold Based Pre Reduction Technique strategy esteems are lower than other. Proposed strategy esteems begins from 0.02 to 0.19.

VI. CONCLUSION

A novel strategy called Threshold Based Pre-decrease Technique utilizing morphological change for the upgrade of mammogram images. This strategy takes out the clamor and also hones the outskirts of the items in the image. This upgrade strategy uses another semi delicate thresholding technique rather than traditional thresholding strategies. As of now there exist settled morphological tasks for wiping out the clamor for upgrading the images. In this chapter we made utilization of the morphological tasks alongside the wavelet change for better upgrade of the mammograms. So as to pick the best thresholding technique, distinctive delicate thresholding calculations presently accessible are assessed alongside certain change likewise made on a portion of the current strategies. Notwithstanding every one of these techniques, we additionally actualized a novel piece plane decay strategy for taking out the commotion substance of the image. The execution of the bit plane decay strategy isn't much huge as all other thresholding technique with the exception of best cap separating.

REFERENCES:

- Zheng, Y., "Breast Cancer Detection with Gabor Features from Digital Mammograms," Algorithms, Vol. 3, No. 1, 44-62, (2010).
- Tan, m., et al., "Assessment of a Four-View Mammographic Image Feature Based Fusion Model to Predict Near-Term Breast Cancer Risk," Ann Biomed Eng. (2015).
- Rangayyan, R. et al., "Computer-aided detection of architectural distortion in prior mammograms of interval cancer," J Digit Imaging, 23(5):611-31 (2010).
- Soriano, D., Aguilar, C., Ramirez-Morales, I., Tusa, E., Rivas, W., Pinta, M., "Mammogram Classification Schemes by Using Convolutional Neural Networks," CITT 2017, Communications in Computer and Information Science, vol. 798, pp 71-85, Springer, Cham (2017)
- Jadoon, M. Mohsin et al., "Three-Class Mammogram Classification Based on Descriptive CNN Features," BioMed Research International 2017: 3640901 (2017).
- Jiang, F., Liu, H., Yu, S., Xie, Y., "Breast mass lesion classification in mammograms by transfer learning," ICBCB '17 Proceedings of the 5th International Conference on Bioinformatics and Computational Biology, Pages 59-62, Hong Kong, 2017.



7. Xie, S., Shan, S. Chen, X., and Chen, J., "Fusing Local Patterns of Gabor Magnitude and Phase for Face Recognition," IEEE Transactions on Image Processing, 19(5), pp.1349-1361 (2010).
8. Krizhevsky, A., Sutskever, I., Hinton, G.E., "Imagenet classification with deep convolutional neural networks", Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1, Pages 1097-1105, Lake Tahoe, Nevada, (2012).
9. Hubel and Wiesel & the Neural Basis of Visual Perception, <https://knowingneurons.com/2014/10/29/hubel-and-wiesel-the-neural-basis-of-visual-perception/> (2014).
10. Simonyan, K., Zisserman, A., "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv technical report, (2014).
11. Russakovsky, O., Deng, J., Su, H., et al., "ImageNet Large Scale Visual Recognition Challenge." International Journal of Computer Vision (IJCV). Vol. 115, Issue 3, pp. 211-252 (2015).
12. Girshick, R., Donahue, J., Darrell, T., and Malik, J., "Rich feature hierarchies for accurate object detection and semantic segmentation," In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 580-587 (2014).
13. Girshick, R., "Fast R-CNN," In Proceedings of the IEEE International Conference on Computer Vision, pages 1440-1448 (2015).
14. Ren, S., He, K., Girshick, R. and Sun, J., "Faster R-CNN: Towards real-time object detection with region proposal networks," In Advances in neural information processing systems, pages 91-99 (2015).
15. Zheng, Y., Yang M C., Merkulov, A., Bandari, M., "Early breast cancer detection with digital mammograms using Haar-like features and AdaBoost algorithm," SPIE Proceedings Vol. 9871, Sensing and Analysis Technologies for Biomedical and Cognitive Applications 2016, 98710D (2016).

AUTHOR PROFILE



Mrs.S.Thilagavathi, is an Assistant Professor of Computer Science in Sri Krishna Adithya College of Arts and Science. She completed her MCA.,M.Phil., (PhD). He has a teaching experience of over thirteen years. She has published three papers in Scopus Indexed Journal, four papers in UGC listed and other journals and also presented papers in both National and International conferences.



S. AHAMED JOHNSHA ALI, I have over 8 years 9 months of teaching experience in the Educational Institutions by handling different subjects like C, C++, Computer Graphics & Multimedia, Java Programming, Visual Programming, Data Mining & Warehousing, Operating System, Data Structures, Multimedia (Photoshop, Corel, and Flash etc), Web designing, HTML, SQL and MS office. I also worked as Corporate

Trainer for 9 months in the software industry ECIL – ECIT (Arwin Varsity S/W Development and Training).



S. Lakshmidevi, I have over 14 years 3 months of teaching experience in the Educational Institutions by handling different subjects like C, C++, computer Graphics & Multimedia, Java programming, Currently I am working as Assistant Professor in Sri Krishna Adithya College of Arts & Science, Coimbatore.



Mrs. K. Brindha, MCA MPhil, Asst. prof in Krishna Adithya CAS has 11 years of teaching experience and 4 years of industrial experience and a member of IAENG. She has published 10 journals in research areas of Data mining and Networking and guided 6 MPhil scholars.