INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881; IMPACT FACTOR :6.514(2021); IC VALUE:5.16; ISI VALUE:2.286
Peer Reviewed and Refereed Journal: VOLUME:10, ISSUE:3(5), March:2021
Online Copy Available: www.ijmer.in

EARLY-STAGE DETECTION AND IDENTIFICATION OF BREAST CANCER (MAMMOGRAPHY) USING EVALUATION OF ARTIFICIAL INTELLIGENCE

[1]Jameel Jeelani and [2]Prof. ShavetaBala
[1]M-tech Scholar UGI Chandigarh and [2]HOD,Departmentof ECE UGI
[1&2]Punjab Technical University

Abstract

Micro-calcifications (MCs) are small calcium sedimentation which regarded as a major radiographic sign of breast cancer. These Individual MCs are basically impalpable and are of different shape and size which makes it very difficult to diagnose in the initial stages. Detection Accuracy plays a major role in efficient diagnosis of MCs in mammogram. The availability of noise/outlier can lead to misclassification, thereby affecting the detection accuracy. These remains to be major issue in various existing technique for early detection of MCs for diagnosis of breast cancer. In this paper, Neuro-fuzzy classifier has been employed, to enhance the classification by employing Cellular Automata segmentation. In this article, the concentration has been given over basic and building block stages. The dataset has been refined up to the mark in filtration and segmentation stages to yield better throughput in later stages. The proposed methodology is implemented over MATLAB platform.

Keywords: Micro-Calicifications, Neuro-Fuzzy Classifier, Cellular Automata, MATLAB.

1. Introduction

There have been several determinations in evolving computer-aided diagnosis (CAD) software to support radiologists in inferring mammograms. The inspiration overdue initial versions of CAD ascended from the perceptive that - even if the detached performance of a CAD system was mediocre to expert humans - its use could still outcome in improved sensitivity when used as a “second look” tool, while early confirmation supported this claim [1].

Digital breast tomosynthesis (DBT) is often deliberated the new, better mammogram based on witnessed increases in specificity and breast cancer detection associated with digital mammography (DM)alone. However, maximum of the distributed studies about DBT, whether from eventual trials or observational studies,3-5 use data from first- or prevalent-round screening moderately than incident-round airing in which breast cancer detection and recall rates are expected to be lower [2].

Mammography technology has progressed considerably over development of digital breast tomosynthesis, an evolving pseudo-three-dimensional mammography technology, also referred to as 3D-mammography. Tomosynthesis has been studied in soon-to-be trials embedded within European population-based screening programs, and in retroactive studies conducted in North America all of which demonstrate that 3D-mammography enhances detection measures [3].

The mainstream of population breast cancer (BC) showing programs, such those in Europe, the UK and Australia, provide breast selection using double-reading (interpretation by two readers) of stock two-dimensional (2D) digital mammography. Though double-reading of broadcast mammography is not a worldwide practice, it was presented into controlled screening programs since it increases BC discovery by an estimated 5%-15% of the fraction of detected cancers relative to single-reading [4].

Substantiation on the detection ability of tomosynthesis for population breast cancer broadcast has grown speedily over recent years and has been the theme of graphic reviews and annotations discussing the advantages and downsides of tomosynthesis. Issued systematic journals on tomosynthesis transmission to date have had a slight scope, have not involved most of the currently available studies, and have not reflected jointly the outcomes of cancer detection and recall, the latter expressive a regular harm of screening given that the mainstream of recalled women are not diagnosed with breast cancer [5].

2. Literature survey

Williamet al. [6],presented an annotation-efficient deep learning method that 1) attained state-of-the-art enactment in mammogram classification,2) magnificently stretched to digital breast tomosynthesis (DBT; “3D mammography”), 3) detected cancers in clinically-negative former mammograms of cancer patients, 4) generalized well to a population with low screening rates, and 5) outclassed five-out-of-five full-time breast imaging experts by improving utter sensitivity by an ordinary of 14%. Their results verified promise towards software that could expand the accuracy of and access to screening mammography worldwide.

Emily F. et. al [7], The purposes of their study were to relate outcomes of breast cancer screening investigations with DM vs DBT and to evaluate whether these results vary by age and breast density. For breast cancers detected from selection, they also estimated cancer features according to TMIST descriptions. These data were also being employed by the Cancer Intervention and

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881; IMPACT FACTOR :6.514(2021); IC VALUE:5.16; ISI VALUE:2.286
Peer Reviewed and Refereed Journal: VOLUME:10, ISSUE:3(5), March:2021
Online Copy Available: www.ijmer.in

Surveillance Modeling Network to model and project population- level DBT selection examination outcomes and cost effectiveness. Among 96 269 women (mean [SD] patient age for all examinations, 55.9 [9.0] years), patient age was 56.4 (9.0) years for DM and 54.6 (8.9) years for DBT of 180 340 breast cancer screening examinations, 129 369 examinations (71.7%) used DM and 50 971 examinations (28.3%) used DBT. Screening examination with DBT (73 of 99 women [73.7%] was associated with the detection of smaller, more often node-negative, HER2–negative, offensive cancers compared with DM (276 of 422 women [65.4%]). Screening examination with DBT was also related with lower recall (odds ratio, 0.64; 95%CI, 0.57-0.72; P < .001) and higher cancer detection (odds ratio, 1.41; 95%CI, 1.05-1.89; P = .02) equated with DM for all age groups smooth when stratified by breast density.

Nehmatet. al [8],performed secondary analysis established on STORM-2 which prospectively related 3D mammography and 2D-mammography in progressive screen-readings. Asymptomatic women 49 years who appeared population-based selection (Trento, 2013–2015) were drafted. Contributors recalled at any screen-read from analogous double-reading arms underwent more testing and/or biopsy. Single reading of 3D-mammography, combined with acquired or synthetized 2D-mammograms, was compared to double-reading of 2D-mammograhy alone for screen-detection measures: number of detected BCs, cancer detection rate (CDR), number and percentage of false-positive recall (FPR). Paired binary data were associated using McNemar's test.

Marinovichet. al [9], A systematic review and random special effects meta-analysis were assumed. Electronic databases were explored for studies comparing tomosynthesis and 2D mammography in asymptomatic women who appeared population breast cancer screening and reporting cancer detection rate (CDR) and recall rate. All statistical tests were two-sided. Results were similar in sensitivity analyses without studies with overlying cohorts.

3-Motivation and proposal: In the literature review we discussed that several research works had been carried out to find the best and suitable approach for Breast cancer detection. In females breast cancer is about 10% among Asian and American countries. Our survey came up with a conclusion that for better detection we need to focus more upon the preprocessing techniques i.e noise mitigation and segmentation. In the world of medical science's Medical image processing forms back bone for detection of diseases like Cancer, Tumor, bone breakages and many other problems. For the detection several steps have to be followed. Several research works have been carried out to perform better at the classification end. Our motive is to enhance the classification by emphasizing fundamental steps like preprocessing, Segmentation and later on classification.

In the proposed plan Cellular automata-based segmentation has been as added innovation technique to improve the segmentation quality and classification. We have included 11+ feature extraction techniques to enhance performance of the classifier. The proposed plan flow diagram has been depicted below in the figure 1. Generally, Micro-calicifications (MCs) are small calcium sedimentation which appears as bright spots in a mammogram image. In women, clustered MCs are regarded as a major radiographic sign of breast cancer. These Individual MCs are basically impalpable and are of different shape and size which makes it very difficult to diagnose in the initial stages. Detection Accuracy plays a major role in efficient diagnosis of MCs in mammogram. The availability of noise/outlier can lead to misclassification, thereby affecting the detection accuracy. These remains to be major issue in various existing technique for early detection of MCs for diagnosis of breast cancer. The mammographic images are collected from below standard website, Database: http://peipa.essex.ac.uk/info/mias.html. From this, there are three types of mammogram breast images are considered Such as Benign, Malignant and Normal.

Database	Total	Training	Testing
Benign	50 images	40 images	10 images
Malignant	50 images	40 images	10 images
Normal	50 images	40 images	10 images
Total	**150 images**	**120 images**	**30 images**

Table 1: Dataset images for training and testing

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881; IMPACT FACTOR :6.514(2021); IC VALUE:5.16; ISI VALUE:2.286
Peer Reviewed and Refereed Journal: VOLUME:10, ISSUE:3(5), March:2021
Online Copy Available: www.ijmer.in

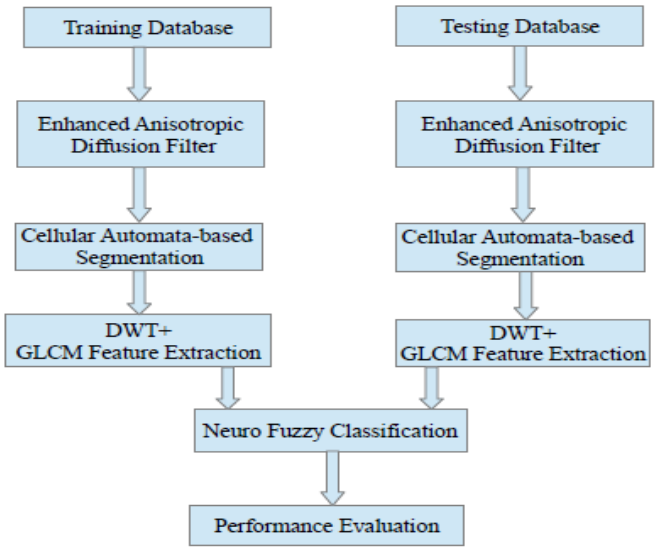


Figure 1: flow diagram of proposed implementation plan.

4-**Research Methodology**: - We have already discussed about the dataset collection in the proposal section of the same article.

4.1- Enhanced Anisotropic Diffusion Filter: - We offered a novel EADF framework with improved γ parameter approximation, based on both a planar region κ and on a set of the strongest edges of the image, denominated edge region ε. We also recommend a novel technique to delineate the planar region that is robust to partial volume effects. Edge detection is executed by means of a Canny proposed methodology in Canny with hysteresis of 80% and 90% of the accumulated histogram of the absolute values of the gradient. The optimal γ parameter is figured according to the desirable conventional or aggressive behavior, and updated at each iteration. The number of iterations $t_{\max}$ is also defined dynamically according to the current SNR, as the image is smoothed.

• To improve image quality by removing noise, here enhanced Anisotropic diffusion filter.
• Enhanced Anisotropic diffusion filter is attained by the combination of Anisotropic diffusion filter with adaptive histogram equalization.

Values considered,
• num_iter = 5; % Number of iterations
• delta_t = 1/10; % Integration constant
• kappa = 30; % gradient modulus threshold
• option = 2; % conduction coefficients

An anisotropic diffusion filter is given by:

$$I_t = div(c(x,y,t)\nabla I)$$
$$= \frac{\partial}{\partial x}(c(x,y,t)I_x) + \frac{\partial}{\partial y}(c(x,y,t)I_y) \quad (1)$$

Where, c is the conduction diffusion coefficient. The diffusion coefficient is considered locally as a function f of magnitude of gradient of brightness to perform region specific smoothing operation which is given by

$$c(x,y,t) = f(\|\nabla I(x,y,t)\|) \quad (2)$$

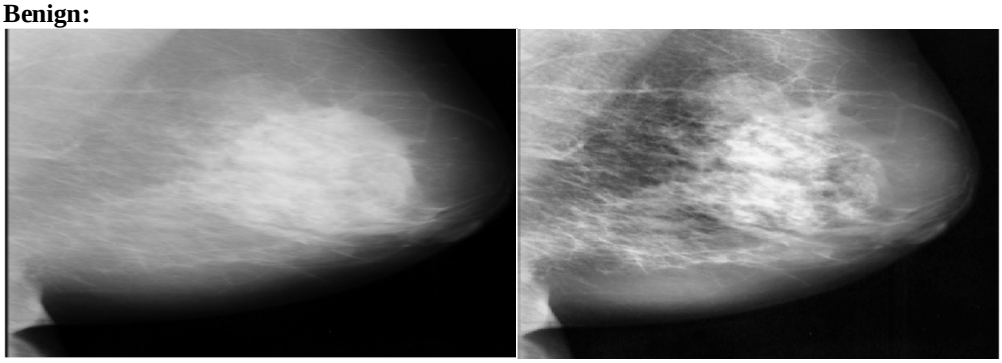
The edges are preserved by proper choice of function f. It also sharpens and brightens the edges. The function f result in low coefficient value at inter-region edges by having large gradient strength and it should have high coefficient value in non-edge regions having low gradient strength. The possible function to preserves sharp edge is

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881; IMPACT FACTOR :6.514(2021); IC VALUE:5.16; ISI VALUE:2.286
Peer Reviewed and Refereed Journal: VOLUME:10, ISSUE:3(5), March:2021
Online Copy Available: www.ijmer.in

$$f(\|\nabla I(x,y,t)\|) = \exp\left(\frac{\|\nabla\|}{k}\right)^2 \quad (3)$$

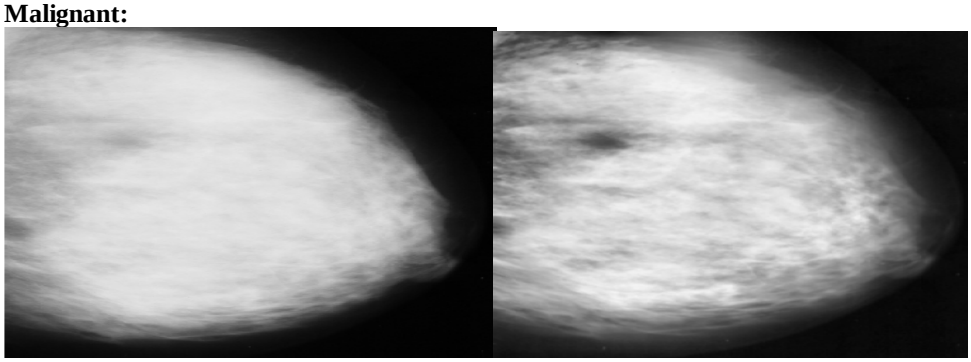
Where k is the edge strength threshold.

Benign:



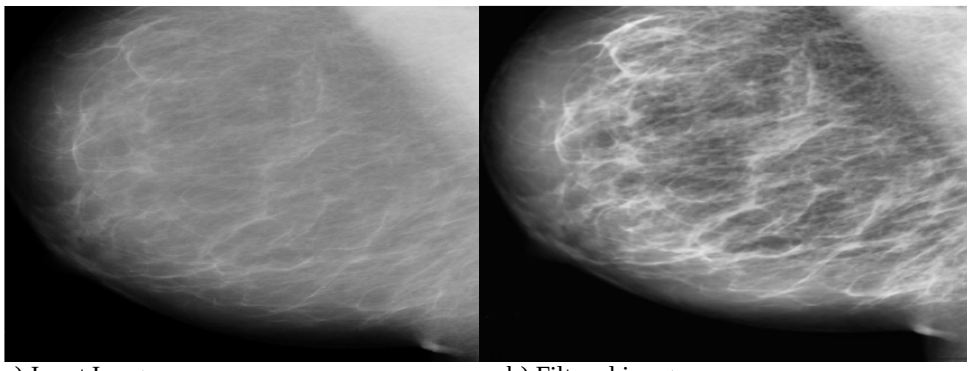
a) Input Image b) Filtered image

Malignant:



a) Input image b) Filtered Image

Normal:



a) Input Image b) Filtered image

Figure2: input and filtered images using Enhanced Anisotropic filter for all 3cases.

4.2-Cellular Automata-based Segmentation:Image segmentation is the heart of any vision system and is a significant step in the image investigation process. Its objective is to source a description of objects confined in the image by the taking out of various visual indications such as the objects edges, homogeneous regions, 3D objects. Then, they will be run for a symbolic explanation of the scene permitting interpretation and possibly decision-making. Until now, there is no universal technique for image segmentation. Any technique is effective only for a given image type, for a given application type, and in a given context. Because of these constraints, the miscellaneous segmentation strategies which were proposed declared their incapacities and their limitations, thus it is necessary to investigate new horizons and to find new more flexible and more effective methods.

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881; IMPACT FACTOR :6.514(2021); IC VALUE:5.16; ISI VALUE:2.286
Peer Reviewed and Refereed Journal: VOLUME:10, ISSUE:3(5), March:2021
Online Copy Available: www.ijmer.in

A cellular automaton is a consistent grid of cells. Each cell has a state chosen among a finite set of states and which can grow in time. The state of a cell at time $t+1$ depends on the state at time t of a limited number of cells called its neighborhood. In every unit of time, the same rules are simultaneously applied to all cells of the grid, producing a new generation of cells depending completely on the previous generation. The cellular automaton is characterized by:
-Dimensions of the grid. Example: 1D, 2D, 3D …
-Number of states or colors. Example: ON or OFF, 256 levels of grey …
-Neighborhood considered by a cell. Examples: Von Neumann: 4 cells in 2D, Moore: 8 cells in 2D …
-Rules defining the automaton and the progress of the generations. Example: Game of Life …

The simplest non-coarse cellular automaton that we can design consists of a 1D grid of cells. Each cell state can be one of only two states: 0 or 1. For every cell, neighborhood is constituted from the cell itself and the two cells which are adjacent. Each cell of the 3 cells can take two states, so there are $2^3 = 8$ possible configurations of such a neighborhood. For each of these configurations, there are $2^8 = 256$ different manners to do that, thus there are 256 different cellular automata from this type. Often, we indicate automata of this family by an integer between 0 and 255 the binary representation of which is the continuation of states taken by the automaton on the successive motives 111, 110, 101, etc.

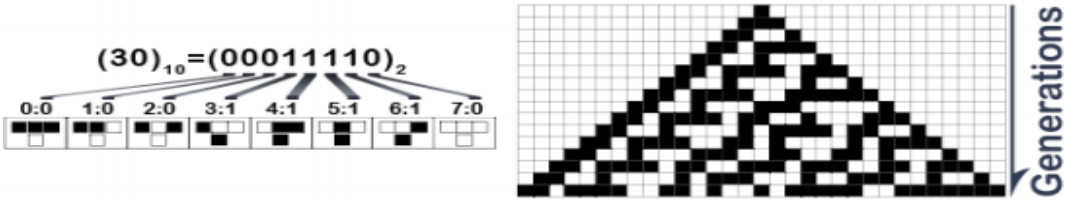
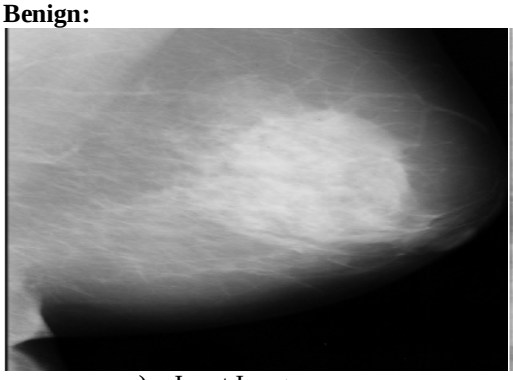


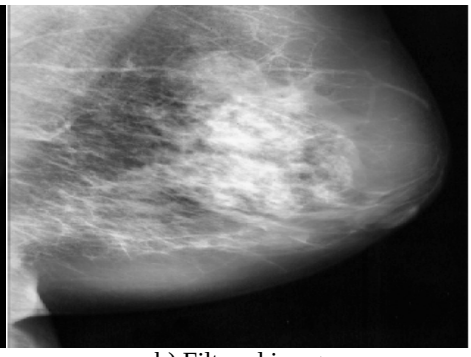
Figure 3:Cellular automata generated by rule n°30

• To segment micro-calcifications from the image, cellular automata-based segmentation is used.
• In this, window with >1 and <8 cells are considered to transform a central pixel to "alive" regardless of its prior state.

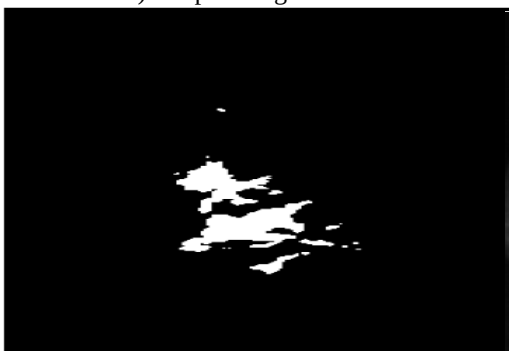
Benign:



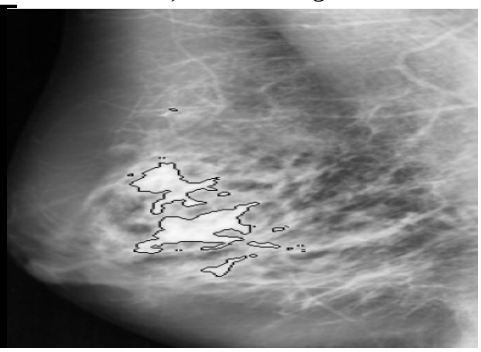
a) Input Image



b) Filtered image



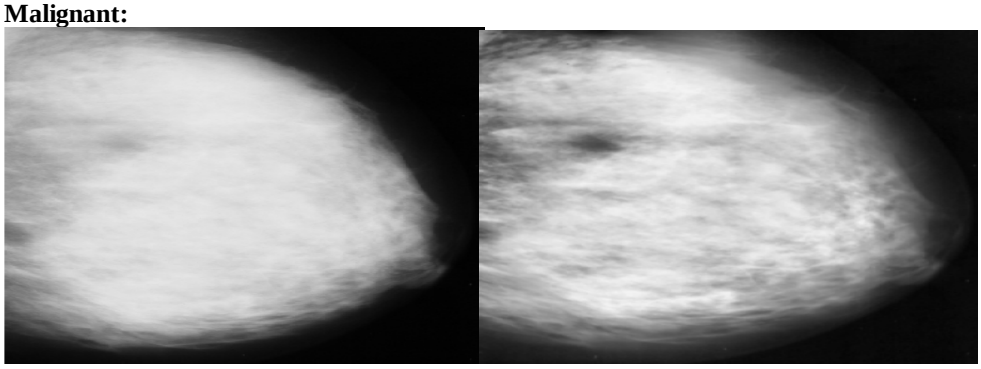
b) Segmented Image



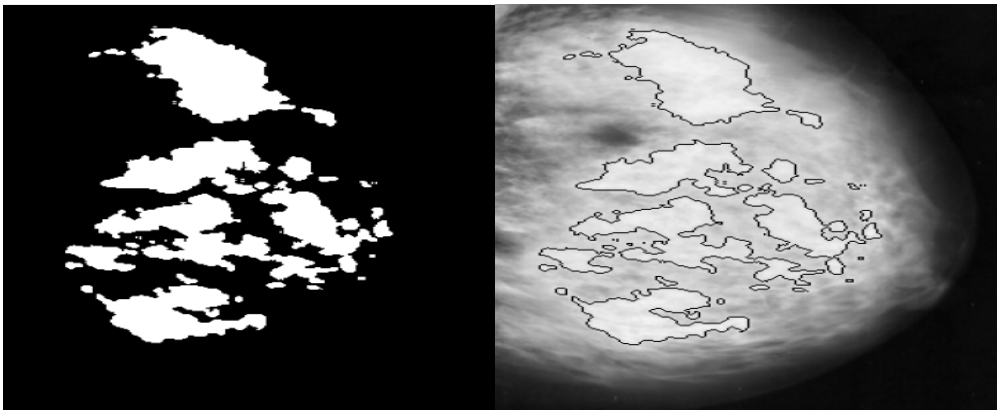
d) Outlined Image

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881; IMPACT FACTOR :6.514(2021); IC VALUE:5.16; ISI VALUE:2.286
Peer Reviewed and Refereed Journal: VOLUME:10, ISSUE:3(5), March:2021
Online Copy Available: www.ijmer.in

Malignant:

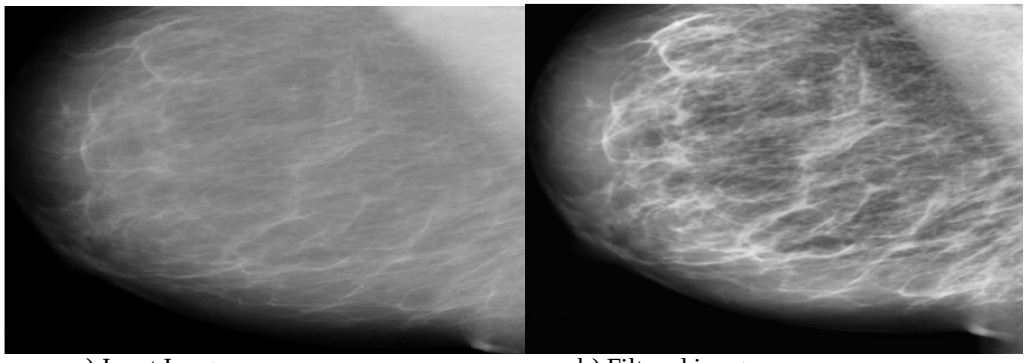


a) Input image b) Filtered Image

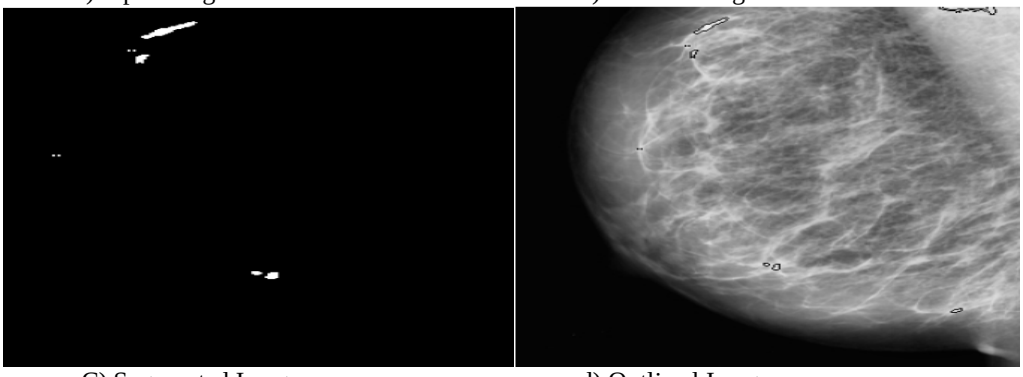


c) Segmented Image d) outlined image

Normal:



a) Input Image b) Filtered image



C) Segmented Image d) Outlined Image

Figure 4- cellular automata segmentation with outlining

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881; IMPACT FACTOR :6.514(2021); IC VALUE:5.16; ISI VALUE:2.286
Peer Reviewed and Refereed Journal: VOLUME:10, ISSUE:3(5), March:2021
Online Copy Available: www.ijmer.in

4.3-Feature extraction

In this process two steps are involved as under.

1. Apply 3-level Discrete Wavelet Transform (DWT) using "db4 wavelet".
2. Then, apply Grey-Level Co-occurance Matrix (GLCM) to extract following statistical features.

Feature Extraction is the process of reducing the size of image data by obtaining necessary information from the segmented image. The visual content of a segmented image can be captured using this process. From the extracted features it is possible to demarcate between normal and abnormal image samples. The reliability of the classification algorithm depends on segmentation method and extracted features. In this work texture features are extracted using Gray Level Co-occurrence Matrix (GLCM) and shape features are extracted using connected regions.

• Contrast • Correlation • Energy • Homogeneity • Mean • Standard Deviation • Entropy • Smoothness • Kurtosis • Skewness

4.4-Neuro-fuzzy classifier: A Neuro fuzzy system is an organization of neural network and fuzzy systems. All the structures are not equally important in indiscriminating all the classes, but thefeature wise belongingness helps in the classification process. The LH-NFC process consists of five phases as discussed below.

1. In the first phase, the input values are fuzzified using Gaussian membership function. The obtained Gaussian membership values provide the membership values for each feature to all the classes. The rows I the output membership matrix represent number of features and the columns represent number of classes.
2. The membership values are redefined using LH.
3. The membership matrix is converted into a vector.The membership vector values are fed to the Neural Network model. The number of nodes created is equal to the product of the features and the number of outputs.
4. Weights are calculated based on fuzzy rules to identify the class.
5. Defuzzification is performed to obtain the crisp output by applying weighted average method.

4.4.1- Membership function plot:

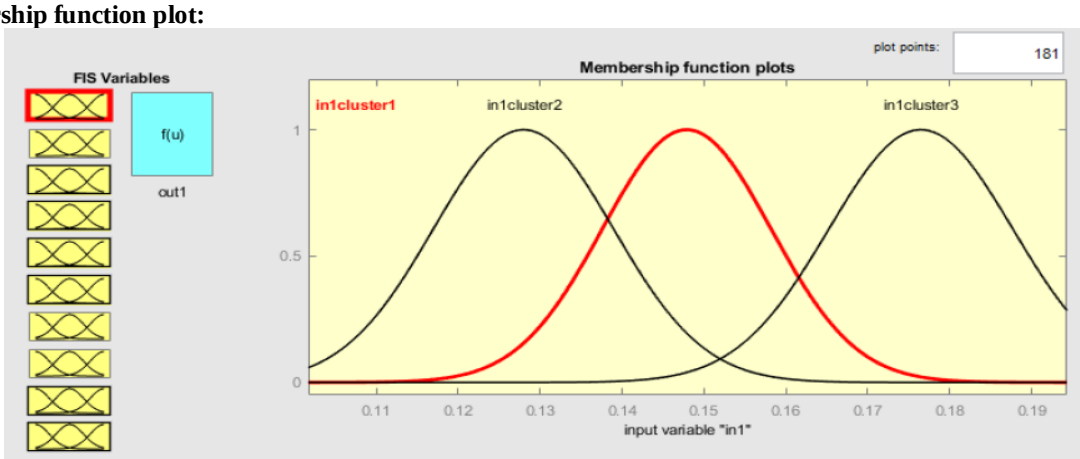


Figure 5: plot for membership function for 10 features.

Fuzzy logic:

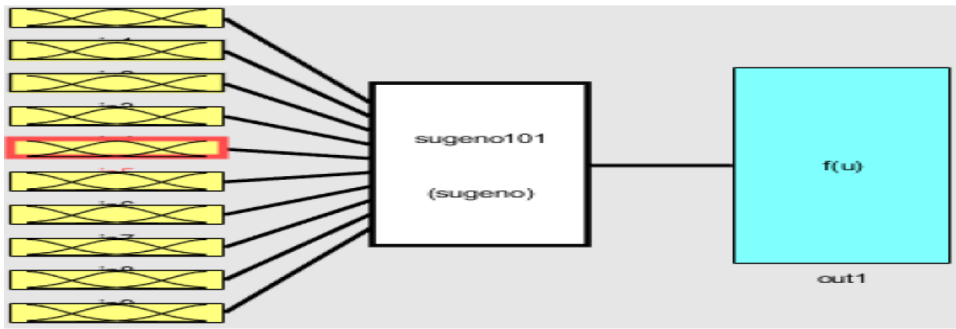


Figure 6: - Fuzzy logic design

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881; IMPACT FACTOR :6.514(2021); IC VALUE:5.16; ISI VALUE:2.286
Peer Reviewed and Refereed Journal: VOLUME:10, ISSUE:3(5), March:2021
Online Copy Available: www.ijmer.in

Fuzzy rule:

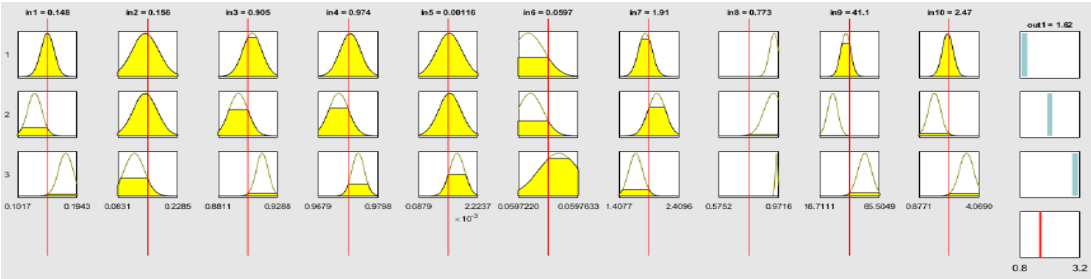


Figure 7:- 10 feature fuzzy rules

ANFIS Model:

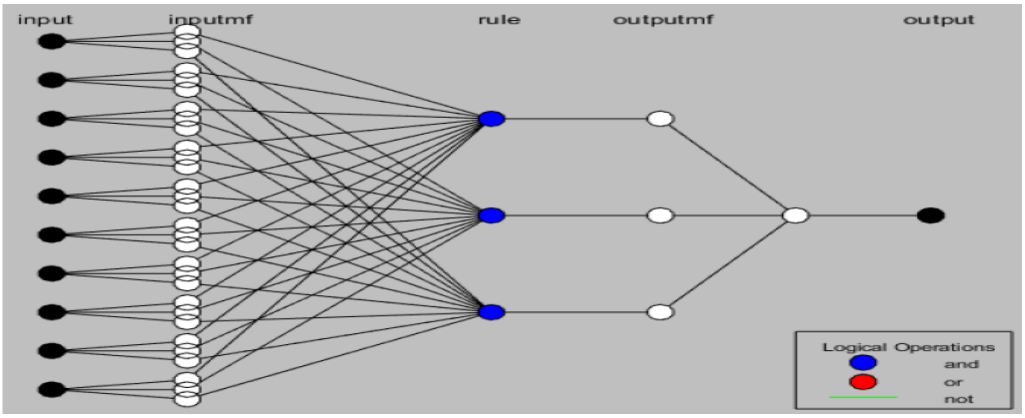
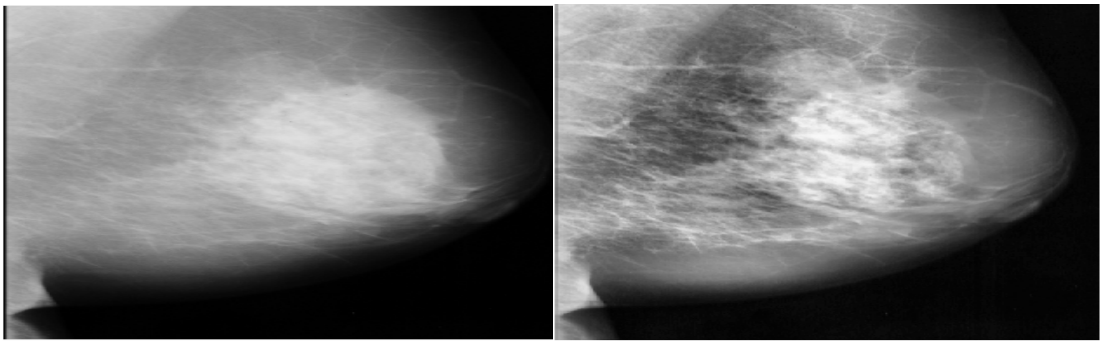


Figure 8: ANFIS Model

• ANFIS model or structure depicts the 10 input features (input layer) is applied input membership function.
• The output of input membership function is matched with 3 fuzzy rules(Hidden layers), then the decision is applied to output membership function.
• Finally, the output membership function classifies the type of category.

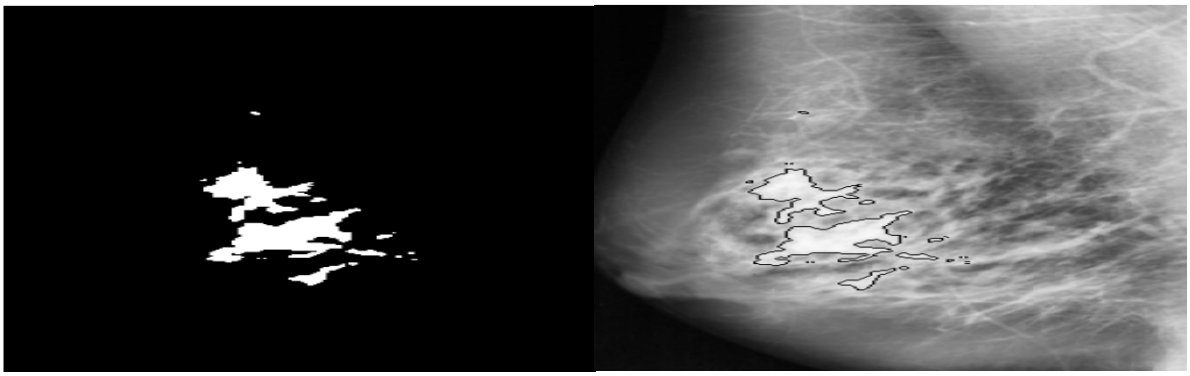
5-Results and discussion

5.1-Type 1-Benign



a) Input Image b) Filtered image

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881; IMPACT FACTOR :6.514(2021); IC VALUE:5.16; ISI VALUE:2.286
Peer Reviewed and Refereed Journal: VOLUME:10, ISSUE:3(5), March:2021
Online Copy Available: www.ijmer.in



c) Segmented Image d) Outlined Image

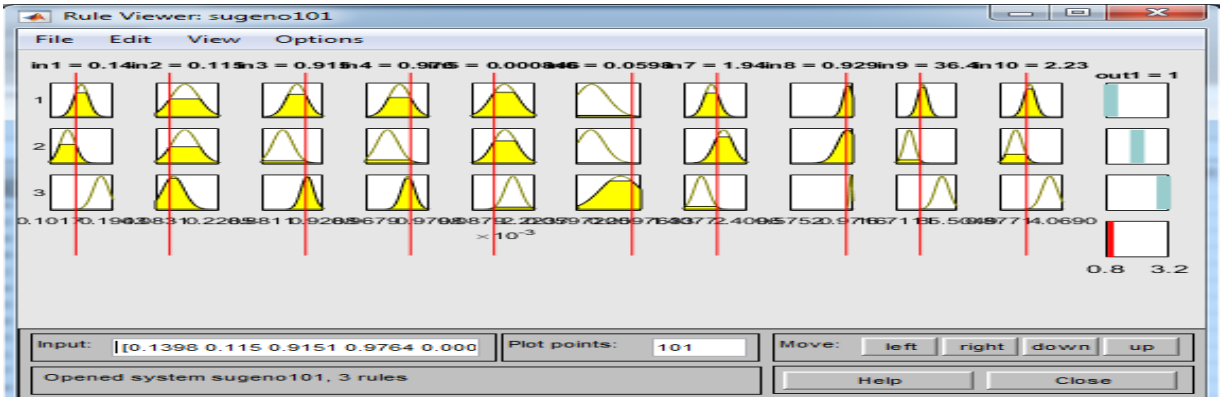
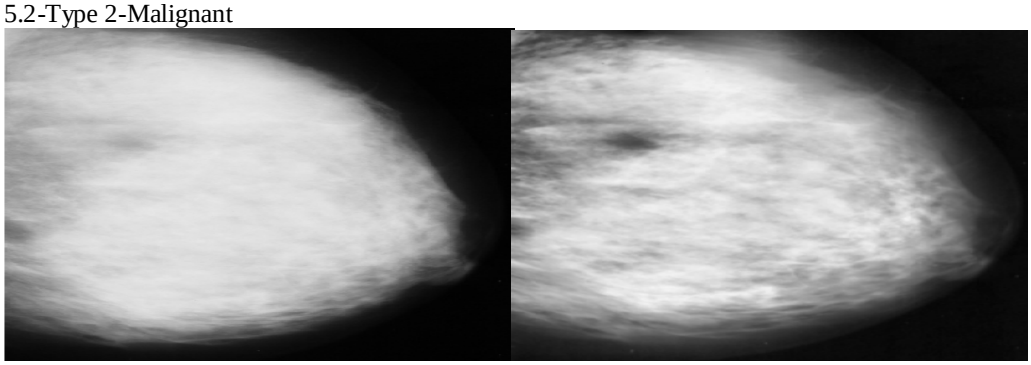
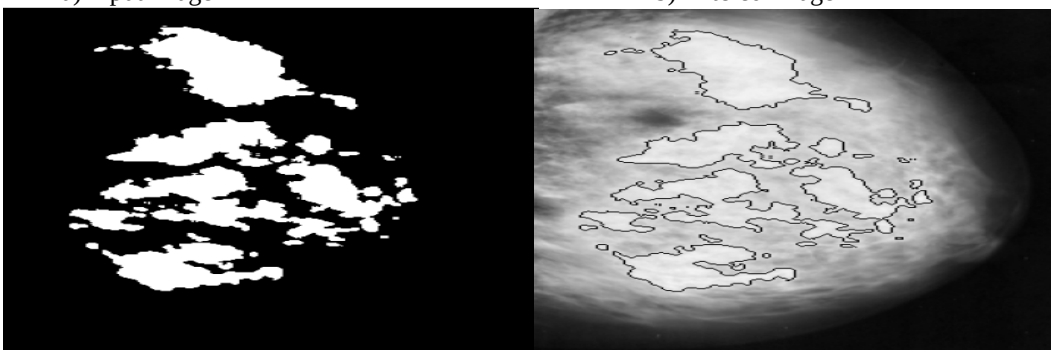


Figure 9:classification of type I cancer detection

5.2-Type 2-Malignant



a) Input image b) Filtered Image



c) Segmented Image d) outlined image

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881; IMPACT FACTOR :6.514(2021); IC VALUE:5.16; ISI VALUE:2.286
Peer Reviewed and Refereed Journal: VOLUME:10, ISSUE:3(5), March:2021
Online Copy Available: www.ijmer.in

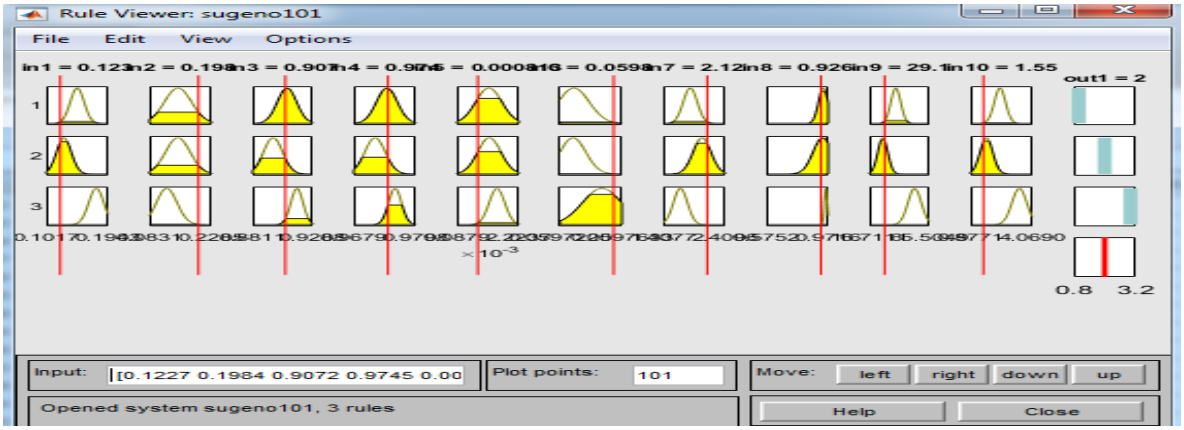
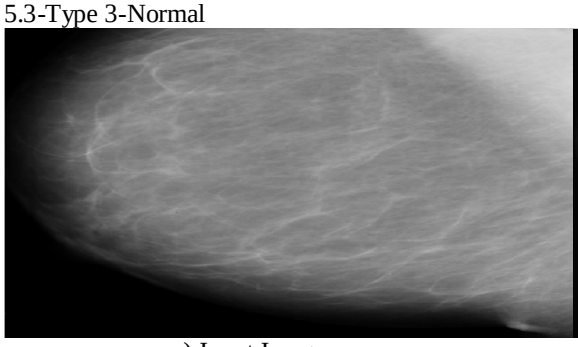
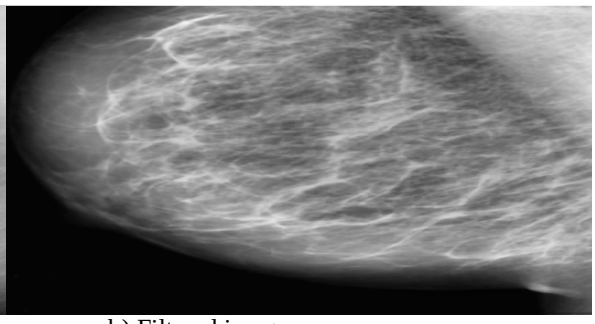


Figure 10: - classification of type II cancer

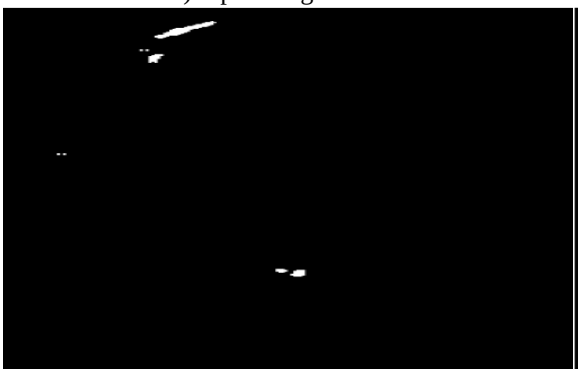
5.3-Type 3-Normal



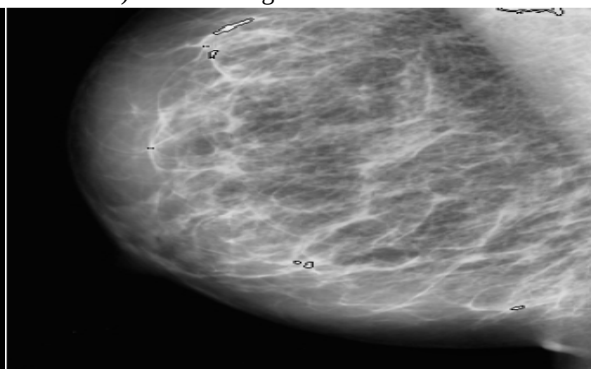
a) Input Image



b) Filtered image



C) Segmented Image



d) Outlined Image

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881; IMPACT FACTOR :6.514(2021); IC VALUE:5.16; ISI VALUE:2.286
Peer Reviewed and Refereed Journal: VOLUME:10, ISSUE:3(5), March:2021
Online Copy Available: www.ijmer.in

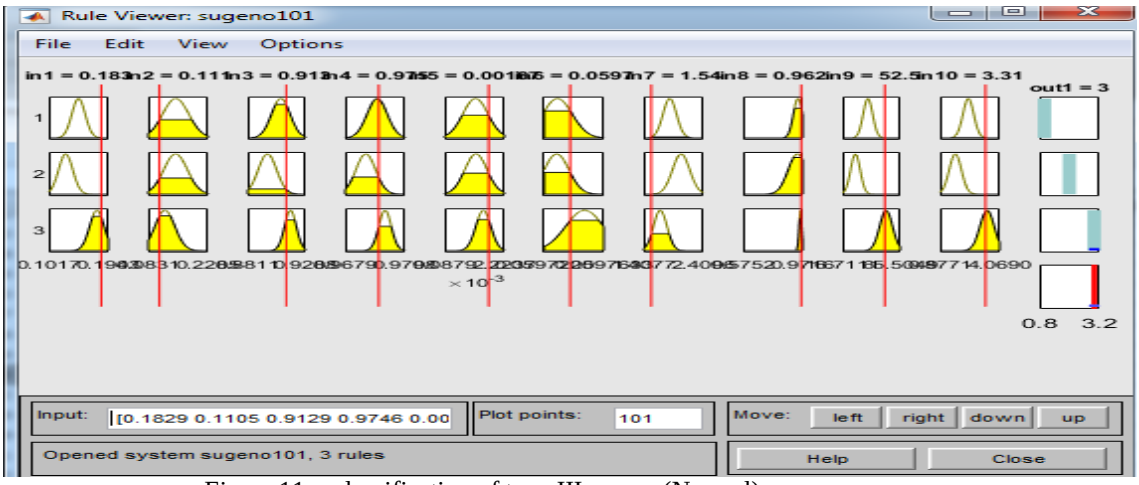
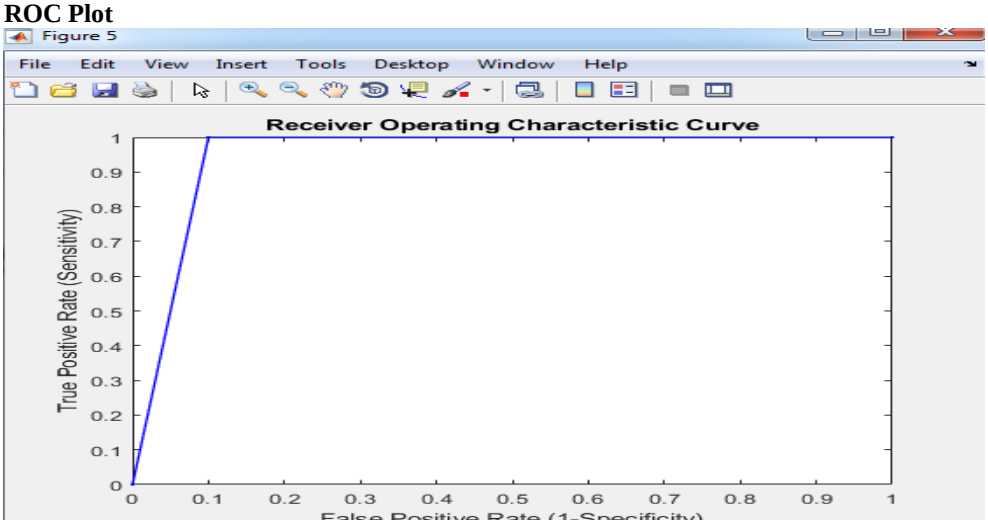


Figure 11: - classification of type III cancer (Normal)

Classifier Neuro-fuzzy has performed better in detection of infected and non-infected images. We assume that our segmentation and filtration techniques have performed very well, due to which we got better classification results from our Neuro-fuzzy classifier. More over our performance evaluation will project better depiction of classifier performance.

Performance evaluation
Accuracy –96.66
Sensitivity –90
Specificity- 100
Execution time-0.86 sec

ROC Plot



Graph 1: Roc plot

6-conclusion: In this article, we have performed below operations.

- Enhanced anisotropic diffusion filtering technique, where we have made fine tuning of parameters to get desired filtered output, free and several noises and distortions to make final throughput flawless.
- Cellular automata segmentation technique, where we formed a segmentation window between 1-8 for suitable segmentation. Suitable segmentation refers to level of segmentation, in which classifier can find sensitive difference while feature extraction.
- Neuro-fuzzy or ANFIS classifier has been employed for classification.

We got better throughput using above implementation plan, the implementation has been done over MATLAB platform. The work can be taken to next level by employing same plan to live project.

References

1. William Lotter, Abdul RahmanDiab, Bryan Haslam, Jiye G. Kim, GiorgiaGrisot, Eric Wu,Kevin Wu, Jorge OnievaOnieva, Jerrold L. Boxerman, Meiyun Wang, Mack Bandler, GopalVijayaraghavan, and A. Gregory Sorensen, "Robust breast cancer detection in mammography and digital breast tomosynthesis using annotation-efficientdeep learning approach", arXiv:1912.11027v2 [eess.IV] 27 Dec 2019.
2. Emily F. Conant,William E. Barlow, Sally D. Herschorn, Donald L.Weaver, Elisabeth F. Beaber, Anna N. A. Tosteson, Jennifer S. Haas, Kathryn P. Lowry, Natasha K. Stout, Amy Trentham-Dietz, Roberta M. diFlorio-Alexander,Christopher I. Li, Mitchell D. Schnall, Tracy Onega, Brian L. Sprague, "Association of Digital Breast Tomosynthesis's DigitalMammography With Cancer Detection and Recall Rates by Age and Breast Density", 2019 American Medical Association.
3. NehmatHoussamia, Daniela Bernardib, Marco Pellegrini, MarviValentini,CarmineFantò, LivioOstillio, PaolinaTuttobene, Andrea Lupariab, Petra Macaskill, "Breast cancer detection using single-reading of breast tomosynthesis(3D-mammography) compared to double-reading of2D-mammography: Evidence from a population-based trial", elsevier, Cancer Epidemiology 47 (2017) 94–99.
4. NehmatHoussamia, Daniela Bernardib, Marco Pellegrini, MarviValentini,Carmine Fantò, LivioOstillio, PaolinaTuttobene, Andrea Lupariab, Petra Macaskill, "Breast cancer detection using single-reading of breast tomosynthesis(3D-mammography) compared to double-reading of2D-mammography: Evidence from a population-based trial", elsevier, Cancer Epidemiology 47 (2017) 94–99.
5. M. Luke Marinovich, Kylie E. Hunter, Petra Macaskill, NehmatHoussami, "Breast Cancer Screening Using Tomosynthesis orMammography: A Meta-analysis of Cancer Detection andRecall", JNCI J Natl Cancer Inst (2018) 110(9): djy121.
6. William Lotter, Abdul RahmanDiab, Bryan Haslam, Jiye G. Kim, GiorgiaGrisot, Eric Wu,Kevin Wu, Jorge OnievaOnieva, Jerrold L. Boxerman, Meiyun Wang, Mack Bandler, GopalVijayaraghavan, and A. Gregory Sorensen, "Robust breast cancer detection in mammography and digital breast tomosynthesis using annotation-efficientdeep learning approach", arXiv:1912.11027v2 [eess.IV] 27 Dec 2019.
7. Emily F. Conant,William E. Barlow, Sally D. Herschorn, Donald L.Weaver, Elisabeth F. Beaber, Anna N. A. Tosteson, Jennifer S. Haas, Kathryn P. Lowry, Natasha K. Stout, Amy Trentham-Dietz, Roberta M. diFlorio-Alexander,Christopher I. Li, Mitchell D. Schnall, Tracy Onega, Brian L. Sprague, "Association of Digital Breast Tomosynthesisvs DigitalMammography With Cancer Detection and Recall Ratesby Age and Breast Density", 2019 American Medical Association.
8. NehmatHoussamia, Daniela Bernardib, Marco Pellegrini, MarviValentini,Carmine Fantò, LivioOstillio, PaolinaTuttobene, Andrea Lupariab, Petra Macaskill, "Breast cancer detection using single-reading of breast tomosynthesis(3D-mammography) compared to double-reading of2D-mammography: Evidence from a population-based trial", elsevier, Cancer Epidemiology 47 (2017) 94–99.
9. NehmatHoussamia, Daniela Bernardib, Marco Pellegrini, MarviValentini,Carmine Fantò, LivioOstillio, PaolinaTuttobene, Andrea Lupariab, Petra Macaskill, "Breast cancer detection using single-reading of breast tomosynthesis(3D-mammography) compared to double-reading of2D-mammography: Evidence from a population-based trial", elsevier, Cancer Epidemiology 47 (2017) 94–99.

Filename:	12
Directory:	C:\Users\DELL\Documents
Template:	C:\Users\DELL\AppData\Roaming\Microsoft\Templates\Normal.dotm
Title:	
Subject:	
Author:	Windows User
Keywords:	
Comments:	
Creation Date:	12/21/2020 11:11:00 AM
Change Number:	18
Last Saved On:	3/24/2021 12:44:00 AM
Last Saved By:	Murali Korada
Total Editing Time:	105 Minutes
Last Printed On:	4/8/2021 8:22:00 PM

As of Last Complete Printing

Number of Pages:	12
Number of Words:	3,524 (approx.)
Number of Characters:	20,092 (approx.)