ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 9, Issue 3, May 2020

Manifold Learning for Breast Cancer Detection using ANFIS

Sarita Chauhan
Assistant Professor, ECE, MLVTEC, Bhilwara, Rajasthan

Abstract: ***Breast cancer is very common and is considered the second most dangerous disease in the world because of its mortality rate. Affected will survive if the doctor detects significant physical changes in the doctor before it occurs . Mammographic (Breast region X-ray) images are commonly used for premature breast discovery. The goal of the proposed program is to develop a method used to distinguish between benign (non-cancerous) and malignant (cancerous) mammograms to help radiologists improve the accuracy of their diagnoses. In the proposed method, texture characteristics were determined from mammograms Use Gray Level Co-occurrence Matrix (GLCM) at 0 °, the most effective features from the calculation. A major contribution to the target performance was chosen and applied to the Artificial Neural Network (ANN) for classification, and the overall sensitivity, specificity. By using the proposed method, the accuracy obtained is 99.3%, 100% and 99.4% respectively.***

***Keywords*: Breast Cancer, entropy, correlation, Neural, network.**

I. INTRODUCTION

Breast cancer research has been tremendous over the last decade. Breast cancer is the second most deadly disease among women in the world[1–4], rising with age. Breast cancer affects not just men and animals but also women. Just 1 per cent is found in men in all situations. The two types of cancerous breast problems are nefarious and benign. Often 10% – 30% of breast cancer lesions are overlooked because of inaccurate human observers. For the patient's prompt recovery, early and accurate diagnosis is required. Identifying at-risk women is an important strategy to reduce the number of women with breast cancer shown in Figure 1. Detecting the possibility of cancer recurrence can save a patient's life. Traditionally, biopsy has been used to diagnose mammography, breast MRI, ultrasonography, BRCA testing, etc. When multiple tests are performed on a patient, it becomes difficult for medical experts to reach a correct conclusion and the methods of investigation produce false positive results. Smart systems are therefore needed to reduce instances of false positives and false negatives. This paper uses soft computing techniques to diagnose breast cancer.

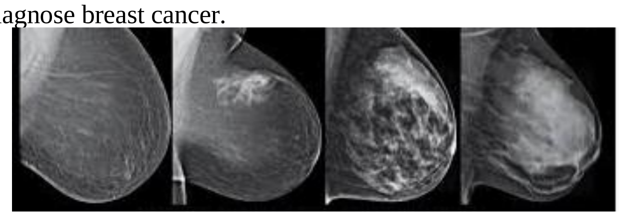


Fig. 1: MRI of breast of affected patients [6]

II. RELATED WORK

Hamza Turabeeh [4] focused on breast cancer recurrence problems to develop a good diagnostic system, hybridizing two methodologies, the Genetic Algorithm (GA) and the Adaptive Neuro Fuzzy Infection System (ANFIS) using classification accuracy , sensitivity and specificity.

Radhanath Patra and Shankh Mitra Sunani [5] reviewed the diagnosis of breast cancer database, and the following few points were strongly emphasized. Some of the review papers have already predicted the accuracy of various machine learning algorithms. Due to the limitation of ANN, various modified forms of machine learning algorithms as well as hybrid processes have been adopted to improve accuracy over the shortest period of time. Various data mining processes, such as decision tree, KNearest Neighbor algorithm, and particle swarm optimization, also considered predicting high-precision breast cancer diagnosis.

Rajamani. R and Rathika. M[6] provided an overview of the analysis of liver cancer using the Adaptive Neuro Fuzzy Infection System (ANFIS) data mining technique. The noise in the computed tomography image, the segmentation process, the morphological operation and the extraction feature techniques were removed in the data pre-processing step. Vibhav Prakash Singh and Rajeev Srivastava [9] used an effective content based mammogram retrieval system. Pre-processing segmentation of the image, selective thresholds based on the growing seed region algorithm have been introduced. The 2-level discrete wavelet transformation (DWT) was applied to the segmented region and the wavelet based center symmetric-local binary pattern (WCS-LBP) feat.

Mugahed A. Al-antari et. al [10] presented a DBN-based CAD system to classify Mammogram Image Retrieval using the IPSO Optimized Anfis Classifier Sonia Jenifer et. Al[11] presented a new method for classifying breast cancer using time series analysis. Birmohan Singh and Manpreet Kaur[12] have developed a methodology for the differentiation of micro-calcification clusters. The performance of shape-based and textured features has been analyzed. It's Ardalan et. Al [12] has presented a method for classifying and detecting masses in mammograms. They performed a tenfold cross-validation and analyzed the complexity of the data on each fold.

III. MATERIAL AND METHODS

ANFIS is used as a predictive model to learn from a set of experimental data that is usually structured to a specific domain issue. Usually, this domain has input parameters and rules that are derived from experimental data. Previous parameters (ANFIS input) are obtained from the experimental data input fuzzified with the appropriate membership function. The resulting parameters are repeated by adjusting the antecedent parameters [7]. The adjustment of the premise parameters and the corresponding parameters are recorded as an epoch.

Layer 1 (input layer) accepts and transits the breast cancer parameters linked to their related values to the next successive layer. Layer 2 (member function) maps specific membership functions for each linguistic input variables. Because of the tendency for huge data set the ANFIS used membership feature. Layer 3 (Rule layer) receives a fuzzified range value from the input layer and calculates the firing strength of each rule. Layer 4 (Normalization layer) calculates the normalized firing intensity of each input signal coming from the law layer. Equation 4 represents the intensity of the firing normalization. Layer 5 (Defuzzification layer) calculates the normalized output parameter function that shapes the data [8]. This overall output is mapped to one of the two value “benign” or “malignant. The ANFIS and Neural network was simulated using MATLAB taking cognizant of numerical data integration and open source capability of MATLAB.

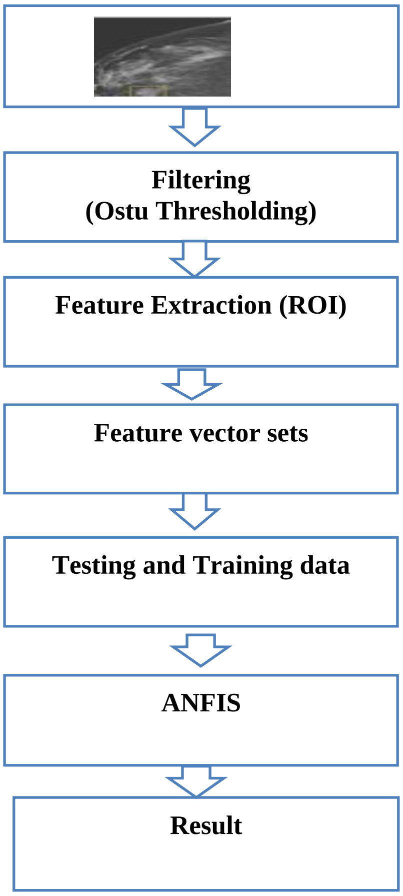


Fig. 2: Flowchart of the proposed work

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 9, Issue 3, May 2020

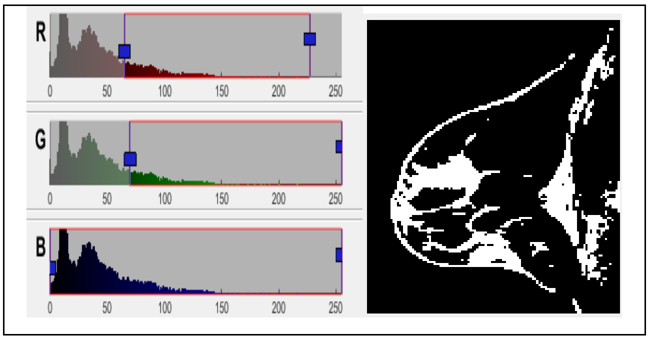
The flow of the method is shown in Figure 2.It starts with the acquisition part from e- database .The thresholding method is used for filtering purpose .The extracted region is used for Feature extraction which is carried by GLCM matrix Then set of vector sets are produced with 6 features which are used for testing and training sets. The hybrid algorithm ANFIS is used to produce confusion matrix and accuracy.

IV. RESULT ANALYSIS

The overall system consists of 4 stages, the first one is image acquisition, and Second mammogram extraction with Optimum features, and classifier to classify appropriate mammographic class with hybrid Fuzzy Inference System (FIS) and ANN .The parts of the mammogram were extracted using the Layer attributes. This data set contains 612 mammograms [6] for this experiment, 470 images are normal (Non cancerous) and 142 (cancerous) images are malignant. The picture is 1024 x 1024 pixels in this folder. Figure 3 shows histogram and binary form of standard sample images and malignant rrespectively.

The experimentation of the entire processing was performed with dataset. This dataset contains the information about benign and malignant tumours [9]. The architecture of the proposed method is figured as follows Fuzzy logic was then employed to capture the broad categories identified during clustering into a Fuzzy Inference System (FIS) and ANN. Texture features are extracted using GLCM [14] along 0 for each texture image. Features represent image in a particular format which on relevant information. Features are in the next stage chosen for training and testing; this is an important stage.

Genfis2 is the function that creates a FIS using subtractive clustering. Genfis2 employs subclust behind the scenes to cluster the data and it uses the cluster centers and their range of influences to build a FIS. The FIS acts as a model that reflects the relationship between the input variables and output variables. The FIS is composed of inputs, outputs and rules. Each input in the FIS represents an input variable in the input dataset and each output in the FIS represents an output variable in the output dataset [10]. Therefore each input and output was characterized by eight membership functions. The membership function type is a Gaussian type membership function and the parameters of the membership function are [0.0051 0.2196] shown in Fig 3.The figure 4 to 8 shows the classification of data into benign and malignant [13].



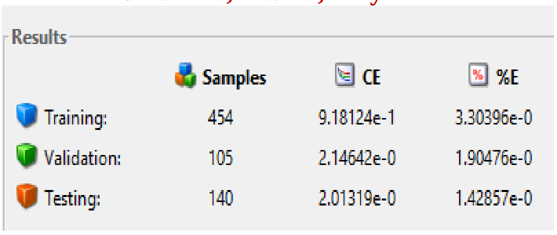
(a) (b)

Fig. 3: (a) histogram (b) binary image

From the selected database, 70 percent of the data is used for training, 15 percent is used for training purposes. The data for testing and the remaining 15% data will be used for validation. The network contains three layers, i.e. the input layer, the hidden layer and Layer output. The parameters used for the artificial neural network are shown in the window of training (Figure 5). Training function Levenberg-Marquardt is used for training purposes. The network shows good results in terms of training and classification. Other Resilient Back Propagation training function, Conjugate Gradient with Powell and so on; all of these training functions are used.

Levenberg-Marquardt is selected by comparing the accuracy of the classification, Training time for convergence and mean square error. Shown in Figure 4-6

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 9, Issue 3, May 2020



Results

	Samples	CE	%E
Training:	454	9.18124e-1	3.30396e-0
Validation:	105	2.14642e-0	1.90476e-0
Testing:	140	2.01319e-0	1.42857e-0

Fig. 4: Error and Validation

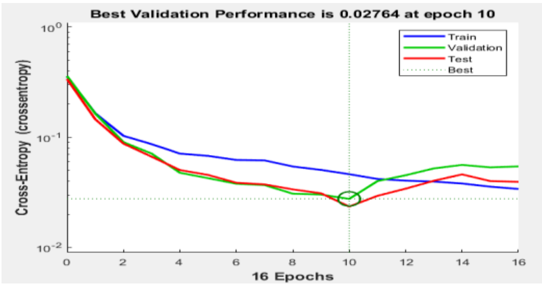


Fig. 5: Cross Entropy

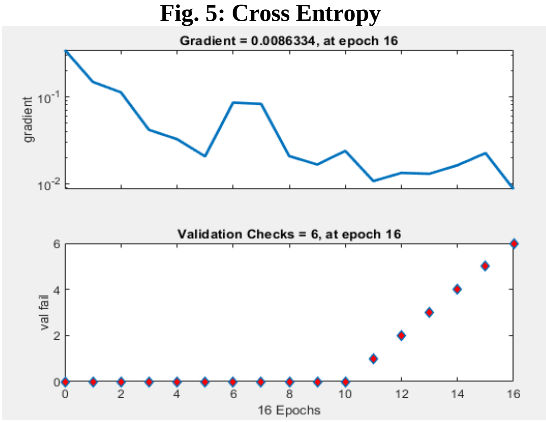


Fig. 6: Gradient plot

The precision of the classification depends primarily on the careful selection of functions. For this research work, mammograms are classified into the other step The neural network is used as a dissection classifier. The problem under evaluation is the binary classification; the parameters used for weighing are precision, specificity and sensitivity. These parameters are defined as the following:

Sensitivity = TP / TP+FN x 100

Specificity = TN / TN+FP x 100

Accuracy = TP+TN / TP+TN+FP+FN x 100

Where TP is positive, TN is negative, and FP and FN are negative.

Table 1 Accuracy analysis

Datasets	Features	Accuracy
612	6	98.7 %

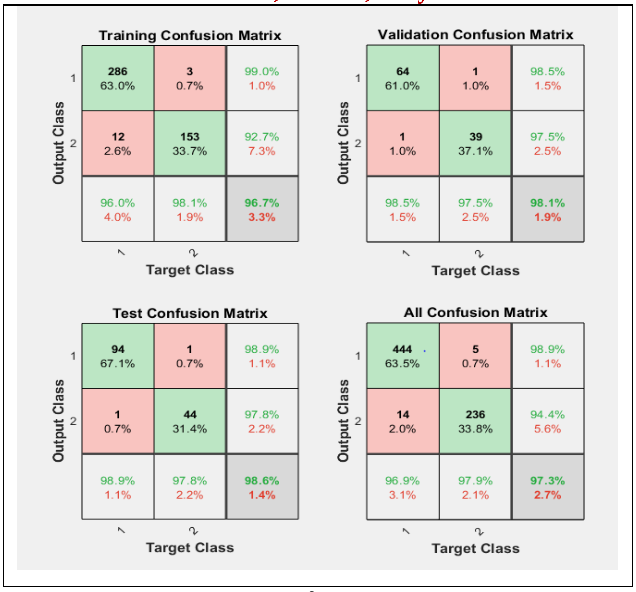


Fig. 7: Confusion matrix

They are false positive and false negative, respectively. Measures on sensitivity ppercentage of really predicted cancer class, specificity measures [14]. The percentage and accuracy of the really expected benign / normal class is Percentage of correctly predicted cancer and normal cases. The data are rotated Five times and the best result out of five is shown in confusion. In the matrix above Figure 7.

V. CONCLUSION

The last decade has witnessed major advancements in the methods of the diagnosis of breast cancer. Only recently the soft computing techniques are being used, hence the body of study in this area is very less. Considering GLCM feature along with fuzzy logic and neural network gives 98.7 accuracy. This classifier overcomes the drawbacks of fuzzy systems and neural networks by combining these two intelligent methods. This can be extended with convolutional network for lung cancer or brain cancer.

REFERENCES

[1] who (2009). "Women's Health". [Online]. Available: http://www.who.int/mediacentre/factsheets/fs334/en/ind ex.html.

[2] Janghel, R.R.; Shukla, A.; Tiwari, R.; Kala, R. (2010). "Breast Cancer Diagnostic System using Symbiotic Adaptive Neuro-Evolution (SANE)". Proceedings International conference of Soft Computing and Pattern Recognition 2010 (SoCPaR-2010), 7th-10th Dec., ABV-IIITM, Gwalior. pp: 326-329.

[3] T. A. ETCHELLS, P. J. G. LISBOA., "Orthogonal Search-based Rule Extraction (SRE) for trained Neural Networks: A Practical and Efficient Approach". IEEE Transactions on Neural Networks, Vol.17, 2006, pp: 374-384.

[4] Hamza Turabieh, "GA-based Feature Selection with ANFIS Approach to BreastCancer Recurrence", IJCSI International Journal of Computer Science Issues, Volume 13, Issue 1, ISSN (Print): 1694-0814. January 2016.

[5] Radhanath Patra and Shankha Mitra Sunani, "A Review On Different Computing Method For Breast Cancer Diagnosis Using Artificial Neural Network And Data mining Techniques", Int. J. Adv. Res. 4(11), pp. 598-610.

[6] https://data.world/datasets/breast-cancer.

[7] Rajamani.R and Rathika.M, "Analysis of Liver Cancer using Adaptive Neuro Fuzzy Inference System (ANFIS)", International Journal of Innovative Research in Computerand Communication Engineering, Vol. 3, Issue 7, July 2015.

[8] Vibhav Prakash Singh, and Rajeev Srivastava, "Automated and effective content-based mammogram retrieval using wavelet based CS-LBP feature and selforganizing map", Biocybernetics and Biomedical Engineering, Vol. 38, No. 1, pp. 90-105, 2018.

[9] Mugahed A. Al-antari, Mohammed A. Al-masni, SungUn Park, JunHyeok Park, Mohamed K. Metwally, Yasser M. Kadah, Seung-Moo Han, and Tae-Seong Kim, "An automatic computer-aided diagnosis system for breast cancer in digital mammograms via deep belief network", Journal of Medical and Biological Engineering, Vol. 38, No. 3, pp. 443-456, 2018.

[10] Syed Jamal Safdar Gardezi, Ibrahima Faye, Jose M. Sanchez Bornot, Nidal Kamel, and Mohammad Hussain, "Mammogram classification using dynamic time warping", Multimedia Tools and Applications, Vol. 77, No. 3, pp. 3941-3962, 2018.

[11] Birmohan Singh, and Manpreet Kaur, "An approach for classification of malignant and benign microcalcification clusters", S?dhan?, Vol. 43, No. 3, pp. 39, 2018.

[12] Ardalan Ghasemzadeh, Saeed Sarbazi Azad, and Elham Esmaeili, "Breast cancer detection based on Gaborwavelet transform and machine learning methods", International Journal of Machine Learning and Cybernetics, pp. 1-10, 2018.

[13] Chisako Muramatsu, "Overview on subjective similarity of images for content-based medical image retrieval", Radiological Physics and Technology, Vol. 11, No. 2, pp. 109-124, 2018.

[14] Isha Suwalka and Navneet Agrawal," A Machine learning technique for tumor detection around edema region" , Proceeding , Second International Conference on Advanced Computational and Communication Paradigms (ICACCP-2019) during 25th to 28th of February, 2019 ,SMIT, Sikkim.

AUTHOR'S PROFILE

Sarita Chauhan is working at MLVTEC ,Bhilwara in the capacity of Assistant Professor in ECE Department and has more than 17 years of experience .She has more than 20 publications in International Journals & Author of 2 books under RTU.Her field of interest are NT, VLSI,VHDL,ANFIS,NN, AE,EDA ,Image processing etc.