International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-11, September 2019

Breast Cancer Detection using Machine Learning

R.Chtihrakkannan, P.Kavitha, T.Mangayarkarasi, R.Karthikeyan

Abstract: ***Cancer is the 2nd source of death in the world. The main reason for this increased death rate is the delayed detection of cancerous tissue growth in a person. Nearly 60% patients with breast cancer are diagnosed in advanced stages. The main objective of our paper is to enhance an image processing algorithm for earlier finding of breast cancer. X-ray mammogram images which have been acquired are used as input Images. [1] The pre-processing of input images are carried out by applying Gaussian Filter and Edge detection techniques to enhance image quality. Wavelet Transform is useful to identified first order features and GLCM based second order features are extracted from the Pre-processed images. The statistical parameters are then used for classification using DNN a Multilayer supervised classifier. Dataset images are created from the training phase. In testing Phase the acquired image from a patient is given as input to the classifier after completing the image processing steps such as Pre-processing and feature extraction. The output of the classifier consists of two classes, normal and abnormal respectively. [2] The entire algorithm is developed in Python language. The Processing time for testing and conformation of Positive cases is very minimum. Using deep learning neural network classifier an accuracy rate of 92% is reached.***

Keywords : Breast Cancer, earlier detection, GLCM, DNN, Wavelet transform, accurate

I. INTRODUCTION

The major health issue that arises these days had led to much advancement in the medical field, yet there are certain diseases which remains incurable. Certain diseases even seem to be more challenging to guess what it is in the earlier stages. Among these many diseases (stroke, cancer, heart attack, viral hepatitis, chronic liver diseases, coronary artery disease etc.), the death rate due to cancer is becoming more and more each year. Once a tissue is affected by radiation, it takes almost 15 years to turn into a cancerous tissue. Even though we have this time period for perfect curability, many of the patients who have this kind of basic abnormalities don't take it as a big deal since the symptoms are not indicating cancerous growth. Nearly 60% patients with breast cancer are diagnosed in advanced stages. [3]

The growing cancer burden is due to a number of factors, including residence growth and ageing as well as the changing commonness of certain causes of cancer connected to social and economic progress. There are so many cancers, among which Breast cancer is the most important cause of cancer death among women. [1]

II.METHODOLOGY

2.1INPUT IMAGE

In thia medical diagnosis systems, X-rays are used as pridiction tool in mammography for the inspection of human breast. These images are recorded as specialized images which are then experiential by radiologists for any possible abnormality.

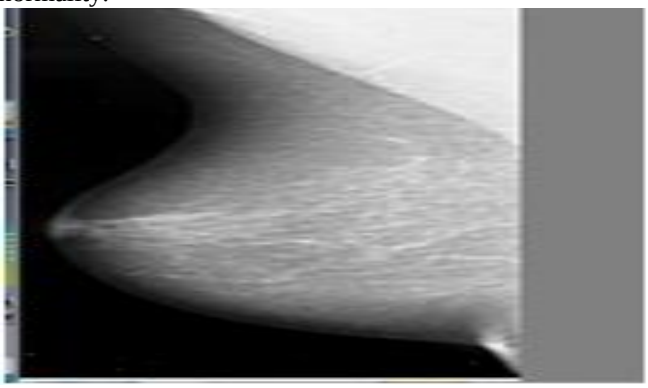


Fig 2.1: Input Image

2.2 BILATERAL FILTER

This filter have Non-linear nature and have fringe protected and noise decreasing flatted types of pictures. It restore the strength of every picture quantified usual strength size of neighbor pixels. This size of pixel represents the Gausian distribution. Critically the quantities relies on not particularly norm closeness of true pictures, still we use radiometric closeness (Ex. Distance such that colour brutality, long range etc.) This method protect keen portion.

The bilateral filter is defined as

$$I^{\text{filtered}}(x) = \frac{1}{W_p} \sum_{x_i \in \Omega} I(x_i) f_r(\|I(x_i) - I(x)\|) g_s(\|x_i - x\|),$$

Where the normalization term

$$W_p = \sum_{x_i \in \Omega} f_r(\|I(x_i) - I(x)\|) g_s(\|x_i - x\|)$$

Manuscript published on 30 September 2019.
*****Correspondence Author(s)
R.Chtihrakkannan, Instrumentation and Control Engineering, Sri Sairam Engineering College, Chennai, Tamilnadu, India
(email: chithrakkannan.ice@sairam.edu.in)
Dr.P.Kavitha, , Electronics and Instrumentation Engineering, RMK Engineering College,Chennai, Tamilnadu, India
(email: kavithathevan@yahoo.com)
T.Mangayarkarasi, Instrumentation and Control Engineering, Sri Sairam Engineering College, Chennai, Tamilnadu, India
(email: hod.ice@sairam.edu.in)
R.Karthikeyan, Instrumentation and Control Engineering, Sri Sairam Engineering College, Chennai, Tamilnadu, India
(email: karthikeyan.ice@sairam.edu.in)

Fig 2.2: Bilateral Filtered Image

Then, the weight assigned for pixel (k,l) to denoise the pixel (i,j) is given by

$$w(i,j,k,l) = \exp\left(-\frac{(i-k)^2+(j-l)^2}{2\sigma_d^2} - \frac{\|I(i,j)-I(k,l)\|^2}{2\sigma_r^2}\right),$$

where σd and σr are smoothing parameters, and I(i, j) and I(k, l) are the intensity of pixels (i,j) and (k,l) respectively.

After calculating the weights, normalize them:

$$I_D(i,j) = \frac{\sum_{k,l} I(k,l)w(i,j,k,l)}{\sum_{k,l} w(i,j,k,l)},$$

Where ID is the denoised intensity of pixel (i,j).

2.3 GAUSSIAN BLURRING

This method provide input-output straight relationship and reshape the picture or reduce the noise. If we proposed for this method we get "unsharp" screening (" bordering discover"). This filter exclusively apply " sharp boundary" and decrease contrast.. [5] The one advantage a Gaussian filter has over a median filter is that it's faster because multiplying and adding is probably faster than sorting.

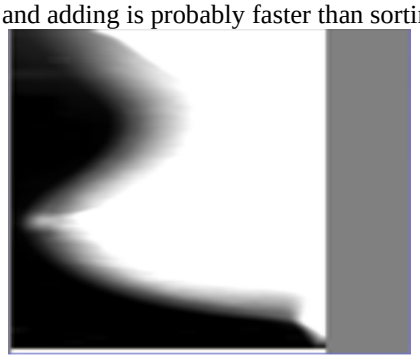


Fig 2.3: Gaussian Filtered image

In fading, we use easy fade pictures and this picture consist of additional keen information or extra interpolated image. If we recognize all the objects and keen bordering are rightly mentioned.[5] The Gaussian image is correctly predicted image due to avoid fate pictures. So we resize and edge information is more transfer from smooth edge information.

2.4 ENHANCED IMAGE:

In this method already handled pictures have several unwanted white parts, this portions of the image consider as dissimilarity pictures is extended. In order to get affected portion are highlighted and matched to severely affected portion and provide more precise information. The unwanted portions are detached. The highlighted portion are enlarged and compare with reference pictures and highlighted pictures

Are resize to byte.

Fig 2.4. Gray scale Enhanced Image

2.5 SEGMENTED PICTURE

In segmentation method transfer digitalize picture into sequence of pictures, that kind of pictures are collection of pixel that we refer as perfect pictures. This method presented in simpler and vary the presentation of a picture into variable form that represent more understandable and handled very lucid way to analyze.[6] That notable pictures are removed from original pictures.

If we notice that after considering segmented process the blur area that is the abnormality exist in the infected portion will appear clearly.

If we consider the depth of subtraction operation, succeeding the blur picture and the threshold is selected.

In this pictures enlargement choose threshold to reduce intra class difference of saturated gray image.

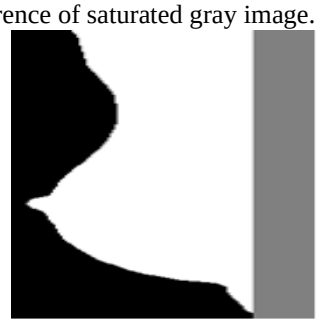


Fig 2.5: Segmented Image

III. MACHINE LEARNING

Machine Learning is a process that machines (computers) are trained with data to make the decision for similar cases. ML is employed in various applications, such as object recognition, network, security, and healthcare. [1] There are two ML types i.e. single and hybrid methods like ANN, SVM, Gaussian Mixture Model (GMM), K-Nearest Neighbor (KNN), Linear Regressive Classification (LRC), Weighted Hierarchical Adaptive Voting Ensemble (WHAVE), etc.

3.1 Feature extraction using GLCM:

The high information contained sub-band is selected from the DWT , this sub-band is taken as a input for feature extraction.

Retrieval Number: K24980981119/19©BEIESP
DOI: 10.35940/ijitee.K2498.0981119
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

In Grey-Level Co-occurrence Matrix (GLCM) ellaporate the co-ordinate difference between each intensity presented in different gray levels i and j at specific angle θ. The following properties of GLCM is estimated

1.Power estimating uniformity of grey scale distribution,

2. Entropy measuring randomness;

3.Standard deviation;

4.Mean;

5. Texture feature.

Finally, 5 The characteristics are removed for the given input pictures and the given values are stored in an array for classification purpose.

Fig 3.1: Extracted Image

IV. MLP CLASSIFIER

- Multi-layer Perceptron (MLP) is used as classifier because it compare input image with target image .
- In this method we use delta learning rule to train the neural network, using that rule convergence time is reduced. [4]

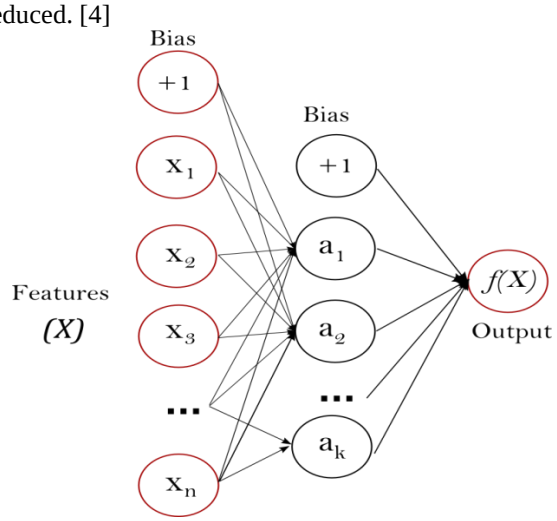


Fig 4.3.1: Multi-Layer Perceptron

V. RESULT

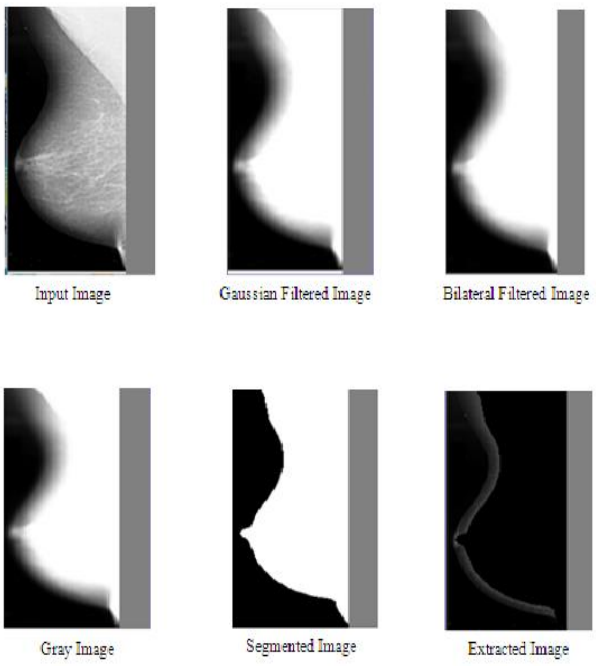
In this paper, detection of breast cancer was done using DNN algorithm. There are several other algorithms are also available such as ANN, SVM, KNN etc. But as per the accuracy level is concerned, DNN provides a better result. The average accuracy was improved to 89.77% and high accuracy changed to 96%.

Comparative study of several algorithms is listed below the table.

METHODOLOGY	DATABASE	SENSITIVITY	PRECISION	ACCURACY
Artificial Neural Network	Mammogram	93.22%	83%	94.3%
Support Vector Machine	Mammogram	86-92%	70%	85-91%
K- Nearest Neighbors	Mammogram	67-87%	70.8%	70-80%
Deep Neural Network	Mammogram	85-93%	86%	96.3%

From the above table it is inferred that, using DNN algorithm accuracy and level of sensitivity is increased than another algorithm. Thus, the earlier detection of breast cancer is detected without any physical contact. This method avoids destruction of the part being tested.

This paper uses the wavelet transformation and GLCM feature extraction. These features were useful to distinguish the maximum number of samples accurately. Finally, the simulated results show that used methodologies provides better classifier rate with minimum error rate for all test samples.



VI. CONCLUSION

This paper focus on the earlier diagnosis of breast cancer, as the detection bring about the success of about 96% accuracy by the use of DNN algorithm for classification. By the use of python as a software platform, training and the computational time has been reduced to a greater extent than others. This project ensures the greater level of detection of breast cancer at earlier stage, by which mortality rate of cancer affected person can be reduced and earlier diagnosis would increase the life time of a patient by giving them a right treatment at a right stage

REFERENCES

1. "Latest Global Cancer Data: Cancer burden rises to 18.1 million new cases and 9.6 million cancer deaths in 2018", International Agency for Research on Cancer, World Health Organization, 12 September 2018.
2. M.M.Mehdy, E.E.Shair and P.Y.Ng, "Artificial Neural Networks in Image Processing for Earlier Detection of Breast Cancer", Hindawi, Computational and Mathematical Methods in Medicine, Volume 2017, Article ID 2610628.
3. Vishnukumar K.Patel, Prof.Syed Uvaid and Prof.A.C.Suthar, "Mammogram of Breast Cancer Detection Based Using Image Enhancement Algorithm", Internationa Journal of Engineering Technology and Advanced Engineering, Volume 2, Issue 8, August 2012.
4. Melanie A. Sutton, "Breast Cancer Detection Using Image Processing Techniques", IEEE International Conference on Fuzzy System- Febraury 2000.
5. Bhagyashri k Yadav, Dr. Prof. M. S. Panse, "Virtual Instrumentation Based Breast Cancer Detection and Classification Using Image- Processing", International Journal of Research and Scientific Innovation (IJRSI), Volume V, Issue IV, April 2018.
6. Moh'd Rasoul A Al-Hadidi, Mohammed Y. Al-Gawagzeh, "Solving Mammography Problems of Breast Cancer Detection Using Artificial Neural Networks and Image Processing Techniques", Indian Journal of Science and Technology, Vol 5, No.4 (April 2012), ISSN: 094-6846.