INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 4.295
(Volume3, Issue4)
Available online at www.ijariit.com

Comparing Three Neural Network Techniques in the Classification of Breast Cancer

Smitha Hallad
Department of Computer Science & Engg, KLE DR M S Sheshgiri College of Engg.& Tech, Belagavi
smitha.hallad@gmail.com

Prof. Roopa Hubballi
Department of Computer Science & Engg, KLE DR M S Sheshgiri College of Engg.& Tech, Belagavi
roopapatil.987@gmail.com

Abstract: Breast cancer is becoming a common disease in women nowadays. Breast cancer is nothing but mass or group of uncontrollable growth of cells in the body is called tumour. Benign is an initial stage of cancer and malignant is the last stage of cancer, these two are the stages of breast cancer. A man can survive or rate of survival is more in benign stage by taking a very good treatment by radiologist where as in malignant stage a man cannot survive directly it leads to death. The neural network is a powerful classifier. Before classifying the neural network is to be trained by collecting the data or images from different datasets. After training the large data set test the new data sample or breast image by extracting the new features from the new image and then classifying the image into cancerous or non-cancerous. Finally, the results are compared with three neural network 1) Radial basis function 2) Feed forward neural network and 3) Back propagation neural network. Accuracy is calculated 95% of RBF, 96% of FFNN and 100% of BPNN.

***Keywords:** NN-Neural Network, RBF-Radial Basis Function, FF-feed forward, BP-Back Propagation*

I. INTRODUCTION

In the present scenario, the most developing countries have breast cancer as a major health issue. Breast cancer is diagnosed by millions women's every day and nearly 500,000 people of them lead to death. Although, the rate of occurring breast cancer increased over the year. Breast cancer is a disease developed from breast cells. Breast cancer takes place in the inner lining of milk duct (tubes which stored the milk) and lobules (glands which create milk). This disease most of time comes to women's then men.

Cancer is nothing but a type of mutation of cells, which involves uncontrolled growth of abnormal cells. Lumps/mass are formed by the cancer cells called a tumour. The swelling or lesion is formed in a body by mutation of abnormal cells. Breast cancer is of two types: Benign tumour and malignant tumour. Benign tumour cells will grow slowly and do not destroy neighbouring tissues and it is less threatening to life but malignant tumour grows faster through the process called metastasis. This tumour cells called malignant can also spread to other body parts. The Breast cancer will start in stromal. If its starts growing in lobules, then it is called as lobules carcinoma.

More death occurs every month attributed to cancer which can be prevented with good support from the government in finding for prevention, detection and treatment program. There are many causes of cancers, they are Carcinogens, age, immune system, genetics. With certain lifestyle and environment factor also lead to cancer, they are Overweight or Obesity, alcohol, workplace hazards, Ionizing radiation, tobacco, Infection.

Improvements can be made in initial stage by detecting of breast cancer which can be diagnosed from a standard procedure called X-ray mammography which is a most reliable method and will also lead to a positive trend in mortality.

II. LITERATURE SURVEY

Proposed an efficient method where computer aided is used for mass classification of digitized mammograms with the use of the artificial neural network. The region of interest contains mass ANN classify into benign and malignant. The texture is main characteristics of mammographic. After extracting these features the ANN will classify the mammogram into benign or malignant on the marked region. The achieved sensitivity is 90% and specificity is 83.87% [1]. Artificial neural network is applied by many researchers in the field of signal analysis, image analysis and in another field of the classification problem. The radial basic function is one among the neural network, multilayer perceptron has single or multiple hidden layers which perform the classification. Many feature methods are used for pre-processing the image to improve training performance and accuracy of the results. Finally, by using RBF the results show higher performance and higher accuracy with minimum time complexity [2].

Proposed an approach to building a Cadx model to extract the features to distinguish a healthy person and a person who is suffering from cancer. Out of 145 features, only 31 features are selected, out of 31 selected features 18 are from SCLGM [3]. Described the technique to detect the mass region in digital mammograms. The technique which is used here is second order statistics. Artificial neural network is one type of classifier which is going to classify the breast image which contains cancer part or not. In depth, it cannot be investigated this is the disadvantages of second order [4].

III. SYSTEM DESIGN

The steps involved in Neural Network is

1) Training
2) Classification

1] Training – In training all the collected images are trained to the model and all six features are extracted and stored in the database.

2] Classification – After training, the neural network will classify the given new input as cancerous or non-cancerous.

The work flow of system design is mainly consisting of

1) Collect the images
2) Preprocessing
3) Image segmentation
4) Feature extraction
5) Train the neural network
6) Classification

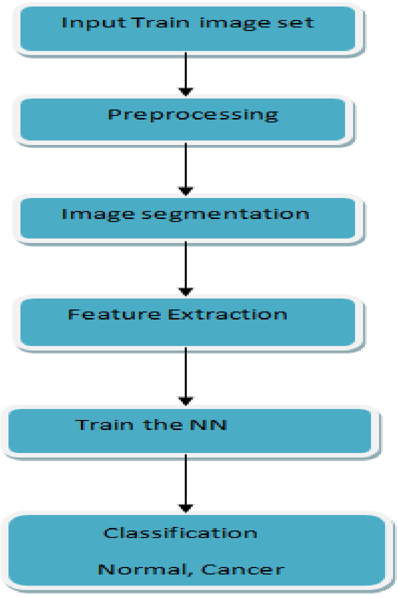

Fig 1: Work flow of the system

A. Collect the Images:

Before train, the neural network first collects the input sample for training. Here input sample set contains 60 images. After collecting all 60 input images the next step is pre-processing. At the first database creation of images is done using which training part of the work is implemented. Lets the user select an image, the format of the image could be any of jpeg, Png or bmp formats. In this work mainly used is the jpeg image format, readily available via camera. All these images are used in training the neural network after that is used to classify the given test images.

B. Pre-Processing

Convert the RGB image format (I) to the gray-scale image, converting the image representation from three dimensional to two-dimensional formats.

Possibility Distribution Algorithm

The functional possibility distribution finds the minimum and maximum values of the image pixels. Also, mean of the pixels intensities id calculated. Using these quantities the total histogram of the image id divided between four different regions. The central part of the image histogram is one which occurs most generally in nature hence that part is said to be of the benign sample, whereas those at extremes are rare to occur hence these are in a different group. This process is the first step of fuzzy logic development.

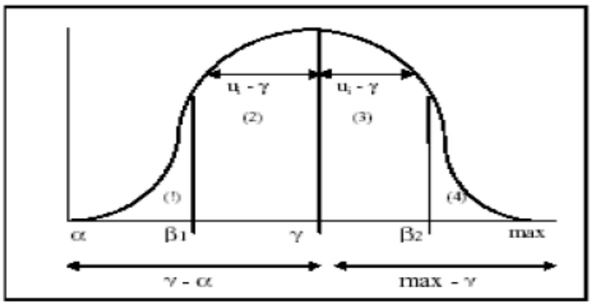

Fig 2: Distribution image

The fuzzy transformation function for computing the fuzzy plane value P is defined as follows:

α= min;
$\beta 1= (\alpha+ \gamma) /2$;
$\beta 2= (max + \gamma) /2$;
γ = mean; max;

The following fuzzy rules are used for contrast enhancement based on Figure (2).

Rule-1: If $\alpha \leq u_i < \beta_1$ then $P = 2 ((u_i - \alpha) / (\gamma -\alpha))^2$ (2)

Rule-2: If $\beta_1 \leq u_i < \gamma$ then $P = 1- 2 ((u_i - \gamma)/(\gamma -\alpha))^2$ (3)

Rule-3: If $\gamma \leq u_i < \beta_2$ then $P= 1- 2((u_i - \gamma) / (max - \gamma))^2$ (4)

Rule-4: If $\beta_2 \leq u_i < max$ then $P = 2 ((u_i - \gamma)/(max - \gamma))^2$ (5)

Where $u_i = f(x,y)$ is the i^{th} pixel intensity.

The possibility distribution algorithm is described as follows:

Step-1: Parameter Initialization

- Set β_1 = (min + mean)/2
- Set β_2 = (max + mean)/2

Step-2: Fuzzification

- For all pixels (i,j) within the image Do
  - If ((data[i][j]>=min) && (data[i][j]< β_1))
    Compute NewGrayLevel =2*(pow(((data[i][j]-min)/(mean- min)),2))
  - If ((data[i][j]> = β_1) && (data[i][j] < mean))
    Compute NewGrayLevel=1-(2*(pow(((data[i][j]-mean)/(mean-min)),2)))
  - If ((data[i][j]>=mean)&&(data[i][j]< β_2))
    Compute NewGrayLevel=1-(2*(pow(((data[i][j]-mean)/(max-mean)),2)))
  - If ((data[i][j] >= β_2) && (data[i][j] <max))
    Compute NewGrayLevel=2*(pow(((data[i][j]-mean)/(max-mean)),2))

Step-3: Modification

- Compute FuzzyData[i][j]= pow(NewGrayLevel,2)

Step-4: Defuzzification

- For all pixels (i,j) within the image Do
  Compute EnhancedData[i][j]=FuzzyData[i][j]*data[i][j];

The process is completed by modification of the pixel values in the two groups such that one with bright value is brighter and one with comparatively dark value is darker.

C. Image Segmentation

The boundaries of this segmented image are identified, for an image with high and low intensities. And the image is set to be bound by top row, right column left column and bottom row. The image enhanced after the fuzzy application is cropped using these boundaries. The image is segmented into two sections. One is black and white and second is a gray image. Statistical properties related to intensities are observed in gray scale format.

D. Feature Extraction

There are six properties in this methodology. They are contrast, mean, standard deviation, variance, kurtosis, and smoothness.

1. Contrast: Contrast is defined as the difference in colour of the image that makes its distinguishable or increases the level of brightness in the image.

2. Mean: Mean is defined as a contribution of the intensity of each pixel of an entire image or average value of the intensity of the entire image. The mean value of mass calcification is higher than the normal image. Mean is represented by μ and it is expressed in term of

$$\mu = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} p(i,j)$$

Where the pixel value is p(i,j) of the image size is of M*N

3. Standard deviation: It is closely defined to mean and its value occurs around the mean. The followed equation is

$$\sigma = \sqrt{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(p(i,j)-\mu)^2}$$

μ is mean of the image pixel p(i,j)

4. Variance: Variance is defined to find how far each pixel varies from mean or central pixel and categorized into two different region.

Variance = standard deviation^2

5. Kurtosis: It is represented as K and is used to detect the peak of the distribution of the image.

$$K = \left\{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left[\frac{p(i,j)-\mu}{\sigma}\right]^4\right\} - 3$$

σ is the standard deviation of the pixel(i,j)

6. Smoothness: It is represented as R and gray level contrast is measured to reduce the noise in the image and less pixelated image.

$$R = 1 - \frac{1}{1+\sigma^2}$$

σ is the standard deviation of the pixel(i,j)

E. Train the Neural Network

Training procedure:

1. Start with images of which classes are known or labeled.
2. Find the property set or feature set for each of them. Label suitably the target variable.
3. Take the next image as input and find features as a new input.
4. Read the entire image set in the database and save the properties and targets.
5. Choose the Neural network of choice.
6. Find the neural network structure based on the target of the image.

F. Classification

1. Take the images to be classified.
2. Extract the property set or features set.
3. Use specific trained knowledge of neural network to the feature set of case image applied.
4. The output will be in the same format as targets.

The neural network is one of the intelligent and powerful classifiers. It is widely used in various field of medical diagnosis and computational model in machine learning also. Training the neural network model is one of the main objective challenges. After training the model the new input for each Neural Network is given and test the given sample into cancerous or non-cancerous.

IV. RESULTS AND PERFORMANCE

The proposed method is developed to train the model to detect and classify the given sample breast image into cancerous or non-cancerous.

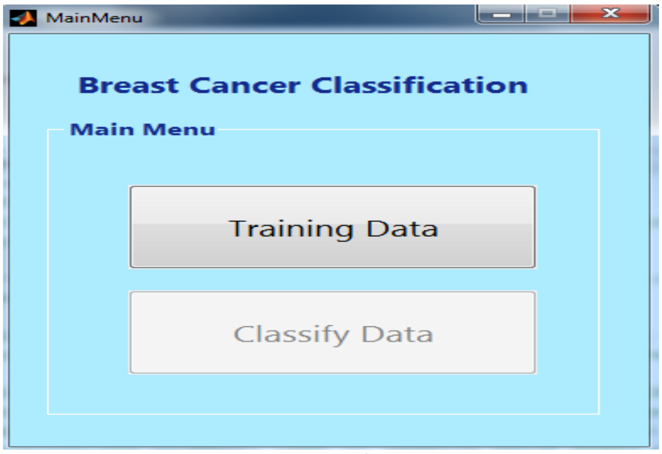

Fig 3: Home page for training

The above-represented fig shows the home page which consists of two bottoms – Training data button and classifies data button. The first training procedure is carried out with 60 images then classify the breast image into normal or cancer. The database consists of 60 images; the model is trained with these 60 images. Then classifying process is carried out the neural network itself has classified then it classified into cancerous or non-cancerous.

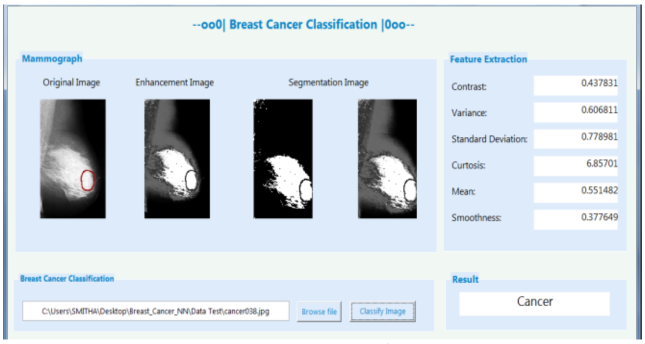

Fig 4: Breast cancer classification page

The above-represented figure shows breast cancer classification page. Here the user can browse an image by clicking on browse file button. The sample image is uploaded in the first section. After browsing the image, next step is to classify the image. First, enhance the image then segment the image into binary and gray scale. After clicking on classify image button the given grayscale image is classified and then shows the results that are it cancer or normal.

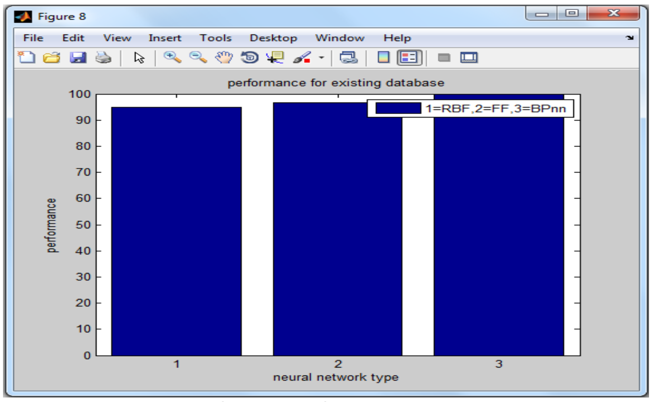

Fig 5: Performance of three neural network

The result obtained is: cancer;
The accuracy for three different neural networks is:
perRBF = 95% for RBF AWGN NN.
perff = 96.6667% for feed forward NN and
person = 100 % RBF back propagation type NN.
Once the model is trained, classify the disease many times by testing the breast images.

V. FUTURE SCOPE

Classification technique can be used for multiple classes also. Discrete wavelet transform can be used for discrete sample input. The neural network can be used for other diseases also like liver disease, brain tumor, and skin disease.

CONCLUSION

The proposed method includes three neural networks they are named as RBF, feed forward and back propagation neural network. Training the neural network model is one of the main objective challenges. After training the model test the given sample into cancerous or non-cancerous. The new input for each Neural Network is given and results observed to be satisfactory. Three different types of neural networks are trained for the same set of inputs and accuracy found to be different. The accuracy of RBF is 95%, feed forward is 96% and back propagation is 100%.The back propagation neural Network gives the better accuracy compared to feed forward and RBF functions.

ACKNOWLEDGEMENT

I wish to express my sincere gratitude to respected Principal and H.O.D of Computer Science Department for providing me an opportunity to present this paper. My sincere thank to my project guide Prof. Roopa Hubballi for guidance and encouragement in paper publishing.

REFERENCES

[1]"An Efficient Automatic Mass Classification Method In Digitized Mammograms Using Artificial Neural Network", Mohammed J. Islam1, Majid Ahmadi2 and Maher A. Sid-Ahmed, International Journal of Artificial Intelligence & Applications (IJAIA), Vol.1, No.3, July 2010.

[2]"Radial Basis Function (RBF) Neural Network Classification based on Consistency Evaluation Measure", Aye Mya Thandar, Myo Kay Khaing, International Journal of Computer Applications (0975 – 8887) Volume 54– No.15, September 2012.

[3]"New Feature Extraction Method For Mammogram Computer Aided Diagnosis", Belal K. Elfarra and Ibrahim S. I. Abuhaiba, International Journal of Signal Processing, Image Processing, and Pattern Recognition Vol. 6, No. 1, February 2013.

[4]"Detection of Masses in Digital Mammogram Using Second Order Statistics and Artificial Neural Network", Al Mutaz M. Abdalla, Safaai Dress, Nazar Zaki, International Journal of Computer Science & Information Technology (IJCSIT), Vol 3, No 3, June 2011.

[5]"Feature Extraction for Classification of Microcalcifications and Mass Lesions in Mammograms", Rabi Narayan panda, Dr.Bijay ketan panigrahi and Dr. Manas Ranjan patro, IJCSNS International Journal of Computer Science and Network Security, VOL.9 No.5, May 2009.

[6]"A Novel Breast Tissue Density Classification Methodology", Arnau Oliver, Jordi Freixenet, Robert Mart´ı, Josep Pont, Elsa P´erez, Erika R. E. Denton, and Reyer Zwiggelaar, Ieee Transactions On Information Technology In Biomedicine, Vol. 12, No. 1, January 2008.

[7]"Breast Cancer Classification Using Neural Network Approach: Mlp And Rbf", Ali Raad, Ali Kalakech, Mohammad Ayache, The 13th International Arab Conference on Information Technology ACIT'2012 Dec.10-13.

