TEST Engineering & Management

May - June 2020
ISSN: 0193-4120 Page No. 4728-4733

Detection of Breast Cancer using Pattern Recognition

Ayesha Banu R[1*], Mallikarjun M. Kodabagi[2]
[1]M.Tech, School of C&IT REVA University, Bangalore, India
[2]Professor, School of C&IT, REVA University, Bangalore, India
[1]r.ayeshabanu@gmail.com, [2]mallikarjunmk@reva.edu.in

Article Info
Volume 83
Page Number: 4728-4733
Publication Issue:
May-June 2020

Article History
Article Received:*19 November 2019*
Revised: *27 January 2020*
Accepted: *24 February 2020*
Publication: *16 May 2020*

Abstract

Computer-aided decision or detection systems which support plan to progress for screening programs of bosom breast cancer by helping radiologists to assess DM (Digital mammography). Regularly such techniques continue in two stages: choice of applicant areas for harm, further characterization as both malignant or benign or normal. The study, utilised as a candidate detection method which is built on deep learning to impulsively identify and also to identify segment soft tissue lesions in Digital Mammogram. A recent study on PNN (probabilistic neural network) training algorithm is predictable. The standard PNN, however requiring an exceptionally short preparing time, when executed displays the downsides of being costly in terms of classification time and of challenging an unimpeded amount of units. The commended alteration disables the concluding disadvantage by introducing a removal measure to evade the storing of excessive patterns. The contortion in the bulk estimation presented by this standard is made up for by a cross-validation technique to adjust the system parameters. The proposed algorithm makes it conceivable to understand the PNN and simultaneously, makes up for certain deficiencies emerging from the hypothetical premise of the PNN, which doesn't perform well with small training sets and classifies the image into malign, benign or not.

Keywords: *Image Classification, Segmentation, GLCM, GMM, PNN.*

1. Introduction

Breast tumour is most popular diseases which is diagnosed and found in women's across world. By detecting breast examination in clinical, but then again the detection rate is very low. Moreover, the anomalous areas which cannot be touched can be pretty interesting to identified by the traditional techniques which certainly seen on a conventional mammography or ultrasound. Timothy de Moor et al [1] the best method for identifying breast cancer at its early stage is Mammogram. Salma Alqazzaz et al [2] identified the problematic with mammogram images which are complicated. As a result, to identify about tumour for radiologist we utilize image processing and features extraction techniques. Y. J. Tan et al [3] concentrated more to help doctors to find out the existence of the tumour using features extracted from suspicious regions in mammography images at real time thus fast in treatment process.

Detecting breast tumour is very challenging. Such as tumour growth which is definitely not a solitary sickness however an assortment of numerous infections. In same way, every malignant growth is not same as each other disease that exist. Any One technique or one algorithm will not give the results to identify breast tumour. Whereas one disease contrast as of another, likewise every single breast tumour looks in a different way as of another. If the patient has undergone some breast surgery identified by Saira Charan et al [4] the digital mammogram image can also be compromised.

Breast Tumour from past two decades it has become very significant disease. Many tumours people have got treated due to early detection. Saira Charan et al [4] says that it was expensive and time consuming during the progress in diagnosis and treatment. Because of the way that each tumour is dissimilar like its host and each necessitates customized medication to be cured by automated detection of mass which is a problematic task.

Shuyi Li et al [5] a dangerous factor that received extensive attention from amount of fibro-glandular lesion in a woman's breast includes mammogram breast density.

May - June 2020
ISSN: 0193-4120 Page No. 4728-4733

The utilization of breast tissue as an intermediary for the nitty gritty data implanted on the mammography which is restricted because of the radiologist the emotional appraisal and varies approximately in breast mass density assessment, and breast density information is enclosed into a single value in the digital images. There can have drastically dissimilar mammography with immensely with different results with same-age patients who are allocated with the same mass density score. While during the past studies [10–12] investigated mechanized techniques to evaluate density of breast, endeavours diminished the mammogram inputs in to a small number of statistical information with a great extent identified with measurements of glandular lesion that were not adequate in recognizing patients who will and who will not create breast cancer.

The organization of the paper is as follows: Section 1 gives the brief introduction about Detection of Breast Cancer using Pattern Recognition. Section 2 related paper and technological issues. Section 3 includes proposed methodologies. Section 4 explains about architectural representation. Section 5 deals with the algorithm of the execution. Section 6 explains the comparison between different algorithms. Section 7 deals with results and discussion mainly deals with the graphical user interface of the project to show the output of the application. Section 8 explains the Conclusion. Finally, all this sections mainly the summary of entire project development and it also suggests some of the enhancement idea which mainly highlights all the journals and IEEE papers referenced for the development of dissertation.

2. Related Work

Timothy de Moor et al [1], utilised DM images for candidate selector to identify and segment cancerous lesions using deep learning neural network.

The main contributions are: Firstly, significantly reduce in the many identified locations which are sensitive near to 100%. Secondly, we removed the false positive and the idea is to keep true positives. The methodology used are patient population, deep learning network and training for preprocessing and image acquisition for Segmentation of ground truth labeling. Technology and parameters used are u-net deep learning, FROC, Sensitivity.

Salma Alqazzaz et al [2], discussed based on CNN to identified three mass detection and evaluate. First, based on heuristic knowledge a suspicious region locating method is employed. Then to categorize the suspicious region as mass or normal three different CNN schemes are considered. The methodology used locate suspicious region, mass detection are Data augmentation, feature Extraction using CNN-1, mass detection based on CNN-3 and CNN-two-view. The skills and limitations are CNN (Convolutional Neural Network), FROC curve.

Y. J. Tan et al [3], have discussed about breast tumor detection using CNN is proposed to classify mammogram image into normal, benign (noncancerous abnormality) and malignant (cancerous abnormality) for mammogram imaging system. The methodology used are image Preprocessing, Tensor Flow, and Convolutional Neural Network.

Saira Charan et al [4], have accomplished the training from scratch for the application of CNN. First layers detect large features using CNN. Next layers detect minor features that are more non-representational. Final layer is for detailed classification. The technology used CNN (Convolutional Neural Network), Morphological operations is to find accuracy.

SHUYI LI et al [5], describe about mass segmentation present in breast, using a U-Net architecture called as attention gates (AGs). The suggested network with an encoder and a decoder has a U-type network. The methodology used are FCN (Fully CNN), U-Net architecture and Attention Dense. The technology used are CNN, U-Net and the parameters are F1-score, mean connection over union, sensitivity, specicity, and accuracy.

3. Methodology

The adaptive mean filters i.e. image enhancement is used to eliminate noise from the mammogram images. Since, it is better among all the spatial filters and recognize adequate details from noise that are identified from the images. For the implementation of spatial processing technique to identify about which pixels have been influenced by impulsive noise of Adaptive Median Filter.

The size of various neighbourhood is adaptable, just as the threshold limit for the correlation comparison. A pixel that is dissimilar of a popular of its neighbours, just as being not fundamentally lined up with those pixels to which it is comparable, is named as impulse noise. For the implementation of spatial processing technique to identify about which pixels have been influenced by impulsive noise of Adaptive Median Filter.

These noise pixel are then substituted by the intermediate pixel value of the pixels in the neighbour-hood that have approved the noise labelling test. Initially, we are altering the images into grayscale image using rgb2gray() purpose and utilizing the adaptive median method to the output image and then converting the image into unsigned integer 8 using unit8() function. By using this way we pre-processed images, and then we performed GMM segmentation on the pre-processed image with number of regions 2 and number of Gaussian Mixture Model components is 2 and max iterations 10. We use k-means segmentation with k=2. HMRF-EM and Expectation-Maximization Algorithm is used for the further execution.

The picture describes the difference between Malignant and Benign tissues in Breast

May - June 2020
ISSN: 0193-4120 Page No. 4728-4733

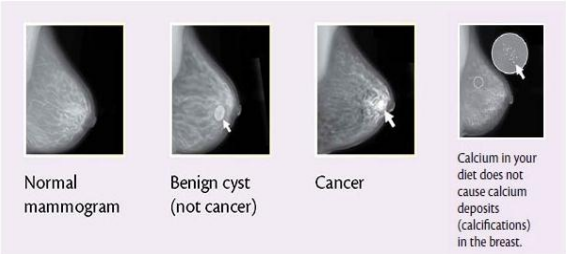


Figure 3.1: Types of cancer

The flow of the project is as shown below

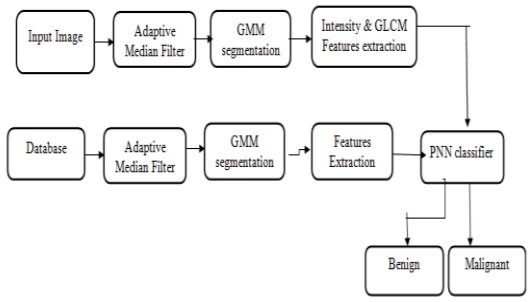


Figure 3.2: Work flow

4. Architectural Representation

This flow diagram is an important diagram which is the Unified Modelling Language (UML) to depict the dynamic portions of the framework. The duplication of the flow from one activity to another activity is known as flow chart.

It demonstrates distinctive streams as parallel, branched, simultaneous and single. Once the mammogram images are input the pre-processing happens using Adaptive median filter by the size of each noise present in the pixels. The segmentation is done utilizing GMM for extraction process using GLCM of the input images and finally it is classified to identify the input image is normal, benign or malignant in the outputted result.

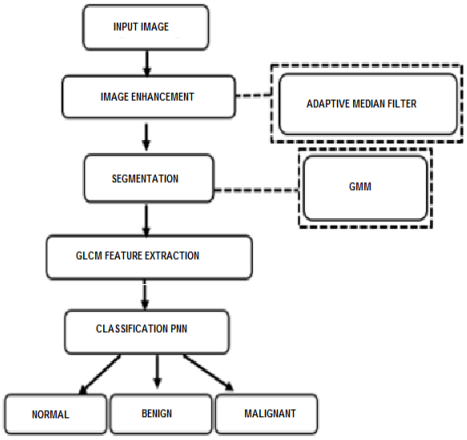


Figure 4.1: Flow of system architecture

Pre-Processing

Pre-processing is estimated as significant advancement in order to notice the direction of the mammogram images and to characteristic of the images. The breast tumour collected together will be exposed to pre-processing for noise removal by utilizing adaptive median filtering process with the goal that the consequence of result is more practical than the initial original picture for a specific presentation. The adaptive Median Filter characterizes picture components as noise by contrasting each pixel inside the picture with its nearby neighbouring pixels. The mechanisms of the neighbourhood area is modify, due to the given threshold. These noise pixels are at that point changes by the median pixel estimation of the pixels within the neighbourhood that have finished the noise marking assessment.

Gaussian Mixture Model

Image pixels are signified as in the form of arrays. A GMM can utilized to divide the pixels into similar segments for further analysis. The mixture models, graphical models, Markov random fields and hidden Markov models are the examples of stochastic models which have importance in probabilistic data analysis. Additionally, picture segmentation intends to partition one picture into various kinds of classes or regions, for instance an image of geometric shapes has a few classes with various colours for example 'circle', 'rectangle', 'triangle' etc.

Gray Level Co-Occurrancy Matrix (Glcm)

The progress between two pixels of gray level is utilized in gray-level co-event grid to separate the surface of the picture. It likewise give joint appropriation of neighbouring pixels of gray level combines inside a picture. For characterization of the breast tumour a descriptors from the grid is separated. For the calculation of GLCM, set up the spatial association between two pixels for example reference pixel and neighbour pixel so it is procedure as a dissimilar mixture of gray pixel standards in an image.

Probabilistic Neural Network

The **Probabilistic Neural Network (PNN)** for identifying classification and pattern recognition of given input problems is applied using a feed-forward method. In the PNN algorithm, of each class is approximately summed up by a parameter window and a non-parametric purpose where the root node i.e., parent node of Probability distribution function. At that point, using Bayesian rule utilization of probability distribution function for each class, where each class probability of a new input data is estimated and by which is then working to assign to class with highest subsequent probability to new inputted data. The chance of misclassification is will be reduced by using this kind of new technique. This form of ANN was successively subsequent from the

TEST Engineering & Management

May - June 2020
ISSN: 0193-4120 Page No. 4728-4733

Bayesian rule and a statistical algorithm is known as Kernel fisher discriminant analysis.

5. Algorithm

a) Algorithm for PNN

Step 1: Start
Step 2: Mark patterns
Step 3: While trainset is unused patterns
Step 4: Set l = random integer (1, n)
Step 5: Set k = class(x (l))
Step 6: Till x(l) used
Step 7: Mark x(l) as used
Step 8: Compute the probability value
Step 9: Choose class c1 with max probability
Step 10: Choose second class c2 with max probability
Step 11: If f (c1) (x(l)) / f (c2) (x(l)) < r (threshold value)
Step 12: and n < n_{max}
Step 12: Add new pattern
Step 14: Select new p = optimal (p)
Step 15: Stop

b) Algorithm for Optimal (P)

Step 1: Start
Step 2: Set random vector
Step 3: If cost (p + random vector) < cost (p)
Step 4: p = p + random vector + bias vector
Step 5: else if cost (p – random vector) < cost (p)
Step 6: p = p – random vector + bias vector
Step 7: else
Step 8: bias vector = ½ (bias vector)
Step 9: end if
Step 10: return p
Step 11: Stop

6. Comparison Tabel

Methods	Strategy	Results	Remarks
KNN	This algorithm is based on feature similarity. It can only classify the data but does not make any kind of statement on the unique data.	Accuracy was reduced to 80% while training on images	Accuracy was efficient but when the tested were done to images the model failed
SVM	This project is based on the dataset used in Wisconsin Breast Cancer detection on the dataset consists of .csv file that consists of recorded breast cancer data	97%. Was available for training for the text data.	Model did not show efficiency with the mammograms and histograms images utilizing in proposed project
Logistic Regression	This model gives results in the form of 0 or 1 and uses sigmoid function to process the final output	90% accuracy was found and this model is good for text data.	This model decreased accuracy and usually Logistic Regression is not proper method for the image data set
Conventional Neural Networks	This model uses breast tumour image that is cropped and the CNN algorithm is applied on them.	85% accuracy is used and the perfect is efficient with the breast image	The Proposed project uses mammography images for the prediction this model is not accurate for DM images.

7. Results and discussion

In this project I have used Matlab, R2015a, 64-bit for the experimental set up. The RGB picture consisting of size 256x256 pixels is used for experimental purpose.

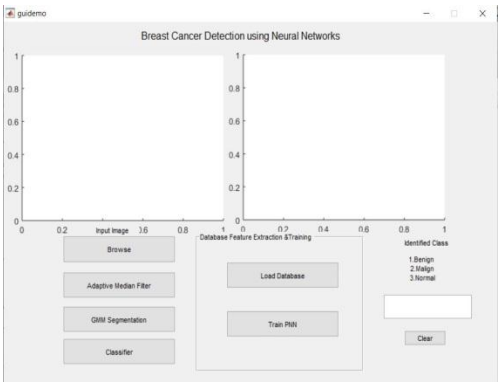


Figure 7.1: The different stages are displayed when the program is executed.

The above figure displays the format of the resulting image display. Firstly, images should be input by browsing from the database and upload the image.

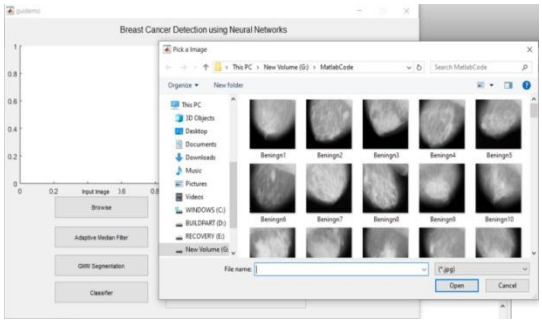
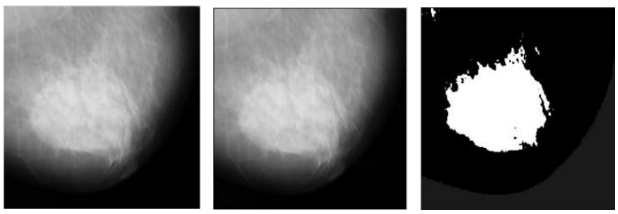


Figure 7.2: Selecting of input image from the database

May - June 2020
ISSN: 0193-4120 Page No. 4728-4733

The above figure depicts the uploading of the image from the Computer for processing and classification of the type of Disease, it will be done after processing using the algorithm Or the code.



(a) (b) (c)

Figure 7.3: a. Benign input image b. Pre-processed image c. Segmented input image

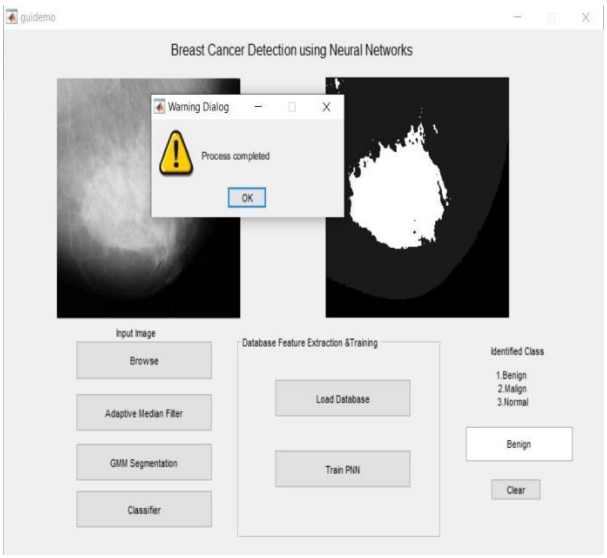


Figure 7.4: This shows the output of PNN

The above figure describes the image loaded from the database after performing segmentation, classification through using PNN. So the predicted result is displayed on the right side describing as benign image.

8. Conclusion

The breast tumour research is very challenging for many scholars for prediction and classification. Initial identification of breast tumour will help in increasing the survival chances. Several NN learning algorithms must be tested to attain the finest key in a specific problem. It is clear that none of algorithm that out-performs all the others NN algorithms in all problems that are the best of the PNN type is serious for delineating. To choose the best implemented PNN each algorithm should be executed several times with different weights. Hence, the PNN classifiers comes with the comparison analysis for the simplification ability. The goal of the classification is to know if it is cancerous (malignant) and non-cancerous (benign) tumours.

References

[1] Timothy de Moor, Alejandro Rodriguez-Ruiz, Albert Gubern Merida, Ritse Mann, and Jonas Teuwen, "Automated soft tissue lesion detection and segmentation in digital mammography using a u-net deep learning network", Conference Paper: IWBI **2018**, At Atlanta, Georgia **,2018**

[2] Salma Alqazzaz, Xianfang Sun, Xin Yang, and Len Nokes, "Mammographic mass detection based on convolution neural network" Computational Visual Media Research Article, Vol. **5(2)** June **2019** - Springer

[3] Y. J. Tan, K. S. Sim, and F. F. Ting, "Breast Cancer detection Using Convolutional Neural Networks for Mammogram Imaging System", International conference on robotics, automation and science (ICORAS),**2017**

[4] Saira Charan, Muham mad Jaleed Khan, Khurram Khurshid, "Breast Cancer Detection in Mammograms using Convolutional Neural Network", **2018** International Conference on Computing, Mathematics and Engineering Technologies-iCoMET,**2018**

[5] Shuyi Li, Min Dong, Guangming Du, and Xiaomin Mu, "Attention Dense-U-Net for Automatic Breast Mass Segmentation in Digital Mammogram", Brain Informatics Journal Springer, **2016**

[6] Abdullah-Al Nahid, Aaron Mikaelian, Yinan Kong, "Histopathological breast-image classification with restricted Boltzmann machine along with backpropagation", Biomedical Research, volume **29 (10) 2018**.

[7] Corey sutphin, Eric Olson, Yuichi Motal, Suk Jin Lee, Jae G. Kim 3, And Kazuaki Takabe, "Elastographic Tomosynthesis From X-Ray Strain Imaging of Breast Cancer", IEEE Journal of Translational Engineering in Health and Medicine, **29th** August **2019**.

[8] R. Vijayarajeswari, P. Parthasarathy, S. Vivekanandan, A. Alavudeen Basha, "Classification of mammogram for early detection of breast cancer using SVM classifier and Hough transform", Elsevier Journal, **2019**.

[9] Nadia Brancati, Giuseppe de Pietro, Maria Frucci, and Daniel Riccio, "A Deep Learning Approach for Breast Invasive Ductal Carcinoma Detection and Lymphoma Multi-Classification in Histological Images", Special Section on deep learning for computer-aided medical diagnosis, **2019**.

[10] Leyli Mahdikhani, Mohammad Reza Keyvanpour, "Challenges of Data Mining Classification Techniques in Mammograms", **5th** Conference on Knowledge-Based Engineering and Innovation, Iran University of Science and Technology, **2019**.

May - June 2020
ISSN: 0193-4120 Page No. 4728-4733

[11] Adam Yala, Constance Lehman, Tal Schuster, Tally Portnoi, Regina Barzilay, "A Deep Learning Mammography-based Model for Improved Breast Cancer Risk Prediction", Department of Electrical Engineering and Computer Science, Radiology **2019**.

[12] Said Pertuz, German F. Torres, Rulla Tamimi, Joni Kamarainen, "Open Framework for Mammography-based Breast Cancer Risk Assessment", IEEE-**2019**.

[13] Naresh Khuriwal, Nidhi Mishra, "Breast Cancer Detection from Histopathological Images Using Deep Learning", **3rd** International Conference and Workshops on Recent Advances and Innovations in Engineering, **22-25** November **2018**.

[14] M. Tahmooresi, A. Afshar, B. Bashari Rad, K. B. Nowshath and M. A. Bamiah, "Early Detection of Breast Cancer Using Machine Learning Techniques", Journal of Telecommunication, Electronic and Computer Engineering e-ISSN: **2289-8131** Vol. **10** No. **3-2**, **2018**.

[15] Samah Elsamaney Mohamed, Talaat M. Wahbi, Mohamed H. Sayed, "Automated Detection and Classification of Breast Cancer Using Mammography Images", International Journal of Science, Engineering and Technology Research (IJSETR) Volume **7**, Issue **4**, April **2018**.

[16] Matthias Kohl, Christoph Walz, Florian Ludwig, Stefan Braunewell and Maximilian Baust, "Assessment of Breast Cancer Histology Using Densely Connected Convolutional Networks", Springer International Publishing AG, **2018**.

[17] Yanfeng Li, Houjin Chen, Linlin Zhang, Lin Cheng, "Mammographic mass detection based on convolution neural network", 24th International Conference on Pattern Recognition (ICPR) **2018**.

[18] Prannoy Giri, K Saravanakumar, "Breast Cancer Detection using Image Processing Techniques", Oriental Journal of Computer Science & Technology,**2017**.

[19] Gustavo Carneiro, Jacinto Nascimento, Andrew P. Bradley, "Deep Learning Models for Classifying Mammogram Exams Containing Unregistered Multi-View Images and Segmentation Maps of Lesions", Deep Learning for Medical Image Analysis, **2017**.

[20] Neeraj Dhungel, Gustavo Carneiro, and Andrew P. Bradley, "The Automated Learning of Deep Features for Breast Mass Classification from Mammograms", Springer International Publishing **2016**.