3rd International Conference and Workshops on Recent Advances and Innovations in Engineering, 22-25 November 2018
(IEEE Conference Record # 43534)

Breast Cancer Detection From Histopathological Images Using Deep Learning

Naresh Khuriwal
Department of Computer Engineering
Poornima University,
Jaipur, India
naresh.khuriwal89@gmail.com

Nidhi Mishra
Department of Computer Engineering
Poornima University Jaipur
Jaipur, India
nidhi.mishra@poornima.edu.in

***ABSTRACT* -- The cancer is the most dangerous diseases in the world, its mainly effective for women. So,our prime target must be curing the cancer through scientific investigation and the second main target should be early detection of cancer because the early detection of cancer can be helpful for remove the cancer completed. After reviewed 41 papers we found that several techniques are available for cancer detection. In this paperwe proposedDeep Leaning algorithm convolutional neural network for diagnosed breast cancer using Mammograph MIAS database. The paper shows how we can use deep learning technology for diagnosis breast cancer using MIAS Dataset. Because deep learning techniques almost used for high task objective Computer Vision, Image processing, Medical Diagnosis, Neural Language Processing. But in this paper, we are applying deep learning technology on the MIAS Database and we have seen that is very beneficial for us for diagnosis breast cancer with accuracy 98%. This paper is divided in three parts first we have collect dataset and applied pre-processing algorithm for scaled and filter data then we have split dataset in training and testing purpose and generate some graph for visualization data. In last implement model on training dataset and achieved accuracy 98%. So, we have seen deep learning technology is a good way for diagnosis breast cancer with MIAS Dataset. This database provides200images and 12 features in the dataset. In this paper we have used 12 features for diagnosis breast cancer that we have got after pre-processing. But before train model we have applied some pre-processing algorithm like Watershed Segmentation, Colour based segmentationand Adaptive Mean Filtersfor scaled dataset then applied model and achieved accuracy. In this paper we also compare deep learning algorithm with other machine learning and seen our proposed system is proved best from others machine learning algorithm.**

Keyboards–Deep Learning, Convolutional Neural Network,Neural Network, Random Forest, Support Vector Machine, MachineLearning,MIASDataset.

I. INTRODUCTION

Breast Cancer is the most common female cancer worldwide. All most 25% of all cancers with an estimated 1.67 million new cancer cases diagnosed in 2012. Breast cancer is the most common type of cancer among women and its incidence is increasing day by days. The life time risk of developing breast cancer in women is approximately 1/8 in USA, 1/ 12 in Europe, 1/40 in Asia (WHO 2008). As per the cancerindia.org one woman dies of cervical cancer every 8 minutes in India, around 2.5 million number of people living with cancer in India and over 7 lakh new cancer patients registered in every year. India also breast cancer is the most common cancer in women and account for 27% of all cancers. As per the previous study 1,44,937 new cases registered and deaths 70,218 in 2012. But in India the incidence rates begin to rise in the early thirties and peak at ages 50-60 years. According to the World Health Organization breast cancer was responsible for 502,000 deaths in 2005 alone and 1,301,867 of new cases of breast cancer resisted. Health case authorities continuously doing efforts to overcome this merciless disease in which one of the efforts is screening. By screening the breast cancer can be detected in early stages and thus the treatment can be more effective. Many other methods also available such as mammography, ultrasound, CT and MRI. Mammography is the most widely used screening method. In our proposed method we are using Break His[A1] dataset that provided Biopsy Images generated by Mammography method. The rest of the paper is organized as follows: Section 2 presents the related work that cover 10 author research work. Section 3 includes the architecture of the proposed work .Section 4 describes the methodology which is used for the proposed work. In section 5 we discuss the model implementation of proposed work. Section 6 discusses the results and Section 7 concludes the proposed work.

II. RELATED WORK

There are many deep learning and machine learning techniques are available for cancer detection and prediction. Some of most used deep learning techniques are for breast cancer diagnosis are Convolutional Neural Network, Recurrent Neural Network and some pre-trained model Alex Net, Google Net, VGG16, VGG19, ResNet. Some most used dataset also available for training and testing are Mammogram image, SEER, UCI, WBCD,The author Abdullah-Al Nahid developed Deep Learning Model utilizing a restricted Boltzmann machine that mainly used back propagation algorithm for classify histography images [1].The author Mohamad Mahmoud Al Rahhal used deep learning algorithm convolutional neural network for classification of histogram images and found 86.60% accuracy [2]. The author Moi Hoon Yap proposed three ultrasound-based method, a patch based LeNet, U-Net and transfer learning model FCN-Alex Net and achieved 98% accuracy [3]. The author Teresa Araujo proposed deep learning-based approach for classification of haematoxylin and eosin stained breast biopsy image using CNN algorithm and achieved 83.30% accuracy [4]. The author Hongchao Song presented the empirical mode decomposition-based feature extraction method and achieved 87% accuracy [5].The author Md. Milon Islam presented a novel modality for the breast cancer detection and discussed about two supervised machine learning algorithm and achieved 95.22% accuracy [6]. The author R.D.Ghongade proposed the RF random and RF-ELM classifier for determine the breast tumour from digital mammograms and achieved 98% accuracy [7]. The author Sudarshan Nayak presented the various machine learning techniques for determine the

978-1-5386-4525-3/18/$31.00 ©2018 IEEE

breast cancer from using 3D images and achieved 98% accuracy with SVM [8]. The author Yinchong Yang proposed encoder and decoder network approaches for create physician therapy for decision [9]. The author YohannesTesehay developed a weakly supervised computer-aided detection system that used for identify prostate cancer on mpMRI [10].

III. FLOW DIAGRAM OF PROPOSED WORK

The Fig1 shows the process flow diagram or proposed work. In this paper we used MIAS Dataset that publicly available for download and used. In the second step we pre-processed the dataset. In the pre-processing we have used three different method for scaled and pre-process dataset. After that applied Deep Learning Neural Network and compute the accuracy. We have seen CNN achieved 98% accuracy.

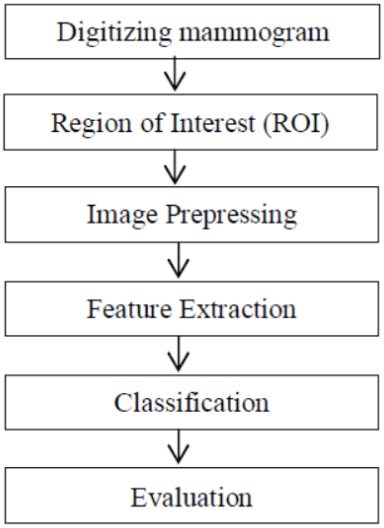

Fig. 1. Process Flow Diagram

IV. EXPERIMENT AND METHODOLOGY

In this paper we have used convolutional neural network for diagnosis breast cancer and also implemented same dataset on other machine learning algorithm such as Neural Network, Support Vector Machine, Random Forest. All paperconsists main two parts: predict models, Pre-processing data. In this paper we have used MIAS Dataset that are publicly available for researcher this database is generated from Histopathology images that having 200images of dataset. In the experiment part first, we have pre-processed the dataset and extracted useful feature using histogram and watershed model then implementsome encoder and found final training dataset. For the trained the dataset we have used deep learning model convolutional neural network algorithm and achieved 98% accuracy.

V. PRE-PROCESSING DATA

As you know that data pre-processing is a data mining technique that used for filter data in a usable format. Because the real-world dataset almost available in different format. Its not available as per our requirement so its must be fitters the dataset in aunderstandable format Data pre-processing is a proven method of resolving such issues. Data pre-processing convert the dataset in to usable format for pre-processing we have used standardization method to pre-process the MIASdataset.

A. Marker Controlled Watershed Segmentation

The watershed transforms find “Catchment basins” and “watershed ridge lines” in an image by treating it as a surface where light pixels are high and dark pixels are low. Segmentation using the watershed transform works better it we can identify or mark foreground object and background locations (Figure 1)..

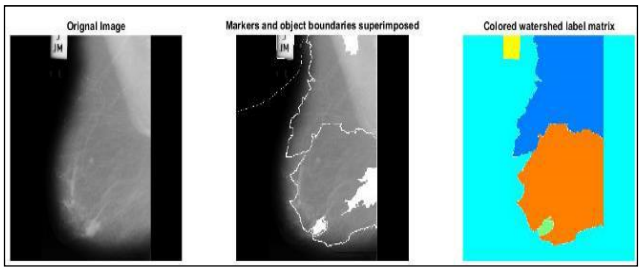

Fig. 2: Marker Controlled Watershed Segmentation.

B. Texture Segmentation Using Texture Filters

In this paper used entropy for creating texture images. Entropy is standard function that return an array value in which each output pixel contains the entropy value of the 9x9 pixel around the corresponding pixel in the input image. The following are some most used filters.:

C. Median Filter

We have used median filter in this pap 12 features that we have used formodel trained.

TABLE I. TOTAL NUMBER OF SELECTED FEATURES

Sr. No	Attributes
1	Mean
2	Standard Deviation
3	Kurtosis
4	skewness
5	Entropy
6	Energy
7	Contrast
8	Correlation
9	Homogeneity
10	Concavity Mean
11	Symmetry Mean
12	Class

We have tried different 2 segmentation algorithm on same images.

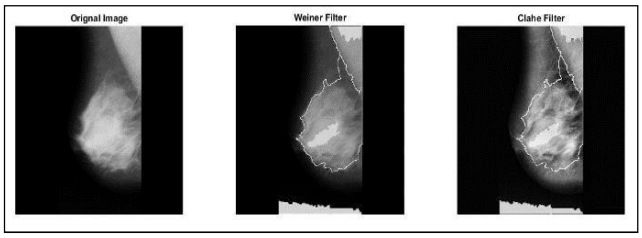

Fig 3: Weiner and Clahe Segmentation

VI. MODEL IMPLEMENTATION

In this stage we have implementedDeep Learning Neural Network algorithm on MIAS dataset.Neural network in

work on a human biological method. In which we have to pass input and get output. But in these two layer some hidden layer are work and some additional process must be added before calculate final output. These additional processes are added unit of bias, add some of additional hidden layer, calculate some activation function then the final output generated. In this paper we have used following parameters for calculating and trained model.

TABLE II. USED PARAMETERS IN CNN MODEL

Number of input	**12**
Number of Neurons	12
Activation function	Sigmoid
Number of Epochs	30

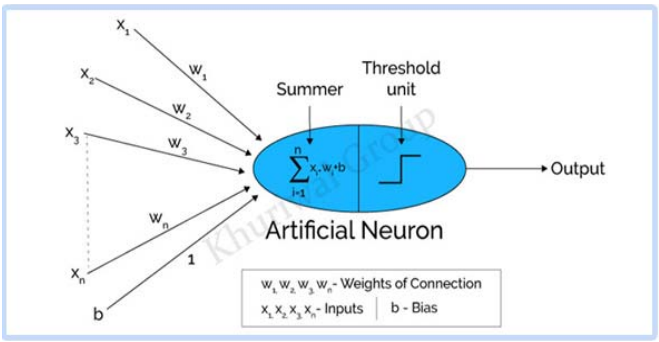

Fig 4. Shows Neural Network Functionality

When we want to use perceptron bias b = threshold then perceptron rule rewritten. Fig show the basic functionality of the neural network. In which after the fixed position the neurons are fired that is called threshold condition.

TABLE III. CNN MODEL SUMMARY

Layer Name	Type	Output Shape	Parameters
Convolutional Layer	Dense	12	496
Hidden Layer	Dense	8	136
Fully Connected Layer	Dense	1	9
Total	params:	641	
Trainable	params:	641	
Non-trainable	params:	0	

In this paper I have create the four-layer convolutional model in which first is input layer. Convolutional layer, hidden layer and fully connected layers. The total number of neurons is 12 for first input layers and 8 neurons for hidden layers and final 1 neurons in output layers. The following step are briefly describing the experimental in this research.

Step 1: the region of interest was extracted from the abnormal images depending on the images.
Step 2: for improve image quality we have used median filter and histogram equalization method.
Step 3: for classification image we have used deep learning algorithm convolutional neural network.

VII. RESULTS AND DISCUSSION

As mention above we have used convolutional neural network algorithm for diagnosis and detection of breast cancer. We have used median filter and histogram equalization method for pre-processing breast cancer dataset. Then we have implemented neural convolutional neural network algorithm on these 12 features and achieved 98% accuracy. MIAS Dataset have contain 200 images in which we have extracted 12 features and achieved 98% accuracy is good from others compared machine learning algorithm..

VIII. CONCLUSION AND FUTURE WORK

As we know that deep learning method convolutional neural network mostly used for image dataset classification that why we used convolutional neural network in this paper. After the implementation this paper we have achieved 98% accuracy. As we mention this paper worked on only 12 features only. In future we will try with new features and also try with the real images dataset so that we can achieved best result and accuracy for diagnosis the cancer. In feature we will also try this method on different type in cancer not only for breast cancer.

REFERENCES

[1] Abdullah-Al Nahid, Aaron Mikaelian and Yinan Kong, Histopathological breast-image classification with restricted Boltzmann machine along with backpropagation, Biomedical Research Volume 29, Issue 10, (2018).

[2] [24] Mohamad Mahmoud Al Rahhal, "Breast Cancer Classification in Histopathological Images using Convolutional Neural Network" International Journal of Advanced Computer Science and Applications (IJACSA), 9(3), 2018.

[3] M. H. Yap et al., "Automated Breast Ultrasound Lesions Detection Using Convolutional Neural Networks," in IEEE Journal of Biomedical and Health Informatics, vol. 22, no. 4, pp. 1218-1226, July 2018.

[4] Araújo T, Aresta G, Castro E, Rouco J, Aguiar P, Eloy C, et al. (2017) Classification of breast cancer histology images using Convolutional Neural Networks. PLoS ONE 12(6): e0177544. https://doi.org/10.1371/journal.pone.0177544.

[5] H. Song, A. Men and Z. Jiang, "Breast tumor detection using empirical mode decomposition features," in IEEE Access.

[6] M. M. Islam, H. Iqbal, M. R. Haque and M. K. Hasan, "Prediction of breast cancer using support vector machine and K-Nearest neighbors," 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), Dhaka, 2017, pp. 226-229.

[7] R. D. Ghongade and D. G. Wakde, "Detection and classification of breast cancer from digital mammograms using RF and RF-ELM algorithm," 2017 1st International Conference on Electronics, Materials Engineering and Nano-Technology (IEMENTech), Kolkata, 2017, pp. 1-6.

[8] S. Nayak and D. Gope, "Comparison of supervised learning algorithms for RF-based breast cancer detection," 2017 Computing and Electromagnetics International Workshop (CEM), Barcelona, 2017, pp. 13-14.

[9] Y. Yang, P. A. Fasching and V. Tresp, "Predictive Modeling of Therapy Decisions in Metastatic Breast Cancer with Recurrent Neural Network Encoder and Multinomial Hierarchical Regression Decoder," 2017 IEEE International Conference on Healthcare Informatics (ICHI), Park City, UT, 2017, pp. 46-55.

[10] Y. Tsehay et al., "Biopsy-guided learning with deep convolutional neural networks for Prostate Cancer detection on multiparametric MRI," 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), Melbourne, VIC, 2017, pp. 642-645.

[11] A. Alzubaidi, G. Cosma, D. Brown and A. G. Pockley, "Breast Cancer Diagnosis Using a Hybrid Genetic Algorithm for Feature Selection Based on Mutual Information," 2016 International Conference on Interactive Technologies and Games (ITAG), Notthingham, 2016, pp. 70-76.

[12] A. I. Pritom, M. A. R. Munshi, S. A. Sabab and S. Shihab, "Predicting breast cancer recurrence using effective classification and feature

selection technique," 2016 19th International Conference on Computer and Information Technology (ICCIT), Dhaka, 2016, pp. 310-314.

[13] A. Osareh and B. Shadgar, "Machine learning techniques to diagnose breast cancer," 2010 5th International Symposium on Health Informatics and Bioinformatics, Antalya, 2010, pp. 114-120.

[14] A. Qasem et al., "Breast cancer mass localization based on machine learning," 2014 IEEE 10th International Colloquium on Signal Processing and its Applications, Kuala Lumpur, 2014, pp. 31-36.

[15] B. M. Abed et al., "A hybrid classification algorithm approach for breast cancer diagnosis," 2016 IEEE Industrial Electronics and Applications Conference (IEACon), Kota Kinabalu, 2016, pp. 269-274.

[16] B. M. Gayathri and C. P. Sumathi, "Comparative study of relevance vector machine with various machine learning techniques used for detecting breast cancer," 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Chennai, 2016, pp. 1-5.

[17] C. Deng and M. Perkowski, "A Novel Weighted Hierarchical Adaptive Voting Ensemble Machine Learning Method for Breast Cancer Detection," 2015 IEEE International Symposium on Multiple-Valued Logic, Waterloo, ON, 2015, pp. 115-120.

[18] D. Bazazeh and R. Shubair, "Comparative study of machine learning algorithms for breast cancer detection and diagnosis," 2016 5th International Conference on Electronic Devices, Systems and Applications (ICEDSA), Ras Al Khaimah, 2016, pp. 1-4.

[19] D. T. Saleh, A. Attia and O. Shaker, "Studying combined breast cancer biomarkers using machine learning techniques," 2016 IEEE 14th International Symposium on Applied Machine Intelligence and Informatics (SAMI), Herlany, 2016, pp. 247-251.

[20] H. R. Mhaske and D. A. Phalke, "Melanoma skin cancer detection and classification based on supervised and unsupervised learning," 2013 International conference on Circuits, Controls and Communications (CCUBE), Bengaluru, 2013, pp. 1-5. [3]