An Improved Prediction of Breast Cancer using Deep Neural Networks (DNN)

K. Naga Durga [1] | Dr. K. V. Krishnam Raju [2]

[1]Department of CSE, S.R.K.R Engineering College, Bhimavaram, India,
[2]Associate Professor, Department of CSE, S.R.K.R Engineering College, Bhimavaram, India,.

Abstract: Cancer is a most dangerous disease that can be affected by any person. There is no specific reason to occur cancer for any human. In recent days one of the tools which are used to predict breast cancer is histopathological imaging. Increasing in abnormal tissues can cause cancer. Various traditional algorithms are present to predict breast cancer. Algorithms such as CNN, ANN, and RCNN are applied to the given dataset. Still, there are a lot of issues in showing accurate results. Among different cancers, breast cancer (BC) is most dangerous cancer that grow the abnormal tissues in the breast. In this paper, an improved prediction of breast cancer is developed to analyze the various patient samples.

KEYWORDS: Breast cancer, DCNN, Break His Dataset.

DOI of the Article: https://doi.org/10.46501/IJMTST0707005

Available online at: http://www.ijmtst.com/vol7issue07.html

As per UGC guidelines an electronic bar code is provided to seure your paper

To Cite this Article: K. Naga Durga and Dr. K. V. Krishnam Raju. An Improved Prediction of Breast Cancer using Deep Neural Networks (DNN). *International Journal for Modern Trends in Science and Technology* 2021, *7*, 0706221, pp. 25-30. https://doi.org/10.46501/IJMTST0707005

Article Info. Received: 2 May 2021; Accepted: 28 June 2021; Published: 4 July 2021

INTRODUCTION

INTRODUCTION

Cancer is a disease that can form abnormal cells within the human body. Every cell in the body manages the cycle of the regeneration process. Woman who crossed years may chances of attacking breast cancer. The death of this disease is very high if proper treatment is not taken. If abnormal situations occur, few cells may convert to cancerous cells. This is more difficult to detect the cancerous cells within the breast [6] [7] [8]. In human body many types of cancers may cause, breast cancer becomes more complex and becomes complicated and it may lead to death. Based on biological organs of the human body, women may have high chances of effecting breast cancer than men. A breast cancer can occur based on the age, history of a family, the density of the breast, intake of alcohol, and other conditions that can cause breast cancer.

For detecting the breast cancer many researchers found various techniques, algorithms and methods are proposed. Among those Deep Learning (DL) and machine learning algorithms are most widely applied on breast cancer. In recent years, various classification algorithms are applied to different medical-related datasets to perform real-time estimated analysis of the patients and their medical reports [9] [10] [11]. Machine learning (ML) Algorithms used to detect the abnormal behaviour of cancer-affected areas based on the tumor size and density. Effective training and testing are very important to detect breast cancer. Previously ML algorithms such as J48, NB, and SVM are used to detect the accuracy of breast cancer [12] [13] [14]. The proposed algorithm is used to improve the classification of breast cancer detection. And this algorithm is most compatible and can process the missing values [15] [16] also.

REVIEW OF LITERATURE

Researchers proposed various classification algorithms of breast cancer prediction. From the research point of view the breast cancer can be predicted by using images and datasets. Different types of methods, algorithms and techniques are proposed by many authors. One of the most complex and tedious task is analysis of medical images and finding the better results. Authors proposed many latest methods to predict the breast cancer and solves different issues that occur at the time of image classification by uisng deep learning algorithms.

Bazazeh et al., [1] proposed SVM, RF and BN to check the breast cancer. The author applied SVM on Wisconsin breast cancer dataset (WBCD) and shown the better performance. The overall accuracy of proposed methodology is 97.5% to 99.2%. It is observed that RF shows the better performance and it is having highest chances of classifying tumour.

Murugan et al.,[2] proposed the new algorithms such as Linear Regression (LR), Decision Tree (DT) and RF to predict the breast cancer and also for classification. These algorithms are applied on UCI machine learning repository. The proposed algorithm analyze the tumor in the breast and knows the status of the cancer weather it is benign or malignant. This system also predicts weather the disease can be cured or uncured stage. The result of this classification is 84.14% and 88.14% is percentage of prediction. In future, the pre-processing of data can show highest performance and also improving the success rate within the prediction and classification of breast cancer.

Ghosh et al.,[3]proposed the various improved approaches to predict the breast cancer. This is efficient breast cancer classification which is applied on WBCD. and this shows the classification based on the different analysis by using deep learning and statistical analysis by using different attributes. The comparison result shows between the NB, SVM, LR, KNN, RF, MLP and CNN. The results show that RF shows the improved performance based on the classifier in terms of better classifier with 50% of testing data and 50% of training data and the accuracy is 96.78%.

Gupta et al.,[4] proposed LR, RF, MLP and DT to diagnosis breast cancer. All these algorithms are applied on WBCD and achieved 97.7%. The comparative results are show on proposed machine learning algorithms. MLP, SVM, KNN and DT are applied on WBCD. The parameters such as accuracy is shown in this, when compare with other algorithms MLP shows the better performance.

Nemissi et al.,[5] developed an enhanced Extreme Learning Machine(ELM) based-Neural Network. To performance is analysed by the ELM and experiments are carried out by using WBCD. The comparison results are shown between the ELM and other traditional algorithms. Accuracy is 97.56% and the main advantage

of this algorithm is ELM consists of huge number of hidden neurons.

Rashed et al.,[6] proposed a new diagnosis approach for detection of breast cancer. The new architecture is developed by adopting various features from the U-net structure which is used to detect the breast cancer in early stages and effectively. Performance is calculated in terms of sensitivity and specificity which can be used for the clinical use. The accuracy of the proposed system which is micro classification and masses are 94.31% and 95.01%. To improve the performance of the proposed system they adopted the better structures such as Alex Net, VGG Net, Google Net. Because of this combination of these algorithms, the results are more improved and performance is high.

Sun et al.,[7] proposed a new mammographic image classification which uses the Convolutional neural networks (CNN). CNN is most widely used to exploit the basic data from various views of mammography. The new system retrieve the various significant features from the proposed method which is evaluated by the comparison between existing method and proposed method applied on open Digital Database for Screening Mammography (DDSM) and Mammographic Image Analysis Society (MIAS) datasets. Accuracy for this method is 72% to 84%. The method obtains the best performance compare with other existing methods.

Senkamalavalli et al.,[8] utilized the hybrid techniques to classify the breast cancer. For the early diagnosis of breast cancer, these algorithms are merged such as AdaBoost, SVM classification algorithm merged with K-means. To get the accurate results the missing data, imbalance of data and other typical cases are considered. By using these hybrid algorithms the accuracy is attained. The overall accuracy is 96% to 99%. This proposed algorithm is focused on applying classifying the images into two types such as tumour and normal images.

Khourdifi et al.,[9] discovered a new algorithm for detecting early stages of cancer cells in human body that uses the features of algorithms such as RF, SVM, KNN, NB. Among all the algorithms SVM achieved huge performance with accuracy of 97.9% and this is applied on University of WBCD. Every sample consists of 11 attributes. From this 2 to 10 attributes are selected for instances. Total number of cases is 699. The author explained the various ML algorithms and their applications based on diagnosing the breast cancer which is applied on WBCD. In future, various deep learning algorithms are combined which is applied on different datasets.

1. DEEP COVOLUTION NEURAL NETWORK ALGORITHM

Improved DCNN

A Deep Learning Model is called deeper when you have a large number of Layers. So, Deep CNN is basically CNN with deeper layers. In Regular CNN, there are usually 5-10 layers. While most modern architectures are 30-100 layers in Deep.

Algorithm

The implementation is totally based on Keras library function that supports python programming language, which is names as "CancerNet"

1. 3×3 CONV filters are used specifically and these are similar to VGGNet.
2. 3×3 CONV filters are on the top and these performs max-pooling for each other.
3. The Depth wise separable convolution can be used in place of standard convolution layers.

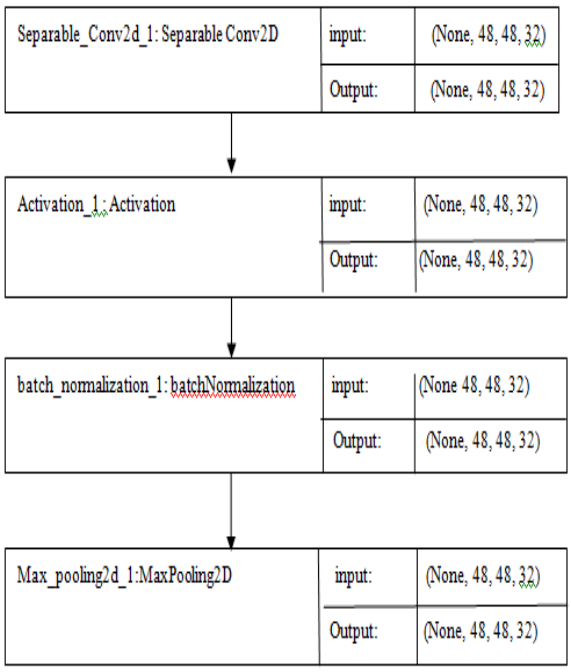

Fig 1: Project Flow by using DCNN

METHODOLOGY

The aim of the proposed algorithm is to classify the breast cancer prediction by using various images that are available from various sources. These images are

used for binary and multiclass classification. The breast cancer is predicted by proposed system and this is applied on histopathological images. The proposed algorithm is applied on BreakHis dataset. The proposed algorithm is as follows:

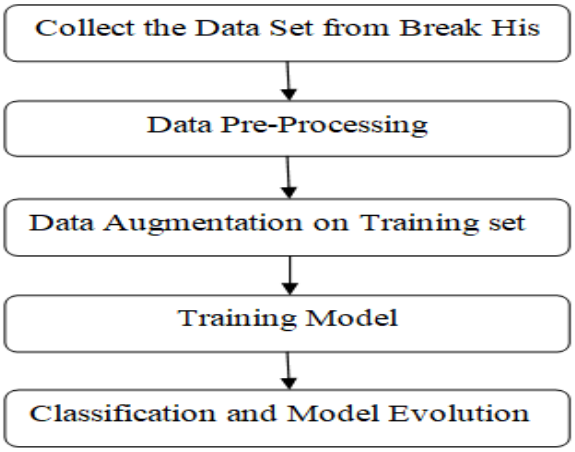

Fig 2: Breast cancer classification framework

The first step starting with the BreakHis dataset is used for the pre-processing of dataset to increase the quality of the result. After this step, the data is randomly selected for training set and for testing set. The next step follows the data augmentation for the training dataset. This step applies various image processing techniques such as image resizing, rotation and reflection images. After the training is completed the testing is also started and the proposed system is applied on text dataset and performance is calculated based on the classification accuracy and confusion matrices.

Architecture Diagram

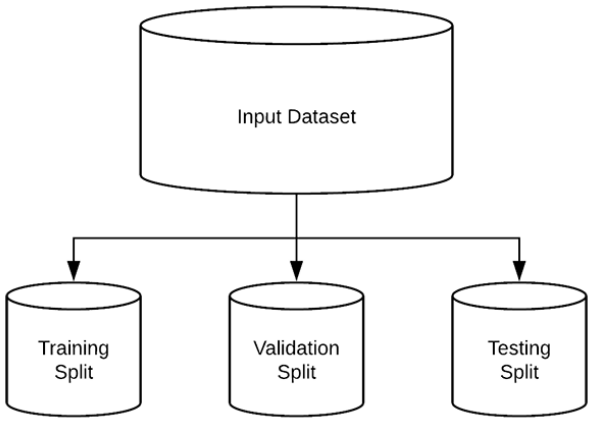

Fig 3: Architecture Diagram

RESULTS AND DISCUSSIONS

Experimental Results

The implementation is done by using a python programming language. The proposed methodology performs effectively to get better results. Experiments are conducted on 2112 images and this is downloaded from Break His Dataset. The following are performance metrics that are calculated as follows:

False Positive Rate (FPR)

The percentage of abnormalities was classified as

$$FPR=\frac{FP}{FP+TN}$$

False Negative Rate (FNR)

The percentage of abnormalities with no predicted cancerous cells

$$FNR=\frac{FN}{FN+TN}$$

Sensitivity

The proportion of original positives that are efficiently used to calculate the sensitivity.

$$Sensitivity=\frac{No.of\ TP}{No.of\ TP+No.of\ TN}$$

Specificity

The proportion of negatives that are correctly identified is the measure of specificity.

$$Specificity=\frac{No.of\ TN}{No.of\ TN+No.of\ FP}$$

Accuracy: This will calculate the overall accuracy of the clusters.

$$Accuracy=\frac{TP+TN}{TP+TN+FP+FN}$$

Accuracy for 40 epochs:

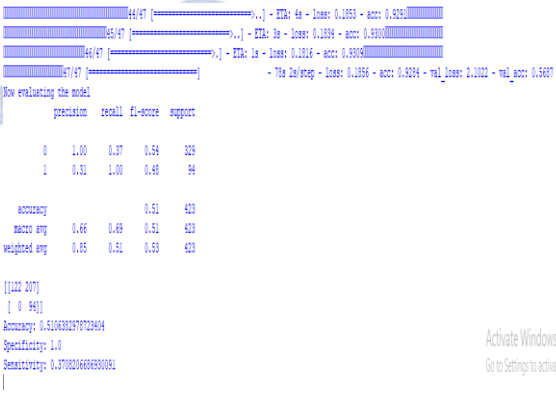

Fig 3: Result analysis based on the proposed methodology

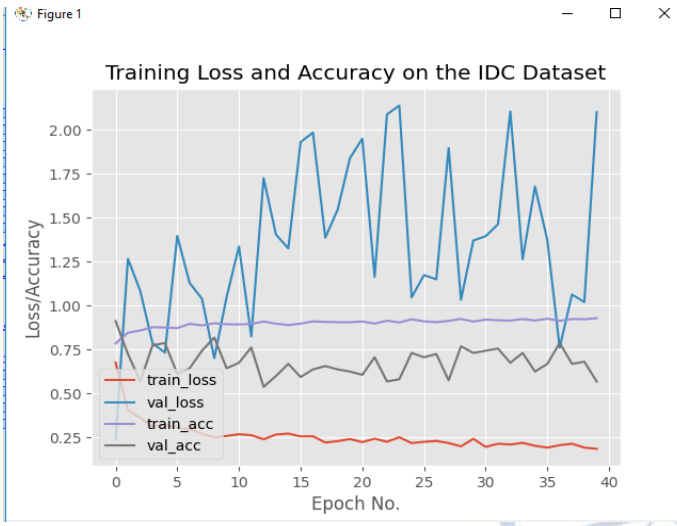

Fig 4: Training loss and Accuracy for 40 epochs is 92.84%

CONCLUSION

Our proposed approach of DCNN classification for breast cancer tendencies in a person could be predicted with higher accuracy than the current standards. The performance is calculated based on improved results and these parameters are accuracy, precision, sensitivity (recall), specificity, and F1-score for every selected input image. For this the confusion matrices is generated and binary and multiclass classification is done. Since the complexity of medical images, we applied the transfer learning technique based on replacing the final three layers of the three pre-trained networks, and the three layers must fine-tune for the new classification with the state-of-the-art results for binary classification, our DCNN achieved an excellent overall average accuracy, and for multiclass classification, ResNet18 achieved the second-highest performance.

FUTURE WORK

In future, it is very important to detect the breast cancer or any other tumor based cancers in the early stages. An Ensemble Classification is required to process the each and every attribute with enhanced feature extraction methods. This system should be more compatible because of this can be applied on different types of cancerous datasets. It will interestingly be to extend our results to patient-based classification and to obtain higher accuracy than the present state-of-the-art techniques. We are also planning to extend our system to make it capable of tackling challenging breast cancer grading problems.

REFERENCES

1. Bazazeh, D., & Shubair, R. (2016, December). Comparative study of machine learning algorithms for breast cancer detection and diagnosis. In *2016 5th ICEDSA,* (pp. 1-4). IEEE.
2. Murugan, S., Kumar, B. M., & Amudha, S. (2017, September). Classification and Prediction of Breast Cancer using Linear Regression, Decision Tree and Random Forest. In *2017 CTCEEC*, pp. 763-766.
3. Shahnaz, C., Hossain, J., Fattah, S. A., Ghosh, S., & Khan, A. I. (2017, December). Efficient approaches for accuracy improvement of breast cancer classification using wisconsin database. In *2017 R10-HTC,* pp. 792-797.
4. Gupta, M., & Gupta, B. (2018, February). A comparative study of breast cancer diagnosis using supervised machine learning techniques. In *2018 ICCMC,* pp. 997-1002.
5. NEMISSI, M., SALAH, H., & SERIDI, H. (2018, November). Breast cancer diagnosis using an enhanced Extreme Learning Machine based-Neural Network. In *2018 SIVA,* pp. 1-4.
6. Rashed, E., & El Seoud, M. S. A. (2019, April). Deep learning approach for breast cancer diagnosis. In *Proceedings of the 2019 8th ICSIE,* pp. 243-247.
7. Sun, L., Wang, J., Hu, Z., Xu, Y., & Cui, Z. (2019). Multi-View Convolutional Neural Networks for Mammographic Image Classification. *IEEE Access*, *7*, 126273-126282.
8. Senkamalavalli, R., & Bhuvaneswari, T. (2017). Improved classification of breast cancer data using hybrid techniques. *IJAERS*, *5*(5).
9. Khourdifi, Y., & Bahaj, M. (2018, December). Applying Best Machine Learning Algorithms for Breast Cancer Prediction and Classification. In *2018 ICECOCS,* pp. 1-5.
10. Khuriwal, N., & Mishra, N. (2018, March). Breast cancer diagnosis using adaptive voting ensemble machine learning algorithm. In *2018 IEEMA eTechNxT,* pp. 1-5.
11. Tahmooresi, M., Afshar, A., Rad, B. B., Nowshath, K. B., & Bamiah, M. A. (2018). Early Detection of Breast Cancer Using Machine Learning Techniques. *JTEC*, *10*(3-2), 21-27.
12. Abed, B. M., Shaker, K., Jalab, H. A., Shaker, H., Mansoor, A. M., Alwan, A. F., & Al-Gburi, I. S. (2016, November). A hybrid classification algorithm approach for breast cancer diagnosis. In *2016 IEEE IEACon,* pp. 269-274.
13. Antony, S. J. S., & Ravi, S. (2013, December). Development of efficient image quarrying technique for Mammographic image classification to detect breast cancer with supervised learning algorithm. In *2013, ICACCS,* pp. 1-7.
14. Hela, B., Hela, M., Kamel, H., Sana, B., & Najla, M. (2013, March). Breast cancer detection: A review on

mammograms analysis techniques. In *10th SSD13*, pp. 1-6.

15. Graja, O., Azam, M., & Bouguila, N. (2018, November). Breast cancer diagnosis using quality control charts and logistic regression. In *2018 9th ISIVC*, pp. 215-220.
16. Bharat, A., Pooja, N., & Reddy, R. A. (2018, October). Using Machine Learning algorithms for breast cancer risk prediction and diagnosis. In *2018 3rd I4C*, pp. 1-4.