International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 06 | June - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

A Review Paper on Breast Cancer Detection Using Deep Learning

Abhishek Borse[1]
Department of Artificial Intelligence & Machine Learning
TGPCET, Nagpur

Prof. Mukul Pande[2]
Department of Artificial Intelligence & Machine Learning
TGPCET, Nagpur

***Abstract*—. According to a study Breast Cancer (BC) is the most common cancer type in the category of cancer, it is the most leading malignancy affecting many women each year. Detection of cancer in the early stage reduces the death rate. There are many different approaches to detect early breast cancer using machine learning algorithms. Some of the common approaches for detection of breast cancer are using SVM, KNN, Naive Bayes etc. to give better performance in their own way. But nowadays, new technology has been introduced to classify breast cancer, i.e deep learning. Deep learning is used to overcome the drawbacks of machine learning. Basically our research is based on the image classification using Convolutional neural networks. For this we can use the biopsy slide images and clinical data of patients for prediction of BC in the early stage. For the prediction of breast cancer in the early stage, one needs to extract features from the patches of biopsy slide images using Deep Convolutional Neural Network. In the present paper, reviews of all authors are conducted.**

***Keywords*—Breast Cancer, Machine Learning, Deep Learning, Convolutional Neural Network, SVM, KNN, Naive Bayes etc.**

I. Introduction

Cancer arises when the abnormal body's cells start to separate and come in contact with normal cells and make them malignant. Breast cancer is one of the major causes of death in women around the world. According to a recent paper published by the American Cancer Society, 41,760 women and more than 500 men died from breast cancer. Breast cancer occurs in four main types: normal, benign, in-situ carcinoma and invasive carcinoma. A benign tumor involves a minor change in the breast structure. In cases of in-situ carcinoma, the cancer is only in the mammary duct lobule system and does not affect other organs. This type is not dangerous and can be treated if diagnosed early. Invasive carcinoma is considered to be the most dangerous type of breast cancer, as it can spread to all other organs. There are several methods available to detect breast cancer like X-ray, Ultrasound, Mammography (UM), Computed Tomography (CT), Mag- netic Resonance Imaging, Biopsy Slides and breast temper- ature measurement. Usually, the golden standard is a pathological diagnosis for detecting breast cancer. This involves an image analysis of the removed tissue, which is stained in the lab to increase visibility. There are four types of breast cancer. First type of cancer is Ductal Carcinoma in Situ that is found in the coating of breast milk ducts and it is pre-stage breast cancer. Second type of breast cancer is the most popular disease and contains up to 70-80% diagnosis. Third type of breast cancer is Inflammatory breast cancer which is forcefully and quickly developing breast cancer in this disease cells penetrate the skin and lymph vessels of the breast. The fourth type of breast cancer is Metastatic breast cancer which spreads to other parts of the body.

There are many diagnosis tests like Mammogram, Ultrasound, MRI and Biopsy that provide the images for the classification. Mammogram is used to test breast cancer with the help of X-ray. While screening mammograms. Ultrasound is done after the mammogram if any suspicious outcomes are found then inform the doctor to test the tissues. When the suspicious site is detected in your breast then the doctor demands an ultrasound test. During symptomatic examination, if the tests are not convincing then the doctor prefers the breast MRI. It shows the picture and point of view of your illness. A biopsy is the main symptomatic system that can decide whether the suspicious region is carcinogenic. Fortunately 80% of ladies who have a breast biopsy don't have breast malignant growth.

There are two approaches, one is Machine Learning and other is Deep Learning. Machine learning plays an important role for the classification of breast cancer. There are many diagnosis processes that have discussed the images. These types of diagnosis images are used for classification using machine learning. Machine learning is a sub-field of AI. Many developers use machine learning to re-train the existing models and for better performance. Machine learning is used for linear data. If the data is small then machine learning gives better results but when the data is too large then it doesn't give the better results. Deep learning is a sub-field of machine learning. Deep is an unsupervised learning that learns from the data. The data may be unstructured or unlabeled. Deep neural network contains more than two hidden layers then it is called a deep network. Basically, the first layer is the input layer and the second is the output layer. The intermediate layer is called a hidden layer that has more layers as compared to a neural network. The node containing the layer is called neurons. The difference between machine learning and deep learning is that deep learning is closer to its goal as compared to machine learning. For the classification of the breast cancer dataset, Convolution Neural Network is used. Convolutional Neural

Network is used to classify the images. It takes the images of the breast cancer dataset as an input. CNN takes the images as an input associated with their corresponding weights. The weights are adjusted to minimize the error and enhance the performance. CNN contains many layers such as convolution layer, pooling layer, ReLU layer and fully connected layer. In the convolution layer, a feature map is used to extract the features of the given image and makes the original image more compact. Pooling layer is used to reduce the dimensions of the image. ReLU layer is used as an activation function in which it checks if the value of the activation function lies in a given range or not. Fully connected layer is the last layer of the model. It combines the results of all layers and applies the softmax function to give the probability to each class of the output.

II. Literature Review

This section gives the information about the related work of the research that has been already done. Basically two techniques are used to detect breast cancer. First one is machine learning and the second is deep learning. There is a lot of research that is conducted through machine learning. But machine learning techniques have some problems that are removed through deep learning.

1. Prediction of Breast Cancer Using SVM Algorithm[1]

In this, author proposed an approach using Support Vector Machine (SVM) on different datasets like Wisconsin Breast Cancer (WBC) dataset the Wisconsin Diagnostic Breast Cancer (WDBC) dataset, the Surveillance, Epidemiology, and End Results (SEER) dataset. The outcomes of SVM consist of accuracy and precision. It is used for classification, which trains models to categorize cancer patients according to their diagnosis. The system also focuses on providing a detailed description of the tumor for each patient diagnosed. Here the author uses K-fold cross validation technique for data validation, where the accuracy of various machine learning algorithms on the data is tested.survival. The proposed system classifies the tumors into malignant or benign using features obtained from cell images with 99% accuracy.

2. Hybrid Approach to predict Breast Cancer using Machine Learning Techniques[2]

In this, the author proposed a model based on a hybrid approach using machine learning. It implemented this approach using MRMR feature selection with four classifiers to find out the best results. The author used the four classifiers SVM, Naïve Bays, Function tree and End Meta and did the comparison between all. It found that SVM was a good classifier. To find out the better results. This model is being trained and tested on datasets like Wisconsin Diagnostic Breast Cancer dataset, Wisconsin Breast Cancer Data and Breast Tissues data. The proposed system classifies the tumors into malignant or benign using features obtained from cell images with 98% accuracy.

3. An Automatic Detection of Breast Cancer Diagnosis and Prognosis Based on Machine Learning Using Ensemble of Classifiers

In this, the author proposed an approach where we build an architecture using 4 different ML models. This architecture is composed of four different ML-based classifiers named SVM, LR, NB, and DT. They are stacked and then further trained as an ensemble. After training, the ANN model is used for the outcome. The performance is compared with the several ML classifiers individually with and without up- sampling techniques. We also compared the performance of the proposed ensemble model with other ensemble models. An ensemble model is stacked, and predictions are concatenated and then fed to the ANN model for final prediction. Here we have analyzed classification results using a 10-fold cross-validation technique. and this experiment is performed on the Breast cancer Wisconsin (Diagnosis) and Breast cancer Wisconsin (Prognosis) databases.

4. Breast Cancer Classification From Histopathological Images Using Patch-Based Deep Learning Modeling

In this study, the author proposed a Pa-DBN-BC model for the classification of breast cancer on the histopathology images. The Pa-DBN-BC model comprises four main phases which are the preprocessing, patch generation, DBN, and classification phase where the model works on equal size patches of images. In preprocessing we can crop our input data around the region of interest (ROI) without worrying about the limitation of keeping the images of equal sizes. The proposed model works in an unsupervised fashion for the extraction of features from the input histopathology image patches in the form of feature vectors. The extracted features matrix is then transferred to the backpropagation neural network which is a supervised learning phase and it comprises the conjugate gradient. A model is formed by the feature matrix of images in the training phase and the final stage is the classification stage which discriminates between cancerous and non-cancerous regions. a publicly available dataset based on histopathology images is used. The dataset includes histopathology images from the four different data cohorts, Hospital of the University of Pennsylvania (HUP), Case Western Reserve University (CWRU), Cancer Institute of New Jersey (CINJ), and The Cancer Genome Atlas (TCGA) and their corresponding binary masks of invasive breast cancer regions annotated by pathologists. In this experiment the model has achieved 86% of the accuracy.

5. Breast Cancer–Detection System Using PCA, Multilayer Perceptron, Transfer Learning, and Support Vector Machine[5]

In this study, the author has proposed an approach of combining 4 different approaches to achieve classification. approaches are Principal Component Analysis (PCA), Multilayer Perceptron (MP), Transfer Learning (TL), and Support Vector Machine (SVM). The experiment performed k-fold cross-validation 50 times on average to enhance the generalizability of the model. and the highest accuracy achieved during 10-fold cross-validation was 86.97%, as applied to the BCCD data set.

6. Optimized Deep Neural Networks Architecture Model For Breast Cancer Diagnosis

In this study, the author proposed Deep Neural Network (DNN) architecture for breast cancer prediction. Here the author has used the Mammographic Mass dataset obtained from UC Irvine Machine Learning Repository .In this experiment, three types of NN models were used: NN1HL (Neural Networks with One Hidden Layer), DNN4HL (Deep Neural Networks with Four Hidden Layers), and DNN8HL (Deep Neural Networks with Eight Hidden Layers) additionally with dropout layer. ReLu is used as an activation function on all the models and sigmoid is used as output activation function. In this experiment the author has achieved 91% accuracy.

III. Discussion of Findings from Literature

From the above information it is concluded that the deep learning technique gives better results as compared to machine learning. The results are calculated in the different dataset that have different outputs. For the tabular dataset, the machine learning techniques give better results but in the images dataset it doesn't give better results. The research work is conducted using some enhancement that is a way for better performance. Some augmentation in the dataset is also leading to better performance. In this section, the research work is conducted using enhancement and augmentation in the dataset. It concluded from the [1] SVM is a good classifier. similarly [2] that SVM was a good classifier and compared with the hybrid techniques. It concluded from [3] that we can enhance the performance, ensembling the classifiers performed better among all used. From [4] and [5], we saw *Histopathological Images Using a Patch-Based Deep Learning Modeling approach and combining PCA, MP, TL and SVM.* The author proposed the model on CNN and compared it with the machine learning techniques and achieved better performance through CNN.

IV. Conclusion

Breast cancer detection is a challenging problem because it is the most popular and harmful disease. Breast cancer is growing every year and there is less chance to recover from this disease. For detection of breast cancer, machine learning and deep learning techniques are used. It is concluded from the previous research, the machine learning techniques give better results in their own field. The previous research is conducted through many machine learning techniques with some enhancement and augmentation in the dataset for better performance. But it is concluded that machine learning gives better results on linear data. It is also concluded from the previous research, when the data is in the form of images where the machine has failed. To solve the problem of machine learning techniques, an innovative technique is used. Deep learning is a recently developed technique that is frequently used in data science. For the classification of the breast cancer images data, a deep learning based technique CNN is used. CNN mostly works on the images dataset. In the previous research, it is also concluded that CNN gives better results as compared to machine learning techniques.

V. References

[1] Prediction of Breast Cancer Using SVM Algorithm Deepika S and Kapilaa Ramanathan Devi N Department of Information Technology Sri Venkateswara College of Engineering, Sriperumbudur, Tamil Nadu, India – 602117 Assistant Professor Department of Information Technology Sri Venkateswara College of Engineering, Sriperumbudur, Tamil Nadu, India - 602117

[2] Hybrid approach to predict breast cancer using machine learning techniques," International Journal of Computer Science Engineering, vol. 5, no. 3, pp. 125- 136, 2016.

[3] An Automatic Detection of Breast Cancer Diagnosis and Prognosis Based on Machine Learning Using Ensemble of Classifiers Usman Naseem, Junaid Rashid Department of Information Technology, Sydney International School of Technology and Commerce, Sydney, NSW 2000, Australia

[4] Breast Cancer Classification From Histopathological Images Using Patch-Based Deep Learning Modeling Irum Hirra, Mubashir Ahmad Department of Computer Science and IT, The University of Lahore, Sargodha Campus, Sargodha 40100, Pakistan

[5] Breast Cancer–Detection System Using PCA, Multilayer Perceptron, Transfer Learning, and Support Vector Machine Huan-Jung Chiu, Tzuu-Hseng,Ping-Huan Kuo AI Robots Laboratory, Department of Electrical Engineering, National Cheng Kung University, Tainan 70101, Taiwan Department of Intelligent Robotics, National Pingtung University, Pingtung 90004, Taiwan

[6] Optimized Deep Neural Networks Architecture Model For Breast Cancer Diagnosis G. Kanimozhi, P. Shanmugavadivu Department of Computer Science and Applications, The Gandhigram Rural Institute (Deemed to be University), Gandhigram, Tamil Nadu, India.

[7] Using convolutional neural networks for the classification of breast cancer images. Eric Bonnet Centre National de Recherche en G´enomique Humaine Institut de Biologie Fran ̧cois Jacob Commissariat `a l'Energie Atomique Universit´e Paris–Saclay, Evry, France eric.bonnet@cea.fr June 2021

[8] Predicting Axillary Lymph Node Metastasis in early breast cancer using deep learning on primary tumor biopsy slides. Feng Xu, Chuang Zhu, Wenqi Tang, Ying Wang, Yu Zhang , arXiv:2112.02222v4 [eess.IV] 8 Jun 2022.

[9] Breast Cancer Prediction using Supervised Machine Learning Algorithms , Mamta Jadhav, Zeel Thakkar, Prof. Pramila M. Chawan, Oct 2019. [4] Using convolutional neural networks for the classification of breast cancer images , Eric Bonnet , 31 Aug 2021.

[10] BCI: Breast Cancer Immunohistochemical Image Generation through Pyramid Pix2pix, ShengJie Liu, Chuang Zhu, Feng Xu, Xinyu Jia, Zhongyue Shi, Mulan Jin 25 Apr 2022.

[11] Breast cancer detection using artificial intelligence techniques: A systematic literature review, Ali Bou Nassif, Manar Abu Talib, Qassim Nasir, Yaman Afadar, Omar Elgendy, Artificial Intelligence in Medicine, Elsevier, Vol 127, May 2022.

[12] Preprocessing of Breast Cancer Images to Create Datasets for Deep-CNN ABHIJITH REDDY BEERAVOLU 1 , SAMI AZAM 1 , MIRJAM JONKMAN 1 , (Member, IEEE), BHARANIDHARAN SHANMUGAM1 , (Member, IEEE), KRISHNAN KANNOORPATTI 1 , AND ADNAN ANWAR 2 , (Member, IEEE) 1College of Engineering, IT, and Environment, Charles Darwin University, Casuarina, NT 0810, Australia 2Centre for Cyber Security Research and Innovation (CSIR), School of IT, Deakin University, Geelong, VIC 3216, Australia Corresponding author: Sami Azam (sami.azam@cdu.edu.au)

[13] Breast Cancer Prediction: A Comparative Study Using Machine Learning Techniques Md. Milon Islam1 · Md. Rezwanul Haque1 · Hasib Iqbal1 · Md. Munirul Hasan2 · Mahmudul Hasan3 · Muhammad Nomani Kabir2 Received: 11 August 2020 / Accepted: 18 August 2020 / Published online: 1 September 2020 © Springer Nature Singapore Pte Ltd 2020

[14] Breast Cancer Detection and Classification using Analysis and Gene-Back Proportional Neural Network Algorithm Amandeep Kaur, Prabhjeet Kaur

[15] Augmentation of Classifier Accuracy through Implication of Feature Selection for Breast Cancer Prediction Deepa B G, Senthil S , Gupta Rahil M, Shah Vishakha R

[16] EARLY DETECTION OF BREAST CANCER USING SVM CLASSIFIER TECHNIQUE Y.Ireaneus Anna Rejani+ , Dr.S.Thamarai Selvi* + Noorul Islam College of Engineering, Tamilnadu, India. *Professor&Head, Department of Information and technology, MIT, Chennai, Tamilnadu, India.

[17] Preprocessing of Breast Cancer Images to Create Datasets for Deep-CNN ABHIJITH REDDY BEERAVOLU, SAMI AZAM, MIRJAM JONKMAN, (Member, IEEE), BHARANIDHARAN SHANMUGAM , (Member, IEEE), KRISHNAN KANNOORPATTI, AND ADNAN ANWAR , (Member, IEEE)

[18] An Automatic Detection of Breast Cancer Diagnosis and Prognosis Based on Machine Learning Using Ensemble of Classifiers USMAN NASEEM , JUNAID RASHID, LIAQAT ALI, JUNGEUN KIM, QAZI EMAD UL HAQ, MAZHAR JAVED AWAN, AND MUHAMMAD IMRAN

[19] Breast Cancer Prediction using Naïve Bayes Classifier Megha Rathi, Arun Kumar Singh

[20] Prediction of Breast Cancer Using SVM Algorithm Deepika S* and Kapilaa Ramanathan Devi N

[21] Breast Cancer Detection using Artificial Neural Networks Md Haris Uddin Sharif1

[22] BREAST CANCER PREDICTION USING MACHINE LEARNING Ramik Rawal