International Journal of Recent Advances in Multidisciplinary Topics
Volume 3, Issue 4, April 2022
https://www.ijramt.com | ISSN (Online): 2582-7839

Breast Cancer Detection using Machine Learning

Richa Sharma[1], Shreya Avindra Borkar[2*], Prajkta Dilip Lokhande[3], Nikita Amar Raut[4], Ruchita Shivaji Khetmalis[5]

[1]*Assistant Professor, Department of Information Technology, JSPM's Rajarshi Shahu College of Engineering, Pune, India*
[2,3,4,5]*Student, Department of Information Technology, JSPM's Rajarshi Shahu College of Engineering, Pune, India*

***Abstract*: Breast cancer is the most frequent cancer in women and the major source of death in women worldwide. Ultrasound imaging is the recommended tool for breast cancer diagnosis in hospitals since it is significantly safer than other imaging modalities. Ultrasound images, on the other hand, are distorted by non-Gaussian regions having additive noise. Medical technicians and clinicians currently diagnose breast cancer by manually reviewing ultrasound images, which is a time-consuming and costly method. This could be a major hurdle to early detection of breast cancer. As a result, early detection of breast cancer can help with not only prescribing medical procedures to prevent the cancer from spreading but also lowering the fatality rate. Automatic detection and diagnosis in ultrasonography is exceedingly difficult due to speckles (noise). In this research, a Convolutional Neural Network (CNN) model for debuler ultrasound pictures is suggested, followed by different CNN model for ultrasound image classification into being and virulent classifications. The proposed models are tested on a Mendeley Breast Ultrasound dataset. Experiments show that the suggested model achieves a classification accuracy of 99.89 percent and that the proposed model(s) surpass other methods proposed in previous publications.**

***Keywords*: CNN, Classifier algorithm, Feature extraction, Breast cancer, Machine Learning.**

1. Introduction

Cancer is one of the deadliest diseases, and recent statistical research in the United States revealed that it is the second greatest cause of death among women. In a natural process, the cells in the human body age, die, and are replaced by new ones. In some cases, the cells might grow abnormally, resulting in a mass known as a tumour. Breast cancer tumours are labeled as either cancerous or non-cancerous. Non-cancerous tumours have little effect on their environment. Cancerous tumours, on the other hand, have the propensity to spread to other organs and tissues. Breasts, lungs, skin, brain, and other organs are some of the most prominent cancer-causing organs in the human body. The International Agency for Research has released a global detailed survey on cancer cases from 185 countries. Breast cancer is the leading cause of ephemerality in women and one of the most commonly diagnosed diseases. As a result, obtaining an early and precise diagnosis, as well as urgent medical treatment, is critical in order to lower the mortality rate due to breast cancer. Among the various imaging techniques available, for medical diagnostic purposes, ultrasound is the most extensively utilised imaging modality. Ultrasound is favoured over other methods because it uses non-ionization sound waves that are regarded medically safe. The ultrasonic transducer has a piezoelectric crystal that sends sound wave pulses and uses the echoes reflected from inside human tissues to create an ultrasound image.

2. Objectives

A lesion that takes up space in one of the three spatial dimensions is known as a mass. When analyzing the lesion in the craniocaudal and oblique midluteal views, it may be assured that it occupies a three-dimensional space. Otherwise, this is referred to as an asymmetry. Malignant masses are somewhat irregular in shape, lack symmetry, and have no fat coverage. Benign masses have a low density, well-defined margins, and a fat covering over the lesion, whereas malignant masses are slightly irregular in shape, lack symmetry, and have no fat coverage. The process of looking for disease indications, such as breast cancer, before symptoms occur is known as screening. Cancer screening tests are necessary to detect cancer early on, allowing it to be treated and cured. A screening test can detect cancers that are very small or slow-growing.

3. Methodology

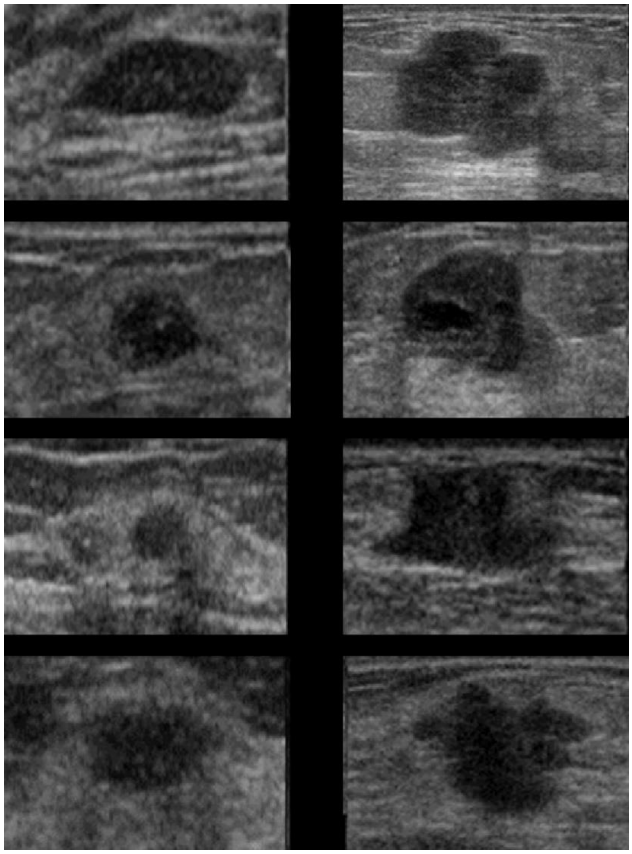
Other breast inspections modalities, such as ultrasound and/or magnetic resonance imaging (MRI) scans, are not typically designed to monitor for breast cancer in many women at average risk. However, for women who are at a high risk of breast cancer, it may be advantageous., have dense breast tissue, or have identified a lump or mass during a breast inspection. Due to the significant likelihood of false positives, the use of ultrasonography screening is questionable. As a result, the test frequently misses the presence of cancer. The use of magnetic resonance imaging (MRI) for routine screening is likewise controversial. Ultrasound or MRI may be beneficial for women who have a concerning breast finding on physical examination or mammography. Additional ultrasound or mammography testing may be required if a lump or mass is identified during a physical examination Women should address the sort of screening that is suggested for them with their doctors, as well as the frequency with which they should be examined. Breast cancer is a form of cancer that begins in the breast cells. A

*Corresponding author: shreyaborkar0201@gmail.com

malignant tumour is a collection of cancer cells that can spread to other parts of the body or infect surrounding organs (metastasize).

4. Relevance

1. Defining the region of interest (ROI) A region of interest (ROI) is a cluster of samples within a data set which have been identified for a specific reason. The concept of a ROI is widely applied in a variety of applications
2. Image preprocessing: the image is upgraded, with noise removed and contrast increased.
3. Segmentation: The lesion is segmented from the area of interest, which has been determined in the first procedure, by determining the contour or pixel region.
4. Feature extraction and selection: Feature extraction is the process of extracting/deriving data from an existing feature set in order to construct a new feature subspace. These strategies, like feature selection techniques, are used to reduce the amount of features from the initial features set in order to reduce model complexity, overfitting, model computation efficiency, and generalisation error.
5. Automatic classification: This final phase is critical for CADx systems, since it determines the lesion's type based on the specified criteria.



(a) (b)

Fig. 1. a) Benign ultrasound sample images, b) Malignant ultrasound sample images

Table 1
Literature survey

Title	Year	Description
Predicting Breast cancer in fine needle aspiration images using machine learning.	2020	We built and deployed a system whose primary objective is to detect the presence of breast cancer lumps in fine needle aspiration images in this project. To identify cancer, they deployed a clustering and feature extraction method such as CNN, followed by a classification method.
Ultrasound Image Denoising and detection of Breast Cancer using Deep CNN	2020	In this paper, a Convolutional Neural Network (CNN) has been proposed for Despeckling (Denoising) the ultrasound images and afterwards another CNN model is proposed for the classification of the ultrasound images into benign and malignant classes.
A study on automatic detection of IDC breast cancer with Convolutional Neural Networks (2020)	2020	The team looked into other CNN architectures for automated diagnosis of breast cancer after being inspired by Cruz's work. They started with a basic CNN that was identical to Cruz's and then expanded it to four alternative architectures.
Accurate Prediction of Neoadjuvant Chemotherapy Pathological Complete Remission (pCR) for the Four Sub-Types of Breast Cancer.	2019	Only the nodal dimensions of the first three treatments were used in this research to see if they could predict pCR. A best feature combination for each breast cancer subtype was screened from the real nodal sizes of the first three treatments and the nodal sizes` of the next three treatments predicted from those of the first three ones.
Towards Breast Cancer Response Prediction using Artificial Intelligence and Radiomics.	2020	Using three basic stages, this research suggests a new methodology for determining breast tumour response to treatment: 1. Tumor segmentation from MR images; 2. Feature extraction from segmented tumours to create a complete and exploitable database; 3. Tumor-response prediction models constructed using deep and machine learning architectures.
An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN)	2021	This report investigates the accuracy of artificial neural networks, Multi-Layer Recurrent neural network Neural Networks (MLP), and Convolutional Neural Networks (CNN) used to detect breast malignancies for early detection of breast cancer in order to determine which method is better for breast cell malignancies diagnosis.
Convolutional neural network-based models for diagnosis of breast cancer	2020	The paper investigates the proposed system that uses various convolutional neural network (CNN) architectures to automatically detect breast cancer, comparing the results with those from machine learning (ML) algorithms

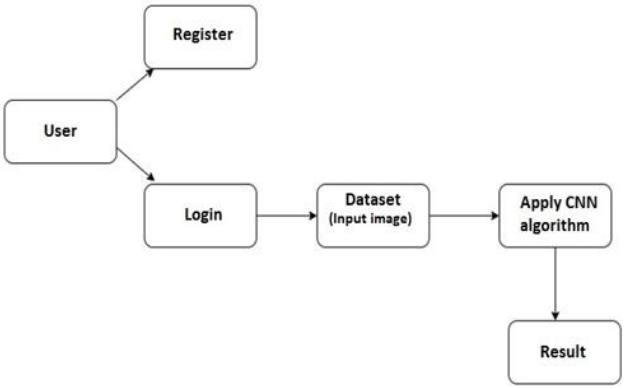


Fig. 2. Architecture diagram

5. Result

IDC tumour grading involves isolating and analyzing malignant tumour tissue out of whole slide images. Because manual isolation by trained and experienced pathologists takes a long time and is vulnerable to bias, an automated method for detecting malignant tissue is available to the overall medical industry. Previous study has determined template matching algorithms based on feature extraction as well as deep learning strategies to do this. We'll look at a few different Neural network architectures in this post. After a series of 10-fold cross validation tests, performance measures can be generated. Compiles the averages of performance metrics for each design trained on the normal dataset and describes the final quality measures for architects trained on the enhanced dataset.

6. Conclusion

The result indicated that the suggested CNN algorithm surpasses all the others. A woman chosen at random has a twelve percent probability of being diagnosed with the condition. kNN is the most successful algorithm for detecting breast cancer, with the highest accuracy, precision, and F1 score among the other algorithms. Thus, in cancer research, supervised machine learning approaches will be highly helpful in early diagnosis and prognosis of a cancer kind. Different mammography reading methods, such as double reading and CAD approaches, have been created due to the need of an accurate medical diagnosis in the case of breast cancer in order to increase the patient's life expectancy. These strategies are inefficient since they increase the number of physicians required and limit the qualities that can be predicted ahead of time. As a result, developing models for automatic feature extraction and categorization of breast lesions on FNAs is advantageous. A mammography lesion classification model was effectively deployed in this study, with an accuracy metric of around 68 percent. The quantity of photos accessible was the main limitation of this work, since it is evident how the models are trained, improving the accuracy metric of training and validation in parallel until a point is reached where the model begins to memories the training images. This is due to a scarcity of new photos in the data set used to train the model.

7. Future Scope

Future research could focus on developing a systematic technique for model selection and implementing a multi-model deep learning approach based on a merging process. In this study, a CNN model for despeckling breast ultrasound images is proposed, as well as another CNN model for classifying breast tumours as benign or malignant. It was discovered that by combining the CNN despeckler and the CNN classifier, a classification accuracy of 88.00 percent was reached, which is more than the accuracies of other systems in the literature. Despeckling strategies that operate on various photos with varying distortions will be identified in the future, as will building deep learning architectures that reach the maximum accuracy for any image quality. In addition, future studies will look into the segmentation of breast cancer.

References

[1] S. G. Durai, S. H. Ganesh, and A. J. Christy, “Novel Linear Regressive Classifier for the Diagnosis of Breast Cancer,” In Computing and Communication Technologies (WCCCT), 2017 World Congress, 2017.

[2] Abdullah-Al Nahid, Aaron Mikaelian and Yinan Kong “Histopathological breast-image classification with restricted Boltzmann machine along with backpropagation.” (2018).

[3] Ram MurtiRawat, Shivam Panchal, Vivek Kumar Singh, Yash Panchal "Breast Cancer Detection Using K-Nearest Neighbors, Logistic Regression and Ensemble Learning" (2020).

[4] Saleem Z. Ramadan “Methods Used in Computer- Aided Diagnosis for Breast Cancer Detection Using Mammograms: A Review” (2020)

[5] T. K. Avramov and D. Si, “Comparison of Feature Reduction Methods and Machine Learning Models for Breast Cancer Diagnosis,” Proc. Int. Conf. Comput. Data Anal. - ICCDA ’17, pp. 69–74, 2017.

[6] V. Bevilacqua, A. Brunetti, M. Triggiani, D. Magaletti, M. Telegrafo, and M. Moschetta, “An Optimized Feed-forward Artificial Neural Network Topology to Support Radiologists in Breast Lesions Classification,” Proc. 2016 Genet. Evol. Comput. Conf. Companion - GECCO ’16 Companion, pp. 1385–1392, 2016.

[7] Shubham Sharma, Archit Aggarwal and Tanupriya Choudhury "Breast Cancer Detection Using Machine Learning Algorithms" (2018).

[8] Mambou, S. J., Maresova, P., Krejcar, O., Selamat, A., & Kuca, K. (2018). Breast Cancer Detection Using Infrared Thermal Imaging and a Deep Learning Model. Sensors (Basel, Switzerland), 18(9), 2799. h

[9] Syed Jamal Safdar Gardezi, Ahmed Elazab, Baiying Lei and Tianfu Wang "Breast Cancer Detection and Diagnosis Using Mammographic Data: Systematic Review" (2019).

[10] S. Karthik, R. Srinivasa Perumal and P. V. S. S. R. Chandra Mouli “Breast Cancer Classification Using Deep Neural Networks” (2019).

[11] Canh Phong Nguyen, Anh Hoang Vo, Bao Thien Nguyen, “Breast cancer histology image classification using deep learning,” 19th International Symposium on Communication and Information Technologies (ISCIT), pp. 366-370, 2019.

[12] M. Desai and M. Shah, “An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN),” Clinical eHealth, vol. 4, pp. 1–11, 2021.