International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 10 Issue: 01 | Jan 2023 www.irjet.net p-ISSN: 2395-0072

A Progressive Review: Early Stage Breast Cancer Detection using Ultrasound Imaging Modality as well as Machine Learning

Jinal Patel[1], Himanshu Patel[2]

[1]Student, Department of Biomedical Engineering, U.V.Patel College of Engineering, Ganpat University, [2]Professor, Department of Biomedical Engineering, U.V.Patel College of Engineering, Ganpat University, Mehsana, India

---

Abstract- *Breast cancer is one of the most commonly diagnosed diseases in women and is the leading cause of cancer-related death in women. The early detection of cancerous tissue can aid in the recovery and treatment of the disease and save more lives. The detection of breast cancer can be challenging using mammography, as dense tissues can overlap in screening. Ultrasound and MRI have also been used for breast cancer growth screening. Numerous methods have been used for the detection of abnormalities in the breast. Here I am only using ultrasound screening modality for the deection of breast cancer at the early stage. For the detection of breast cancer at the early stage, many research works have been carried out. Machine learning is frequently used for detection. Machine learning is a type of artificial intelligence (AI) that provides computers with the ability to learn without being explicitly programmed. For the classification of images, Machine learning algorithms like Convolutional neural networks can be used. In this study, we have presented a comprehensive review of Machine learning algorithms for early-stage breast cancer detection using imaging modalities.*

Key Words: Breast cancer, image processing, ultrasound, early stage detection, machine learning, deep learning.

1. INTRODUCTION

Breast cancer is the most common invasive cancer in women, and is the leading cause of cancer death in women worldwide. Breast cancer is a type of cancer that begins in the breast. Most breast cancers start in the ducts or glands, but some start in other parts of the breast. Breast cancer cells form a tumour or lump. There are two types of tumour 1. Benign Tumour 2. Malignant Tumour.

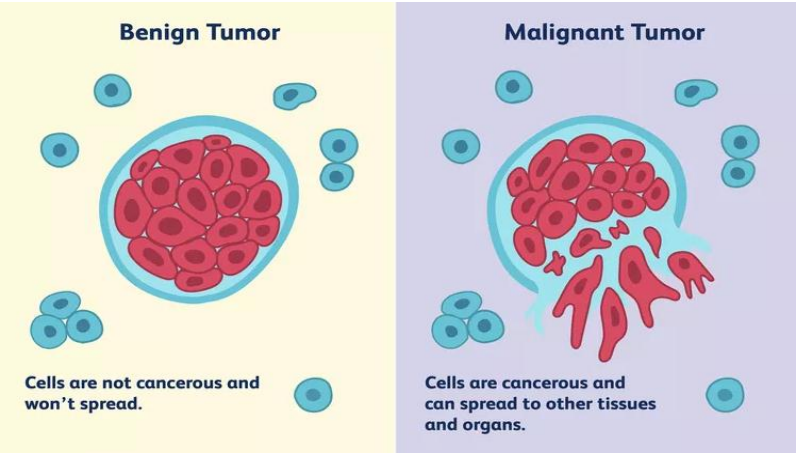

Fig. Types of Tumour

But most breast lumps are benign and not malignant (cancer). Benign tumours are tumours that do not spread to other parts of the body. Breast cancer begins in the ducts or glands as per the type of cancer cells. According to the World Health Organization (WHO), breast cancer is the most common cancer in women, and is the leading cause of cancer death in women worldwide. In 2020, there were an estimated 5 million cases and 2.3 million deaths from breast cancer. The number of cancer cases in India is expected to increase significantly in the next few years, according to a recent report by the National Cancer Registry Programme (NCRP). This increase underscores the need for effective diagnostic like ultrasound, mammogram, MRI, histopathology and Biopsy and classification techniques for breast cancer. There are many deep learning and machine learning procedures available for the detection and classification of breast cancer. In this paper, we review the work of different techniques used for the detection and classification of breast cancer.

2. LITERATURE REVIEW

Early stage breast cancer detection is a vital topic in the medical field and various methods have been proposed for detection and diagnosis. In this progressive review section, the authors have discussed different segmentation processes, features and classifiers along with their merits and demerits for early stage breast cancer detection. Machine learning and deep learning approaches have been discussed for early-stage breast cancer detection and their role for detection and diagnosis of breast cancer and assistance in the treatment, recovery and decrease the chances of critical situations has been discussed.

Narumi Harada-Shoji et al. [1] have discussed in this research work, the use of Ultrasonography in combination with mammography screening is explored as a means to detect early-stage and invasive breast cancers in asymptomatic women with both dense and nondense breasts. It is suggested that this combination of screening methods may be more accurate in detecting breast cancers than mammography alone.

Kushangi Atrey et al. [2] have discussed in this research work, to investigate a multimodal approach to the detection and validation of breast cancer images, combining both ultrasonography and mammography. segmentation approach that can accurately detect breast lesions using a dual-modality system. To this end, experiment with and evaluate the performance of three popular techniques, namely Fuzzy-c-Means (FCM), K-means (KM) and Darwinian Particle Swarm Optimization (DPSO), for detecting lesions using the combined ultrasonography and mammography. The fuzzy C-means algorithm gives accurate results.

Eduardo Reyes et al. [3] have discussed in this research work, the radiologic characteristics, histology, and prognostic factors of breast cancer in pregnant women. It aims to assess the impact of these factors on the prognosis of the disease.

Ul Haq, Amin et al. [4] have discussed in this research work, machine learning-based diagnosis technique to address the challenge of early diagnosis of breast cancer. The SVM model was utilized to classify malignant and benign cases. To enhance accuracy, the Minimal Redundancy Maximal Relevance and Chi-square algorithms were employed to select relevant features from the breast cancer dataset.

Kiran Jabeen et al. [5] have discussed in this research work, This proposed approach is evaluated on a publicly available ultrasound image dataset and results demonstrate its effectiveness for breast cancer classification. The proposed framework is composed of five key steps: (i) data augmentation is used to increase the size of the original dataset for improved training of Convolutional Neural Network (CNN) models; (ii) a pre-trained DarkNet-53 model is selected and its output layer modified according to the classes in the augmented dataset; (iii) the modified model is trained using transfer learning, with features extracted from the global average pooling layer; (iv) the optimal features are identified using two improved optimization algorithms, namely reformed differential evaluation (RDE) and reformed gray wolf (RGW); and (v) the best selected features are fused together using a novel probability-based serial approach and classified by machine learning algorithms.

Gelan Ayana et al. [6] have discussed in this research work, Transfer Learning (TL) is a method of leveraging a pre-trained model to construct an efficient model using limited training data. In this work, propose a multistage Transfer Learning (MSTL) algorithm which utilizes three pre-trained models: EfficientNetB2, InceptionV3, and ResNet50 with three optimizers: Adam, Adagrad, and Stochastic Gradient Descent (SGD), to improve the performance of classification of microscopic cancer cell line images in ultrasound breast cancer image classification. The MSTL algorithm combines knowledge from both natural and medical datasets.

Roxana Iacob et al. [7] have discussed in his research work, in order to achieve the most beneficial outcome for the patient, we should use all available imaging techniques to the fullest extent. Mammography and 3D tomosynthesis should be employed for the screening and initial diagnosis, with the assistance of ultrasound. Additionally, ultrasound, MRI, and PET-CT can be used together for the assessment of the tumor's stage.

R. Karthiga et al. [8] have discussed in his research work, In this work, the performance of a machine learning (ML) classifier is evaluated using deep convolutional neural networks (CNNs) and transfer learning models (Alexnet, VGG-16, VGG-19, Resnet-50, and Resnet-101). Before the ML classification process, an anisotropic diffusion filter is used to extract tumoral features from the data.

Mehedi Masud et al. [9] have discussed in his research work, This research implements pre-trained convolutional neural networks for identifying breast cancer through ultrasounds. The models were fine-tuned to detect key features from the images and a classifier was applied to the top layer. Fivefold cross validation was used to measure the accuracy of seven popular models with different optimizers and hyper-parameters. In addition, Grad-CAM and occlusion mapping were used to evaluate how well the models captured the relevant features in the images. After fine tuning, DenseNet201 and ResNet50 achieved perfect accuracy with Adam and RMSprop optimizers. Stochastic Gradient Descent optimizer enabled VGG16 to have 100% accuracy.

Pratiksha Joshi et al. [10] have discussed in his paper reviews the use of machine learning and deep learning techniques for the early detection of breast cancer. It is found that deep learning is more effective than machine learning when it comes to detecting tumors and classifying breast cancer images. The convolutional neural network (CNN) model is used for classification and provides more accurate outcomes than other machine learning approaches. Thus, deep learning can be a reliable tool for detecting breast cancer in its early stages, which can be beneficial for successful treatment and recovery.

Yaghoub Pourasad et al [11] have discussed in his research work, to design a method for identifying and diagnosing breast tumors based on ultrasound images. For this purpose, six techniques have been performed to detect and segment ultrasound images. Features of images are extracted using the fractal method. Moreover, k-nearest neighbor, support vector machine, decision tree, and Naïve Bayes classification techniques are used to classify images. Then, the convolutional neural network (CNN) architecture is designed to classify breast cancer based on ultrasound images directly.

Viswanatha Reddy Allugunti et al [12] have discussed in his research work, a Computer-aided Diagnosis (CAD) for categorizing patients into 3 distinct classes (cancer, no cancer, and non-cancerous) utilising a database. CAD is a system for computer-assisted diagnosis. They explore Convolutional Neural Networks (CNNs), Support Vector Machines (SVMs), and Random Forests (RFs), which are all powerful classifiers, for the classification stage. The three robust classifiers that investigate and evaluate are Convolution Networks, Support Vector Machines, and Random Forest.

Mahmoud Ragab et al [13] have discussed in his research work, an Ensemble Deep-Learning-Enabled Clinical Decision Support System for Breast Cancer Diagnosis and Classification (EDLCDS-BCDC) was designed to detect the presence of breast cancer through the use of USIs. The EDLCDS-BCDC consists of two pre-processing stages, Wiener filtering and contrast enhancement, followed by image segmentation with Chaotic Krill Herd Algorithm (CKHA) and Kapur's Entropy (KE). Lastly, an ensemble of three deep learning models, VGG-16, VGG-19, and SqueezeNet, were employed for feature extraction.

Epimack Michael et al [14] have discussed in his research work, an automated CAD system that formulates an optimized algorithm. Thirteen features were chosen from the total 185 features to train machine learning. Five classifiers were used to distinguish between malignant and benign tumors. The results of 10-fold cross-validation revealed Bayesian optimization with a tree-structured Parzen estimator. LightGBM classifier showed the best performance with 99.86% accuracy, 100.0% precision, 99.60% recall, and 99.80% FI score.

Ye-Jiao Mao et al [15] have discussed in his review paper, to analyze the use of machine learning models in ultrasound elastography systems for the classification of breast tumours. Literature databases PubMed, Web of Science, CINAHL and EMBASE were searched and thirteen studies were found to be eligible for review. Of these, six studies investigated shear-wave elastography, while seven studies focused on strain elastography (five freehand and two Acoustic Radiation Force). Traditional computer vision workflows were commonly used in strain elastography which included image segmentation, feature extraction and classifier functions implemented by means of different algorithms, neural networks or support vector machines (SVM). Shear-wave elastography, on the other hand, tended to adopt the deep learning model of convolutional neural networks (CNN), which combines multiple functional tasks. All the articles included in the review achieved sensitivity of at least 80%, though only half of them yielded acceptable specificity of 95%. Surprisingly, deep learning models did not necessarily outperform traditional computer vision workflows.

3. PROPOSED BLOCK DIAGRAM

Digital Image processing has been a great benefit to the medical profession, enabling more accurate diagnosis and treatment of diseases. It has been especially effective in early breast cancer detection, making it possible to identify the disease in its earliest stages. This review will cover the most recent advances in the field of breast cancer diagnosis, such as

segmentation methods, feature selection and classifier. These advancements have made it possible to enhance the accuracy and effectiveness of early detection and treatment.

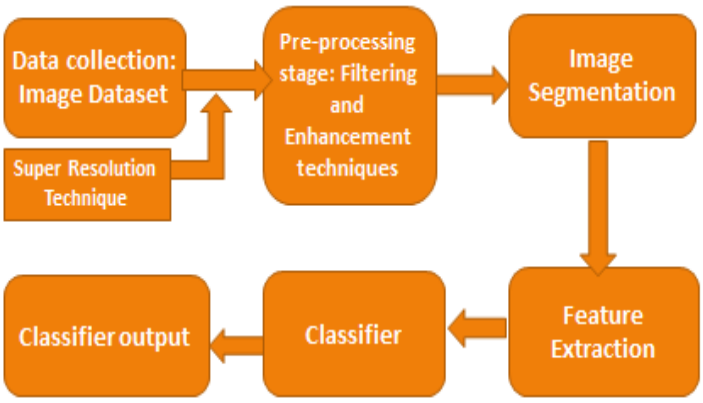

Fig. Proposed methodology

4. TECHNIQUES

The current section of the progressive review paper outlined various segmentation, feature extraction and classifier methods, evaluating their respective strengths and weaknesses in terms of noise sensitivity, accuracy, sample size and simplicity. Table I then summarised the significant contributions to the field made between 2018 and 2022.

Table -1: Summary of segmentation methods, feature extraction and classifiers method and some of the key merits and demerits

Segmentation methods		
Techniques	**Merits**	**Demerits**
Threshold-based segmentation methods	Uncomplicate, effortless, and economical.	It is not suitable for multi-channel images and can be easily affected by noise in the images.
Region basedsegmentation method	It is easy to obtain results with fewer noise when applying multiple criteria	A time-consuming method, this one cannot be used for noisy images and does not provide accurate shading results for actual mammograms.
Edge basedsegmentation methods	This technique is suitable for images having prominent edges.	Images with smooth edges are not suitable for applying edge detection techniques, which can also lead to a decrease in the image's contrast.
Fuzzy theory basedimage segmentation	This technique can be used to process images and eliminate any unwanted noise.	It cannot be applied directly to grayscale images.
Artificial neuralnetwork based imagesegmentation	Fewer data are needed and implementation is straightforward.	A quantity of pixels should be specified.
K-means clusteringmethod	Simplicity, fast and if data is recognizable, provides better result.	The number of clusters must be predetermined.
Active Contours Algorithm	Computationally efficient, versatile, robust, and capable of handling multiple objects	Requires good initial contour, sensitive to errors, not suitable for segmenting objects with large intensity variations, and computationally expensive for large images.
Darwinian Particle Swarm Optimization (DPSO)	For optimizing a specific function	Less accurate

Feature extraction/selection		
Techniques	**Merits**	**Demerits**
Gray level Co-occurrence matrix(GLCM)	Second-order statistical texture characteristics can provide enhanced performance on basic textures	Very sensitive to the dimensions of the texture samples.
Principal ComponentAnalysis	Gives clearer visuals and boosts performance.	Suitable for standardizeddata, Not easy to interpret
Linear DiscriminantAnalysis	Supervised learning, Simple to execute and rapid categorization	sample size
Wavelet Transform	Wavelet Transform is able to represent signals in both time and frequency domains	Wavelet Transform is computationally intensive and requires more memory for storage
Classifiers		
Techniques	**Merits**	**Demerits**
Convolutional Neural Network	Supervised learning, High accuracy	It necessitates a vast amount of data
Support Vector Machine	Suitable for binary data	It cannot be used for multiple data simultaneously
Artificial Neural Network	Provide high accuracy	Data needs to be converted into numerical format.
Random Forest classifier	Provides higher accuracy than Decision tree classifier	Complicated as it generates multiple trees
Decision Tree classifier	Categorizes the data set into pre-determined classes	Less accurate compare to Random forest classifier
K-Nearest Neighbor algorithm	Simple and supervised learning	Time consuming algorithm

5. PROBLEM STATEMENT

Early-stage breast cancer screening is an important step to preventing potentially fatal consequences. There are various diagnostic imaging techniques available to detect breast tumours, such as ultrasound, mammography (X-ray), magnetic resonance imaging (MRI), and computed tomography (CT scan). Additionally, physical examinations are also conducted to identify any abnormalities. Mammography is particularly effective in detecting breast cancer, but overlapping tissues can obscure the results. An ultrasound can provide a more accurate result, but has a lower resolution and contrast.

For early-stage breast cancer detection, mammography is the preferred screening test, however, it can have some potential harms. False positives can lead to an erroneous diagnosis of breast cancer in a healthy person and false negatives can lead to a misdiagnosis of a patient having cancer as healthy. Furthermore, mammography can be uncomfortable and cause pain. In contrast, ultrasound imaging is a non-invasive technique with no radiation risk, however, the resolution of the images is not as accurate as that of a mammogram. Mammography is quite expensive and absorbs more radiation, making it take longer to diagnose. To improve the accuracy of the results, we suggest a method wherein only ultrasound screening modality is used and the resolution of the image is increased by using a super resolution technique.

6. CONCLUSION

The purpose of this research is to detect breast cancer at its early stage and reduce the number of fatalities. Therefore, ultrasound imaging is used as the modality for initial diagnosis and image processing techniques are utilized for further

assessment. This study examined the use of machine learning and deep learning techniques to identify early-stage breast cancer, thereby providing the opportunity for early treatment and improved recovery chances. The results revealed that machine learning techniques, such as artificial intelligence, provide improved accuracy and precision. Specifically, Convolutional Neural Networks (CNN) were used to classify breast cancer images, which showed a greater degree of accuracy than other machine learning techniques. In conclusion, machine learning is a viable approach to early detection and diagnosis of breast cancer, thus aiding in reducing fatalities.

REFERENCES

[1] Harada-Shoji, N., MD PhD. (2021, August 18). *Evaluation of Adjunctive Ultrasonography for Breast Cancer Detection Among Women Aged 40-49 Years With Varying*. https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2783177

[2] *Breast cancer detection and validation using dual modality imaging*. (2020, January 1). IEEE Conference Publication | IEEE Xplore.

[3] Reyes, E., Xercavins, N., Saura, C., Espinosa-Bravo, M., Gil-Moreno, A., & Cordoba, O. (2020). Breast cancer during pregnancy: matched study of diagnostic approach, tumor characteristics, and prognostic factors. *Tumori Journal*, *106*(5), 378–387.

[4] Haq, U. A. (2020, January 1). *A novel integrated diagnosis method for breast cancer detection - IOS Press*.

[5] Jabeen, K. (n.d.). *Breast Cancer Classification from Ultrasound Images Using Probability-Based Optimal Deep Learning Feature Fusion*. MDPI.

[6] Ayana, G. (n.d.). *A Novel Multistage Transfer Learning for Ultrasound Breast Cancer Image Classification*. MDPI.

[7] Iacob, R., Manolescu, D. L., Stoicescu, E. R., Fabian, A., Malita, D., & Oancea, C. (2022, June 22).

[8] Narasimhan, K. (2022). Automated diagnosis of breast cancer from ultrasound images using diverse ML techniques. *Multimedia Tools and Applications*, *81*(21), 30169–30193.

[9] Masud, M., Hossain, M. S., Alhumyani, H., Alshamrani, S. S., Cheikhrouhou, O., Ibrahim, S., Muhammad, G., Rashed, A. E. E., & Gupta, B. B. (2021). Pre-Trained Convolutional Neural Networks for Breast cancer detection using

Cancer Detection Using Ultrasound Images. *ACM Transactions on Internet Technology*, *21*(4), 1–17. https://doi.org/10.1145/3418355

[10] Journal, I. (2022, April 22). *A Progressive Review on Early Stage Breast Cancer Detection*.https://www.academia.edu/77263822/A_Progressive_Review_on_Early_Stage_Breast_Cancer_Detection

[11] Pourasad, Y., Zarouri, E., Salemizadeh Parizi, M., & Salih Mohammed, A. (2021). Presentation of Novel Architecture for Diagnosis and Identifying Breast Cancer Location Based on Ultrasound Images Using Machine Learning. *Diagnostics*, *11*(10), 1870. https://doi.org/10.3390/diagnostics11101870

[12] Reddy Allugunti, V. (2022). Breast cancer detection based on thermographic images using machine learning and deep learning algorithms. *Researchgate*, 49–51.

[13] Ragab, M. (n.d.). *Ensemble Deep-Learning-Enabled Clinical Decision Support System for Breast Cancer Diagnosis and Classification on Ultrasound Images*. MDPI.

[14] Michael, E., Ma, H., Li, H., & Qi, S. (2022). An Optimized Framework for Breast Cancer Classification Using Machine Learning. *BioMed Research International*, *2022*, 1–18.

[15] Mao, Y. J., Lim, H. J., Ni, M., Yan, W. H., Wong, D. W. C., & Cheung, J. C. W. (2022). Breast Tumour Classification Using Ultrasound Elastography with Machine Learning: A Systematic Scoping Review. *Cancers*, *14*(2), 367. https://doi.org/10.3390/cancers14020367

[16] https://www.who.int/news-room/fact-sheets/detail/breast-cancer

[17] https://www.indiatoday.in/information/story/world-cancer-day-2021-breast-cancer-scenario-in-india-1765496-2021-02-03

[18] https://www.verywellhealth.com/what-does-malignant-and-benign-mean-514240