International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Breast Cancer Detection Using Deep Learning

Keshav Tambre [a], Khushi Dhake [b], Needhi Patil [c], Dnyaneshwar Pawa r[d], Pranav Mahadik [e]

[a,b,c,d,e] *Department of Information Technology, International Institute of Information Technology(I^{2I}T),Pune-411057,India*

A B S T R A C T

One of the top causes of death for women is breast cancer, which is challenging to prevent because the primary causes of breast cancer are still unknown. For women who may be at risk of developing malignant tumours, breast cancer characteristics, such as masses and microcalcifications detectable in mammograms, can be used to aid in early identification. X-ray mammography is the main screening tool for early diagnosis, and accurate clinical report interpretation is crucial for breast cancer prediction, though the choice may be subject to error. Mammograms are difficult to interpret, especially in the context of screening. The level of skill of the radiologist and the image quality has an impact on the sensitivity of screening mammography. In this project, digital mammography pictures are subjected to image pre-processing, feature extraction, image segmentation, and image classification in order to enable early and automated identification of breast cancer. The project's goal is to use image processing techniques such image segmentation, image pre-processing, features extraction and selection, and picture classification to detect early-stage malignancies that are not subject to human error.

Keywords: CAD (computer- aided detection), Breast Cancer, Mammography, CNN (Convolutional neural network),Deep Learning.

1. Introduction

Breast cancer occurs when the breast cells multiply uncontrollably. Cancers of the breast can be classified into various types. In the US, breast cancer is the second-leading cause of mortality from cancer for females. Over 39 million diagnostic and screening mammography procedures were carried out in the US in 2014. According to estimates, 232,000 women were given breast cancer diagnoses in 2015, and 40,000 of them passed away as a result. Despite the fact that only mammography has been demonstrated to reduce breast cancer mortality, concerns have been raised about the potential negative effects of screening, for example, biopsies that resulted in false positive recalls. Majority of the 10-15% of females whose screening mammography results were unclear underwent a second mammogram or/and ultrasound for more clarification. Many of these abnormalities are found to be benign after the subsequent imaging tests, and only 10–20% are advised to have a needle biopsy for further investigation. Among them, only 20–40% result in a cancer diagnosis. Evidently, there is a gap in the literature about the need to balance the benefits and risks of routine breast cancer screening. Despite the fact that multicenter studies show that current CAD applications do not improve radiologists' diagnostic performance, radiologists frequently use traditional computer-aided detection, also known as CAD in mammography to help analyze the images. These algorithms frequently employ hand-crafted characteristics to identify areas on a mammogram that stand out from healthy tissue. In deciding whether to recall these findings, the radiologist weighs their clinical importance and practicability.

2. Literature Survey

2.1 Name: An Automatic Detection of Breast Cancer Diagnosis and Prognosis Based on Machine Learning Using Ensemble of Classifiers, 2022.

Algorithms:

Decision tree, Random Forest, Logistic regression, Support vector machine, Naive Bayes.

Result:

The combination of Support Vector Machine, Logistic Regression, Naive Bayes and Decision Tree along with ANN provides most accuracy of 98.835%.

2.2 Name:Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening, 2020

Algorithms:

Adam optimization algorithm, L2 regularization.

2.3 Name:A Novel Hybrid K-Means and GMM Machine Learning Model for Breast Cancer Detection, 2021.

Algorithm:

Gaussian mixture model K-Mean

Result:

Accuracy of 95.5%, an error rate of 18.64%, and signal-to-noise of 13.05

Limitations/Drawbacks:

K-means and GMM have 93.8% and 65% accuracy with high error rates of 29.47% and 24.35%.

2.4 Name:Breast Cancer Malignancy Prediction Using Deep Learning Neural Networks, 2020.

Algorithm:

Decision Trees, Support Vector Machine, AdaBoost, Random Forest and Proposed Neural Network

Result:

Decision Trees 0.946, Support Vector Machine 0.963, AdaBoost 0.967, Random Forest 0.990, Proposed Neural Network 0.990.

Limitations/Drawbacks:

The validation loss is very high. Over-fitting of neural networks.

2.5 Name:Breast Cancer Classification from Histopathological Images Using Patch-Based Deep Learning Modelling, 2021.

Algorithm:

A novel patch-based deep learning method called Pa-DBN-BC Deep Belief Network (DBN).

Result:

Accuracy of 86%

Limitations/Drawbacks:

Only classifies between the cancer regions from the background regions.

2.6 Name:The Machine Learning based Optimized Prediction Method for Breast Cancer Detection, 2020.

Algorithm:

Logistic Regression, SVM, KNN and Naive Bayes.

Result:

LR 96.49%, SVM 98.24%, KNN 97.20% and Naive Bayes 94.74%.

2.7 Name:A Deep CNN Technique for Detection of Breast Cancer Using Histopathology Images, 2020.

Algorithm:

Convolutional Neural Network (CNN).

Result:

Accuracy is 95.58%, precision 0.90 recall 0.99 F1-score 0.89.

2.8 Name:Breast Cancer Detection Using Extreme Learning Machine Based on Feature Fusion with CNN Deep Features, 2019.

Algorithm:

Computer-aided diagnosis (CAD), Convolutional Neural Network (CNN), Unsupervised Extreme Learning Machine (US-ELM) clustering.

Result:

Outperformed other existing methods.

Limitations/Drawbacks:

The accuracy of existing CAD systems remains unsatisfactory.

2.9 Name:Classification of static infrared images using pre-trained CNN for breast cancer detection,2021.

Algorithm:

CNNs (VGG-16, Densenet201, and Resnet50) combined with transfer learning.

Result:

F1-score of 0.92, 91.67% for accuracy, 100% for sensitivity, and 83.3% for specificity.

Limitations/Drawbacks:

VGG and Resnet did not have significant results for the experiments with all poses.

2.10 Name:Convolutional Autoencoder Application for Breast Cancer Classification, 2020.

Algorithm:

Convolutional Autoencoder.

Result:

84.72% accuracy. sensitivity 86.87% precision 80.23%.

Limitations/Drawbacks:

2.11 Name:Breast Cancer detection Using Convolutional Neural Networks for Mammogram Imaging System, 2017.

Algorithm:

CNN

Result:

Accuracy of 0.8271.

Limitations/Drawbacks:

In this experiment, we have limitation to get more data.

3. Proposed system

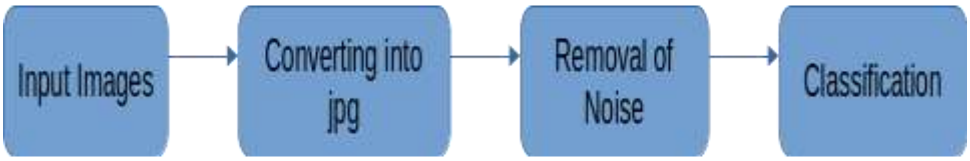

Fig 3.1 : Proposed system

1. Input the mammography image:

The first step is to obtain the mammography image that you want to classify. This image is typically obtained through medical imaging techniques such as X-rays or digital mammography.

2. Converting the input image to JPG format:

If the input image is not already in JPG format, it may need to be converted to this format. Converting the image to JPG ensures compatibility with various image processing algorithms and tools commonly used for analysis.

3. Removing noise from the image:

Mammography images often contain noise, which can interfere with the clarity and accuracy of the image. Noise can be caused by various factors, including equipment artifacts, patient movement, or environmental factors. To enhance the visibility of the image and improve the accuracy of subsequent analysis, it is important to remove noise. This can be done through various image processing techniques such as filtering, denoising algorithms, or adaptive noise reduction methods.

4. Classifying the image:

Once the image has been preprocessed and the noise has been removed, the next step is to classify the image. Mammography image classification involves determining whether the image is normal, benign (non-cancerous), or malignant (cancerous).

This classification task is typically performed using machine learning or deep learning algorithms. These algorithms are trained on a large dataset of mammography images that have been annotated with their corresponding labels (normal, benign, malignant). During the training process, the algorithm learns to identify patterns and features in the images that are indicative of different classes. The classification algorithm takes the preprocessed mammography image as input and generates a prediction or probability score for each class. The algorithm analyzes various features in the image, such as the shape, texture, and density of structures, to make its prediction. The output of the classification algorithm can be used to determine the likelihood of the image being normal, benign, or malignant. The algorithm can provide a confidence score or a probability distribution over the classes, indicating the level of certainty in its prediction .The final result of the classification is presented in a user-friendly format, such as a textual output indicating the predicted class (normal, benign, malignant), or it can be visualized with additional information overlaid on the image itself. It's important to note that the accuracy of the classification depends on the quality of the training data, the performance of the algorithm, and the preprocessing steps applied to the image. Regular validation and testing of the algorithm with independent datasets are crucial to ensure reliable and accurate results. Additionally, the classification output should always be interpreted by a medical professional or radiologist, as they have the expertise to make a final diagnosis based on the image analysis.

4. Conclusion

The objective of the CAD system is to enhance the effectiveness and precision of breast cancer detection by offering an supplementary tool for radiologists. It aids in the timely identification of tumours, which can facilitate prompt interventions and potentially result in better patient outcomes. Nevertheless, it is crucial to emphasize that the CAD system should always be utilized as a support to medical professionals, and the ultimate diagnosis should be made by an experienced radiologist who combines thorough analysis with clinical expertise and judgment.

5. References

U. Naseem et al., "An Automatic Detection of Breast Cancer Diagnosis and Prognosis Based on Machine Learning Using Ensemble of Classifiers," in IEEE Access, vol. 10, pp. 78242-78252, 2022.

N. Wu et al., "Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening," in IEEE Transactions on Medical Imaging, vol. 39, no. 4, pp. 1184-1194, April 2020.

P. E. Jebarani, N. Umadevi, H. Dang and M. Pomplun, "A Novel Hybrid K-Means and GMM Machine Learning Model for Breast Cancer Detection," in IEEE Access, vol. 9, pp. 146153-146162, 2021.

S. S. Prakash and K. Visakha, "Breast Cancer Malignancy Prediction Using Deep Learning Neural Networks," 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), 2020.

I. Hirra et al., "Breast Cancer Classification From Histopathological Images Using Patch-Based Deep Learning Modeling," in IEEE Access, vol. 9, pp. 24273-24287, 2021.

N. Kumar, G. Sharma and L. Bhargava, "The Machine Learning based Optimized Prediction Method for Breast Cancer Detection," 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2020.

G. Wadhwa and A. Kaur, "A Deep CNN Technique for Detection of Breast Cancer Using Histopathology Images," 2020 Advanced Computing and Communication Technologies for High Performance Applications , 2020.

Z. Wang et al., "Breast Cancer Detection Using Extreme Learning Machine Based on Feature Fusion With CNN Deep Features," in IEEE Access, vol. 7, pp. 105146-105158, 2019.

C. B. Gonçalves, J. R. Souza and H. Fernandes, "Classification of static infrared images using pre-trained CNN for breast cancer detection," 2021 IEEE 34th International Symposium on Computer-Based Medical Systems , 2021

M. Naderan and Y. Zaychenko, "Convolutional Autoencoder Application for Breast Cancer Classification," 2020 IEEE 2nd International Conference on System Analysis & Intelligent Computing (SAIC), 2020.

X. Zhang et al., "Classification of Whole Mammogram and Tomosynthesis Images Using Deep Convolutional Neural Networks," in IEEE Transactions on NanoBioscience, vol. 17, no. 3, pp. 237-242, July 2018.