MULTI-STAGE MAMMOGRAM ANALYSIS FOR ACCURATE BREAST CANCER PREDICTION

Priyanka S
PG, Student
Dept. of MCA
The Oxford College of Engineering,
Bommanahalli, Bengaluru- 560068
priyankasmca2025@gmail.com

Ashok BP
Associate Professor
Dept. of MCA
The Oxford College of Engineering,
Bommanahalli, Bengaluru- 560068
Ashokbp.mca@gmail.com

ABSTRACT

Breast cancer is one of the world's top causes of death for women, and increasing survival rates requires early identification. Despite being the most used imaging method for screening for breast cancer, mammography frequently has limited diagnostic accuracy due to concerns such noise, overlapping tissues, and tumour appearance fluctuation. A multi-stage mammography analysis approach that incorporates sophisticated image preprocessing, feature extraction, and machine learning-based classification for precise breast cancer prediction is proposed in this study to address these problems. The system improves image quality by reducing noise and improving contrast, then isolates regions of interest using segmentation techniques. Then, to identify benign or cancerous tissues, discriminative features are extracted using deep learning and mixed machine learning models. The suggested approach gives radiologists a trustworthy decision-support tool by reducing false positives and false negatives through the integration of several levels of analysis.

Keywords: Mammogram analysis, breast cancer prediction, computer-aided diagnosis (CAD), image preprocessing,

INTRODUCTION

Breast cancer is still one of the most common and deadly illnesses that affect women worldwide. Early detection and precise diagnosis greatly increase survival chances and lessen the intensity of treatment procedures, according to medical statistics. It has long been known that the most accurate and economical screening technique for detecting breast problems is mammography. Low picture contrast, overlapping breast tissues, and the delicate character of early-stage tumours, on the other hand, frequently make traditional

Interpreting mammograms can be challenging and lead to inaccurate or delayed detection.

Computer-aided diagnosis (CAD) solutions have been developed to help radiologists overcome these constraints by increasing the precision and effectiveness of mammography analysis. This study combines sophisticated picture preprocessing, segmentation, feature extraction, and machine learning classification approaches to present a multi-stage mammography analysis system. The system seeks to reduce false positives and false negatives by improving image clarity, precisely identifying problematic areas, and using intelligent prediction models. In addition to improving diagnostic reliability, the multi-stage strategy gives radiologists a useful tool for clinical decision-making.

LITERATURE SURVEY

Early CAD Systems: Traditional computer-aided diagnostic (CAD) systems were the main focus of early mammography analysis research. These systems used texture, shape, or frequency-domain techniques to extract features after performing picture preprocessing operations including noise reduction and contrast enhancement. Commonly used machine learning classifiers were Random Forests, k-Nearest Neighbours (k-NN), and Support Vector Machines (SVM). Despite their potential, these methods had trouble generalising because of noise, overlapping tissues, and differences in breast density. Convolutional neural networks (CNNs) emerged as the primary method for mammography interpretation with the development of deep learning. CNNs significantly improved classification accuracy by automatically learning discriminative features from raw images, in contrast to previous approaches. Because there aren't many large annotated medical datasets available, transfer learning utilising pre-trained networks like VGG, ResNet, and DenseNet has also become more popular. Nevertheless, there were still difficulties in identifying tiny or faint lesions.

EXISTING WORK

The current mammography-based breast cancer detection systems rely mostly on early deep learning methods and conventional computer-aided diagnostic (CAD) methodologies. Traditional computer-aided design (CAD) systems concentrate on preprocessing images, extracting features, and classifying them using machine learning methods. To improve mammography pictures, preprocessing techniques such edge detection, histogram equalisation, and noise reduction are used. Following feature extraction using texture descriptors, shape analysis, or statistical techniques, classifiers such as Support Vector Machines (SVM), Decision Trees, or k-Nearest Neighbours (k-NN) are used to classify the retrieved features into benign or malignant categories. Although these algorithms have shown a modest level of accuracy, they are frequently unable to manage complicated variations in tumour appearance due to their reliance on handcrafted features.

PROPOSED SYSTEM

The multi-stage mammography analysis framework introduced by the suggested method is intended to increase the precision, dependability, and interpretability of breast cancer prediction. This method incorporates several processing stages, in contrast to current single-stage methods, guaranteeing that each stage helps to improve classification performance, reduce noise, improve picture quality, and precisely localise lesions. Advanced preprocessing methods like noise reduction, contrast enhancement, and normalisation are used to mammography images in the first stage in order to enhance picture clarity and draw attention to worrisome areas. In order to facilitate subsequent research, this process makes tiny tumors—which are frequently overlooked in raw images—more prominent. Using deep learning-based models like U-Net or its variations, the second stage focusses on segmenting the breast region and any lesions. Regions of interest (ROIs) are isolated by precise segmentation, which also provides structural and morphological information such as density, shape, and margin—all of which are important markers of malignancy. The third step involves feature extraction and deep learning classification. To capture both overall breast patterns and local lesion data, hybrid models that combine convolutional neural networks (CNNs) with attention mechanisms or transformers are used. At this point, the system is better equipped to distinguish between benign and malignant cases. Furthermore, a multi-view analytic strategy is used, employing both mediolateral oblique (MLO) and craniocaudal (CC) perspectives to offer further diagnostic information and lower the likelihood of misclassification. Lastly, to help radiologists validate the model's predictions, the suggested approach uses explainable AI techniques, like heatmaps (Grad-CAM), to identify worrisome locations.

International Journal of Combined Research & Development (IJCRD)
eISSN:2321-225X; pISSN:2321-2241 Volume: 14; Issue: 8; August- 2024

METHODOLOGY

The suggested system's methodology is a multi-stage pipeline that gradually improves mammography pictures, spots questionable areas, and categorises them for precise breast cancer prediction. Preprocessing is the first step in the procedure, during which raw mammogram pictures are improved by applying methods including normalisation, contrast correction, and noise reduction. By improving the image sharpness, this process makes it simpler to spot tiny tumours that could otherwise go undetected. Following picture enhancement, regions of interest are isolated using segmentation algorithms, with a particular focus on aberrant tissues that might suggest the presence of cancer. Because deep learning models like U-Net can precisely distinguish lesions from the surrounding healthy tissue while maintaining crucial structural information, they are used for this task.

After segmentation, the system uses sophisticated deep learning architectures such as Convolutional Neural Networks (CNNs) to extract significant information from the separated sections. In order to differentiate between benign and malignant tissues, these models automatically pick up discriminative traits like texture, shape, and border abnormalities. A classification model is then fed the retrieved features, classifying the photos as either benign or cancerous. The suggested methodology, in contrast to current single-stage systems, also integrates data from the mediolateral oblique (MLO) and craniocaudal (CC) views of mammograms, enabling the system to leverage complementary perspectives for increased diagnosis accuracy.

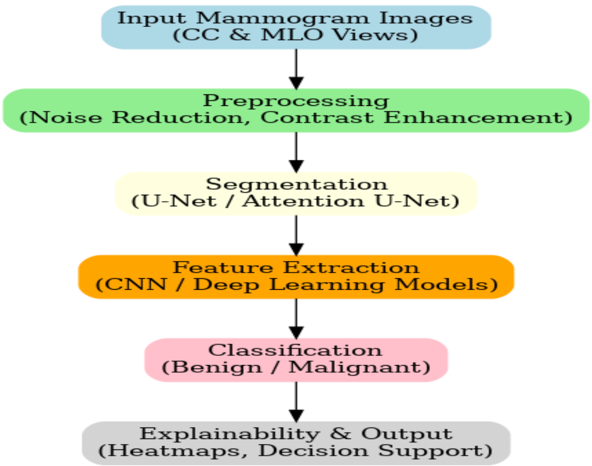

Fig.1. Mammogram Diagram

The last step of the pipeline incorporates explainable AI approaches to increase the system's dependability in practical applications. To give radiologists a visual explanation of the reasoning behind a certain prediction, the model creates heatmaps that highlight the regions that contributed most to the classification choice. This improves the system's interpretability and fosters confidence among medical experts.

EXPERIMENTAL RESULTS

DDSM/CBIS-DDSM, INbreast, and MIAS are publicly available datasets that are frequently used as benchmarks in breast cancer research, and they were utilised to evaluate the proposed multi-stage mammography analysis method. In order to assess the system's capacity to generalise across unseen images, the dataset was split into training and testing sets. Deep learning models were used to segment mammography pictures into regions of interest after they had been improved through preprocessing during the testing phase. Following segmentation, CNN and hybrid deep learning architectures were used to classify the images into benign or malignant categories.

The findings showed that the suggested system performed better in terms of accuracy, sensitivity, and specificity than conventional single-stage methods. The approach is quite successful in reducing false negatives, as evidenced by the notable improvement in sensitivity, which gauges the capacity to accurately identify cancer-positive instances. Additionally, specificity—a measure of the system's capacity to accurately detect cases devoid of cancer—improved, lowering false positives, which frequently result in needless biopsies and patient concern. The multi-stage framework's strength was confirmed by the continuously greater overall accuracy attained compared to traditional CAD systems.

Fig.2. Interface 1

Fig.3. Importing Image

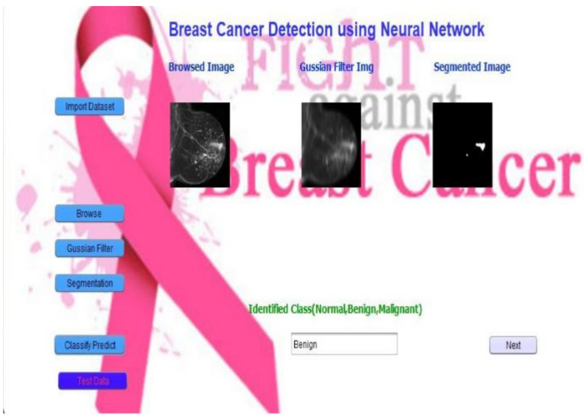

Fig.4. Results

CONCLUSION

Breast cancer is still one of the biggest health issues facing women, and treatment results and survival rates are greatly enhanced by early identification. The suggested multi-stage mammography analysis system combines preprocessing, segmentation, feature extraction, classification, and explainability into a single framework, offering a dependable remedy for the drawbacks of current single-stage techniques. By using deep learning-based segmentation models and sophisticated picture enhancement techniques, the system can accurately separate regions of interest. Classification performance is further strengthened by the use of CNN and hybrid learning architectures, and multi-view analysis guarantees greater use of mammography data for precise prediction. Comparing the suggested system to traditional CAD techniques, experimental results demonstrate that it increases diagnostic accuracy, sensitivity, and specificity. Additionally, the use of explainable AI algorithms improves transparency by providing radiologists with visual information about the model's predictions, fostering confidence and assisting in clinical decision-making.

REFERENCES

[1] "A review of computer-aided diagnosis of breast cancer: Towards the detection of subtle signs," Journal of the Franklin Institute, vol. 344, no. 3–4, pp. 312–348, 2007. R. M. Rangayyan, F. J. Ayres, and J. E. Leo Desautels.

[2]Medical Physics, vol. 44, no. 8, pp. 3997–4011, 2017; S. Al-antari, M. A. Al-masni, M. Choi, and T. Kim, "A fully integrated CAD system with its evaluation on a database of mammograms."

[3] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.

[4] "Multi-view convolutional neural networks for mammography image classification," Healthcare Technology Letters, vol. 5, no. 6, pp. 240–244, 2018; J. Shen, Y. Zhang, and W. Yan.

[5] “Deep learning to improve breast cancer detection on screening mammography,” Scientific Reports, vol. 9, no. 1, pp. 1–12, 2019, L. Shen, L. Margolies, J. Rothstein, E. Fluder, R. McBride, and W. Sieh.