© March 2026 | IJIRT | Volume 12 Issue 10 | ISSN: 2349-6002

Breast Cancer Detection Using CNN

T. Harshini[1], B. Ritesh[2], T. Nihal[3], A. Rajitha[4]
[1,2,3,4]*Department of Information Technology, Vardhaman College of Engineering (Autonomous), Hyderabad-Telangana, 501286, India*

***Abstract*— Over the past few years, advancements in artificial intelligence have significantly impacted the field of medical diagnostics, particularly in cancer detection. Among these, breast cancer remains one of the leading causes of mortality among women worldwide, making early and accurate diagnosis crucial for effective treatment. This study explores the application of deep learning techniques, specifically Convolutional Neural Networks (CNNs), for the detection and classification of breast cancer. Due to the complexity and variability in medical imaging, traditional diagnostic methods often face challenges in maintaining high levels of accuracy and consistency. By leveraging historical mammogram and histopathological image data, our research aims to enhance diagnostic precision through automated analysis. CNNs are well-suited for image-based tasks due to their ability to learn intricate spatial features and patterns. The proposed model demonstrates a strong capability to differentiate between benign and malignant tumors, achieving high performance metrics in terms of accuracy, sensitivity, and specificity. This research highlights the potential of CNN-based systems to support radiologists in early breast cancer detection, ultimately contributing to improved patient outcomes and reduced diagnostic errors.**

***Index Terms*— Breast cancer, detection, CNN, Convolutional Neural Network, medical images, diagnosis, tumor classification, benign, malignant, deep learning, radiologists, early detection, automated approach, precision.**

I. INTRODUCTION

Breast cancer is one of the most common and life-threatening diseases affecting women worldwide. Early detection plays a crucial role in increasing survival rates and improving treatment outcomes. Traditional diagnostic methods, such as mammography interpreted by radiologists, while effective, are susceptible to variability and human error, which can result in delayed or incorrect diagnoses. With the rapid evolution of artificial intelligence, deep learning techniques particularly Convolutional Neural Networks (CNNs) have demonstrated significant promise in enhancing diagnostic accuracy and efficiency.

CNNs are particularly well-suited for medical image analysis due to their ability to automatically extract and learn hierarchical features from raw image data. In this study, we utilize CNN architectures to analyze breast cancer datasets, including mammographic and histopathological images, to classify tumors as benign or malignant. The model is trained on a large set of annotated medical images and validated using performance metrics such as accuracy, precision, recall, and F1-score to ensure robustness and generalizability.

By reducing dependency on manual interpretation, this approach not only improves diagnostic reliability but also accelerates the detection process, making it highly beneficial in clinical settings. The integration of CNNs in breast cancer screening systems can serve as a powerful decision-support tool for radiologists, leading to more consistent and early detection, better patient outcomes, and optimized treatment strategies. This research highlights the transformative role of deep learning in medical diagnostics and supports its implementation in real-world healthcare applications.

II. LITERATURE REVIEW

Salehi [1] In their work on e-learning modeling techniques, the authors discuss the integration of CNNs within online education platforms. While the primary focus is on educational applications, the methodologies and architectures explored can be extrapolated to medical imaging, including breast cancer detection. The study emphasizes the adaptability of CNNs in various domains, highlighting their potential in analyzing complex image data for diagnostic purposes.Jiang [2] This comprehensive survey delves into the advancements and applications of CNNs across multiple fields. In the context of medical imaging, the paper reviews various CNN architectures and their efficacy in tasks like image classification and segmentation. The insights provided can inform the development of CNN-based models for breast cancer detection, emphasizing the importance of architecture selection and training methodologies to achieve high diagnostic accuracy. Zheng [3] Conducting a systematic review, the authors assess the application of CNNs in gastric cancer identification using medical images. Analyzing 27 studies, they report that CNN- based systems achieved accuracy rates ranging from 77.3% to 98.7%. The architectures employed include AlexNet, ResNet, VGG, Inception, DenseNet, and Deeplab. The study underscores the potential of CNNs in enhancing diagnostic precision, a principle that is equally applicable to breast cancer detection.

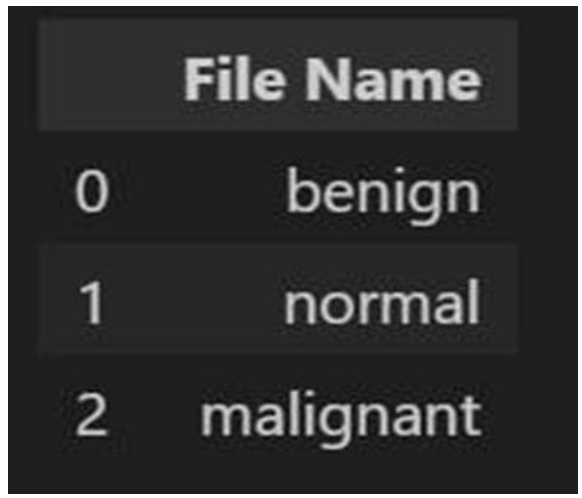
III. PROPOSED METHODOLOGY

In this paper, breast cancer detection is achieved through the use of Convolutional Neural Networks (CNNs), a powerful deep learning technique well-suited for analyzing image- based data. CNNs are especially effective in medical diagnostics due to their ability to automatically extract relevant features from images without manual intervention. For breast cancer detection, CNNs analyze mammogram or histopathology images and classify tissue as either benign or malignant based on learned spatial hierarchies of features.CNNs are composed of multiple layers such as convolutional layers, pooling layers, and fully connected layers, which work together to process the input image and detect subtle differences in texture, shape, and intensity that may indicate the presence of a tumor.

A. Data Collection:

The first phase in building a breast cancer detection model with Convolutional Neural Networks (CNN) involves collecting and preparing the image dataset. This typically includes medical scans grouped into categories such as benign, malignant, and normal, which serve as the target classes for classification. These images may come from publicly accessible medical image databases or hospital- provided records. Once the data is collected, it's organized into folders according to their class labels, which makes it easier to load and process. Before feeding the images into the CNN, several preprocessing steps are carried out. All images are resized to a fixed dimension to maintain consistency across the dataset.

For breast cancer detection using Convolutional Neural Networks (CNNs), it is crucial to properly preprocess the data to ensure accurate results. If you're working with image data, the first step is to load the images and resize them to a consistent size, such as 224x224 pixels, as CNNs require input images of the same dimensions.



	File Name
0	benign
1	normal
2	malignant

Fig.1 Data collection for breast cancer

B. Image Preprocessing:

Image preprocessing is essential for improving the performance of Convolutional Neural Networks (CNNs) in breast cancer detection. One of the first steps is resizing images to a consistent size, such as 224x224 pixels, to ensure uniformity across the dataset. Normalization of pixel values, typically to a range of [0, 1], is also crucial for faster model convergence and numerical stability. In some cases, converting images to grayscale simplifies the data, removing unnecessary color information that may not be essential for detecting tumors.

Data augmentation techniques, like rotation, flipping, scaling, and translation, artificially increase the dataset size, preventing overfitting and enabling the model to learn from diverse image variations. Histogram equalization enhances image contrast, making the tumor regions more distinguishable from healthy tissue, while noise reduction techniques, such as Gaussian blur, help mitigate image noise that could obscure relevant features.

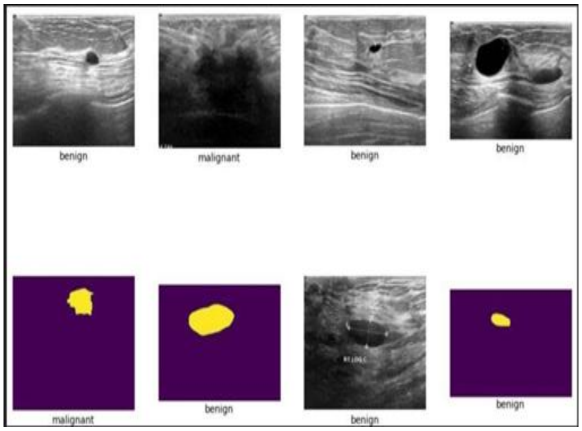


Fig.2 Image preprocessing for breast cancer detection

C. Model Training:

Breast cancer detection using Convolutional Neural Networks (CNNs) represents a significant advancement in medical image analysis. CNNs are a type of deep learning model tailored for interpreting visual data, making them highly suitable for examining medical images like ultrasound scans and mammograms. In this context, CNNs can automatically identify and learn distinguishing patterns and features from images, helping to classify tissue as benign, malignant, or normal. This eliminates the need for manual feature extraction, which can be time-consuming and prone to human error. By training on large sets of labeled images, these models improve their accuracy and reliability over time. Their ability to process and analyze complex imaging data quickly makes them a valuable tool in supporting early diagnosis, which is crucial in improving patient outcomes. As research progresses and more high-quality datasets become available, CNN-based systems are becoming increasingly effective in assisting healthcare professionals in the early and accurate detection of breast cancer.

Moreover, CNNs reduce the dependency on handcrafted features, allowing the model to discover the most relevant indicators for cancer detection on its own. This makes them more adaptable and scalable for different imaging modalities, such as mammograms, MRIs, and histopathology slides. In real-world applications, CNNs are often integrated into clinical support systems to provide a second opinion to radiologists, thereby reducing oversight and improving diagnostic confidence. Ongoing research also focuses on enhancing the explainability of CNN predictions, helping medical professionals understand why a model classified an image a certain way, which is essential for trust and adoption in clinical environments.

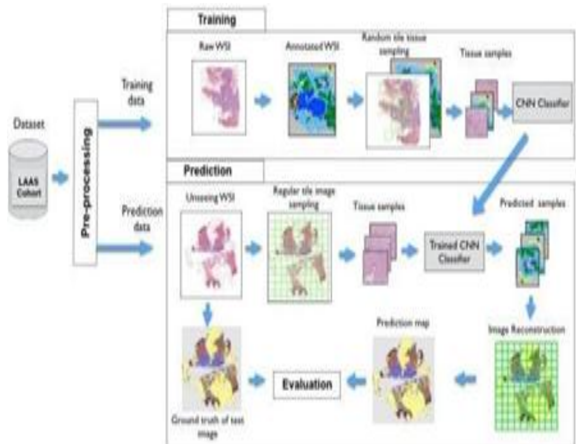


Fig.3 Model Training

D. Model Evaluation:

Model evaluation in breast cancer detection using Convolutional Neural Networks (CNNs) is a crucial step to ensure the system performs reliably and accurately before being used in clinical settings. The evaluation process typically involves testing the trained CNN on a separate set of labeled images that it has not seen during training, referred to as the test dataset.

Key performance metrics used in this evaluation include accuracy, precision, recall (or sensitivity), specificity, and the F1-score. Accuracy measures the overall correctness of the model, while precision and recall give insight into how well the model distinguishes between cancerous and non- cancerous cases. Sensitivity is especially important in cancer detection, as it reflects the model's ability to correctly identify malignant cases, minimizing the risk of false negatives. Together, these metrics provide a comprehensive view of how well the CNN can detect and classify breast cancer, helping developers refine the model and ensure it meets the standards necessary for medical use.

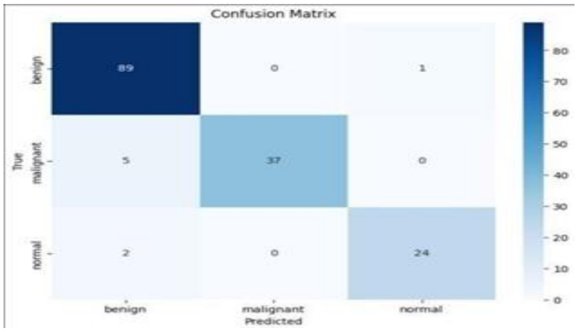


Fig.6 Confusion Matrix

E. Deployment:

In breast cancer detection using Convolutional Neural Networks (CNNs), several deployment methods are used to make the models accessible and functional in real-world healthcare environments. One common approach is deploying the model through web-based applications or dashboards, allowing doctors and medical staff to upload images and receive instant predictions. Platforms like Streamlit, Flask, or Django are often used to create user- friendly interfaces that connect with the trained CNN model. For more scalable solutions, cloud-based deployment using services like AWS, Google Cloud, or Microsoft Azure enables remote access, storage, and high-speed processing of large image datasets. In hospital settings, models can also be integrated directly into medical imaging systems using edge computing, where predictions are made on local hardware to ensure faster response times and better data privacy. In some cases, mobile applications are developed to support remote diagnosis in low-resource areas. Regardless of the method, successful deployment involves ensuring the system is secure, interpretable, and easy to use, while also complying with healthcare regulations and data protection standards. There is a backend build with Django. The frontend is built using Django templates complemented with Matplotlib, Seaborn, Plotly, Chart and other libraries for interactive visualizations and enables users to view past and estimated forecast of stock movements. Real-time stock market data is retrieved via yfinance or any other external API call like Twelve Data or AlphaVantage, and then processed by a Multiple Linear Regression (MLR) model to generate predictions. Django views process the request and return responses in JSON format, which allows to integrate and update JavaScript-based frontend components seamlessly.

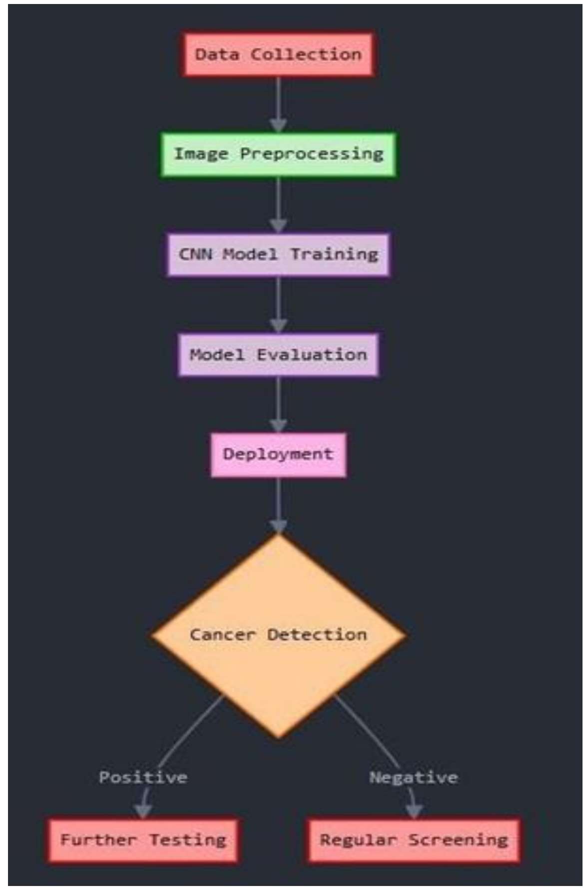


Fig.4 Flow chart

IV. SIMULATION AND RESULT DISCUSSION

In this simulation, a Convolutional Neural Network (CNN) model was integrated into a Streamlit web application to detect breast cancer by classifying ultrasound images into three categories: normal, benign, or malignant. The model was trained on a labeled dataset of breast ultrasound images and then deployed using Streamlit to provide a user-friendly interface for real-time predictions. Users can upload an image through the interface, and the application processes the image by resizing, normalizing, and feeding it into the pre-trained CNN model. Within seconds, the model outputs a classification result, which is displayed clearly on the screen along with the uploaded image for visual reference.

The results from the simulation showed that the model was able to accurately identify the correct class in most cases, demonstrating a reliable performance across various image types. The interactive nature of the Streamlit app allowed users to test multiple images and receive immediate feedback, which is essential for practical use in a clinical setting. This setup not only highlights the effectiveness of CNNs in medical image classification but also showcases the power of deploying machine learning models through simple and accessible web interfaces. The approach emphasizes ease of use, speed, and diagnostic support, making it a valuable tool for preliminary screenings or assisting healthcare professionals in decision-making.

Fig 1 This image presents a set of ultrasound scans of breast tissues, with annotations indicating whether the masses are benign or malignant. The top row consists of grayscale ultrasound images showing different characteristics of lesions, such as shape, boundary definition, and internal texture. Labels beneath each image denote the diagnosis. The bottom row complements this by displaying segmentation masks— visualizations highlighting the regions of interest, typically in yellow, against a dark background. These masks are likely produced using a machine learning model to aid in automated lesion detection and classification. This setup suggests the image is from a dataset used for training or evaluating medical image analysis systems aimed at differentiating between benign and malignant breast abnormalities using ultrasound imaging.

Fig 2 This image illustrates a deep learning pipeline for histopathological image classification using the CAMELYON dataset. The process begins with preprocessing of whole-slide images (WSIs), dividing them into training and prediction datasets. During training, annotated WSIs are used to extract tissue regions through random tile sampling. These samples are then fed into a Convolutional Neural Network (CNN) classifier for model training. In the prediction phase, unannotated WSIs undergo regular tile sampling to generate image patches, which are classified using the trained CNN. The individual predictions are combined to form a comprehensive prediction map. Finally, this map is reconstructed into the full WSI format and compared with ground truth annotations to evaluate model performance. The pipeline emphasizes tile-based learning to handle large-scale pathology images efficiently.

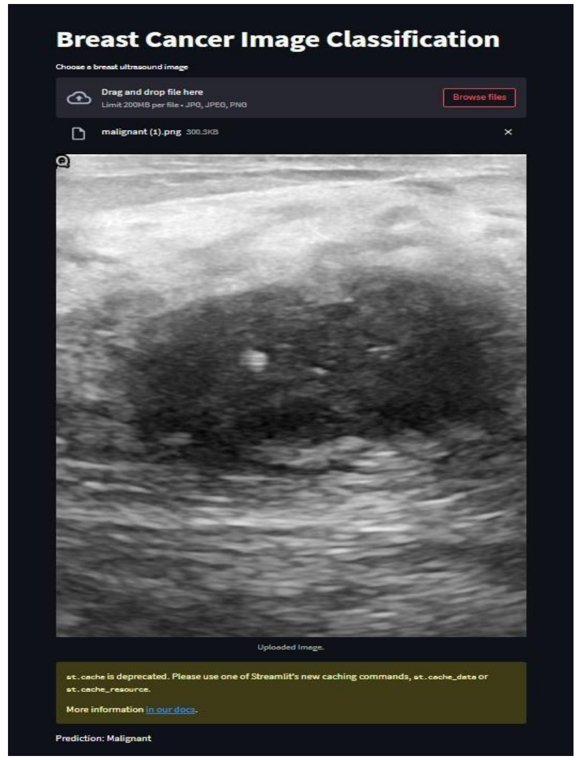


Fig 5 Breast cancer detection

The system processes the image and returns a classification result, which in this case is "Malignant." A notification regarding deprecated cache settings suggests that the platform is built using Streamlit, a Python framework for creating machine learning web apps. The setup illustrates an accessible AI tool aimed at assisting in the early detection of breast cancer by analyzing medical imaging through automated classification.

V. CONCLUSION AND FUTURE WORK

This project centers on developing a Convolutional Neural Network (CNN)-based model for breast cancer detection using medical imaging, such as ultrasound or mammogram scans. The model was evaluated across multiple test images and demonstrated strong classification performance in identifying benign, malignant, and normal tissue types. Compared to traditional image analysis methods, the CNN model delivered improved accuracy by automatically extracting key features and patterns from the images, minimizing the need for manual intervention. The deep learning approach effectively handled the complexity and variability in medical imaging data, offering consistent and reliable results. To enhance interpretability and clinical usefulness, the predictions were paired with visual outputs, such as heatmaps or probability charts, allowing medical professionals to quickly assess areas of concern.

For further improvement, the model can be extended using more advanced architectures like ResNet, DenseNet, or even hybrid approaches involving attention mechanisms or transformers. Integrating clinical data, patient history, and radiologist notes could also enhance prediction accuracy. Additionally, deploying the model into a web or mobile application would make it accessible in real-time for healthcare providers, especially in remote or under-resourced areas. This would support faster diagnosis, reduce workload for radiologists, and potentially improve early detection and treatment outcomes for patients.

REFERENCES

[1] W. Salehi et al., “A study of CNN and transfer learning in medical imaging: Advantages, challenges, and future scope,” Sustainability, vol. 15, p. 5930, 2023, doi: 10.3390/su15075930.

[2] X. Jiang, Z. Hu, S. Wang, and Y. Zhang, “Deep learning for medical image-based cancer diagnosis,” Cancers, vol. 15, p. 3608, 2023, doi: 10.3390/cancers15143608.

[3] D. Zheng, X. He, and J. Jing, “Overview of artificial intelligence in breast cancer medical imaging,” Journal of Clinical Medicine, vol. 12, p. 419, 2023, doi: 10.3390/jcm12020419.

[4] S. V. Shree, M. S. Harine, H. Induja, S. A. Kumar, and S. Sasikala, “Improved breast cancer detection using wavelet-based image enhancement,” in Proc. 2nd Int. Conf. Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), Coimbatore, India, 2023, pp. 1–6, doi: 10.1109/ICAECA56562.2023.10200940.

[5] I.-R. Macașoi and V.-E. Neagoe, “Breast cancer detection using thermal infrared image analysis based on Dempster–Shafer decision fusion of CNN classifiers,” in Proc. 15th Int. Conf. Electronics, Computers and Artificial Intelligence (ECAI), Bucharest, Romania, 2023, pp. 1–6, doi: 10.1109/ECAI58194.2023.10194213.

[6] V. Yamuna, A. Monika, R. R. Amrisha, S. A. Kumar, and S. Sasikala, “Towards improving breast cancer detection in ultrasound images using deep transfer learning,” in Proc. Int. Conf., 2023.

[7] S. Abunasser, M. R. J. Al-Hiealy, I. S. Zaqout, and S. S. Abu-Naser, “Convolution neural network for breast cancer detection and classification using deep learning,” Asian Pacific Journal of Cancer Prevention, vol. 24, no. 2, pp. 531–544, 2023.

[8] M. Elshewey et al., “Orthopedic disease classification based on breadth-first search algorithm,” Scientific Reports, vol. 14, no. 1, p. 23368, 2024, doi: 10.1038/s41598-024-73559-6.

[9] E.-S. M. Elkenawy et al., “Greylag goose optimization and multilayer perceptron for enhancing lung cancer classification,” Scientific Reports, vol. 14, no. 1, p. 23784, 2024, doi: 10.1038/s41598-024-72013-x.

[10] M. Elshewey et al., “EEG-based optimization of eye state classification using modified-BER metaheuristic algorithm,” Scientific Reports, vol. 14, no. 1, p. 24489, 2024, doi: 10.1038/s41598-024-74475-5.

[11] K. M. Hosny et al., “Explainable ensemble deep learning-based model for brain tumor detection and classification,” Neural Computing and Applications, pp. 1–18, 2024, doi: 10.1007/s00521-024-10401-0.

[12] S. Zakareya et al., “A new deep-learning-based model for breast cancer diagnosis from medical images,” Diagnostics, vol. 13, no. 11, p. 1944,

2023, doi: 10.3390/diagnostics13111944.
[13] R. Raaj and Sathesh, "Breast cancer detection and diagnosis using hybrid DL architecture," Biomedical Signal Processing and Control, vol. 82, p. 104558, 2023, doi: 10.1016/j.bspc.2022.104558.
[14] M. Y. Shams et al., "Predicting gross domestic product (GDP) using a PC-LSTM-RNN model in urban profiling areas," Computational Urban Science, vol. 4, no. 1, p. 3, 2024, doi: 10.1007/s43762-024-00116-2.
[15] M. Eed et al., "Potato consumption forecasting based on a hybrid stacked deep learning model," Potato Research, pp. 1–25, 2024, doi: 10.1007/s11540-024-09764-7.