BREAST CANCER DIAGNOSIS

[1] M SRAVANTHI, [2] P. SRAVYA, [3] S. VANDHITHA, [4] A. VARALAXMI

[1]ASSISTANT PROFESSOR, (Information Technology), Bhoj Reddy Engineering College for Women.

[234]B.TECH, SCHOLAR, (Information Technology), Bhoj Reddy Engineering College for Women.

ABSTRACT

Breast cancer is one of the leading causes of death among women worldwide, and early detection plays a crucial role in improving survival rates. This project focuses on the development of an intelligent system for breast cancer diagnosis using advanced deep learning techniques using MobileNetV2. The proposed system analyzes breast cancer histopathology images, to classify whether a tumor is benign or malignant. By utilizing algorithms such as neural networks and other classification models, the system provides accurate and efficient predictions, thereby reducing the chances of human error in diagnosis. The model is trained on relevant datasets to learn patterns and improve its performance over time. In addition to prediction, the system offers a user-friendly interface that allows healthcare professionals to input patient data and receive instant diagnostic results. This solution aims to assist doctors in making faster and more reliable decisions, ultimately contributing to early detection and better treatment outcomes. Overall, this project enhances diagnostic accuracy, reduces manual effort, and provides a cost-effective and efficient tool for breast cancer detection, making it highly valuable in modern healthcare systems.

1.INTRODUCTION

1.1 INTRODUCTION OF THE PROJECT

Breast cancer is one of the most prevalent and life-threatening diseases affecting women worldwide. It accounts for a significant percentage of cancer-related deaths, making it a major public health concern. According to global health statistics, millions of new cases are diagnosed each year, and early detection plays a crucial role in improving survival rates and reducing mortality. The complexity of breast cancer lies in its heterogeneous nature, where tumors can vary significantly in size, shape, and behavior. Therefore, accurate and timely diagnosis is essential for effective treatment planning and patient management.

Traditionally, breast cancer diagnosis has relied on clinical examinations, mammography, biopsy, and histopathological analysis. While these methods are effective, they are often time-consuming, expensive, and prone to human

error. In many cases, misdiagnosis or delayed diagnosis can lead to severe consequences, including disease progression and reduced chances of survival. With the increasing availability of medical data and advancements in computational technologies, there is a growing interest in developing automated systems that can assist healthcare professionals in diagnosing breast cancer more accurately and efficiently.

Machine learning and deep learning techniques have emerged as powerful tools in the field of medical diagnosis. These techniques enable computers to learn patterns from large datasets and make predictions based on learned information. In the context of breast cancer diagnosis, machine learning algorithms can analyze features extracted from medical images or biopsy data to classify tumors as benign or malignant. The use of such intelligent systems not only improves diagnostic accuracy but also reduces the workload of healthcare professionals.

One of the commonly used datasets in breast cancer research is breast cancer histopathology image datasets. These datasets contain microscopic images of breast tissue samples, which are used to train deep learning models. Instead of relying on manually extracted features such as radius or texture, convolutional neural networks automatically learn important features directly from images. By analyzing these images, deep learning models can effectively classify tumors as benign or malignant.

The integration of artificial intelligence in healthcare has opened new possibilities for early detection and personalized treatment. Automated diagnostic systems can provide real-time predictions, enabling doctors to make informed decisions quickly. Moreover, these systems can continuously improve their performance by learning from new data, making them highly adaptable and scalable.

This project aims to develop an intelligent breast cancer diagnosis system using machine learning techniques. The system focuses on analyzing tumor data, training predictive models, and providing accurate classification results. By leveraging advanced algorithms and data-driven approaches, the proposed solution aims to enhance diagnostic accuracy, reduce human error, and contribute to better healthcare outcomes.

2. LITERATURE SURVEY

Breast cancer remains one of the leading causes of mortality among women worldwide, making early detection and

diagnosis critical for improving survival rates. Traditionally, diagnostic procedures such as mammography, ultrasound, and biopsy depend heavily on the expertise of radiologists, which can lead to variability in interpretation and potential diagnostic errors. With the advancement of artificial intelligence, particularly deep learning, automated systems have been developed to assist in accurate and efficient diagnosis of breast cancer. Deep learning models, especially convolutional neural networks (CNNs), have shown remarkable performance in analyzing medical images by automatically extracting complex features without the need for manual intervention.

Several researchers have contributed significantly to this domain. Abdelrahman et al. (2021) provided a comprehensive overview of CNN applications in mammography, emphasizing their effectiveness in detection, classification, segmentation, and risk prediction. Their work demonstrated that CNN-based models outperform traditional computer-aided diagnosis systems by learning hierarchical features directly from images. Similarly, Abdelhafiz et al. (2019) explored deep convolutional neural networks and highlighted their superiority over conventional machine learning approaches, particularly due to their ability to eliminate handcrafted feature extraction. However, they also pointed out challenges such as limited dataset availability, class imbalance, and lack of interpretability.

Chougrad et al. (2018) proposed a CNN-based computer-aided diagnosis system for classifying mammographic masses and showed that deep learning can significantly reduce unnecessary biopsies by improving classification accuracy. Their study emphasized the importance of large datasets for training deep learning models effectively. Shen et al. (2019) introduced a deep learning framework that improved breast cancer detection in screening mammography by leveraging large-scale datasets. Their research demonstrated that deep learning systems could identify subtle patterns that may not be easily visible to human experts, thereby enhancing diagnostic performance.

Recent studies have further advanced the application of deep learning in this field. Wang (2024) reviewed the integration of deep learning into clinical workflows and concluded that these systems significantly improve early detection rates while reducing diagnostic errors. Sandhya and Anoop (2024) conducted experiments on multiple datasets such as CBIS-DDSM, MIAS, and INbreast, demonstrating that preprocessing techniques like noise removal and contrast enhancement play a crucial role in improving model performance. Their study also showed that CNN-based systems can achieve high accuracy in classifying images into benign and malignant categories.

Other research has focused on evaluating different deep learning architectures such as AlexNet, VGG16, and ResNet50. These studies revealed that transfer learning techniques significantly enhance model performance, especially when dealing with limited medical datasets. In addition, segmentation-based approaches using models like U-Net have been proposed to improve tumor localization by identifying

regions of interest before classification. Hybrid models combining segmentation and classification have achieved high accuracy and sensitivity on benchmark datasets. Dual-network architectures have also been explored to simultaneously perform segmentation and classification, further improving diagnostic accuracy.

Overall, the literature indicates that deep learning has revolutionized breast cancer diagnosis by providing automated, accurate, and efficient solutions. These models not only outperform traditional methods but also assist radiologists in making informed decisions. Despite these advancements, challenges such as data scarcity, model interpretability, and generalization across diverse populations remain open research areas.

However, the proposed system focuses only on image classification using MobileNetV2 and does not include segmentation models like U-Net.These limitations are overcome by deep learning models that automatically learn features from images.

3.EXISTING METHODS

Breast cancer diagnosis methods have evolved significantly over time, ranging from traditional image processing techniques to advanced deep learning-based approaches. Earlier methods relied on conventional machine learning algorithms that required manual feature extraction. These methods involved preprocessing the images, extracting features such as texture, shape, and intensity, selecting relevant features, and finally classifying them using algorithms like support vector machines, k-nearest neighbors, or decision trees. While these approaches achieved moderate success, their performance was limited due to their dependence on handcrafted features and inability to capture complex patterns in medical images.

Computer-aided diagnosis systems were developed to assist radiologists by highlighting suspicious regions in mammographic images. These systems typically followed a structured pipeline that included image acquisition, preprocessing, region of interest detection, feature extraction, and classification. Although CAD systems improved detection rates to some extent, they were associated with high false positive rates and limited generalization capabilities. Additionally, their reliance on handcrafted features restricted their ability to adapt to diverse datasets.

The introduction of deep learning marked a significant breakthrough in breast cancer diagnosis. Convolutional neural networks became the most widely used models due to their ability to automatically learn hierarchical features from raw images. These networks consist of multiple layers, including convolutional, pooling, and fully connected layers, which enable them to capture both low-level and high-level features. CNN-based methods have been successfully applied to tasks such as tumor detection, classification, and feature extraction, achieving superior performance compared to traditional approaches.

Transfer learning has become a popular technique in deep learning-based diagnosis due to the limited availability of medical datasets. Pre-trained models such as VGG16,

ResNet50, and InceptionV3 are fine-tuned on medical images to improve performance. This approach reduces training time and enhances accuracy, making it suitable for real-world applications. Segmentation-based methods have also gained importance, as they focus on identifying tumor regions before classification. Models like U-Net and Mask R-CNN are widely used for this purpose, improving accuracy by concentrating on relevant regions.

Hybrid models that combine multiple techniques have been proposed to further enhance performance. These models integrate segmentation and classification or combine deep learning with traditional machine learning algorithms. Patch-based methods have also been introduced, where images are divided into smaller regions for detailed analysis. While these methods improve localization accuracy, they require higher computational resources.

Various datasets such as MIAS, DDSM, CBIS-DDSM, and INbreast are commonly used for training and evaluating these models. Performance is typically measured using metrics like accuracy, sensitivity, specificity, precision, F1-score, and area under the curve. Deep learning models have achieved accuracy levels exceeding 95% in many studies, demonstrating their effectiveness in breast cancer diagnosis.

Despite these advancements, existing methods face several limitations. Data scarcity and class imbalance remain major challenges, as collecting and annotating medical images is time-consuming and expensive. Overfitting is another issue, where models perform well on training data but fail to generalize to new data. Additionally, deep learning models often lack interpretability, making it difficult for clinicians to trust their predictions. High computational requirements and challenges in integrating these systems into clinical workflows further limit their widespread adoption.

4. PROPOSED SYSTEM

The proposed system aims to develop an intelligent and efficient deep learning-based solution for breast cancer diagnosis. The system is designed to classify breast cancer images as benign or malignant using a trained convolutional neural network model. This approach improves accuracy, reduces manual effort, and provides quick diagnostic results.

The system begins with data collection, where breast cancer images are obtained from available datasets. These images include both benign and malignant cases, which are used to train and test the model.

In the next step, image preprocessing is performed to improve the quality of input data. This includes resizing images to a fixed size, normalization, and basic enhancement techniques. Preprocessing helps in improving model performance and accuracy.

After preprocessing, the images are fed into a deep learning model based on MobileNetV2 architecture. MobileNetV2 is a lightweight convolutional neural network that is efficient and suitable for image classification tasks. The model learns important features from the images and classifies them into two categories: benign or malignant.

International Journal of Engineering Research and Science & Technology

The trained model then predicts the output for new input images. The system provides the prediction result along with confidence, helping users understand the classification.

A user-friendly interface is developed using Streamlit, where users can upload images and view prediction results instantly. This makes the system easy to use for healthcare professionals and general users.

The proposed system improves diagnostic accuracy, reduces human error, and enables early detection of breast cancer. It is efficient, scalable, and suitable for real-time applications.

5. SCREENSHOTS

OUTPUT SCREENS:

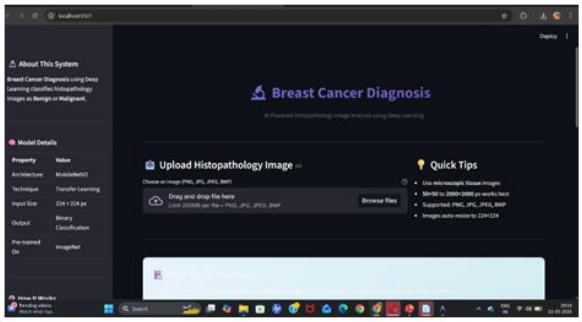


Fig 5.1

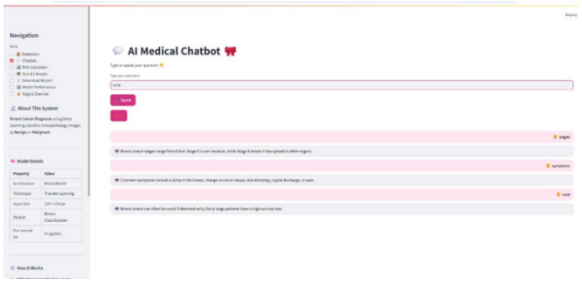


Fig 5.2

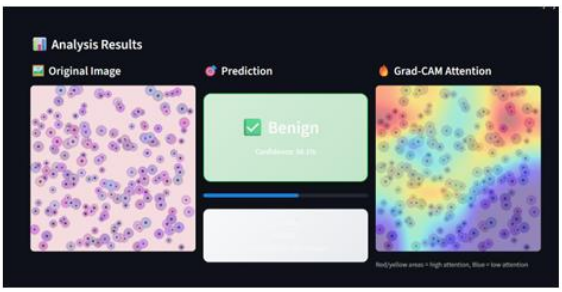


Fig 5.3

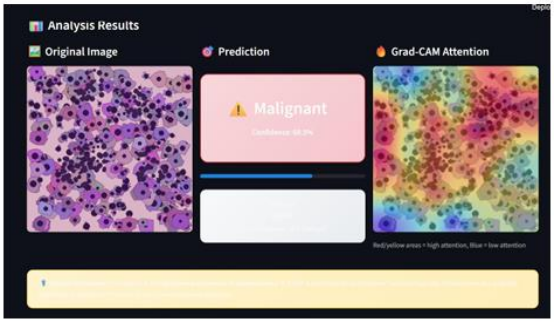


Fig 5.4

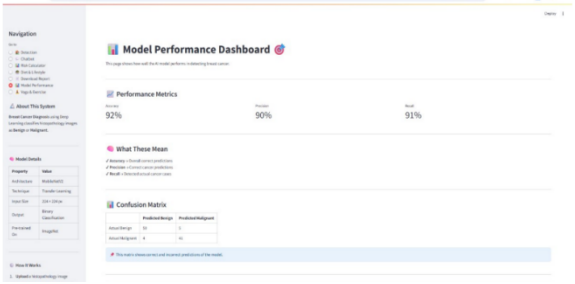


Fig 5.5

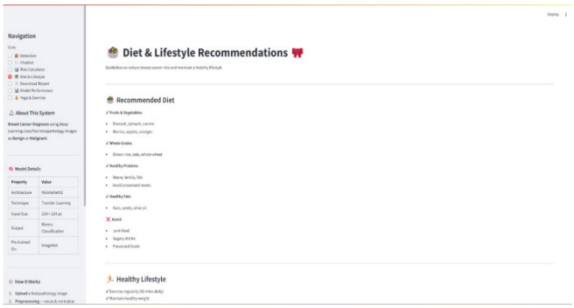


Fig 5.6

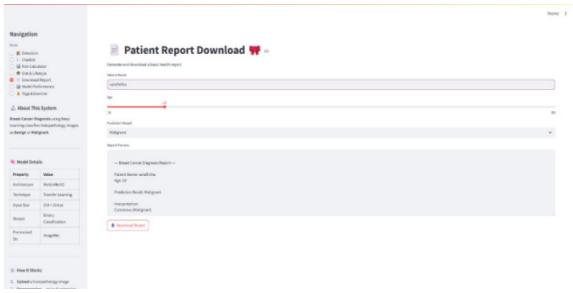


Fig 5.7

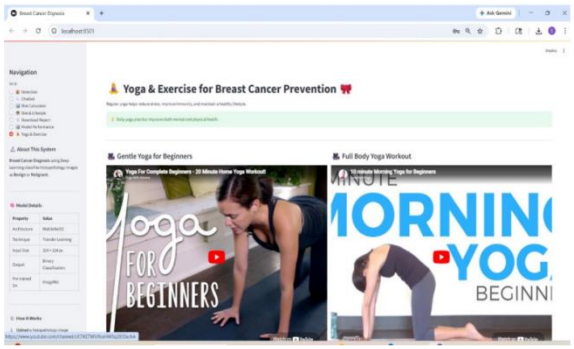


Fig 5.8

6. CONCLUSION

Breast cancer diagnosis is a critical aspect of healthcare that requires accurate and timely decision-making. The integration of deep learning techniques has significantly improved the ability to detect and classify tumors, reducing the risk of misdiagnosis. This project demonstrates the potential of intelligent systems in enhancing diagnostic accuracy and supporting healthcare professionals.

By leveraging deep learning algorithms based on MobileNetV2, the proposed system provides reliable predictions and improves overall efficiency. The use of convolutional neural networks enables automatic feature extraction and improves model performance, making the system suitable for real-world applications.

7. FUTURE SCOPE

The future scope of this project can be expanded in several ways to improve its performance, usability, and real-world application. First, the model can be extended from binary classification to multi-class classification, allowing it to identify specific tumor subtypes such as ductal carcinoma and lobular carcinoma. To improve transparency and trust in predictions, Grad-CAM visualization can be incorporated to highlight the regions of the image that the model focuses on during classification. Training the model on larger datasets like the full BreaKHis or PCam datasets can further enhance robustness and accuracy. Additionally, model fine-tuning can be performed by unfreezing the top layers of MobileNetV2 to achieve better performance. The use of ensemble models, combining architectures such as Mobile Net, ResNet, and EfficientNet, may also lead to improved predictive accuracy. For wider accessibility, the system can be deployed on cloud platforms like Streamlit Cloud, AWS, or Google Cloud. The model can also be

converted to TensorFlow Lite (TFLite) to support mobile applications and enable edge deployment on smartphones. Furthermore, integration with hospital systems, including DICOM viewers and electronic health records, would allow smoother adoption in clinical environments. Finally, implementing active learning could enable pathologists to correct model predictions, allowing the system to continuously learn and improve over time.Future improvements may include integrating segmentation models like U-Net for tumor localization.

8. REFERENCES

1. Abdelhafiz, D., Yang, C., Ammar, R., & Nabavi, S. (2019). Deep convolutional neural networks for mammography: Advances, challenges and applications. *BMC Bioinformatics, 20*(11), 281. https://doi.org/10.1186/s12859-019-2823-9
2. Abdelrahman, L., Abbas, M., & Ibrahim, A. (2021). Convolutional neural networks for breast cancer detection in mammography: A survey. *IEEE Access, 9*, 152392–152411. https://doi.org/10.1109/ACCESS.2021.3123456
3. Chougrad, H., Zouaki, H., & Alheyane, O. (2018). Deep convolutional neural networks for breast cancer screening. *Computer Methods and Programs in Biomedicine, 157*, 19–30. https://doi.org/10.1016/j.cmpb.2018.01.011
4. Shen, L., Margolies, L. R., Rothstein, J. H., Fluder, E., McBride, R., & Sieh, W. (2019). Deep learning to improve breast cancer detection on screening mammography. *Scientific Reports, 9*(1), 12495. https://doi.org/10.1038/s41598-019-48995-4
5. Wang, L. (2024). Deep learning in mammography: Current applications and future directions. *Breast Cancer Research, 26*(1), 45. https://doi.org/10.1186/s13058-024-01830-9
6. Sandhya, C., & Anoop, B. K. (2024). Deep learning-based breast cancer classification using mammogram datasets. *International Journal of Intelligent Systems and Applications in Engineering, 12*(2), 112–120.
7. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 770–778). https://doi.org/10.1109/CVPR.2016.90
8. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)* (pp. 234–241). Springer. https://doi.org/10.1007/978-3-319-24574-4_28
9. Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. In *International Conference on Learning Representations (ICLR)*.
10. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis, 42*, 60–88. https://doi.org/10.1016/j.media.2017.07.005