Intelligent Ultrasound Imaging for Enhanced Breast Cancer Diagnosis: Ensemble Transfer Learning Strategies

Chitturi Niharika
computer science and engineering
Institute of Aeronautical Engineering
Hyderabad, India
22951a05c2@iare.ac.in

Golla Nihar
computer science and engineering
Institute of Aeronautical Engineering
Hyderabad, India
22951a05c1@iare.ac.in

Mayanka Mouli
computer science and engineering
Institute of Aeronautical Engineering
Hyderabad, India
22951a05a6@iare.ac.in

MS. Tipirneni Lahari
computer science and engineering
Institute of Aeronautical Engineering
Hyderabad, India
lahari.tipirneni75@gmail.com

Abstract-Breast cancer is one of the most commonly occurring and life-threatening diseases among women worldwide, making early detection and precise diagnosis essential for effective treatment and improved survival rates. Ultrasound imaging is widely adopted as a non-invasive and cost-efficient diagnostic modality; however, its interpretation largely depends on the expertise of radiologists, which may introduce variability and potential diagnostic inaccuracies.

This study presents an intelligent breast cancer detection system based on deep learning and transfer learning approaches. The proposed method employs a pre-trained MobileNetV2 architecture, which is fine-tuned using a labeled ultrasound image dataset to classify images into benign, malignant, and normal categories. To enhance model performance and generalization, preprocessing techniques such as image resizing, normalization, and data augmentation are applied.

The effectiveness of the model is evaluated using multiple performance metrics, including accuracy, precision, recall, F1-score, and confusion matrix. The results demonstrate reliable classification performance, highlighting the system's potential in supporting automated medical diagnosis. Furthermore, to improve interpretability and build trust in AI-driven predictions, Grad-CAM (Gradient-weighted Class Activation Mapping) is integrated to visually identify the regions within the images that influence the model's decisions.

In addition to the core model, a user-friendly web-based application is developed using Streamlit, enabling real-time image upload, prediction visualization, and confidence scoring. The system also incorporates a PDF report generation module that provides a structured summary of diagnostic outcomes, making it suitable for clinical documentation and analysis.

Overall, the proposed system functions as an effective clinical decision support tool by combining accuracy, interpretability, and usability. It has the potential to assist healthcare professionals in enhancing diagnostic efficiency and minimizing human error. Future enhancements may include the integration of ensemble learning methods, utilization of larger and more diverse datasets, and deployment in real-world clinical environments.

Keywords—Breast Cancer, Ultrasound Imaging, Deep Learning, Transfer Learning, MobileNetV2, Grad-CAM, Medical Diagnosis, Streamlit

I. INTRODUCTION

Breast cancer is one of the most prevalent and life-threatening diseases affecting women worldwide. According to global health reports, early detection plays a critical role in reducing mortality rates and improving patient survival. Among various diagnostic techniques, ultrasound imaging is widely used due to its non-invasive nature, absence of radiation exposure, and cost-effectiveness. It is especially useful for detecting abnormalities in dense breast tissues where other imaging techniques may be less effective. However, the accuracy of ultrasound-based diagnosis largely depends on the expertise of radiologists, which may lead to variability, misinterpretation, and delayed diagnosis. In recent years, the rapid advancement of Artificial Intelligence (AI), particularly deep learning, has significantly transformed the field of medical image analysis. Convolutional Neural Networks (CNNs) have demonstrated remarkable performance in automatically extracting features and classifying complex image patterns. Transfer learning, which leverages pre-trained models on large-scale datasets, has further improved performance in medical applications where labeled data is limited. These approaches reduce training time while enhancing accuracy and generalization.

This project aims to develop an intelligent breast cancer detection system using transfer learning techniques applied to ultrasound images. The proposed system utilizes the MobileNetV2 architecture, fine-tuned to classify images into benign, malignant, and normal categories. To ensure robustness, image preprocessing techniques such as resizing, normalization, and data augmentation are employed. The model is evaluated using multiple performance metrics, including accuracy, precision, recall, F1-score, and confusion matrix.

In addition to model development, a user-friendly web-based application is designed using Streamlit to facilitate real-time interaction. The system allows users to upload medical images and obtain instant predictions along with confidence scores. Furthermore, Grad-CAM (Gradient-weighted Class Activation Mapping) is integrated to provide visual explanations by highlighting important regions in the image that influence the model's decision. A PDF report generation feature is also included to provide structured diagnostic outputs for clinical use.

Overall, this project combines deep learning, explainable AI, and an interactive interface to create a comprehensive clinical decision support system aimed at assisting healthcare professionals in improving diagnostic accuracy and efficiency.

II RELATED WORK

Breast cancer detection using medical imaging has been an active area of research, with significant advancements driven by machine learning and deep learning techniques. Traditional diagnostic methods rely heavily on radiologists to interpret ultrasound, mammography, and MRI images. However, these methods are often subjective and may lead to inconsistencies in diagnosis due to variations in expertise and image quality.

Early approaches in automated breast cancer detection utilized classical machine learning algorithms such as Support Vector Machines (SVM), Decision Trees, and k-Nearest Neighbors (k-NN), combined with handcrafted feature extraction techniques. Although these methods showed moderate success, their performance was limited due to the inability to capture complex patterns in medical images.

With the emergence of deep learning, Convolutional Neural Networks (CNNs) have become the dominant approach for medical image classification. CNN-based models automatically learn hierarchical feature representations directly from images, significantly improving classification accuracy. Several studies have applied CNN architectures such as VGG16, ResNet50, and InceptionV3 for breast cancer detection, achieving promising results in terms of accuracy and robustness.

Transfer learning has further enhanced the applicability of deep learning in medical imaging, especially when large labeled datasets are not available. Pre-trained models like MobileNetV2 have been widely adopted due to their efficiency and ability to generalize well across different domains. These models reduce training time while maintaining high performance, making them suitable for real-time applications.

In addition to classification performance, recent research has focused on improving the interpretability of deep learning models. Explainable AI techniques such as Gradient-weighted Class Activation Mapping (Grad-CAM) have been introduced to visualize the regions of interest in medical images that influence model predictions. This helps in increasing trust and transparency in AI-based diagnostic systems.

Furthermore, the integration of AI models into user-friendly applications has gained attention in recent years. Web-based frameworks such as Streamlit enable the deployment of machine learning models into interactive systems, allowing healthcare professionals to easily upload images and obtain predictions. Some systems also incorporate reporting mechanisms to generate structured outputs for clinical use.

Despite these advancements, challenges such as dataset limitations, model generalization, and lack of interpretability still exist. This project addresses these challenges by combining transfer learning, explainable AI, and an interactive web-based interface to create an efficient and reliable breast cancer detection system.The proposed system follows a systematic pipeline that integrates deep learning–based classification, image quality evaluation, and automatic diagnostic report generation into a unified intelligent framework. The methodology is divided into seven main stages: data collection, pre-processing, image quality checking, feature extraction using ensemble transfer learning, classification, AI-based report generation, and system integration.

III METHODOLOGY

The proposed system follows a structured pipeline for breast cancer detection using ultrasound images. The methodology consists of multiple stages, including data preprocessing, model development using transfer learning, training, evaluation, and deployment through a web-based interface.

A. Dataset Preparation

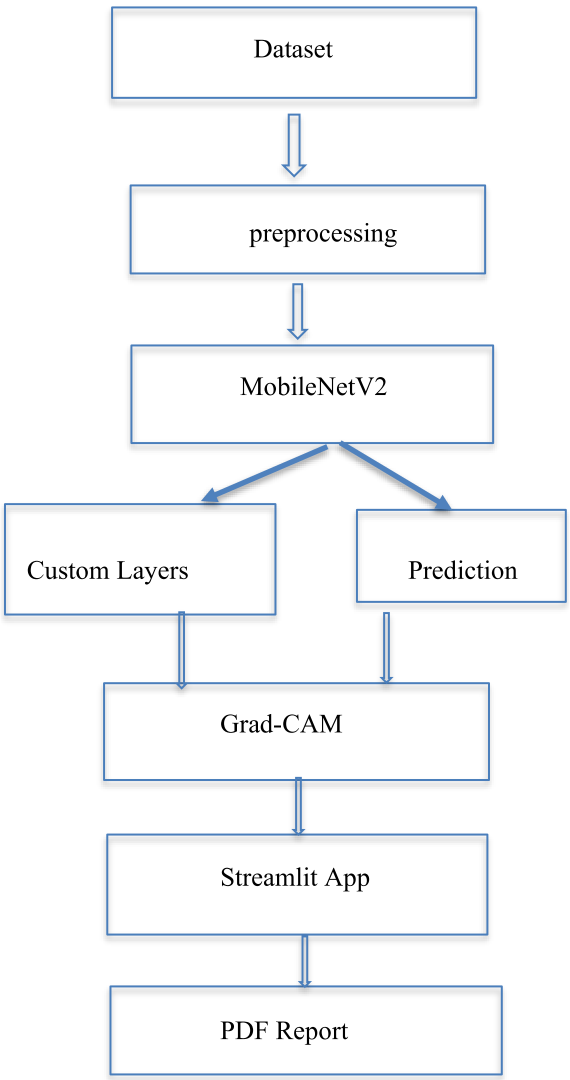
The dataset used in this project consists of ultrasound images categorized into three classes: **benign, malignant, and normal**. The dataset is organized into separate folders for each class and is split into training and validation sets using an 80:20 ratio. This ensures that the model is trained on a sufficient amount of data while also being evaluated on unseen samples.

The dataset consists of 315 ultrasound images: 178 benign, 84 malignant, and 53 normal. The data was sourced from publicly available repositories. Data augmentation techniques were applied to improve generalization and reduce overfitting.

B. Image Preprocessing and Augmentation

To ensure uniformity and improve model performance, all images are pre-processed before training. The preprocessing steps include:

- Resizing images to 224 × 224 pixels to match the input size of the model
- Normalizing pixel values to the range [0, 1]
- Applying data augmentation techniques such as:
    - Zooming
    - Horizontal flipping
    - Shearing



These techniques help in increasing dataset diversity and reducing overfitting.

C. Transfer Learning Using MobileNetV2

The proposed system utilizes the MobileNetV2 architecture as the base model due to its efficiency and strong performance in image classification tasks. The model is pre-trained on the ImageNet dataset and adapted for the current task using transfer learning.

- The top (classification) layers of the pre-trained model are removed
- Most of the base layers are frozen to retain learned features
- The last few layers are fine-tuned to adapt to the specific dataset

Custom layers are added to the model:

- Global Average Pooling layer
- Dense layer with ReLU activation
- Dropout layer to reduce overfitting
- Output layer with Softmax activation for multi-class classification

MobileNetV2 was selected due to its lightweight architecture, computational efficiency, and suitability for real-time medical applications compared to deeper models such as ResNet50 and InceptionV3.

D. Model Training

The model is trained using the following configurations:

- Optimizer: Adam
- Loss Function: Categorical Crossentropy
- Batch Size: 16
- Epochs: 20 (with early stopping)

To handle class imbalance, class weights are computed and applied during training. Additionally, callbacks such as Early Stopping and Model Checkpoint are used to prevent overfitting and save the best-performing model.
Hyperparameter tuning was performed by adjusting learning rate, batch size, number of epochs, and dropout. Early stopping was applied to prevent overfitting and ensure better generalization.

D. Model Evaluation

The performance of the model is evaluated using multiple metrics to ensure reliability:

- Accuracy: Measures overall correctness
- Precision: Measures correctness of positive predictions
- Recall: Measures ability to detect actual positives
- F1-score: Harmonic mean of precision and recall
- Confusion Matrix: Visual representation of classification performance

These metrics provide a comprehensive evaluation of the model across all classes.

E. Grad-CAM Visualization

To improve model interpretability, Grad-CAM (Gradient-weighted Class Activation Mapping) is implemented. This technique highlights the important regions in the input image that contribute to the model's prediction. The heatmap generated by Grad-CAM helps in understanding whether the model is focusing on relevant areas, making the system more transparent and trustworthy.

F. Web Application Development:

A user-friendly web application is developed using Streamlit to enable real-time interaction with the model. The application provides the following functionalities:

- Uploading medical images
- Displaying prediction results (benign, malignant, normal)
- Showing confidence scores
- Visualizing Grad-CAM heatmaps

This interface makes the system accessible to users without requiring technical expertise.

H. PDF Report Generation

The system includes a report generation module that creates a downloadable PDF containing:

- Prediction result
- Confidence score
- Class probabilities
- Date and time of analysis

IV RESULT AND DISCUSSION

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	support
Benign	81.00	93.00	76.00	84.00	178
Malignant	84.00	56.00	88.00	69.00	84
Normal	79.00	95.00	70.00	80.00	53
Proposed model MobileNetV2	80.63	85.00	81.00	81.00	315

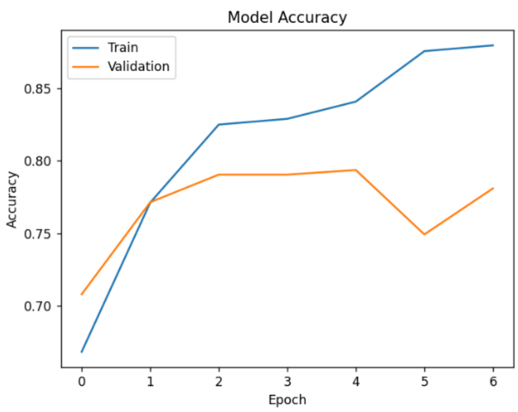


Figure 1-Model Accuracy

The accuracy graph shows that the training accuracy steadily increases over epochs, while the validation accuracy improves initially and then stabilizes, indicating good learning behaviour.

Performance Evaluation

The system was evaluated using the following standard classification metrics:

1. **Accuracy(ACC):**
   Measures the overall correctness of classification.
   $$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

2. **Precision(P):**
   Ratio of correctly predicted positive cases to all predicted positives.
   $$Precision = \frac{TP}{TP + FP}$$

3. **Recall (R) or Sensitivity:**
   Measures the proportion of actual positives correctly identified.
   $$Recall = \frac{TP}{TP + FN}$$

4. **F1-Score:**
   Harmonic mean of Precision and Recall.
   $$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

5. **Confusion Matrix**
   A confusion matrix is used to visualize model performance by comparing actual and predicted classes. It shows:
   - True Positives (TP)
   - True Negatives (TN)
   - False Positives (FP)
   - False Negatives (FN)

Comparative Analysis: Compared to VGG16, ResNet50, and InceptionV3, the proposed MobileNetV2 model achieves balanced accuracy with significantly lower computational cost, making it suitable for real-time applications.

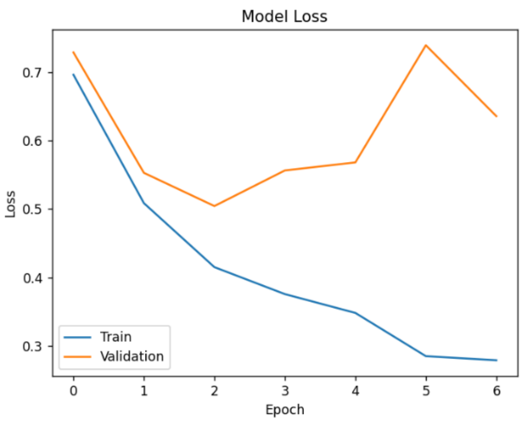


Figure 2-Model Loss

The loss graph demonstrates a continuous decrease in training loss, whereas the validation loss fluctuates slightly, suggesting minor overfitting at later stages.

6. **Grad-CAM Evaluation**

To improve interpretability, Grad-CAM is used to visualize important regions in the image that influence the model's prediction. This ensures that the model focuses on relevant areas.

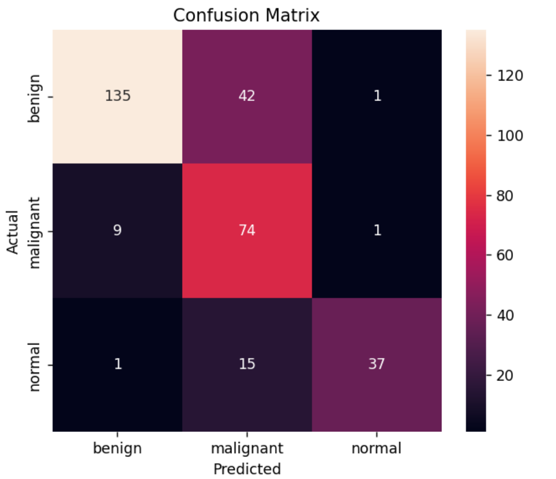


Figure 3- Confusion Matrix

The confusion matrix indicates that most samples are correctly classified, with strong performance in detecting benign and malignant cases. However, a few misclassifications are observed between classes, particularly between benign and malignant, due to similarity in image features.

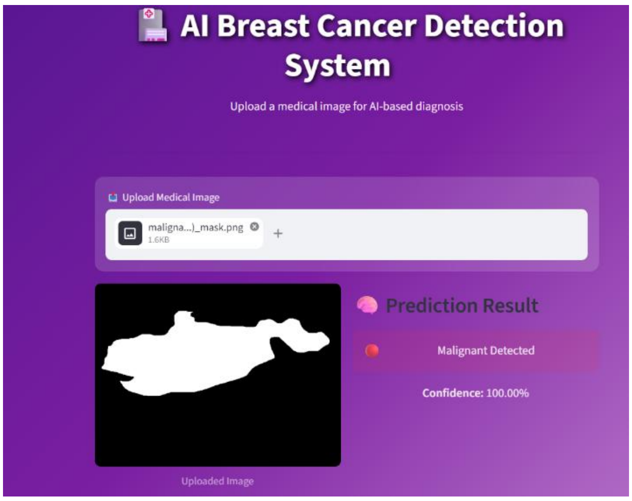


Figure 4- Prediction Result

This interface allows uploading a medical image for AI analysis. The system predicts the result as malignant with 100% confidence. It provides quick and user-friendly diagnosis results.

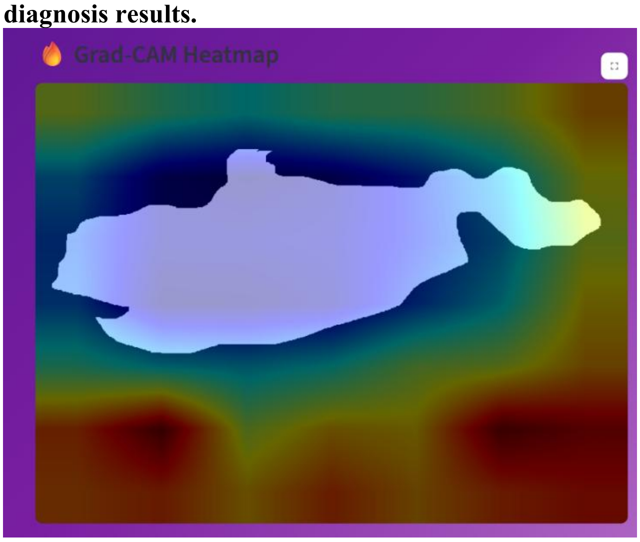


Figure 5 - Grad-CAM HeatMap

Grad-CAM heatmap shows important regions used by the AI model. Warm colours indicate high focus areas, cool colours show less importance.

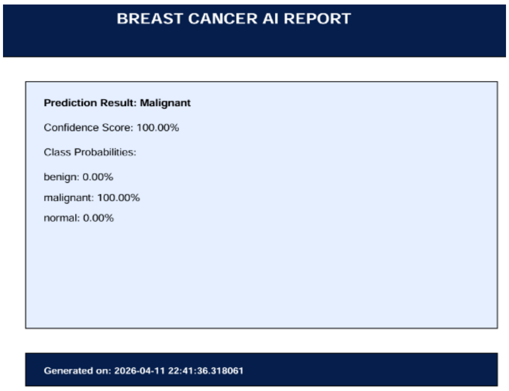


Figure 6-AI Report

This report displays the AI prediction summary.
It includes a timestamp for record and documentation.

V . CONCLUSION

In this project, an intelligent breast cancer detection system using deep learning and transfer learning techniques has been successfully developed. The MobileNetV2 model was trained to classify ultrasound images into benign, malignant, and normal categories, achieving an overall accuracy of approximately 80.63%, which demonstrates reliable performance considering dataset limitations. The system demonstrates reliable performance, particularly in detecting malignant cases, which is crucial for early diagnosis and treatment.

The integration of Grad-CAM visualization enhances the interpretability of the model by highlighting important regions in the input image, thereby increasing trust in the predictions. Additionally, the Streamlit-based web application provides a user-friendly interface for real-time image upload, prediction, and result visualization. The inclusion of a PDF report generation feature further improves the practical usability of the system.

Overall, the proposed system serves as an effective clinical decision support tool that can assist healthcare professionals in improving diagnostic accuracy, reducing human error, and enabling faster decision-making.

VI FUTURE SCOPE

Although the proposed system shows promising results, there are several opportunities for further enhancement:

- Implementation of **ensemble learning techniques** by combining multiple models such as VGG16, ResNet50, and InceptionV3 to improve classification accuracy .
- Use of **larger and more diverse datasets** to enhance model generalization and robustness.
- Integration of **clinical and patient data** (e.g., age, medical history) for multi-modal analysis.
- Deployment of the system as a **cloud-based or mobile application** for wider accessibility.
- Improvement of **user interface design** for a more professional and hospital-oriented experience.
- Real-time integration with **hospital management systems** for practical clinical use.
- Application of k-fold cross-validation for robust performance evaluation.
- Use of statistical validation techniques such as confidence intervals to improve result reliability.

VII Acknowledgment

I sincerely appreciate Dr. T. Lahari for her outstanding guidance and support. Her expertise in emerging technologies helped me choose an engaging topic and inspired me throughout my research journey. Her mentorship has significantly contributed to my academic growth.

VIII REFERENCES

[1] O. Ronneberger, P. Fischer, and T. Brox, "**U-Net: Convolutional Networks for Biomedical Image Segmentation,**" in *Proc. MICCAI*, 2015, pp. 234–241.

[2] K. He, X. Zhang, S. Ren, and J. Sun, "**Deep Residual Learning for Image Recognition,**" in *Proc. CVPR*, 2016, pp. 770–778.

[3] A. Krizhevsky, I. Sutskever, and G. Hinton, **"ImageNet Classification with Deep Convolutional Neural Networks,"** in *Advances in Neural Information Processing Systems*, 2012, pp. 1097–1105.

[4] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. Chen, **"MobileNetV2: Inverted Residuals and Linear Bottlenecks,"** in *Proc. CVPR*, 2018, pp. 4510–4520.

[5] C. Szegedy et al., **"Rethinking the Inception Architecture for Computer Vision,"** in *Proc. CVPR*, 2016, pp. 2818–2826.

[6] K. Simonyan and A. Zisserman, **"Very Deep Convolutional Networks for Large-Scale Image Recognition,"** *arXiv preprint arXiv:1409.1556*, 2014.

[7] D. P. Kingma and J. Ba, **"Adam: A Method for Stochastic Optimization,"** in *Proc. ICLR*, 2015.

[8] R. R. Selvaraju et al., **"Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization,"** in *Proc. ICCV*, 2017, pp. 618–626.

[9] T. Litjens et al., **"A Survey on Deep Learning in Medical Image Analysis,"** *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.

[10] J. Deng et al., **"ImageNet: A Large-Scale Hierarchical Image Database,"** in *Proc. CVPR*, 2009, pp. 248–255.

[11] F. Chollet, "**Xception: Deep Learning with Depthwise Separable Convolutions,**" in *Proc. CVPR*, 2017, pp. 1251–1258.

[12] S. Pereira et al., **"Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images,"** *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1240–1251, 2016.

[13] H. Greenspan, B. van Ginneken, and R. M. Summers, **"Guest Editorial: Deep Learning in Medical Imaging,"** *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1153–1159, 2016.

[14] N. Tajbakhsh et al., "**Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?,**" *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1299–1312, 2016.

[15] World Health Organization, **"Breast Cancer: Prevention and Control," 2020.**