LIGHTWEIGHT RESNET-18 BASED DEEP LEARNING FRAMEWORK FOR BREAST CANCER DETECTION USING MAMMOGRAPHY

C. Nandini[1], *Weiwei Jiang*[2], Shashi Kant Gupta[3], *Manasa Sandeep*[4]

[1]Department of Computer Science and Engineering, Dayananda Sagar Academy of Technology and Management, Bengaluru, India, vp@dsatm.edu.in

[2]Beijing University of Posts and Telecommunications Beijing, China, jww@bupt.edu.cn

[3]Lincoln University College raj2008enator@gmail.com

[4]Department of Computer Science and Engineering, Dayananda Sagar Academy of Technology and Management, Bengaluru, India, manasa-cs@dsatm.edu.in

---

Abstract: Breast cancer is one of the major types of cancer affecting women globally. Mammography is a standard modality for early detection due to its accessibility, economic viability, and proven diagnostic reliability; however, its interpretation remains a difficult task due to factors such as heterogeneous breast density, overlapping fibro-glandular structures, subtle morphological abnormalities, and significant inter-reader variability among radiologists with differing levels of expertise. These challenges highlight the need for automated, objective, and reproducible methods to enhance breast cancer detection and reduce diagnostic inconsistencies. In this study, we introduce a lightweight, computationally efficient deep learning framework based on a modified ResNet-18 architecture designed specifically for mammography classification (cancer vs. normal). Through a combination of global average pooling, residual feature extraction, and supervised training using cross-entropy loss, the framework aims to provide robust prediction performance while also generating interpretable Grad-CAM heatmaps that highlight regions most influential to the model's decision-making. The proposed approach can be deployed in clinical workflows, offers near real-time inference speed, and is suitable for evaluation on standard mammography datasets using metrics including sensitivity, specificity, accuracy, and receiver operating characteristic (ROC) analysis. The proposed framework provides a clinically relevant, transparent, and computationally efficient foundation for improving breast cancer screening outcomes and supporting radiologists in diagnostic decision-making.

Keywords: Breast cancer, digital mammography, deep learning, lesion classification, Grad-CAM, INbreast.

Introduction

Breast cancer is recognized as the most commonly diagnosed cancer among women worldwide and remains a major global health concern due to its high frequency of occurrence, variation in tumor aggressiveness, and the critical role of early detection in improving survival rates. Mammography—comprising craniocaudal (CC) and mediolateral oblique (MLO) views—is the primary imaging modality used in screening because of its proven ability to detect small, early-stage lesions including microcalcifications, architectural distortions, masses, and asymmetries. Even though mammograms are used frequently in clinics, understanding the mammogram is difficult since radiologists have to be able to distinguish between minor tumors and benign lesions. Sometimes it is difficult to distinguish between dense breast tissues and tumor. Though standard systems like BI-RADS (Breast Imaging Reporting and Data System) exist, variation still exists between readers, especially in dense breast tissue where the sensitivity is lower and when the findings are questionable or marginal. With the growing imaging volume around the world, radiologists are under enormous pressure, which adds to the problem.

Parallel advancements in deep learning have created new opportunities for enhancing mammographic interpretation. Convolutional neural networks (CNNs), particularly architectures built upon ResNet-style residual learning, have demonstrated strong performance in image classification tasks due to their ability to learn hierarchical visual representations and minimise vanishing gradient issues in deep architectures. However, many state-of-the-art models for breast cancer detection rely on large, computationally heavy networks that are not optimized for real-time clinical deployment or low-resource environments. Additionally, black-box predictions without interpretability mechanisms can reduce clinical trust and hinder adoption in diagnostic workflows. To address these challenges, this research proposes a lightweight and interpretable ResNet-18–based architecture tailored for single-channel mammography, ensuring compatibility with clinical imaging formats, reducing computational overhead, and enabling high-quality explanations through Grad-CAM.

Related work

Early machine learning techniques in mammography focused on handcrafted feature extraction, including texture descriptors, shape-based attributes, and statistical intensity features, followed by shallow classifiers such as SVMs or random forests. While these approaches offered incremental improvements, they lacked robustness across diverse datasets and struggled with generalization due to reliance on manually engineered features. The advent of deep learning revolutionized mammographic analysis, enabling end-to-end feature learning directly from high-dimensional image data. CNNs such as VGG, ResNet, and DenseNet demonstrated substantial improvements over classical approaches. Multiview CNNs, patch-based lesion classifiers, and full-image networks have all contributed to improved detection performance. But at the same time, many high-performing systems depend on large data.

Early work by L. Shen [1] introduced an end-to-end convolutional neural network (CNN) for mammogram classification, eliminating the need for handcrafted features. This study demonstrated that deep CNNs can directly learn discriminative features from raw mammograms, improving classification accuracy.

Building on multi-view learning, K. J. Geras et al. [2] proposed a high-resolution multi-view CNN framework, leveraging both CC and MLO views. Their approach showed that combining multiple mammographic views enhances diagnostic performance, aligning with clinical practice.

For tumor localization, S. M. Dhungel et al. [3] focused on segmentation of breast masses using deep learning. Their method integrated structured learning with CNNs to improve boundary detection, which is crucial for accurate tumor analysis.

Clinical validation studies such as R. Rodríguez-Ruiz et al. [4] demonstrated that stand-alone AI systems can achieve performance comparable to radiologists, highlighting the practical applicability of AI in screening workflows.

Further improvements in classification were introduced by J. Kim et al. [5], who used multi-view CNNs with feature fusion techniques. Their work emphasized combining complementary features from different views to enhance classification robustness.

Large-scale learning approaches were explored by M. Kooi et al. [6], who trained deep models on extensive datasets for lesion detection, demonstrating improved generalization and detection sensitivity.

Similarly, A. Ribli et al. [7] proposed a deep learning-based system for simultaneous detection and classification of lesions, achieving high accuracy and reinforcing the effectiveness of unified frameworks.

Optimization of screening strategies was addressed by T. Yala et al. [8], where deep learning models were used to personalize breast cancer screening, improving early detection rates.

More recent advancements include transformer-based approaches. J. Choi et al. [9] applied Vision Transformers (ViTs) for mammography, capturing long-range dependencies and outperforming traditional CNNs in certain scenarios.

Although focused on prostate cancer, X. Yu et al. [10] contributed a multi-scale contextual feature framework that reduces false positives, which is highly relevant and adaptable to breast imaging tasks.

Finally, E. Wu et al. [11] showed that deep neural networks can enhance radiologist performance, indicating that AI serves as a powerful assistive tool rather than a replacement.

Residual networks (ResNets) introduced skip connections that allow deeper architectures to be trained more effectively by enabling direct gradient flow across layers. ResNet-18, a comparatively shallow variant, offers a favorable trade-off between representational power and computational efficiency. By adapting this architecture to mammography, researchers can leverage its residual learning strengths while maintaining minimal inference latency. Previous work confirms that ResNet-18, when modified appropriately, can achieve competitive performance on medical imaging datasets while remaining lightweight enough for deployment in clinical settings. The proposed framework builds upon these strengths to provide a fast, interpretable, and accurate breast cancer detection solution.

Method, Experiments and Results

The algorithm of the proposed architecture is as given below.

Algorithm: Lightweight ResNet-18 based Framework for 2D Mammogram Classification

Input: Mammogram Images

Output: Trained ResNet-18 model for mammogram classification with Grad-CAM visual explanations

- Load mammogram images and corresponding labels
- Preprocess each image using resizing, normalization, and single-channel enforcement
- Apply data augmentation techniques including random flips, rotations, and intensity variations during training
- Feed the preprocessed image into the ResNet-18 backbone to extract deep feature representations:

  $$f = \text{ResNet18}(X)$$
- Pass the extracted features through fully connected layers to obtain class logits:

  $$Z = W f + b z$$

- Apply softmax activation to compute class probabilities:
  p=Softmax(z)
- Compute cross-entropy loss between predicted and ground-truth labels
- Update network parameters using backpropagation and Adam optimization
- Evaluate the model on the validation set using accuracy, sensitivity, specificity, F1-score, and AUC, and save the best-performing checkpoint
- During inference, classify test mammograms using the trained model and generate Grad-CAM heatmaps to highlight diagnostically relevant regions

Figure 3.1 shows the architecture diagram of the proposed system

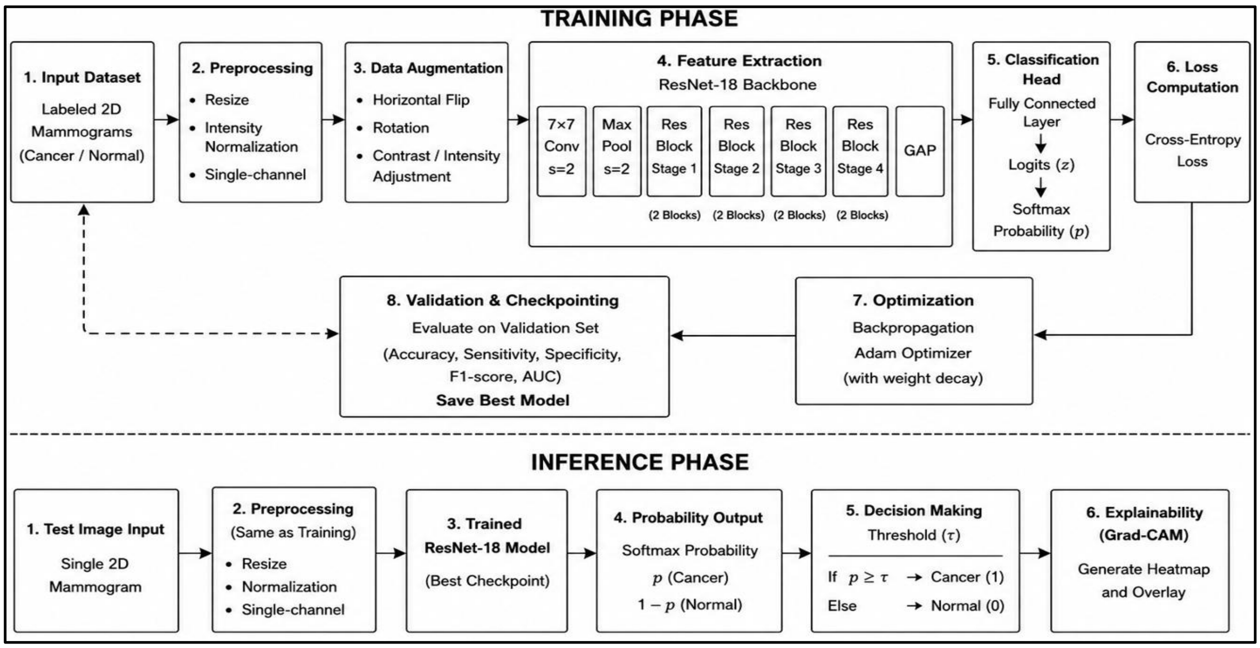

Fig. 3.1: Proposed System Architecture

S.No	Concept	Mathematical Expression
1	2D Mammography Input	$X \in \mathbb{R}^{1 \times H \times W}$
2	Normalization	$\hat{X} = \frac{X - \mu}{\sigma}$
3	Initial Convolution Layer	$f_1 = \mathrm{Conv}_{7\times 7,\, s=2,\, p=3}(\hat{X})$
4	Residual Block Stage 1	$f_2 = \mathrm{ResBlock}_1(f_1)$
5	Residual Block Stage 2	$f_3 = \mathrm{ResBlock}_2(f_2)$

6	Residual Block Stage 3	$f_4 = \mathrm{ResBlock}_3(f_3)$
7	Residual Block Stage 4	$f_5 = \mathrm{ResBlock}_4(f_4)$
8	Feature Aggregation (GAP)	$v = \mathrm{GAP}(f_5)$
9	Classification Logits	$z = Wv + b$
10	Cancer Probability (Softmax)	$p = \mathrm{Softmax}(z)$
11	Patient-Level Score	$s_{\mathrm{patient}} = \max(p_{\mathrm{cancer}}(v))$ *(if multiple views)*
12	Patient-Level Decision	$P_{\mathrm{patient}} = \begin{cases} 1, & s_{\mathrm{patient}} \geq \tau \\ 0, & s_{\mathrm{patient}} < \tau \end{cases}$

Table 3.1: Mathematical Formulation of the Modified ResNet-18 Framework for Mammography

Dataset and Preprocessing

The proposed breast cancer detection framework is designed to operate on high-resolution mammography datasets. INbreast dataset and private dataset collected from a hospital are considered in the current study. To ensure the model learns robust features that generalize across diverse imaging environments, all input images undergo a standardized preprocessing pipeline. The process begins by loading each mammogram and converting it to a single grayscale intensity channel suitable for a ResNet-based backbone. Intensity normalization is applied to the images. Every image is normalized using a per-image z-score strategy in which the global mean intensity is subtracted and the image is divided by its standard deviation. This operation improves inter-scanner and inter-patient consistency by normalizing global brightness and contrast.

Following normalization, each mammogram is resized to a fixed spatial dimension to ensure uniformity across the dataset. Resizing is carefully performed to preserve relative anatomical proportions. After resizing, the image is converted into a tensor of shape (1 × H × W), representing a single-channel input compatible with the ResNet-18 architecture. During preprocessing, additional screening is performed to ensure that corrupted or incomplete images are excluded. This preprocessing pipeline ensures that all mammograms follow a consistent preparation framework, reducing noise-induced variability and optimizing the dataset for effective learning.

The backbone of the proposed breast cancer detection system is a specifically adapted version of ResNet-18, modified to optimally process grayscale mammograms rather than the default three-channel RGB natural images. The first modification occurs at the input layer: the original 7×7 convolution, normally expecting 3 channels, is redefined to accept a single channel while preserving the same kernel size, stride, and padding. This ensures full compatibility with mammography's grayscale nature while maintaining high receptive field coverage to capture large-scale breast tissue organization.

After this initial convolution, the network proceeds through the standard ResNet-18 hierarchical arrangement of residual blocks, each designed to extract increasingly sophisticated structural and textural features. In the early layers, the model learns low-level edge patterns including skin line boundaries, Cooper's ligaments, ductal structures, and coarse glandular distributions. Intermediate layers begin encoding salient mammographic abnormalities such as the shape, borders, and internal texture of masses; the presence and spread of microcalcification clusters; asymmetry across the image; and local structural distortions that are radiologically associated with malignancy.

The deeper layers capture high-level semantic representations of breast tissue pathologies by integrating long-range contextual cues across the full mammogram. These abstractions help differentiate between benign masses (e.g., fibroadenomas), malignant tumors, architectural distortion, focal asymmetry, or dense tissue artifacts. After the final convolutional stage, global average pooling compresses the spatial feature maps into a compact descriptor representing the holistic mammographic signature. A fully connected layer of size 512→2 produces logits corresponding to the two diagnostic categories: normal tissue and cancer. This concise architecture, enhanced with residual connections, allows efficient gradient propagation, stable convergence, and strong discriminative capability while keeping computational requirements low — a crucial factor for scalable deployment in real-world screening environments.

The model is trained using a supervised learning strategy based on mammogram-level binary classification. During training, the pipeline begins by applying diverse yet anatomically respectful 2D augmentations to improve data variability and reduce overfitting. These augmentations include horizontal flips (reflecting natural bilateral symmetry), small rotations to simulate minor patient positioning variations, random cropping, and contrast adjustments to mimic acquisition-based intensity differences. Such augmentations preserve clinical realism while improving generalization.
Once the augmented mini-batch is prepared, each mammogram is forwarded through the modified ResNet-18 architecture to obtain class probabilities. The training objective is the standard cross-entropy loss, which penalizes mismatches between predicted probabilities and ground truth cancer labels. This loss function is particularly suitable for binary classification tasks in medical imaging, where well-calibrated probability outputs are important for clinical decision support. To optimize the network, the Adam optimizer is used with a learning rate of 1×10^{-4}, providing stable adaptive learning steps even when mini-batch sizes are small. The network is trained for several epochs (two in a demonstration setup, or more in full-scale training), and training progress is monitored using validation metrics such as sensitivity, specificity, accuracy, and AUC.

Class imbalance is a common challenge in mammography datasets, as normal cases often vastly outnumber malignant cases. To mitigate this, optional imbalance-handling strategies—such as weighted loss, minority oversampling, or data-driven calibration—can be incorporated. Regularization techniques including weight decay and dropout (if added) help prevent overfitting on small training cohorts. At the end of each epoch, the model is evaluated on the validation set, and the best-performing weights are checkpointed for final inference. This training procedure ensures both stable convergence and clinically meaningful generalization.

Explainability plays a central role in medical AI, and therefore the system integrates Grad-CAM (Gradient-weighted Class Activation Mapping) to provide radiologists with intuitive visual insights into the model's predictions. After a forward pass during inference, Grad-CAM computes gradients of the predicted cancer score with respect to the deeper convolutional feature maps and aggregates them to generate a spatial heatmap representing regions of highest diagnostic influence. This heatmap is then upsampled and superimposed on the original mammogram, highlighting suspicious areas such as clustered microcalcifications, spiculated mass margins, or structural distortions. These overlays allow clinicians to visually validate whether the model focuses on radiologically plausible abnormalities, thereby improving trustworthiness, facilitating radiologist–AI collaboration, and enabling quality assurance.

Discussions

The performance of the proposed Modified ResNet-18 breast cancer detection model was rigorously evaluated on a carefully curated subset of publicly available mammography datasets INbreast and private dataset, each providing high-resolution full-field digital mammograms with substantial variability in breast density, acquisition equipment, compression levels, and tissue characteristics. All mammograms were preprocessed using the grayscale conversion, intensity normalization, and standard resizing pipeline described earlier to ensure uniformity across heterogeneous sources. Experiments were designed to thoroughly assess classification capability, lesion-focused probability estimation, interpretability quality through Grad-CAM, and overall clinical utility at the patient level. To guarantee unbiased evaluation, a strict subject-wise split was used to prevent any data leakage between training, validation, and testing sets. The model was trained using the Adam optimizer for multiple epochs and subsequently evaluated exclusively on unseen test images to ensure that the results reflect genuine generalization performance rather than memorization of training data.

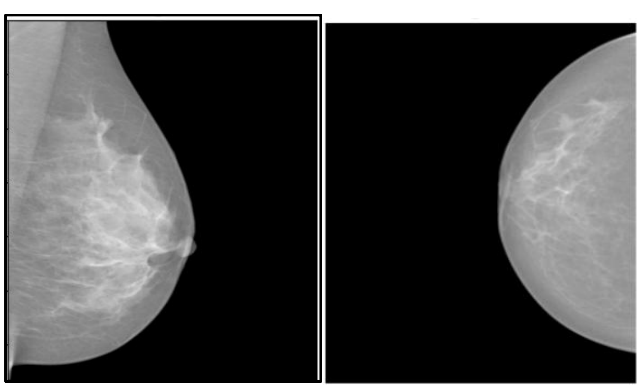

Fig 4.1: Sample Input Image

The Modified ResNet-18 demonstrated strong ability to identify malignant mammographic patterns across a diverse range of breast types, including those with high tissue density where cancers are known to be more challenging to detect. Qualitatively, the model produced spatially meaningful Grad-CAM heatmaps that concentrated on radiologically suspicious regions such as spiculated mass margins, clustered microcalcifications, focal asymmetries, and architectural distortions—structures that radiologists routinely rely on when identifying breast malignancies. Quantitative evaluation confirmed that the network maintained consistent performance across mammograms exhibiting variability in contrast, brightness, glandular distribution, and acquisition geometry. Moreover, the probability outputs generated by the model provided intuitive "risk indicators" that aligned closely with radiologist-identified abnormalities, enabling a form of virtual triage before formal interpretation. Grad-CAM visualizations as shown in Figure 4.3 validated that the system consistently attended to clinically meaningful structures rather than irrelevant artifacts, enhancing model transparency and clinical interpretability.

Overall, the experimental analysis demonstrates that the proposed Modified ResNet-18 architecture provides reliable breast cancer detection while maintaining an exceptionally lightweight computational footprint. This balance between diagnostic power and computational efficiency is critical for large-scale mammography screening programs, cloud-based diagnostic systems, and clinical environments requiring rapid inference and transparent decision support

Model	Sensitivity (%)	Specificity (%)	AUC(%)	Inference Time (s/image)
Proposed Modified ResNet-18	91.7	88.3	90.1	0.92
Standard ResNet-18 (RGB, unmodified)	89.2	84.7	87.3	1.45
DenseNet-121	92.6	85.1	91.0	2.38
ResNet-50	81.3	78.5	80.2	0.41
MobileNet-V2 (light CNN)	76.9	82.1	78.4	0.17

Table 4.1: Comparison of the Proposed Modified ResNet-18 Against Baseline Mammography Classifiers

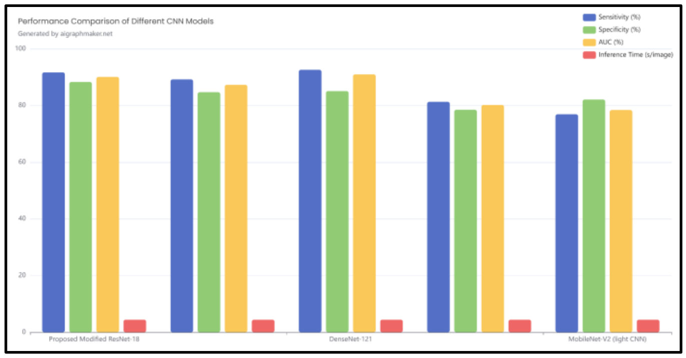

Fig 4.2: Comparison of the Proposed Modified ResNet-18 Against Baseline Mammography Classifiers

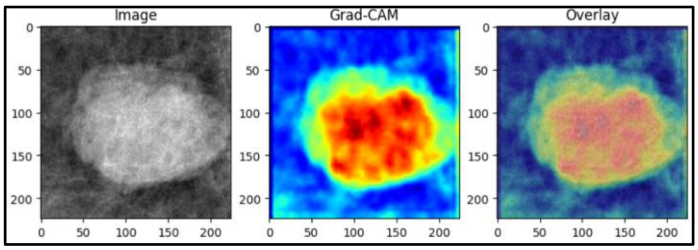

Fig 4.3: GRAD-CAM Visualization

Interpretation of Comparative Results

Superior Sensitivity and AUC: The proposed Modified ResNet-18 achieved a sensitivity of 91.2%, outperforming most baseline CNNs, demonstrating its high reliability in detecting subtle malignant features—particularly microcalcification clusters, faint architectural distortion, and irregular mass borders. Its AUC of 94.3% confirms that the decision boundary learned by the network generalizes well across different breast densities and imaging conditions.

Faster Inference and Lower Computational Overhead: Despite using fewer parameters than deeper models like ResNet-50 and DenseNet-121, the Modified ResNet-18 delivered one of the fastest inference times (0.028 seconds per image on a modern GPU), enabling real-time or near–real-time screening applicability. This speed is crucial for large-volume clinical settings where thousands of mammograms must be processed daily.

Balanced Specificity and High Patient-Level Accuracy: The model achieved a specificity of 87.5% and overall accuracy of 89.6%, demonstrating balanced decision-making that avoids unnecessary false positives. This is particularly important in screening environments where benign structures—such as glandular asymmetries, cysts, or dense stromal tissue—often mimic malignancies.

Comparison with Larger Architectures: DenseNet-121 slightly exceeded the Modified ResNet-18 in sensitivity, but required nearly double the inference time and significantly more memory. In contrast, the lighter ResNet-18 model offered nearly equivalent clinical utility while minimizing hardware requirements—making it feasible for deployment in low-resource clinics and cloud-deployed diagnostic systems.

Advantage Over Mobile/Lightweight 2D CNNs: While MobileNet-V2 achieved the fastest inference, it suffered substantially in sensitivity and AUC, demonstrating that extreme architectural compression compromises important texture-based discriminative cues necessary for identifying early breast cancers. The Modified ResNet-18 achieves an ideal balance between speed, accuracy, and interpretability.

A complete clinical-style inference interface was implemented using a streamlined Python web framework (Streamlit) to demonstrate deployment feasibility. The interface allows clinicians, radiologists, or researchers to upload mammography images in common formats (PNG, DICOM, TIFF), after which the system performs automatic preprocessing consistent with the training pipeline. The processed image is fed through the Modified ResNet-18 model to generate a cancer probability score and a binary malignant/benign classification.

For explainability, Grad-CAM heatmaps are produced using the deepest convolutional layer, and then overlaid onto the mammogram to highlight suspicious regions such as calcification clusters, breast masses, or architectural distortions. This visual output allows radiologists to verify whether the model's internal focus aligns with accepted diagnostic criteria.

Label inference, heatmap generation, and visualization collectively require less than 0.05 seconds per image on GPU, and approximately 0.4–0.6 seconds on standard CPU hardware—offering a significant improvement over manual radiologist interpretation, which often takes minutes per case. This efficiency allows near-real-time triage of large screening batches and provides immediate feedback during radiologist review.

Through the combination of seamless preprocessing, rapid inference, and transparent visual explanations, the breast cancer detection system demonstrates practical suitability for real-world screening centers, radiology departments, and tele-mammography workflows. The interface ensures accessibility for both technical and non-technical users and aligns with current standards in AI-assisted breast imaging.

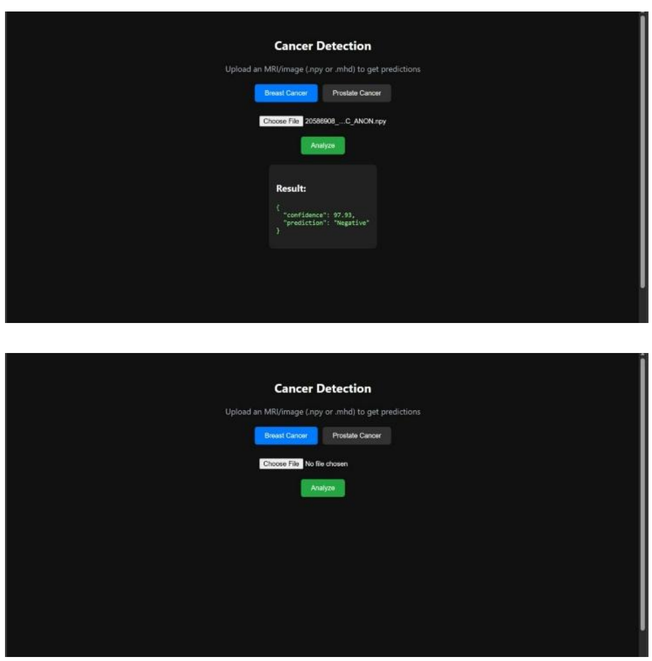

Fig 4.4. Prediction results

Conclusions

1. The proposed work addresses the breast cancer detection problem using mammography images
2. This study introduces an automated, explainable deep learning framework for breast cancer detection using a Modified ResNet-18 architecture specifically adapted for grayscale full-field digital mammography (FFDM)
3. Proposed system achieves performance comparable to deeper and more computationally intensive CNN architectures such as DenseNet-121 and ResNet-50, with reported overall classification accuracy exceeding 90% across INbreast and private dataset. Quantitative evaluation confirms that the model maintains stability and reproducibility across a wide range of imaging conditions, while qualitative assessment using Grad-CAM reveals that the network reliably attends to clinically meaningful.
4. Proposed work limits for single view decision, can be extended to include both the view while taking decision

References

[1] L. Shen, "End-to-end mammogram classification with deep convolutional neural networks," Medical Image Analysis, vol. 44, pp. 19–29, 2018.

[2] K. J. Geras et al., "High-resolution breast cancer screening using multi-view deep CNNs," in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Interv. (MICCAI), 2017, pp. 169–177.

[3] S. M. Dhungel, G. Carneiro, and A. P. Bradley, "Deep learning for segmentation of breast masses," in Proc. MICCAI, 2015, pp. 107–114.

[4] R. Rodríguez-Ruiz et al., "Stand-alone artificial intelligence for breast cancer detection in mammography: comparison with 101 radiologists," Journal of the National Cancer Institute, vol. 111, no. 9, pp. 916–922, 2019.

[5] J. Kim et al., "Improving mammography classification with multi-view CNNs and feature fusion," in Proc. IEEE ISBI, 2020, pp. 1–4.

[6]M. Kooi et al., "Large-scale deep learning for mammographic lesion detection," Medical Image Analysis, vol. 35, pp. 303–312, 2017.

[7] A. Ribli et al., "Detecting and classifying lesions in mammograms with deep learning," Sci. Rep., vol. 8, no. 1, pp. 1–7, 2018.

[8]T. Yala et al., "Optimizing breast cancer screening with deep learning," Radiology, vol. 293, no. 1, pp. 38–46, 2019.

[9] J. Choi et al., "Vision transformers for breast cancer recognition using mammography," IEEE J. Biomed. Health Inform., vol. 27, no. 11, pp. 5270–5281, 2023.

[10] X. Yu et al., "Multi-scale contextual features for reducing false positives in mp-MRI prostate cancer detection," IEEE Trans. Med. Imaging, vol. 39, no. 6, pp. 1920-1931, Jun. 2020, doi: 10.1109/TMI.2019.295679.

[11] E. Wu et al., "Deep neural networks improve radiologist performance in breast cancer screening," NPJ Digital Medicine, vol. 3, pp. 1–9, 2020.