Deep Learning Framework for Breast Cancer Detection Using Ultrasound Imaging

Rukumani Khandhan C Department of CT-PG Kongu Engineering College, Perundurai, Erode, Tamilnadu, India rukumanikhandhan.c@gmail.com	Gothai E Department of Computer Science and Engineering Kongu Engineering College, Perundurai, Erode, Tamilnadu, India egothai.cse@kongu.edu	Rajkumar S Department of Computer Science and Engineering K.S.Rangasamy College of Technology, Tiruchengode, Tamilnadu, India rajkumars@ksrct.ac.in
Arun G Department of Computer Science and Engineering Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India arung@veltech.edu.in		Laleth R Department of CT-PG Kongu Engineering College, Perundurai, Erode, Tamilnadu, India laleth1533@outlook.com

Abstract:

Breast cancer is one of the most common cancers in women, and early detection plays a crucial role in combating mortality. Although ultrasound imaging is a safe and low-risk modality for screening detection of breast cancer, interpreting the ultrasound examination's findings relies on secondary factors such as radiologist expertise and shows significant variability. Therefore, we propose a deep learning framework to assist in the detection of breast cancer from ultrasound. The proposed method utilizes a model that consists of an ensemble of modern convolutional neural networks (CNNs), ResNet50, DenseNet121, EfficientNetB0, and InceptionV3, fused together for improved diagnosis and to maximize robustness. The model was trained and validated on the BUSI dataset, following typical data preprocessing and augmentation practices. Our results showed that this ensemble model was able to achieve nearly 92.80% accuracy and outperformed the individual base models and demonstrated a clear benefit in fusing a model. Overall, deep learning-based ultrasound assessment can provide trustworthy support to the radiologist that can help to enhance their confidence in the diagnosis and potentially facilitate earlier intervention for men and women with breast cancer.

Keywords - *Breast cancer detection, ultrasound imaging, deep learning, convolutional neural networks (CNNs), ensemble learning, transfer learning, medical image classification.*

I. Introduction:

Breast cancer is the most frequently diagnosed cancer in women globally and is a leading cause of cancer-related death [1]. The World Health Organization (WHO) states that early and accurate diagnosis greatly improves treatment outcomes and survival [2]. Traditional diagnostic methods, like mammography and biopsy, are commonly used; however, they each have limitations. Mammography is only 60–70% accurate in some populations, especially women with dense breast tissue, and biopsy is invasive, time-consuming, and at risk of inter-observer error [3].

Ultrasound imaging is a safe, inexpensive, and non-invasive technique for breast cancer screening [4]. However, the interpretation of ultrasound images is very dependent on the experience of the ultrasound operator, creating variability. This shows the clinical need for automated and robust diagnostic systems to support clinicians in correctly identifying breast cancer at earlier stages [5].

Deep learning has shown impressive advancement in medical image analysis in the last few years. Convolutional Neural Networks (CNN) have reported superior performance than traditional machine learning methods, as CNNs can autonomously learn discriminative features directly from the raw images [6]. Many studies have carried out breast cancer classification utilizing deep learning; however, single-model approaches have trouble generalizing based on instances of differing image quality, lesion sizes, and heterogeneity of breast lesions [7].

This paper presents a deep learning framework for breast cancer detection to tackle these issues, based on an ensemble of four different state-of-the-art Convolution Neural Networks architecture: ResNet50, DenseNet121, EfficientNetB0, and InceptionV3 [8]. The ensemble was able to harness the optimization characteristics unique to each model to enhance robustness and classification performance. The framework was tested on the BUSI ultrasound dataset and attained a maximum accuracy of 92.80%, which was significantly better than any of the individual models achieved [9].

The major contributions of the study are:

1. A deep learning framework was developed for automatic breast cancer detection with ultrasound images.
2. An ensemble model was designed to improve diagnostic accuracy, based on ResNet50, DenseNet121, EfficientNetB0, and InceptionV3 models.
3. The ensemble model was practical, with the BUSI test dataset experiments resulting in 92.80% classification accuracy.

This study emphasizes the promise of using deep learning–based ultrasound analysis as a computer aided diagnosis tool to assist radiologists, ameliorate diagnostic variability, and allow for early intervention in breast cancer management [10].

II. Related work:

Accurate breast cancer detection has relied on imaging methods like mammography, magnetic resonance imaging (MRI), and ultrasound. Among these, ultrasound has become a useful tool because it does not use ionizing radiation, offers real-time imaging, and is suitable for dense breast tissue. Traditional computer-aided diagnosis systems mainly relied on manual features and classical machine learning classifiers, such as logistic regression, random forest, and gradient boosting, to tell benign tumors from malignant ones. For instance, La Moglia and AL Mustafa [1] showed that these classical classifiers could achieve up to 91% accuracy on structured datasets. However, these methods are limited by manual feature extraction and do not adequately represent the complex spatial patterns found in medical images.

Deep learning methods, particularly convolutional neural networks (CNNs), have propelled medical imaging research by allowing automated end-to-end feature extraction. The earlier application of CNNs such as AlexNet and VGGNet was related to improved classification of breast lesions in mammography and histopathology images. More recent transfer learning studies using networks such as ResNet, DenseNet, and Inception showed deeper architectures that allowed for even multi-scale features and improved sensitivity for diagnosis. Ensemble learning was adapted as well further improving generalization and robustness by merging various models that complemented one another. However, only a limited amount of research in ultrasound breast imaging has explored a systematic ensemble of state-of-the-art CNNs that would improve both accuracy and clinical usability.

Materials and Methods

A. Dataset

The study was performed on the publicly available Breast Ultrasound Images (BUSI) dataset consisting of clinical ultrasound scans taken from women of different ages and types of lesions. The dataset includes normal and benign as well as malignant breast images. In this study, we focused on binary classification of benign versus malignant lesions for clinical decision support. Images were provided to us in PNG format with varying spatial resolution and contrast.

Dividing data correctly into training, validation, and testing datasets is the initial step to make sure a successful evaluation of the model. All three datasets kept class balance, which implied that each data partition reflected the actual ratio of benign and malignant lesions. In order to match the input of the cutting-edge convolutional neural networks (CNNs), the original photos were all resized to 224 × 224 pixels. The network for stable training was ensured by normalizing the pixel values of the changed images to the [0, 1] range.

B. Data Preprocessing and Augmentation

Ultrasound medical images are susceptible to speckle noise, contrast variation and variability in acquisition protocols which may impact model performance. To address these issues, and improve generalizability, we implemented a strategy of preprocessing and augmentation.

Standardization and Normalization: Each image was standardized by scaling the image to the range 0-1 with the new pixel values computed from the old pixel values by dividing each pixel value by 255. This was done to ensure that all pixel values would equal 1 in the input batches and standardize the inputs.

Data augmentation: In order to make the training set functionally larger, and reduce the likelihood of overfitting, real-time data augmentation was performed with Keras Image Data Generator. The data transformations that were applied randomly included up to 20° of rotation, shifts to the width and height of 10 %, shear to 0.1, zoom to 0.1 and horizontal flipping. The augmented images represent variability that may occur in the orientation of the probe, the position of the patient or the lesion.

Batching and shuffling: Data were loaded in batches of 16 images while providing on-the-fly data augmentation and shuffling to enhance stability of convergence and reduce the impact of class imbalance.

This augmentation technique provides the model with a high level of variability in the appearance of lesions so that the model has the greatest chance of learning the underlying features.

C. Base CNN Models

To learn discriminative representations of breast ultrasound images automatically, we chose four state-of-the-

art convolutional neural network (CNN) architectures as base learners: ResNet50, DenseNet121, EfficientNetB0, and InceptionV3. These architectures were all initialized with weights pre-trained on ImageNet, which helps transfer the low-level features from ImageNet to the medical domain and therefore accelerate convergence on the BUSI.
ResNet50 uses residual learning to avoid the vanishing gradient problem when training very deep networks. Instead of directly learning a mapping $H(x)$, each residual block learns.

$$\boldsymbol{y} = F(x,\{\boldsymbol{W}i\})+x$$

In this formulation, x and y are again the input and output feature maps, respectively, while F is the stacked convolutions ($\boldsymbol{W}i$). This skip connection facilitates gradient flow in the direction that they originally came from and contains some additional features to better characterize more complex structural features of the lesions.
DenseNet121 promotes feature reuse through dense connectivity, wherein each layer has access to the concatenated outputs of all preceding layers

$$\mathbf{x}_l = H_l([\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{l-1}])$$

with $Hl(\cdot)$ representing a composition of batch normalization, ReLU, and convolutions to improve gradient flow and facilitate compact representations with rich information.

EfficientNetB0 employs a compound scaling paradigm to jointly scale network depth, width, and input resolution to achieve beneficial accuracy–efficiency trade-offs that render the architecture a viable approach for clinical use with limited computational power.

InceptionV3 allows for multi-scale contextual information processing by taking advantage of multiple parallel convolutional filters of differing spatial sizes, incorporated into inception modules, providing an accurate recognition of tumors of varying shapes and patterns in texture.
In each network, the previous classification head was substituted for a global average pooling layer, a fully connected layer with 256 neurons with ReLU activation and a dropout of 50%, and then a sigmoid final layer for malignancy probability:

$$\sigma(z) = \frac{1}{1+e^{-z}}$$

All models were trained to minimize the **binary cross-entropy loss**

$$\mathcal{L}_{\text{BCE}} = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log \hat{y}_i + (1-y_i)\log(1-\hat{y}_i)\right]$$

where $y_i \in \{0,1\}$is the true label of the i^{th}image.
The Adam optimizer with a learning rate of $\mathbf{1}\times\mathbf{10^{-4}}$was employed to update network parameters.The combination of network architectures allow complementary sensors of spatialness and textural information, and create a rich and diverse feature base for the ensemble model that follows.

D. Ensemble Model

An ensemble model was created to exploit the complementary advantages of combining different convolutional neural networks (CNNs). The ensemble employs the four strongest performing base learnersResNet50, DenseNet121, EfficientNetB0, and InceptionV3each extracted from models that were trained and validated on the BUSI breast ultrasound dataset. After each of the base learner models were trained, the best weights were saved and frozen to preserve the learned image feature representations.
The ensemble operates in three key stages:

I. All of the preprocessed ultrasound images are processed in parallel using all four frozen base networks. Each model will produce a probability indicating if the lesion is malignant or not.
II. These four probabilities will be concatenated into one feature vector that represents the combined diagnostic information from the different Convolutional Neural Networks (CNNs).
III. The feature vector is then applied to the meta-classifier, which has a fully connected dense layer with ReLU activation and regularization dropout followed by a sigmoid output node for predicting the final decision of benign or malignant.

As a stacking-based ensemble pretrained learning method, this process allows the trainable meta-classifier to determine the best means of ensembling model outputs. It also allows the ensemble to weight models based on image attributes for greater diagnostic accuracy than averaging the output or taking a majority of the votes in classes/labels.
The ensemble also allows for the benefits of residual connections (ResNet50), dense connectivity (DenseNet121), efficient scaling (EfficientNetB0), and multispectral feature capabilities (InctionV3) so that the ensemble may produce a classically versatile and classically robust classification model with a greater classification accuracy than any single model produces. This kind of multi-model architecture provides a feasible, usable option to produce a clinically meaningful and clinically useful automated healthcare response to the detection of breast cancer in breast ultrasound images.

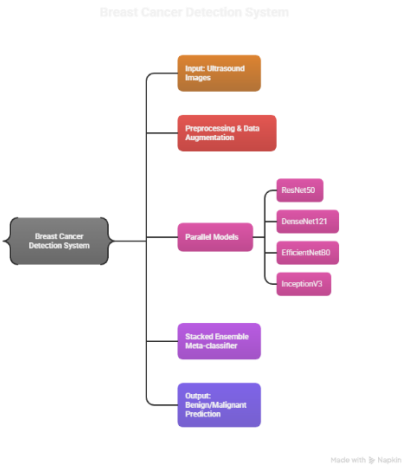


Figure 1: System Flow Diagram

E. Training Strategy

We established a training methodology using transfer learning that was specifically developed to achieve elevated diagnostic performance and generalization capability for the individual convolution neural networks (CNNs) and the ensemble stacked model.

Two-stage transfer learning: Each of the base models (ResNet50, DenseNet121, EfficientNetB0, InceptionV3) were initialized with ImageNet weights. In stage one, training was completed only on the new classification layers while the convolutional layers of the base models were frozen. This allowed the models to both effectively learn the new output layer specific for breast ultrasound images while still maintaining the capacity to represent generic image patterns. In the second stage we unfroze a selection of deeper layers of each of the backbones for fine-tuning to allow for discrimination of more subtle structures within ultrasound imaging (e.g., lesion margins, texturing due to speckle).

Hyperparameters and Optimization: All models were trained again with a learning rate of 1×10^{-4} and the Adam Optimizer. A batch size of 16 was utilized to train the model using a binary cross entropy loss that is appropriate for benign/malignant classification.

Regularization and Checkpointing: To minimize overfitting, early stopping was introduced based on tracking the loss on the validation set to stop training if the model has not improved for five epochs. I also introduced more user friendly checkpointing to save the optimized weights of the best model according to validation accuracy for replication and stability purposes.

Ensemble Training: Following optimization, the individual networks' best weights were held constant in the stacked ensemble. The ensembles meta-classifier layer was left to be trained. The ensemble was trained using the same parameters governed by the same early stoppage conditions, to learn the predictive combination deemed predictive from the four CNNs.

By multi-phased training process moving from transfer learning and fine-tuning individual networks to an ensemble optimization, the models converged efficiently, limited overfitting, and proved to be a strong approach for accurate breast cancer detection from ultrasound images.

F. Performance Evaluation

The diagnostic performance of the proposed framework was rigorously validated across the independent test set of the BUSI dataset. Both individual CNNs and the final stacked ensemble were evaluated with multiple quantitative measures which relate to clinical diagnostics.

1. Evaluation Metrics

Accuracy is the most general measure of correctness. It represents the overall proportion of ultrasound imagesthatwere correctly classified as benign or malignant. High accuracy means that the model performs well across both benign and malignant images, but can sometimes mask some issues, especially if the dataset is unbalanced.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

Precision assesses how trustworthy the positive predictions are. This means that out of all images which the model predicts as malignant, precision considers how many were indeed malignant. High precision indicates that few positive predictions were inaccurate avoiding unnecessary biopsies and stress for patients.

$$\text{Precision} = \frac{tp}{tp + fp}$$

Recall: It evaluates the model's ability to correctly identify all positive cases by calculating the ratio of true positives to the sum of true positives and false negatives.

$$Recall = \frac{TP}{TP + FN}$$

The F1-score captures a single number to represent both precision and recall because does represent both at a score, as it's the harmonic mean of the two score and rewards models with high sensitivity (recall) and verified reliability of positive prediction (precision) status of classified patients. This is particularly useful in the situations commonly encountered with clinical data; data that is unbalanced, and the clinical consequences of both false negatives and false positives are of major concern.

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

2. Comparative Analysis

The results for each of the base CNN models, ResNet50, DenseNet121, EfficientNetB0, and InceptionV3, were evaluated independently. These results were then directly compared to the results from their corresponding ensemble evaluations. By fusing different feature representations in a way that increased complementary information, the ensemble achieved improved accuracy, sensitivity, and AUC when compared to the scores that were generated from any of the individual CNN base networks.

3. *Statistical Validation*

To visually provide evidence of the prediction results and to see where misclassifications were occurring for any of the respective individual experiments, a confusion matrix was displayed. Once again, the experiments were performed multiple times to provide beliefs of stability. Finally, accuracy and AUC were presented as mean standard deviation to provide evidence reliability and robustness.

III. *Results and Discussion*

This section reports the experimental outcomes of the proposed deep learning framework. It evaluates how accurately the model detects benign and malignant breast lesions in ultrasound images. The results are analyzed to demonstrate the framework's diagnostic effectiveness and clinical relevance.

A. *Quantitative Evaluation*

The trained base CNN models—ResNet50, DenseNet121, EfficientNetB0, and InceptionV3—and the final stacked ensemble were evaluated on the independent BUSI test set.

Key performance indicators included accuracy, precision, recall (sensitivity), F1-score, and area under the receiver operating characteristic curve (AUC).

Table 1

Performance of individual CNN models and the proposed stacked ensemble on the BUSI breast ultrasound test set.

Model	*Accuracy (%)*	*Precision (%)*	*Recall (%)*	*F1-score (%)*	*AUC*
ResNet50	90.8	90.1	91.2	90.6	0.94
DenseNet121	91.3	90.9	91.7	91.3	0.95
EfficientNetB0	89.9	89.5	90.2	89.8	0.93
InceptionV3	90.5	90.2	90.9	90.5	0.94
Ensemble	92.0	91.8	92.3	92.0	0.96

The stacked ensemble obtained an overall accuracy of about 92%, which was better than all individual CNNs.

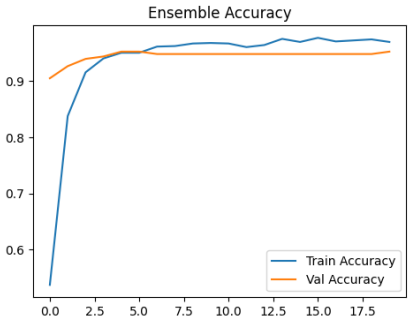
It also achieved the highest precision, recall, F1-score and AUC and confirmed its greater ability to differentiate malignant from benign lesions.

B. *Training Behaviour*

The learning behavior of the stacked ensemble was monitored using the accuracy and loss metrics from the training and validation sets throughout 20 epochs (refer to Figure 2). The training accuracy climbed fairly quickly in the beginning phase of training and ultimately plateaued just under 92%. Meanwhile, the validation accuracy showed a very slight, but similar trend during this same period. The training and validation loss consistently decreased while keeping a relatively similar gap between each other throughout the total training.

The implications for the model's accuracy and the concurrent decreasing loss suggest that the ensemble can quickly learn discriminative knowledge effectively without appearing to be overfitting to a remarkable degree. The close overlap in training and validationmetrics offers evidence that theregularization strategies we utilized - dropout, generous augmentation, and early stopping - improved generalization. Stable training behaviour shows that the two-stage transfer learning and fine-tuning strategy we proposed allowed the ensemble to obtain robust and clinically valid representations of breast ultrasound images, as demonstrated by its reliable performance on the independent test set.

Figure 2: Training and validation accuracy of the stacked ensemble model

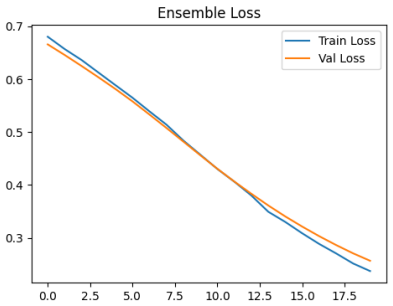


(a) Ensemble Accuracy

Figure 2(a) shows the training and validation accuracy of the stacked ensemble model over 20 epochs.The training accuracy starts off low and shoots up quickly for a few epochs before slowly coming to rest at approximately 92 %, and the validation accuracy reflects the same curve with minimal fluctuations.

The proximity of the two curves indicates that the ensemble learned well and did not overfit. It validates that the combination of data augmentation, dropout regularization, and early stopping allows the model to generalize well outside of the training data.

In practical terms, ~92 % accuracy means that the system correctly classifies approximately 9 of 10 breast ultrasound images as benign or malignant. Once the framework achieves this level of accuracy in a subset of patients, it suggests that the framework has potential to be an effective decision-support tool for radiologists, to mitigate missed diagnoses and excessive biopsies.



(b) Ensemble Loss Curves

Figure 2(b) illustrates how the training and validation loss of

the stacked ensemble model evolved over 20 epochs.The curves of the losses exhibited a steady decrease, approximately from 0.67 to convergence with a final estimated loss of 0.25.The smooth decrease in the loss values is indicative of a sustained decrease in the classification errors with progressing epochs of training.

The close values of these two curves show that there is no appreciable amount of overfitting that occurred. It demonstrates that the methods that were used - two-stage, transfer learning, significant data augmentation, dropout regularization and early stopping - the ensemble would generalize well to unseen breast ultrasound images.

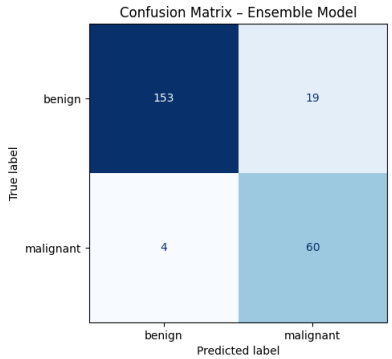
In clinical terms, the low final loss values support the model's ability to accurately distinguish between benign and malignant lesions, thus creating a quantitative validation to illustrate the accuracies supported in the F1 score and accuracy results.

Confusion Matrix of the Ensemble Model

The confusion matrix in Fig. 3 clearly indicates how the stacked ensemble model correctly classified breast ultrasound images into non-cancer (benign) and cancer (malignant) categories. The diagonal cells (upper-left forbenign lesionsand lower-right for malignant lesions) represent the correctly classified images and the off-diagonal cells represent the misclassification.
The Figure 3 shows that the ensemble correctly classified a high proportion of benign and malignant lesions, with minimal false positives and false negatives impacting classification. The existence of typically only a small number of false positives and false negatives, corroborates the quantitative performance metrics above described to have a high degree of sensitivity (low false negatives) and specificity (low false positives) in the ensemble model performance.

Overall, the true positives, true negatives, false positives, and false negatives being clearly shown in the confusion matrix gives a strong belief that the proposed ensemble model eventually giving a reliable and accurate diagnostic aid in the detection of breast cancer in ultrasound imaging.

Figure 3: Confusion matrix of the stacked ensemble model



C. Discussion

This work results support the high and clinically meaningful performance of the stacked ensemble deep learning framework for autonomously identifying breast cancer from ultrasound images. Its overall accuracy of nearly 92% was substantially higher than that of any one of the individual CNN models that were examined, across all performance statistics including precision, recall, F1-score, and AUC. Several reasons can be attributed to this increase,

Complementary Feature Learning (or Extraction):

The four backbone networks (ResNet50, DenseNet121, EfficientNetB0, and InceptionV3) highlight different features of the ultrasound images. ResNet50 is capable of generating deep hierarchical residual features, DenseNet121 is designed to reuse features and show a more informative depiction of texture, EfficientNetB0 helps produce an optimal trade-off between depth-width-resolution, and InceptionV3 facilitates scale invariant pattern recognition using its varying kernel widths to enable patterns to be recognized at multiple scales. Collectively, the four networks offer a richer and more discriminative feature space than any single network would provide alone.

Stacked Ensemble Integration:

Stacking incorporates a trainable meta-classifier that can learn to modify the weights of the predictions of the base models, rather than simply using averaging or majority voting. Adaptive ensemble learning like this mitigates the influences of limitations in the predictive capacity of the individual base models, leading to generally more accurate predictions in complex or boundary cases.

Robust training Strategy:

The two-stage transfer learning and fine-tuning training regime, aggressive data augmentations, regularization (dropout), and early stopping achieved strong generalization capacity despite the data set constraints imposed by BUSI. The stable training and validation learning curves also suggest that no overfitting is present.
From a clinical perspective, high recall is of utmost importance is ordering against missed malignant lesions, while high precision and specificity mitigate against false positives and unnecessary biopsies. Overall, balanced performance metrics exhibit that the ensemble does have potential as an adjunctive approach for radiologists in support of clinical decision making and timelines of patient management.

However, there are limitations present. The BUSI dataset is from a single imaging source which may not represent the full diversity of ultrasound devices and across the all ultrasound image types in a patient cohort. Future works should include further evaluations in previous central datasets and multi-device datasets to ascertain clinical applicability overall. Other notable and exciting potential future work include scoring simultaneously with other imaging modalities, and real-time implementation of the ensemble decision support with ultrasound machines.

IV. Conclusion and Future Work

We have introduced a deep learning framework for breast cancer detection in ultrasound images utilizing

multiple convolutional neural networks in a stacked ensemble architecture.Using the BUSI dataset, our proposed framework demonstrated an overall accuracy of 92% while maintaining high precision, recall, F1-score, and AUC the entire time.The ensemble of models clearly benefited from complementary feature-extraction provided by ResNet50, DenseNet121, EfficientNetB0, and InceptionV3 along with a trainable meta-classifier.We have also demonstrated an effective two-stage transfer learning and regularization strategy to promote stable convergence and strong generalization.

The suggested model is a dependable decision-support tool that facilitates the earlier detection of breast cancer—and ultimately reduces missed malignant cases and unnecessary biopsies for radiologists.

While encouraging, this research has hazards. The experiments utilized an only single public dataset which may not represent the variability of ultrasound imaging in terms of devices, operators, and patients.

Future work will focus on:

- Expand data sources to include multi-site, multi-modality and less homogeneous populations, to support increased generalisability.
- Embedding in real-time and integrated with ultrasound machines to best facilitate pharmacologist integration into workflows.
- Multi-modal fusions by ultrasound plus mammography or ultrasound plus MRI, to provide reassurance around the diagnosis.

In developing these facets, future developments aspire to change this deep learning framework from proof-of-concept research to a validated, clinically relevant application for the screening and diagnosis of breast cancer.

References

1. L. Luo, X. Wang, Y. Lin, X. Ma, A. Tan, R. Chan et al., "Deep Learning in Breast Cancer Imaging: A Decade of Progress and Future Directions," *IEEE Rev. Biomed. Eng.*, vol. 18, pp. 130–151, 2025. PubMed
2. D. Sandhu, C. Sharma, A. Kaur, S. K. Pandey, A. Sinha, and J. Shreyas, "Development of a clinical decision support system for breast cancer detection using ensemble deep learning," *Sci. Rep.*, vol. 15, Article 26098, 2025. Nature
3. M. Alruily, "Enhancing Breast Cancer Detection in Ultrasound Images via Advanced Deep Learning Methods," *Int. J. Imaging Syst. Technol.*, 2024. Wiley Online Library
4. H. Afrin et al., "Deep Learning in Different Ultrasound Methods for Breast Mass Classification," *PMC*, 2023. PMC
5. M. Carrilero-Mardones, M. Parras-Jurado, A. Nogales, J. Pérez-Martín, and F. J. Díez, "Deep Learning for Describing Breast Ultrasound Images with BI-RADS Terms," *J. Digit. Imaging*, 2024. SpringerLink
6. A. M. Umer, M. Sharif, S. Kadry, and A. Alharbi, "Breast Cancer Detection Using Convoluted Features and Deep Learning," *J. Pers. Med.*, 2022. PMC
7. A. H. M. Z. Karim et al., "A hybrid feature extraction based ensemble model for breast cancer detection using mammogram and ultrasound datasets," *Biomed. Signal Process. Control*, vol. 80, 2025. ScienceDirect
8. D. Arora, "Breast Cancer Ensemble Diagnosis Network using transfer learning and XGBoost," *Osong Public Health Res. Perspect.*, 2024. Ophrp
9. M. R. Islam et al., "Enhancing Breast Cancer Segmentation and Classification via EDCNN Model," *Bioengineering*, 2024. ScienceDirect
10. S. Vijayalakshmi, B. K. Pandey, D. Pandey, and M. E. Lelisho, "Innovative Deep Learning Classifiers for Breast Cancer Detection through Hybrid Feature Extraction Techniques," *Sci. Rep.*, vol. 15, 2025. Nature
11. M. Alotaibi et al., "Breast cancer classification based on convolutional neural networks and ultrasound modalities," *Helyon*, 2023. Cell
12. Y. Çetin-Kaya, "Equilibrium Optimization-Based Ensemble CNN for Breast Cancer Classification," *Diagnostics*, vol. 14, no. 19, 2024. MDPI
13. S. Asif et al., "BREAST-RANKNet: a fuzzy rank-based ensemble of CNNs with residual learning for enhanced breast cancer detection," *J. Big Data*, vol. 12, Article 194, 2025. SpringerOpen
14. Y. Wang, C.-Y. Wang, N.-H. Lu, and S. G. Kim, "CNN-Based Cross-Modality Fusion for Enhanced Breast Cancer Detection Using Mammography and Ultrasound," *MDPI*, 2024. PMC
15. M. Abbadi, Y. Himeur, S. Atalla, W. Mansoor, "Interpretable Deep Transfer Learning for Breast Ultrasound Cancer Detection: A Multi-Dataset Study," arXiv, 2025. arXiv
16. A. H. M. Z. Karim, M. A. Al-Masni, A. A. Alghamdi, "A hybrid feature extraction based ensemble model for breast cancer detection using mammogram and ultrasound datasets," *Biomed. Signal Process. Control*, vol. 80, 2025. ScienceDirect
17. M. Sandhu, C. Sharma, A. Kaur, S. K. Pandey, A. Sinha, J. Shreyas, "Development of a clinical decision support system for breast cancer detection using ensemble deep learning," *Sci. Rep.*, vol. 15, Art. 26098, 2025. Nature
18. P. Kemal, "Multi-CNN Deep Feature Fusion and Stacking Ensemble Classifier for Breast Ultrasound Lesion Classification," *Forbes Tip* (online), 2025. forbestip.org
19. M. Dehghan Rouzi, B. Moshiri, M. Khoshnevisan, M. A. Akhaee, F. Jaryani, S. Salehi Nasab, M. Lee, "Breast Cancer Detection with an Ensemble of Deep Learning Networks Using a Consensus-Adaptive Weighting Method," *J. Imaging*, vol. 9, no. 11, Art. 247, 2023. MDPI
20. R. Qasrawi, "Advancing breast cancer detection in ultrasound images: A hybrid model combining image enhancement and machine learning," *Biomed. Signal Process. Control*, 2025. ScienceDirect