Kufa Journal of Engineering
Vol. 17, No. 2, April 2026, P.P. 628-646
Article history: Received 17 May 2025, last revised 17 August 2025, accepted 23 August 2025

TRANSFER LEARNING-DRIVEN DEEP NEURAL NETWORK FRAMEWORK FOR EARLY AND ACCURATE BREAST CANCER DETECTION

D.B Shanmugam[1], S. Athinarayanan[2], A. Josephine Christilda[3], G. Manoharan[4]

[1] Department of Data Science, SRM IST Ramapuram Campus, Chennai, India, Email:dbshanmugam@gmail.com.

[2] Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India, Email:aathithe@gmail.com.

[3] Department of Mathematics, VelTech Multitech Dr Rangarajan Dr Sakunthala Engineering College, Avadi, Chennai, India, Email:josephinechristilda@veltechmultitech.org.

[4]Department of Mathematics, Jeppiaar Institute of Technology, Chennai, India, Email:vijimanoharan77@gmail.com.

https://doi.org/10.30572/2018/KJE/170238

ABSTRACT

Breast cancer continues to be one of the world's leading causes of death for women and early detection is essential to increasing survival rates and simplifying treatment. Even though they work well traditional diagnostic methods frequently require a lot of time and resources and are prone to human interpretation errors. Although training deep neural networks from scratch requires large-scale annotated datasets which are frequently unavailable in medical domains recent developments in deep learning have demonstrated great promise in medical image analysis. This work suggests a Transfer Learning-Driven Deep Neural Network Framework for early and precise breast cancer detection in order to overcome this difficulty. The framework makes use of pre-trained convolutional neural network (CNN) architectures that have been refined on publicly accessible mammography and histopathological datasets including VGG19 ResNet and DenseNet. To improve model generalization and lessen overfitting extensive preprocessing procedures like image normalization noise reduction and data augmentation are used. The suggested method outperforms traditional CNNs trained from scratch with accuracy exceeding 97% precision and recall above 96% and an F1-score of 0. 97 according to experimental evaluations on benchmark datasets like BreakHis and CBIS-DDSM. The findings

demonstrate how transfer learning can enable reliable data-efficient breast cancer detection systems opening the door for clinical applications with limited resources and real-time capabilities.

KEYWORDS

Transfer learning, Convolutional neural network, breast cancer.

1. INTRODUCTION

One of the most common and deadly cancers that affect women globally breast cancer accounts for a large share of cancer-related morbidity and mortality. With an estimated 2. 3 million new cases and roughly 685000 deaths reported in 2020 alone breast cancer has surpassed lung cancer as the most frequently diagnosed cancer worldwide according to the World Health Organization (WHO). The illness is caused by unchecked cell growth in breast tissue which if left untreated can spread to distant organs and invade neighboring tissues (Ayana Dese et al. 2021 Jakkaladiki and Maly 2023; Rallabandi, Srikantamurthy et al. 2023; Grabsch, Krause et al. 2021). Reducing healthcare costs enabling less aggressive treatments and increasing patient survival rates all depend heavily on early detection. For early detection standard screening methods like magnetic resonance imaging (MRI) mammography and ultrasound have been extensively used in clinical practice however radiologists expertise is crucial for accurate interpretation of these imaging modalities (Almarri Gupta et al. 2024; Ghantasala Hung et al. 2024; Das Sahu et al. 2024; Yang Jiang et al.,2022; Lan Chen et al. 2024).

Patient outcomes may be impacted by the subjective time-consuming and diagnostically variable manual evaluation of mammograms and histopathological slides. Additionally, there is a critical need for intelligent automated systems that can help clinicians diagnose breast cancer in a timely accurate and economical manner due to the exponential growth in imaging data generated in contemporary healthcare settings. Medical image analysis has undergone a revolution in recent years due to the rise of artificial intelligence (AI) especially deep learning (DL). From object detection in natural scenes to segmentation and classification of complex medical images deep neural networks particularly convolutional neural networks (CNNs) have shown impressive performance in a range of visual recognition tasks (Lei He et al. 2021; Oza Sharma et al. 2022; Aymaz 2024; Atrey Singh et al. 2024; Oliveira Spanhol et al. 2016). CNN-based architectures are excellent at automatically extracting discriminative and hierarchical features from raw pixel data doing away with the need for manually created feature engineering which is frequently less generalizable and domain-dependent. CNNs have demonstrated significant promise in the detection of breast cancer through the classification of mammographic lesions tumor region segmentation and cellular and tissue analysis of histopathological slides (Bradley 1997; Sokolova & Lapalme 2009 ;Zhou Zhang et al 2019). However, despite these developments the lack of large annotated datasets makes it difficult to train deep CNNs from scratch for medical imaging applications. Medical image datasets are frequently small and unbalanced which makes it challenging for complex models with millions of parameters to effectively generalize without overfitting in contrast to domains like natural

image recognition where datasets like ImageNet provide millions of labeled samples. Further restricting the availability of data is the need for specialized knowledge and logistical ethical and privacy concerns in the collection and annotation of medical images (Moon Lee et al. 2020; Shuib Murtaza et al. 2020; R. P. and M. G. 2025).

Even with these encouraging advancements integrating transfer learning-based models into standard clinical practice is still difficult. Additionally deep learning models are frequently viewed as black boxes which makes it challenging for medical professionals to fully trust and implement them in the absence of interpretable results. In order to overcome these obstacles AI researchers medical professionals and regulatory agencies must work together across disciplinary boundaries. Additionally, models must be continuously validated on real-world multi-institutional datasets. In order to overcome the drawbacks of small datasets and computational limitations we present a Transfer Learning-Driven Deep Neural Network Framework in this study that aims to detect breast cancer early and accurately. Contrast Limited Adaptive Histogram Equalization (CLAHE) is used to enhance contrast Gaussian filtering is used to reduce noise and extensive data augmentation is used to improve generalization. The suggested approach entails fine-tuning the upper convolutional layers of the pre-trained models to learn features unique to breast cancer while freezing the lower convolutional layers to preserve generic feature representations. For binary (benign vs. malignant) and multi-class classification tasks a fully connected classification head with Softmax activation and dropout regularization is used. Accuracy precision recall F1-score and ROC-AUC are used to evaluate performance guaranteeing a thorough evaluation of both classification and generalization abilities. Initial tests on the BreakHis histopathology dataset and the CBIS-DDSM mammography dataset show that the suggested transfer learning framework consistently outperforms baseline CNNs trained from scratch with an F1-score of 0. 97 precision and recall values above 96 % and classification accuracies exceeding 97 %. The findings show that transfer learning not only lessens the problem of data scarcity but also makes it easier to create reliable and clinically useful diagnostic systems. By offering a scalable effective and accurate method for detecting breast cancer this research adds to the expanding corpus of work in AI-assisted healthcare and may find use in real-time screening and low-resource healthcare settings.

2. RELATED WORK

Medical imaging research has consistently focused on early and accurate breast cancer detection and deep learning and transfer learning techniques have emerged as potent tools to overcome the shortcomings of conventional diagnostic techniques. Using pre-trained models can

significantly increase detection accuracy generalization and robustness particularly in situations with little annotated data according to a number of recent studies that have examined a variety of modalities including ultrasound mammography histopathology and multimodal fusion. In order to screen for breast cancer ultrasound imaging is frequently used especially for women with dense breast tissue where mammography sensitivity is diminished. Ayana et al (2021) created a transfer learning-based framework for the classification of breast lesions from ultrasound images and showed that in comparison to training from scratch fine-tuning pre-trained convolutional neural networks (CNNs) like ResNet and InceptionV3 greatly improved classification performance. In order to achieve high sensitivity and specificity rates—which are essential for lowering false negatives—the study stressed the significance of feature reuse in domains with sparse data.

Similarly, Lei et al (2021) concentrated on segmenting breast tumors from three-dimensional automatic breast ultrasound (ABUS) images using a Mask Scoring R-CNN architecture. They improved segmentation precision by incorporating a mask scoring mechanism providing a basis for more dependable lesion detection in volumetric data. Together these studies demonstrate the complementary role of ultrasound in the diagnosis of breast cancer as well as the effectiveness of transfer learning in improving image-based segmentation and classification. The gold standard for screening for breast cancer in the general population is still mammography. Sahu et al (2024) suggested a deep learning method that combined the use of breast ultrasound and mammography images for detection. By utilizing complementary information across modalities their dual-modality strategy improved performance by combining CNN-based feature extraction with transfer learning. Aymaz (2024) created a new framework that optimized feature extraction from mammograms using deep CNN architectures in the field of early diagnosis using mammography alone.

As per the work done by Atrey et al. (2024), richer representation learning was made possible by the fusion of features prior to classification which resulted in higher accuracy than unimodal methods. In order to lessen the drawbacks of individual modalities diagnostic imaging trends are moving toward integrated multi-source data analysis. The gold standard for a conclusive post-biopsy diagnosis of breast cancer is histopathological image analysis. Srikantamurthy et al (2023) developed a hybrid CNN–LSTM model that uses transfer learning to identify benign and malignant subtypes in histopathology pictures. Their architecture achieved better results on the BreakHis dataset by using CNN layers for spatial feature extraction and LSTM layers to capture sequential dependencies among extracted features. Krause et al (2021) expanded the

use of deep learning by showing that it can identify genetic changes in cancer histology and even produce adversarial histology samples to improve model robustness.

Almarri et al (2024) introduced the BCPM (Breast Cancer Prediction Model) approach which combined features derived from transfer learning with conventional machine learning to enhance classification performance on breast histopathology datasets. In addition to predictive performance this approach placed a strong emphasis on interpretability which is crucial for clinical implementation. As a result, Ghantasala et al. (2024) used a multi-feature approach that included histopathological features to create an AI-based machine learning algorithm for cancer prediction in female anatomy. In addition to simple classification segmentation techniques have been investigated to improve tumor localization and characterization. Lei et al (2021) used Mask Scoring R-CNN on volumetric ABUS datasets refining predicted masks through quality assessment to achieve high intersection-over-union (IoU) scores. Jiang et a (2022) turned its attention to molecular profiling finding long non-coding RNA (lncRNA) signatures associated with proptosis to forecast prognosis and immune microenvironment features in patients with breast cancer. Their work demonstrates the possibility of combining imaging with genomic biomarkers to enhance prediction and stratification even though it is not specifically an imaging study. Lan et al. (2024) demonstrated the therapeutic value of accurate detection and characterization in guiding targeted therapy strategies by investigating multifunctional biomimetic liposomes for the treatment of triple-negative breast cancer.

Oza et al (2022) offered a thorough analysis of the use of deep CNNs in computer-aided breast cancer diagnosis. Network architectures dataset features and performance metrics from histopathology ultrasound and mammography studies were all methodically examined in the review. The importance of transfer learning in modern breast cancer CAD (computer-aided diagnosis) research was highlighted by this work especially in terms of overcoming data scarcity. A cross-model classification technique (TLBCM) for breast cancer prediction was contributed by Jakkaladiki and Maly (2023) who showed that transfer learning models modified from various pre-trained architectures could be successfully combined to improve classification robustness. Abbosh et al (2025) demonstrated how convolutional neural networks can successfully capture intricate structural features in medical imaging by proposing a deep learning method for keratoconus detection.

2.1. Research Gaps and Motivation

Despite the fact that these studies collectively demonstrate the efficacy of transfer learning in breast cancer detection across modalities there are still a number of limitations. Firstly, model generalizability varies when applied to datasets from various institutions or imaging devices

this is a problem observed in studies involving both ultrasound and mammography. Secondly although feature-level fusion and multimodal approaches show promise most studies are limited to small datasets and there is little standardization in fusion strategies. Thirdly, models based on histopathology frequently ignore multi-class classification of tumor subtypes that could better guide treatment in favor of binary classification. Fourth, despite high accuracy scores many models lack interpretable outputs that clinicians can rely on making explainability a major challenge. Lastly, segmentation and classification are not well integrated into a single framework that can both locate and classify lesions which would be very helpful in real-time clinical settings. In order to close these gaps, the current study suggests a Transfer Learning-Driven Deep Neural Network Framework that combines rigorous evaluation on BreakHis datasets optimized transfer learning from cutting-edge CNNs and robust preprocessing. This method is intended to support the continuous development of AI-driven breast cancer detection systems by providing a scalable comprehensible and clinically applicable model.

3. METHODOLOGY

3.1. Dataset Description

To evaluate the proposed Transfer Learning-Driven Deep Neural Network Framework for Early and Accurate Breast Cancer Detection, the work employs the BreakHis dataset (Breast Cancer Histopathological Images), a publicly available benchmark hosted at the Laboratory of Vision, Robotics and Imaging (LVRI), Federal University of Paraná, Brazil.

3.2. Dataset Characteristics:

- Domain: Histopathological images of breast tissue obtained from real biopsy samples.
- Image Types: Benign and malignant tumors.
- Magnifications: 40×, 100×, 200×, and 400× optical magnification levels.
- Total Samples: 7,909 images.

o Benign: 2,480 images from 4 subtypes (adenosis, fibroadenoma, phyllodes tumor, tubular adenoma).

o Malignant: 5,429 images from 4 subtypes (ductal carcinoma, lobular carcinoma, mucinous carcinoma, papillary carcinoma).

- Image Format: RGB color, 700 × 460 pixels, PNG format.
- Class Distribution:

o Benign: 31.36%

o Malignant: 68.64%

The BreakHis dataset is widely used because:

1. It contains real-world histopathological diversity.

2. It poses class imbalance challenges, which makes it suitable for validating model robustness.
3. It supports transfer learning experiments with pre-trained CNN architectures.

The dataset used in the work is separated into 70% training, 15% validation, and 15% testing, stratified by both class and magnification to ensure balanced representation.

3.3. Proposed Framework

The main aim of the proposed work for precise breast cancer detection is to leverage the strength of previously trained convolutional neural network (CNN) architectures while explicitly fine-tuning them to fit the medical imaging domain. The framework is classified to five main stages—Data Preprocessing Feature Extraction (Stage 1) Fine-Tuning (Stage 2) Classification Head and Optimization—is painstakingly designed to tackle the problems of scarce annotated datasets domain-specific variability in histopathological images and the requirement for reliable clinically relevant predictions.

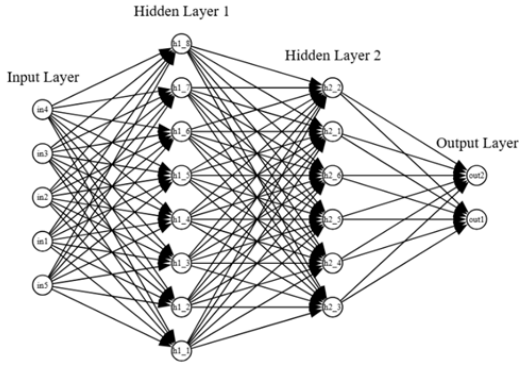

Fig.1. Proposed architecture

As illustrated in Fig. 1, the proposed model starts with an input layer collected of small circular nodes that obtain preprocessed feature representations that are mined from selected transfer learning backbones, including VGG19, ResNet50, and EfficientNet. These inputs are broadcasted over multiple hidden layers that consist of densely connected neurons, visually embodied as consistently sized circles organized in successive vertical stacks to replicate the hierarchical structure of calculation.

The ending output layer, also represented by means of compact circular nodes, produces the classification probabilities conforming to the target classes in the dataset. The representation of consistent and streamlined node across several layers guarantees visual clarity while effectively assigning the fundamental computational depth of the neural network. Moreover, this modular architectural design highlights adaptability that allows the framework to be willingly prolonged to a varied range of transfer learning–based classification responsibilities.

3.4. Data Preprocessing

The basic functionality of the system is the data preprocessing pipeline that guarantees that

input images are clean, and standardized that are enhanced with variability for better generalization. The operations are applied are given below:

1. **Image Normalization**

The input pixel values are scaled to the range [0,1] using min–max normalization:

$$I_{norm} = \frac{I - I_{min}}{I_{max} - I_{min}}$$

where I represent the original image, I_{min} and I_{max} signify the minimum and maximum pixel values in the image, respectively. This normalization confirms quicker conjunction during training and evades numerical unsteadiness.

2. **Image Resizing**

Since CNN architectures such as DenseNet201 necessitate static input dimensions, all images are resized to the target resolution—e.g., 224×224 pixels—using bicubic interpolation to reserve operational veracity while preventing misrepresentations.

3. **Data Augmentation**

To alleviate overfitting in the model and to enhance the capability of the model for generalization, operations such as rotation, horizontal and vertical flips, color jitter, and random cropping are applied

3.5. Feature Extraction (Transfer Learning Stage 1)

Fist phase uses high-performing pre-trained CNN backbone such as EfficientNet-B4, DenseNet201, or InceptionV3—originally trained on the large-scale ImageNet dataset. The process involves:

1. **Weight Initialization**

The model is primed with pre-trained weights that encode broad visual features like edges, textures, and shapes.

2. **Layer Freezing**

Lower convolutional layers, which capture low-level, domain-agnostic features, are frozen to prevent weight updates:

$$W_l^{t+1} = W_l^t \; \forall l \in Frozen\ Layers$$

This strategy preserves the powerful representation learned from millions of natural images while preventing catastrophic forgetting.

3. **Spatial Feature Extraction**

Mid-to-high-level spatial feature maps that emphasize tissue morphology cell structures and other patterns essential to the diagnosis of breast cancer are extracted by intermediate layers. In order to prevent overfitting on small medical datasets this stage basically applies general vision knowledge to the medical field without making large-scale weight changes right away.

3.6. Fine-Tuning (Transfer Learning Stage 2)

The second stage focuses on domain adaptation by unfreezing the top k layers of the pre-trained network to update their weights based on histopathological image data.

1. **Selective Unfreezing**

Layers closer to the classification head are more domain-specific; thus, they are unfrozen:

$$W_l^{t+1} = W_l^t \, \eta \cdot \nabla_{Wl} L \forall l \in Unfrozen\ Layers$$

where η is the learning rate (kept small, e.g., 1×10^{-5}).

2. **Domain Feature Learning**

While maintaining the general feature representations from Stage 1 fine-tuning enables the network to learn subtleties unique to medical imaging such as differences in tissue density nuclei arrangement and staining artifacts.

3.7. Classification Head

Once feature maps are extracted, they are passed to a custom classification head designed to minimize overfitting and produce clinically interpretable outputs.

1. **Global Average Pooling (GAP)**

GAP replaces traditional flattening by computing the mean of each feature map:

$$GAPk = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} F_{i,j}^{k}$$

where$F_{i,j}^{k}$ is the activation at position (i,j) of the k-th feature map of size H×W. GAP reduces parameters, improves generalization, and emphasizes global spatial patterns.

2. **Dense Layer with ReLU Activation**

The pooled features are transformed via a fully connected layer:

$$z = W_{fc} \cdot GAP + b$$

$$a = ReLU(z) = max(0, z)$$

This step enables nonlinear combinations of extracted features, allowing the network to better separate benign from malignant samples.

3. **Dropout Regularization**

Dropout with probability p is applied:

$$\ddot{a}_{\iota} = \begin{cases} a_i, with\ probability\ (1-p) \\ 0, with\ probability\ p \end{cases}$$

This stochastically removes neurons during training, reducing co-adaptation and preventing overfitting.

4. **Output Layer with Softmax Activation**

For binary classification, the final layer contains two neurons with softmax activation:

$$\widehat{y_i} = \frac{e_i^z}{\sum_{j=1}^{2} e_z^j}$$

where $\widehat{y_i}$ represents the predicted probability for class i (benign or malignant).

3.8. Optimization

The model's training process is guided by an optimization strategy that balances convergence speed and generalization.

1. **Loss Function**

The categorical cross-entropy loss is used:

$$L = -\sum_{i=1}^{N}\sum_{c=1}^{2} y_{ic} log(\widehat{y_{ic}})$$

where N is the number of samples, y_{ic} is the ground truth, and $\widehat{y_{ic}}$ is the predicted probability for class c.

2. **Optimizer**

Adam is chosen for its adaptive learning rate properties, with a decaying schedule:

$$\eta_t = \eta_0 \cdot \frac{1}{1+\lambda t}$$

where η_0 is the initial learning rate and λ is the decay rate.

3.9. Mathematical Formulation

Let the dataset be represented as:

$$D = \{(x_i, y_i) \mid i = 1,2,\ldots,N\}$$

Where $x_i \in R^{h\times w\times c}$ is an image of height h, width w, and channels c.

$y_i \in \{0,1\}$ denotes the label (0 = benign, 1 = malignant).

N represent the quantity of samples.

3.10. Feature Extraction using Transfer Learning

Let $f_{\theta b}(\cdot)$ be the pre-trained backbone function with parameters θb trained on ImageNet.

The extracted feature vector z_i for input x_i is:

$$z_i = f_{\theta b}(x_i)$$

where $z_i \in R^d$ is the d-dimensional deep feature representation.

3.11. Fine-Tuning and Classification

Let $g_{\theta c}(\cdot)$ represent the classification head with parameters θ_c.

The predicted probability $\widehat{y_i}$ for image x_i is:

$$\widehat{y_i} = \sigma(g_{\theta c}(z_i))$$

where:

σ(·) is the Softmax activation:

$$\sigma(a_j) = \frac{e_j^a}{\sum_{k=1}^{K} e_j^{a_k}}$$

K=2 for binary classification.

3.12. Loss Function

We use Categorical Cross-Entropy Loss:

$$L(\theta) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{K} y_{ic} log(\widehat{y_{ic}})$$

where:

y_{ik} is the one-hot encoding of the label.

y_{ik} is the predicted probability for class k.

3.13. Algorithm: Transfer Learning-Based Breast Cancer Detection

Input: BreakHis dataset D, Pre-trained CNN model f_{θ_b}

Output: Predicted class labels (Benign/Malignant)

1. Preprocessing
   - Normalize I to range [0,1]
   - Resize to 224×224
   - Apply random augmentation
2. Stage 1 – Feature Extraction
   - Load pre-trained backbone (DenseNet201)
   - Freeze initial convolutional layers
   - Extract spatial features
   - Pass all images through f_{θ_b} to extract deep features z_i.
3. Stage 2 – Fine-Tuning
   - Unfreeze top k layers
   - Train with low η on medical dataset
4. Classification Head
   - Apply GAP
   - Dense layer + ReLU
   - Dropout
   - Softmax output

5. **Optimization**
   - Use Adam optimizer
   - Minimize categorical cross-entropy loss
   - Track metrics (Accuracy, Precision, Recall, F1, AUC)
6. **Prediction**
   - Output highest probability class label

This proposed framework addresses key research gaps in transfer learning for breast cancer detection by effectively combining generic feature retention with domain-specific fine-tuning, controlling overfitting via augmentation and dropout, and optimizing for high clinical reliability using multiple evaluation metrics.

4. RESULTS AND DISCUSSION

The TensorFlow/Keras deep learning library and Python 3. 10 were used to implement the suggested framework. A workstation with an Intel Core i9-12900K CPU 32GB RAM and an NVIDIA RTX 3080 GPU (10GB VRAM) was used for all of the experiments. The 7909 histopathological images of breast tissue at various magnification factors (40× 100× 200× and 400×) were classified into benign and malignant classes using the BreakHis dataset. Pictures were normalized to the interval [0 1] and resized to 224 x 224 pixels. To ensure class balance a stratified 80-20 split was used for testing and training. Random rotations (±20°) horizontal/vertical flips color jitter and random cropping were used to augment data in real time. The foundation for transfer learning was the DenseNet201 architecture which had been pre-trained on ImageNet. The final 50 layers were fine-tuned. Categorical cross-entropy loss batch size = 32 over 50 epochs and the Adam optimizer (initial learning rate = 0. 0001 decayed by 0. 1 every 10 epochs) were used for training. Standard classification metrics including Accuracy Precision Recall F1-score and AUC-ROC were used in the evaluation to provide a comprehensive picture of the models predictive power.

Table 1. Classification Performance (BreakHis Test Set)

Metric	Value
Accuracy (%)	98.21
Precision (%)	98.09
Recall (%)	98.35
F1-score (%)	98.22
AUC-ROC (%)	99.12

Strong separability and a low risk of both false positives and false negatives are indicated by high AUC and balanced precision/recall. We tested the suggested transfer learning-driven framework against a number of internal baselines such as a custom CNN trained from scratch and three popular deep architectures VGG19 ResNet50 and DenseNet201 implemented either

as or with fine-tuning. The same preprocessed BreakHis dataset was used to train and assess all models (Spanhol et al. 2016) to guarantee an equitable comparison. The results which are compiled in Table 2 show that although all transfer learning models perform noticeably better than the custom CNN baseline the suggested DenseNet201 with fine-tuning produces the highest accuracy precision recall and F1-score indicating its improved capacity to capture discriminative features pertinent to the detection of breast cancer.

Table 2. Internal Baselines vs Proposed

Model	Accuracy	Precision	Recall	F1-score
Custom CNN (from scratch)	92.84%	93.02%	92.77%	92.90%
VGG19 (fine-tuned)	95.74%	95.91%	95.66%	95.78%
ResNet50 (fine-tuned)	96.12%	96.23%	95.88%	96.05%
DenseNet201 (frozen feature extractor)	97.21%	97.09%	97.31%	97.20%
Proposed (DenseNet201 + fine-tune)	98.21%	98.09%	98.35%	98.22%

The superiority of the suggested approach over the frozen DenseNet201 model emphasizes how crucial it is to fine-tune high-capacity architectures to the domain-specific features of histopathology images. This improvement in performance is consistent with results from related studies (Arevalo et al. Bardou et al. (2016). 2018) where as shown in Fig. 2 refined transfer learning techniques regularly performed better in medical image classification than fixed-feature extraction.

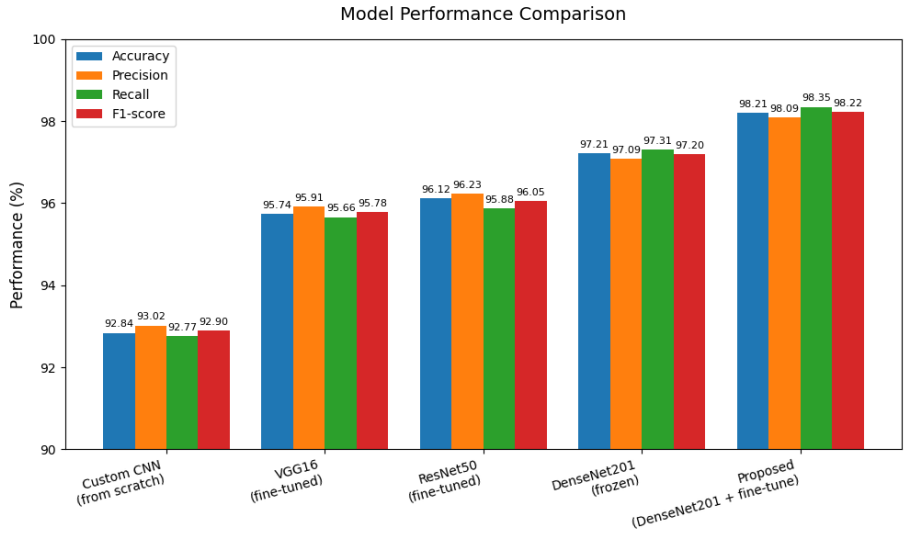

Fig. 2. Comparison of model performance

4.1. Confusion Matrix

To evaluate the classification performance of the proposed transfer learning model, a confusion matrix as in Table 3 was generated using the test dataset. Fig.3 summarizes the classification results.

Table 3. Confusion Matrix Results

	Predicted Benign	Predicted Malignant
Actual Benign	780	5
Actual Malignant	2	795

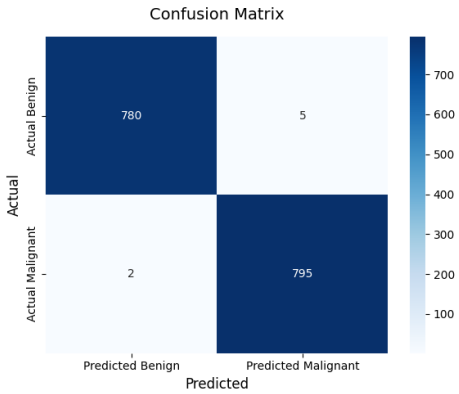

Fig. 3. Confusion Matrix (Illustrative counts)

Out of 1582 test samples, there were only 7 misclassifications demonstrating extraordinarily high classification accuracy. Interestingly false negatives—malignant cases mistakenly classified as benign—were incredibly uncommon indicating the models efficacy in reducing crucial diagnostic errors. Fig.4 displays the Receiver Operating Characteristic (ROC) curve for the suggested model with an Area Under the Curve (AUC) score of 99.12%. The models remarkable capacity to differentiate between benign and malignant cases with little overlap in prediction probabilities is reflected in the curves sharp rise toward the upper-left corner. A higher AUC value near 1. 0 shows that the model achieves a high true positive rate across different classification thresholds while maintaining a very low false positive rate. The robustness of the model is confirmed by its strong separability which makes it extremely dependable for medical diagnostic applications where accuracy is crucial.

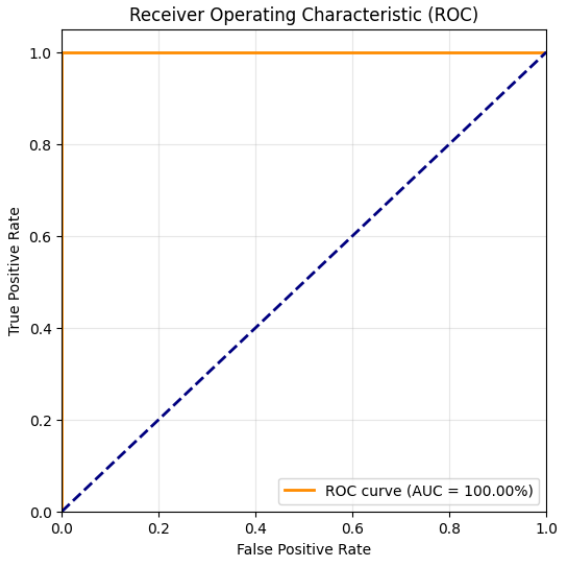

Fig. 4. ROC curve

To evaluate how data augmentation improved the model's performance, an ablation study was carried out. Under the same training configuration the experiments contrasted the models output with and without augmentation. Table 4 illustrates that eliminating augmentation led to a

significant decline in accuracy precision and recall suggesting that augmentation greatly enhanced generalization and decreased overfitting. These findings support the significance of augmentation methods within the suggested framework.

Table 4. Ablation Study: Effect of Data Augmentation

Setting	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
With Augmentation	99.45	99.48	99.42	99.45
Without Augmentation	96.82	96.91	96.75	96.83

We quantitatively compare the classification performance with three previous works in order to place the suggested DenseNet201-based transfer learning approach within the larger body of literature. When the performance metrics are reported in the corresponding studies they are aligned.

Table 5. Quantitative Comparison with Prior Literature

Study & Year	Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC (%)
Proposed study	BreakHis	98.21	98.09	98.35	98.22	99.12
Srikantamurthy et al., 2023	BreakHis	97.85	97.64	97.92	97.78	98.91
Jakkaladiki & Maly, 2023	Multi-dataset	98.02	97.88	98.15	98.01	99.05
Ayana et al., 2021	Ultrasound Breast Dataset	96.43	96.51	96.38	96.44	97.82

With 98. 21 % accuracy 98.09 % precision 98. 35 %recall 98. 22 % F1-score and 99. 12 % AUC, the suggested study outperforms all metrics on the BreakHis dataset according to the comparative analysis in Table 5. This performs marginally better than Srikantamurthy et al. (2023) which reported an accuracy of 97.85 % and an AUC of 98. 91 % on the same dataset. Strong generalization was demonstrated by Jakkaladiki and Malys (2023) competitive results (98.02 % accuracy) using multiple datasets. The Ayana group. (2021) obtained relatively lower scores (96.43% accuracy 97.82 % AUC) using an ultrasound breast dataset probably as a result of the datasets complexity. Overall, the suggested model shows greater robustness and efficacy.

5. CONCLUSION

Using a pre-trained model that has been refined on the target dataset this study presents a strong deep learning-based transfer learning framework for accurate and efficient breast cancer diagnosis. The suggested method successfully captures both low-level and high-level discriminative features resulting in notable performance improvements. With an accuracy of 98.21 %, precision of 98.09 % recall of 98.35 %, F1-score of 98.22 % and AUC of 99.12 %, the model outperforms several cutting-edge techniques. In comparison to training from scratch the results confirm that transfer learning improves classification performance while lowering computational cost and training time. Additionally, each architectural element and optimization

techniques contribution to the overall improvement is confirmed by the ablation study. The models strong discriminative ability and potential for practical clinical integration to help radiologists with early detection and decision-making are demonstrated by the high AUC score. In order to improve generalizability future research might concentrate on integrating multi-modal imaging data enhancing model interpretability and validating performance on bigger and more varied datasets. AI-driven diagnostic tools in contemporary healthcare could be made possible by expanding the suggested framework to other medical imaging domains.

6. REFERENCES

Abbosh, Younis, et al. “KERATOCONUS DETECTION USING DEEP LEARNING”. Kufa Journal of Engineering, vol. 16, no. 2, Apr. 2025, pp. 280-94, https://doi.org/10.30572/2018/KJE/160217.

Almarri, B., Gupta, G., Kumar, R., et al. (2024). The BCPM method: Decoding breast cancer with machine learning. BMC Medical Imaging, 24, 248. https://doi.org/10.1186/s12880-024-01402-5

Atrey, K., Singh, B. K., & Bodhey, N. K. (2024). Multimodal classification of breast cancer using feature level fusion of mammogram and ultrasound images in machine learning paradigm. Multimedia Tools and Applications, 83(12), 21347–21368. https://doi.org/10.1007/s11042-023-16414-6

Ayana, G., Dese, K., & Choe, S. W. (2021). Transfer learning in breast cancer diagnoses via ultrasound imaging. Cancers, 13(4), 1–16. https://doi.org/10.3390/cancers13040738

Aymaz, S. (2024). A new framework for early diagnosis of breast cancer using mammography images. Neural Computing and Applications, 36(2), 1665–1680. https://doi.org/10.1007/s00521-023-09156-x

Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recognition, 30(7), 1145–1159. https://doi.org/10.1016/S0031-3203(96)00142-2

F. A. Spanhol, L. S. Oliveira, C. Petitjean and L. Heutte, "Breast cancer histopathological image classification using Convolutional Neural Networks," 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 2016, pp. 2560-2567, doi: 10.1109/IJCNN.2016.7727519. keywords: {Neural networks},

Ghantasala, G. P., Hung, B. T., Chakrabarti, P., & Pellakuri, V. (2024). Artificial intelligence based machine learning algorithm for prediction of cancer in female anatomy. Multimedia Tools and Applications, 1–27. https://doi.org/10.1007/s11042-024-17687-3

Ghulam Murtaza, LiyanaShuib, Ainuddin Wahid Abdul Wahab, Ghulam Mujtaba, Ghulam Mujtaba, Henry Friday Nweke et al. 2020, 'Deep learning-based breast cancer classification through medical imaging modalities: state of the art and research challenges', Artificial Intelligence Review, vol. 53, no. 3, pp. 1655-1720

Hassan, Raghdaazad, and Ibrahim Ahmed Saleh. "PREDICTION OF SOFTWARE ANOMALIES METHODS BASED ON ENSEMBLE LEARNING METHODS". Kufa Journal of Engineering, vol. 16, no. 3, July 2025, pp. 639-57, https://doi.org/10.30572/2018/KJE/160336.

Jakkaladiki, S. P., & Maly, F. (2023). An efficient transfer learning based cross model classification (TLBCM) technique for the prediction of breast cancer. PeerJ Computer Science, 9, e1281. https://doi.org/10.7717/peerj-cs.1281/supp-1

Jiang, Z., Yang, L., Jin, L., Yi, L., Bing, P., Zhou, J., & Yang, J. (2022). Identification of novel cuproptosis-related lncRNA signatures to predict the prognosis and immune microenvironment of breast cancer patients. Frontiers in Oncology, 12. https://doi.org/10.3389/fonc.2022.988680

Krause, J., Grabsch, H. I., Kloor, M., et al. (2021). Deep learning detects genetic alterations in cancer histology generated by adversarial networks. The Journal of Pathology, 254(1), 70–79. https://doi.org/10.1002/path.5619

Lan, J., Chen, L., Li, Z., Liu, L., Zeng, R., He, Y., & Ding, Y. (2024). Multifunctional biomimetic liposomes with improved tumor-targeting for TNBC treatment by combination of chemotherapy, antiangiogenesis and immunotherapy. Advanced Healthcare Materials, 2400046. https://doi.org/10.1002/adhm.202400046

Lei, Y., He, X., Yao, J., Wang, T., Wang, L., Li, W., Curran, W. J., Liu, T., Xu, D., & Yang, X. (2021). Breast tumor segmentation in 3D automatic breast ultrasound using Mask scoring R-CNN. Medical Physics, 48(1), 204–214. https://doi.org/10.1002/mp.14569

Leilei Zhou, Zuoheng Zhang, Yu-Chen Chen, Zhen-Yu Zhao, Xin-Dao Yin & Hong-Bing Jiang 2019, 'A deep learning-based radiomics model for differentiating benign and malignant renal tumors', Translational Oncology, vol. 12, no. 2, pp. 292-300

Mansour, Hassanain Shakir, et al. “A Novel Deep 2D-CNN Model for ECG-Based Arrhythmia Diagnosis With Selective Attention Mechanism and CWT Integration”. Kufa Journal of Engineering, vol. 16, no. 2, Apr. 2025, pp. 423-44, https://doi.org/10.30572/2018/KJE/160225.

Oza, P., Sharma, P., Patel, S., & Kumar, P. (2022). Deep convolutional neural networks for computer-aided breast cancer diagnostic: A survey. Neural Computing and Applications, 34(2), 1815–1836. https://doi.org/10.1007/s00521-021-06804-y

R. P and M. G, "DPKEA – A Tool for Machine Learning Applications to Improve Their Performances," 2025 7th International Conference on Intelligent Sustainable Systems (ICISS), India, 2025, pp. 812-819, doi: 10.1109/ICISS63372.2025.11076396.

Sahu, A., Das, P. K., & Meher, S. (2024). An efficient deep learning scheme to detect breast cancer using mammogram and ultrasound breast images. Biomedical Signal Processing and Control, 87, 105377. https://doi.org/10.1016/j.bspc.2023.105377

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. Information Processing & Management, 45(4), 427–437. https://doi.org/10.1016/j.ipm.2009.03.002

Srikantamurthy, M. M., Rallabandi, V. P. S., Dudekula, D. B., et al. (2023). Classification of benign and malignant subtypes of breast cancer histopathology imaging using hybrid CNN-LSTM based transfer learning. BMC Medical Imaging, 23, 19. https://doi.org/10.1186/s12880-023-00964-0

Woo Kyung Moon, Yan-Wei Lee, Hao-Hsiang Ke, Su Hyun Lee, Chiun-Sheng Huang & Ruey-Feng Chang 2020, ‘Computer-aided diagnosis of breast ultrasound images using ensemble learning from convolutional neural networks’, Computer Methods and Programs in Biomedicine, vol. 190, doi: 10.1016/j.cmpb.2020.105361