RESEARCH PAPER

An Innovative Deep Learning Framework Leveraging Transfer Learning for Breast Cancer Classification

Laxmi Yadav[1], Divakar Yadav [2*], Girish Chandra [3]

[1]*2200520107102@ietlucknow.ac.in*, [2]*divakaryadav@ignou.ac.in* [3]girish.chandra@ietlucknow.ac.in

[1, 3]Department of Computer Science & Engineering, IET Lucknow (UP), India

[2]School of Computer and Information Science, Indira Gandhi National Open University (IGNOU), New Delhi, India

ORCID: [2]0000-0001-6051-479X

Abstract

One of the most prevalent medical conditions along with one of the main causes of fatalities remains breast cancer among women worldwide. Early detection and accurate classification are vital to improving results of treatment and rates of patient survival. In this study, we propose a new deep learning framework that employs an attention-based fusion of multiple convolutional neural network (CNN) architectures, specifically DenseNet121, InceptionV3, and EfficientNetB7, for the diagnosis and categorization of breast cancer in histopathological biopsy images. Experimental results in benchmark breast cancer datasets that show how well our suggested model performs, attaining accuracy of 99. 83%, precision is 99.73%, recall of 99.7 3 %, and an F1 score is 99.7 3 %. The proposed model outperforms the individual models with an accuracy of DenseNet121 (98. 76%), InceptionV3 (98. 47%) and EfficientNetB7 (98. 82%) highlighting the effectiveness of our fusion strategy.

Keywords: Deep Learning, Medical Image Analysis, Breast Cancer, Transfer Learning

How to cite this article: Yadav L, Yadav D, Chandra G. An Innovative Deep Learning Framework Leveraging Transfer Learning for Breast Cancer Classification. Int J Drug Deliv Technol. 2026;16(65s):141-150. DOI: 10.25258/ijddt.16.65s.17

1 INTRODUCTION

Breast cancer (BC) is the most prevalent disease reported in women and accounts for the majority of fatalities from cancer worldwide [1]. According to the WHO statistics report [2], approximately 2.3 million women were diagnosed with BC and 6,70,000 deaths worldwide in 2022. Early detection of BC can improve survival rates, which prompts the development of innovative methods for quick diagnosis. Furthermore, continuous improvements in treatment methods have resulted in a fall in the death rate [3]. Mammography and biopsy are two frequently used diagnostic techniques for BC detection. Mammography is a low-dose X-ray technique used for screening and is the initial process for detecting suspicious areas in the breast. And biopsy is a highly accurate and successful way to diagnose BC. To diagnose and classify the tumor, a pathologist will take a tissue sample from the suspected breast area and look at it under a microscope [4]. Malignant or cancerous cells are those that start to divide and grow unconditionally. Manual microscopic image analysis is very challenging and complex because normal and malignant cells have different appearances [5].

To ensure accurate diagnoses, Methods that improve the consistency and quality of diagnostic results are essential. Complex tasks that are difficult for traditional machine learning approaches to handle include preprocessing, segmentation, and feature extraction. These tasks can have a detrimental effect on the accuracy and efficiency of the system. In order to overcome these drawbacks, Deep Learning (DL) has become a potent substitute that can automatically extract significant features straight from unprocessed photos and use them efficiently for classification tasks [6]. These studies show the way TL may greatly improve model performance across a range of medical specialties and overcome the difficulties presented by paucity datasets. Additional developments by [6] reveal AI's ability to surpass conventional detection techniques and, in certain situations, radiologists, highlighting the innovative potential of DL in medical diagnostics, particularly in breast pathology and illness categorization.

CNNs, a type of DL, have made great strides in biological image processing in recent years. Numerous tasks, such as mitosis detection in microscopy pictures, tumor recognition, skin disease classification, immune cell identification and classification, and mass quantification in mammograms, have been effectively handled by CNN-based techniques. Even when CNN models perform well on larger datasets, they occasionally

fail to yield noticeable improvements on smaller datasets. TL has emerged as a practical method to get around this restriction, enhance recognition precision, and reduce computational expenses. CNN performance can be improved by applying pre-trained models and incorporating knowledge from different architectures [7].

The following summarizes this work's main contributions:

- Development of a DL-based method for the effective diagnosis and categorization of BC.
- Transfer learning (TL) is implemented and assessed using three different CNN architectures.
- A comprehensive comparative analysis of each architecture's performance as classification accuracy within the context of transfer learning.

The remainder of the document is structured as follows: Section two includes the literature review and A thorough study of the suggested strategy is provided in Section 3, which includes subsections on TL, pre-trained CNN architectures, and data preprocessing and augmentation. Section 4 focuses on the Dataset and Evaluation Parameters. Section 5 presents Experimental findings and insights and finally, Section 6 contains the Conclusion and future directions.

2 LITERATURE REVIEW

This literature review's objective is to provide a thorough overview of the most recent developments in DL methods for recognizing and classifying BC, emphasizing important research, recent developments, and new trends in this field. DL has transformed MIA by facilitating the development of computer-aided diagnosis (CAD) systems that help medical professionals in recognizing and classifying a range of illnesses [8]. An analysis technique for haematoxylin and eosin (H&E) stained breast biopsy images was presented in the work for the classification of BC histology images using CNNs. CNNs' effectiveness in distinguishing between benign and malignant histological features was confirmed by the method's accuracy of 83.3% in carcinoma/non-carcinoma classification and 77.8% in four-class classification [9]. In this context, pre-trained deep CNN models are often used to extract features from large-scale, generic image datasets, and the extracted features are subsequently applied to domain-specific tasks using smaller datasets [10]. Context-based learning has advanced TL and demonstrated encouraging results in tasks involving the identification and classification of BC by training CNNs in two stages using both single and overlapping picture patches. With 99.02% accuracy on IDC and 99.12% accuracy on BreaKHis datasets, this study [11] proposes a magnification-independent AI system for BC classification utilizing TL.

By combining several CNN designs, TL efficacy is greatly enhanced, possibly outperforming the drawbacks of conventional single-model CNN techniques, a quick and precise framework for cell-based picture classification has been produced by combining InceptionV2, ResNet50, and InceptionV3, which were pre-trained on the ImageNet dataset [12]. The ResNet architecture is particularly attractive due to its depth and skip connections, which enable the capture of intricate patterns and enhanced feature extraction [13]. This study [14] BreNet, an attention-enhanced multi-scale CNN, classifies BC in histopathological images, achieving high accuracies of 99.96% on BreakHis and 88.26% on the IDC dataset, demonstrating strong generalization potential.

CNNs and TL have become increasingly popular for medical imaging applications such as breast cancer identification, as they leverage pre-trained models on large datasets like ImageNet to initialize weights for specific tasks with limited data [13]. The suggested model uses TL to address the noted shortcomings of current techniques for the identification and categorization of malignant tumors. The system improves accuracy and efficiency by using pre-trained DL models, especially when applied to small medical imaging datasets [10, 15, 13, 14].

3 PROPOSED METHODOLOGY

This section introduces the suggested methodology for identifying and categorizing cancerous cells in breast images using CNN architectures. In order to separately extract low-level features from the input photos, the framework makes use of three well-known pre-trained CNN models: DenseNet121, InceptionV3 and EfficientNetB7 [15, 16, 17]. The last step in the classification process involves concatenating these extracted features and passing them through a fully linked layer. The proposed method's general structure is depicted in the basic diagram in Figure 1. In the next subsections, each element of the architecture is

explained in detail.

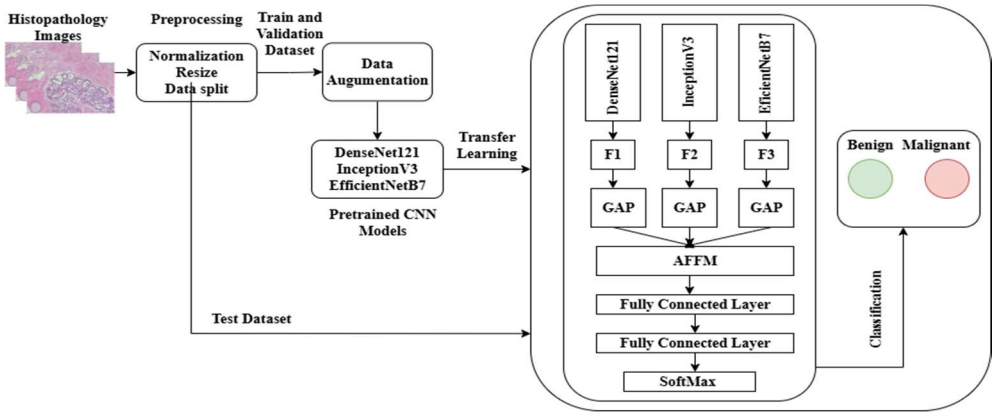

Figure 1: Architecture of the Proposed Deep Learning Attention based feature fusion Model.

3.1 Pre-Processing

It plays a vital role in improving the quality and consistency of microscopic tissue images by eliminating various types of noise and artifacts. In the proposed research, we apply a normalization technique specifically designed for H&E-stained images, as outlined in [12]. Input image size is 128x128 and 70:30 split protocol used here.70% data is used for the training, 15% data used for validation and 15% data used for testing. Without generalization, CNNs may perform exceptionally well on training data but fail to deliver accurate predictions on unseen test data.

3.2 Augmentation

To address the limitations of small datasets and reduce overfitting, data augmentation methods are applied to artificially expand the training data through techniques such as rotation, scaling, flipping, and color adjustments. The dataset was highly imbalanced, with 2,480 non- cancerous and 5,429 cancerous images, totaling 7,909 samples. This imbalance could lead to biased model training, favoring the malignant class. The augmentation process not only mitigated the class imbalance but also enriched the data. The dataset after augmentation is shown in Table 1.

3.3 CNN architecture with pre-trained features for feature extraction

Features are extracted using CNN models, followed by their combination and classification using a fully connected layer. This combined feature set could include a number of descriptors, including characteristics relating to shapes, like compactness, roundness, and circularity. In the next subsections, each chosen CNN architecture's structural overview is covered in detail.

3.3.1 DenseNet121

DenseNet121 is a DL architecture that connects every layer to every subsequent layer in a feed-forward fashion [15]. It promotes feature reuse rather than learning redundant feature maps, which increases efficiency and lowers the number of parameters. The model is composed of dense blocks separated by transition layers that include batch normalization and pooling. DenseNet121 has 121 layers and is known for its compactness and high performance on image classification tasks.

3.3.2 InceptionV3

InceptionV3 is an advanced version of the Inception architecture that captures features at various scales by using a single module with several convolutional filter sizes[16]. It introduces factorized convolutions (e.g., 1×7 followed by 7×1) and auxiliary classifiers to lower the level of computing complexity while preserving accuracy. The model is 48 layers deep and includes techniques such as batch normalization and label smoothing to improve generalization.

3.3.3 EfficientNetB7

EfficientNetB7 is the most extensive and most accurate variant in the EfficientNet family, it employs a compound scaling strategy that simultaneously and uniformly adjusts its width, depth, and input resolution [17]. In order to

maximize feature extraction and model efficiency, EfficientNetB7 uses squeeze-and-excitation modules, swish activation, and MBConv blocks.

3.4 Transfer Learning

A big dataset is usually needed to train a CNN from scratch however, acquiring extensive and relevant data is often challenging in many real-world applications. Unlike ideal scenarios, obtaining well-matched training and testing datasets can be difficult or impractical. To deal with this problem, the concept of TL has been introduced. It is a widely adopted DL strategy that uses acquired knowledge from solving one problem and applies it to related tasks. A base network is initially trained on a large, relevant dataset for a specific task. Subsequently, the learned parameters are transferred to a target task where the model is fine-tuned using a smaller target dataset [18].

3.5 Proposed Model

The proposed model leverages the strengths of three robust CNNs, DenseNet121, InceptionV3, and EfficientNetB7, to perform breast cancer classification. Each of these pre-trained networks receives input images after they have been resized and normalized and acts as a feature extractor, independently processing the input histopathology images to extract high-level feature representations. To efficiently combine the extracted features, an AFFM is used. In order to ensure that the most informative features from each model contribute more significantly to the final representation, this mechanism adaptively weighs and integrates the feature maps based on their relevance.

We formally define the mathematical formulation of the proposed Attention-Based Feature Fusion Method (AFFM). Let $F_1 \in \mathbb{R}^{d_1}$, $F_2 \in \mathbb{R}^{d_2}$, and $F_3 \in \mathbb{R}^{d_3}$ represent the deep feature vectors that were taken from EfficientNetB7, DenseNet121, and InceptionV3, respectively. To guarantee dimensional compatibility, each feature vector is first projected using a linear transformation into a shared embedding space of dimension d:

$$\hat{F}_i = W_i F_i + b_i, i \in \{1,2,3\} \quad (1)$$

where $W_i \in \mathbb{R}^{d \times d_i}$ and b_i are learnable parameters. After that, an attention scoring function is used to each projected feature vector to ascertain its relative importance for the classification task. In particular, the attention score e_i is calculated as follows:

$$e_i = \mathbf{w}^T \sigma(\hat{F}_i) + b \quad (2)$$

where σ (·) is a non-linear activation function and w and b are shared learnable parameters.

The SoftMax function was then used to obtain the attention weights α_i are in order to guarantee normalization:

$$\alpha_i = \frac{\exp(e_i)}{\sum_{j=1}^{3} \exp(e_j)}, \sum_{i=1}^{3} \alpha_i = 1 \quad (3)$$

Finally, the fused feature representation F_{fused} is calculated as a weighted sum of the individual feature vectors:

$$F_{\text{fused}} = \sum_{i=1}^{3} \alpha_i \hat{F}_i \quad (4)$$

For the final prediction, the fused feature vector is then sent into the fully linked classification layers. In the fully connected layer, dense layer followed by Batch normalization, followed by dropout layer. There two Fully connected layers used. Final output layer is used to classify class into benign and malignant with a SoftMax.

4 DATASET AND EVALUATION PARAMETERS

This section outlines the datasets utilized to train and validate the proposed DL model, along with the key parameters and metrics used to evaluate its performance.

4.1 Dataset

The BreakHis dataset serves as a publicly accessible benchmark dataset [19]. The dataset comprises 7,909 RGB images of histopathological breast tissue, stained with Hematoxylin and Eosin (H&E), collected from 82 female patients at the Pathological Anatomy Laboratory (PAL) in Brazil. Each image, with a resolution of 700 × 460 pixels, was taken using a microscope at one of four magnification levels: 40×, 100×, 200×, and 400×. The dataset is divided into two main categories: benign and cancerous. To improve data diversity, clinically appropriate data augmentation was used, specific justification and methods are explained in subsection 3.2.

Table 1: Benign (B) and malignant (M) distribution of the dataset at different magnification levels and split (70:15:15) protocol and train dataset augmentation

[illegible]	Without	Train-Test-Validation split (70:30) Protocol	Train set

	Augmentation		Test dataset (15 %)		Val dataset (15 %)		Train dataset (70 %)		with Augmentation	
	B	M	B	M	B	M	B	M	B	M
100X	644	143 7	96	21 5	96	21 5	452	100 7	1500	1500
200X	625	137 0	93	20 5	93	20 5	439	960	1500	1500
400X	588	123 2	88	184	88	184	412	864	1500	1500
400X	623	139 0	93	166	93	166	437	1058	1500	1500
Total	2480	5429	370	770	370	770	1740	3889	6000	6000

4.2 Evaluation Parameters

A set of hyperparameters are carefully chosen through iterative experimentation and fine-tuning in order to optimize the suggested framework. The Adam optimizer was used to train the model for a maximum of 100 epochs with a batch size of 32 and a learning rate of 0.0001. ReduceLROnPlateau used validation loss to alter learning rates, and He initialization and categorical cross-entropy ensured stable training. These hyperparameters are chosen to balance the trade-off between model generalization and training efficiency.

4.3 Hardware and Implementation Details

All experiments were carried out on a workstation with an NVIDIA GPU to speed up the model training. Specially, the system included an Intel (R) Xeon (R) w5-2465X i9 CPU, 128 GB of system RAM, and an NVIDIA RTX A4000 GPU with 16 GB VRAM. Model training was conducted using TensorFlow/PyTorch in Python 3.11 and vscode-x64 1.102.2 on a 64-bit Windows operating system.

5 Experimental Findings and Insights

The experimental results of the proposed DL model for BC categorization are shown in this section. The model's efficiency is examined by analyzing the data and then its performance advantages are highlighted by comparing it with current SOTA methods.

5.1 Ablation Study:

An ablation study was carried out to examine the individual and combined contributions of the CNN architectures utilized in the suggested framework. To ensure a fair comparison, all models were trained with the same dataset split, preprocessing pipeline, hyperparameters, and evaluation protocol shown in Table 2.

Table 2: Performance of the individual and combined CNN architectures

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**
DenseNet121	98.76	99.17	97.72	98.44
Inception V3	98.47	98.52	97.80	98.16
EfficientNetB7	98.82	98.67	98.67	98.95
DenseNet121 + InceptionV3	99.21	99.07	99.36	99.22
InceptionV3 + EfficientNetB7	99.32	99.29	99.36	99.32
EfficientNetB7 + DenseNet121	99.43	99.42	99.42	99.42
Proposed Model	**99.83**	**99.73**	**99.73**	**99.73**

5.2 Results

The proposed model is trained separately on three CNN architectures such as DenseNet121, InceptionV3, and EfficientNetB7 and their learned features are subsequently combined through TL for feature extraction. The results obtained from each individual CNN model are compared with those from the combined feature set, as well as with other existing methods.

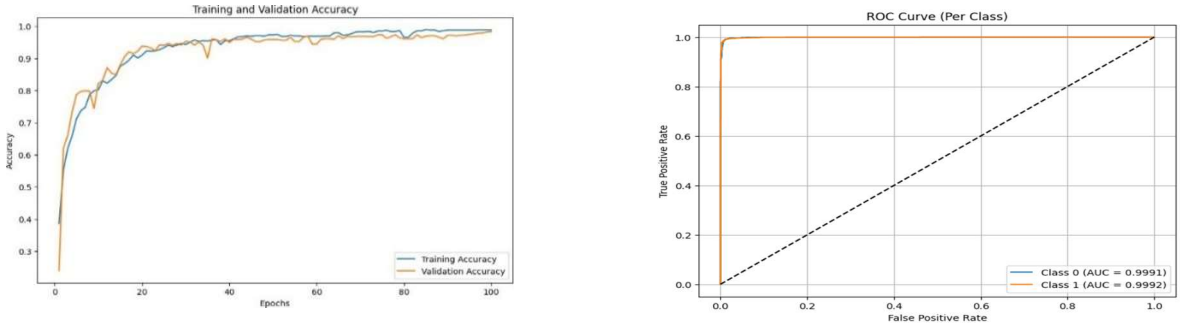

Figure 2: (a) Training and validation curve of accuracy (b) ROC curve with AUC value

Table 3: Magnification dependent and independent performance of the proposed model on BreakHis Dataset

Evaluation Protocol	Magnification	Accuracy	Precision	Recall	F1-Score
Magnification dependent	100X	99.33	99.33	99.33	99.33
	200X	99.67	99.66	99.66	99.66
	400X	99.00	99.01	99.00	98.99
	40X	99.00	99.33	98.68	99.00
Magnification Independent	Combined	99.83	99.73	99.73	99.73

The training and validation accuracy of the proposed model shown in Figure 2(a) and ROC curve for proposed model shown in Figure 2(b). Table 3 shows the performance of the proposed model on magnification dependent and independent level. The suggested method achieves over 99% in all assessed parameters, continuously outperforming the baseline models.

5.3 Performance Comparison with Existing Approaches

Comparative analysis was also conducted between the proposed framework state-of-the-art

(SOTA) methods, as summarized in Table 4. In terms of accuracy, precision, recall, and F1-score, the suggested model outperforms the SOTA approaches with different preprocessing and augmentation methods.

Table 4: Performance Comparison with Existing Approaches

Methods	**Accuracy**	**Precision**	**Recall**	**F1-score**
Nguyen et al. [10]	92.64%	91.89%	92.48%	-
Kabir et al. [11]	99.12%	99.08%	99.09%	99.08%
Mehmood et al. [12]	95.63%	94.79%	96.42%	95.57%
Kensert et al. [14]	97.1 %	97.01%	96.9%	96.9%
Sajiv et al. [20]	98.28%	97%	98%	99%
Proposed Model	99.83 %	99.73%	99.73%	99.73 %

6 CONCLUSION

In this work, three pre-trained CNN models such as InceptionV3, DenseNet121, and EfficientNetB7 are used to extract valuable features from breast histopathology images. The system's performance was evaluated against individual CNN models and compared to various existing methods. Specially, it attained the maximum accuracy of 99. 83%, precision of 99.73%, recall of 99. 73%, and an F1 score of 99. 73%, outperforming DenseNet121, InceptionV3, and EfficientNetB7. The results demonstrate how our approach increases classification efficiency by attaining more accuracy without the need for initial training.

The limitation of this work is that it focuses on attention-based feature fusion, which incurs a higher computational cost

and classifies only benign and malignant classes. In Future, a lightweight framework can be made for multiclass classification of breast cancer, and incorporating multimodal data and explainable AI techniques can enhance diagnostic accuracy and model interpretability.

References

[1] Freddie Bray, Jacques Ferlay, Isabelle Soerjomataram, Rebecca L Siegel, Lindsey A Torre, and Ahmedin Jemal. Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 68(6):394–424, 2018.

[2] World Health Organization et al. Who position paper on mammography screening. In *WHO position paper on mammography screening*. 2014.

[3] Alison Chetlen, Julie Mack, and Tiffany Chan. Breast cancer screening controversies: who, when, why, and how? *Clinical imaging*, 40(2):279–282, 2016.

[4] Andrei Chekkoury, Parmeshwar Khurd, Jie Ni, Claus Bahlmann, Ali Kamen, Amar Patel, Leo Grady, Maneesh Singh, Martin Groher, Nassir Navab, et al. Automated malignancy detection in breast histopathological images. In *Medical Imaging 2012: Computer-Aided Diagnosis*, volume 8315, pages 332–344. SPIE, 2012.

[5] Carlos López, Maryle`ne Lejeune, Ramon Bosch, Anna Korzynska, Marcial García-Rojo, Maria-Teresa Salvadó, Toma´s Á lvaro, Cristina Callau, Albert Roso, and Joaquín Jae´n. Digital image analysis in breast cancer: an example of an automated methodology and the effects of image compression. In *Perspectives on Digital Pathology*, pages 155–171. IOS Press, 2012.

[6] Yoshua Bengio, Aaron Courville, and Pascal Vincent. Representation learning: A review and new perspectives. *IEEE transactions on pattern analysis and machine intelligence*, 35(8):1798–1828, 2013.

[7] Ali Sharif Razavian, Hossein Azizpour, Josephine Sullivan, and Stefan Carlsson. Cnn features off-the- shelf: an astounding baseline for recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 806–813, 2014.

[8] Yuliana Jime´nez-Gaona, María Jose´ Rodr´ıguez-Á lvarez, and Vasudevan Lakshminarayanan. Deep- learning-based computer-aided systems for breast cancer imaging: a critical review. *Applied Sciences*, 10(22):8298, 2020.

[9] T. Araújo et al., "Classification of breast cancer histology images using convolutional neural networks," PLoS ONE, vol. 12, no. 6, e0177544, 2017.

[10] Long D Nguyen, Dongyun Lin, Zhiping Lin, and Jiuwen Cao. Deep CNNs for microscopic image classification by exploiting transfer learning and feature concatenation. In *2018 IEEE international symposium on circuits and systems (ISCAS)*, pages 1–5. IEEE, 2018.

[11] V. Kumari and R. Ghosh, "A magnification-independent method for breast cancer classification using transfer learning," Healthcare Analytics, vol. 3, p. 100207, 2023.

[12] Ruqayya Awan, Navid Alemi Koohbanani, Muhammad Shaban, Anna Lisowska, and Nasir Rajpoot. Context-aware learning using transferable features for classification of breast cancer histology images. In *Image Analy. and Recognition: 15th International Conference, ICIAR 2018, Po´voa de Varzim, Portugal, June 27–29, 2018, Proceedings 15*, pages 788–795. Springer, 2018.

[13] Alexander Kensert, Philip J Harrison, and Ola Spjuth. Transfer learning with deep convolutional neural networks for classifying cellular morphological changes. *SLAS Discovery: Advancing Life Sciences R&D*, 24(4):466–475, 2019.

[14] H. M. Alshehri, "BreNet: Attention-enhanced multi-scale CNN framework for breast cancer classification in histopathological images," IEEE Access, 2025.

[15] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017.

[16] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826, 2016.

[17] Mingxing Tan and Quoc Le. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pages 6105–6114. PMLR, 2019.

[18] Liu Yang, Steve Hanneke, and Jaime Carbonell. A theory of transfer learning with applications to active learning. *Machine learning*, 90:161–189, 2013.

[19] Fabio A Spanhol, Luiz S Oliveira, Caroline Petitjean, and Laurent Heutte. A dataset for breast cancer histopathological image classification. *Ieee transactions on biomedical engineering*, 63(7):1455–1462, 2015.

[20] G Sajiv, G Ramkumar, S Shanthi, Arunachalam Chinnathambi, and Sulaiman Ali Alharbi. Predicting breast cancer risk from histopathology images using hybrid deep learning classifier. *Medical Engineering & Physics*, page 104149, 2024.