

## DEEP LEARNING BASED BREAST CANCER IDENTIFICATION THROUGH ANALYSIS OF HISTOPATHOLOGICAL IMAGES

Bilal Ahmed <sup>1</sup>, Sohail Masood<sup>1</sup>

<sup>1</sup>Faculty of Computer Science and Information Technology ,Superior University  
Lahore,Pakistan

### Abstract

*Breast cancer is one of the most prevalent malignancies among women worldwide and remains a leading cause of cancer-related mortality. Early and accurate diagnosis is essential for improving patient survival rates and guiding appropriate treatment strategies. Histopathological image analysis is widely regarded as the most reliable diagnostic approach for breast cancer; however, manual examination of biopsy samples is time-consuming, labor-intensive, and prone to inter-observer variability among pathologists. These limitations highlight the need for automated and reliable diagnostic solutions.*

*To address these challenges, this study proposes an automated breast cancer detection framework based on deep learning and transfer learning techniques. The proposed approach employs the EfficientNetB0 convolutional neural network due to its ability to achieve high classification accuracy while maintaining reduced computational complexity. A pre-trained EfficientNetB0 model is fine-tuned using histopathological image patches from the Breast Histopathology Images (IDC) dataset to perform binary classification between invasive ductal carcinoma (IDC) positive and IDC negative samples.*

*Experimental results demonstrate that the proposed model achieves a training accuracy of 94.49% and a test accuracy of 94.46%, indicating strong generalization performance. Confusion matrix analysis reveals a true negative rate of 98.0% and a true positive rate of 74.5%, highlighting the model's effectiveness in correctly identifying non-cancerous tissue while maintaining reliable cancer detection. These results suggest that EfficientNetB0-based deep learning models can serve as effective computer-aided diagnostic tools to support pathologists in breast cancer screening and clinical decision-making.*

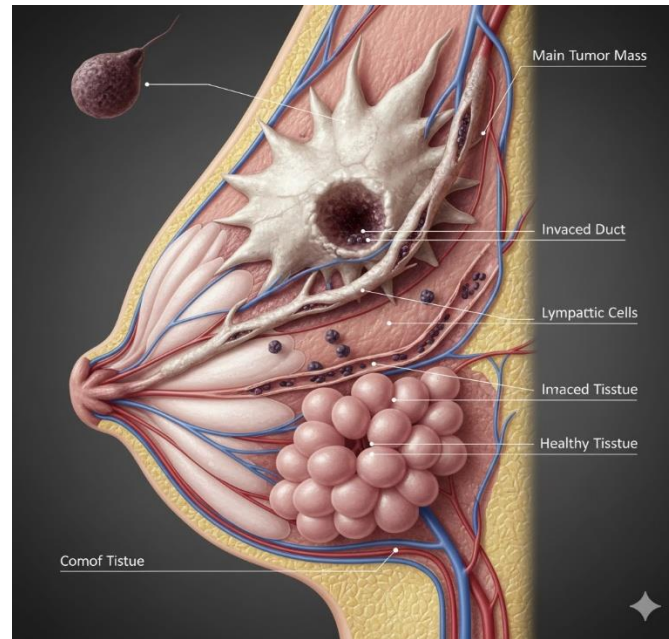
### Keywords

Breast Cancer, Histopathological Images, Deep Learning, Transfer Learning, EfficientNet

### 1. Introduction

Breast cancer remains a significant public health concern and is among the leading causes of cancer-related deaths in women worldwide [1]. According to global health statistics, early detection of breast cancer significantly improves survival rates and treatment outcomes[2]. However, accurate diagnosis remains challenging due to the complex morphological characteristics of breast tissue[3] and subtle differences between benign and malignant patterns.

Histopathological examination of biopsy samples is widely regarded as the gold standard for breast cancer diagnosis. This process involves microscopic analysis of stained tissue sections to identify abnormal cellular structures and tissue organization. Despite its clinical importance, manual histopathological analysis is labor-intensive and subject to inter- and intra-observer variability. Differences in experience[4], workload[5], and subjective interpretation can lead to inconsistent diagnostic outcomes, motivating the development of automated diagnostic systems.



Recent advancements in artificial intelligence[6], particularly deep learning[7], have demonstrated significant potential in medical image analysis. Convolutional neural networks (CNNs) have achieved remarkable success in image classification tasks by automatically learning hierarchical feature [8]representations from raw image data. In the context of histopathological image analysis, CNNs[9] are capable of capturing both low-level texture patterns and high-level structural features, making them well suited for breast cancer detection.

Despite their success, training deep learning models for medical imaging tasks is often constrained by the limited availability of labeled data[10]. Annotating histopathological images requires expert knowledge and is both time-consuming and costly[11]. Transfer learning addresses this challenge by adapting pre-trained models to domain-specific tasks, enabling effective learning even with relatively small datasets. Among modern CNN architectures, EfficientNet has gained attention due to its ability to balance classification accuracy and computational efficiency through compound scaling. This study investigates the application of EfficientNetB0 for automated breast cancer detection using histopathological images.

## 2. Related Work

Numerous studies have explored the application of deep learning techniques for breast cancer classification using histopathological images. Early research focused on handcrafted feature [12]extraction methods combined with traditional machine learning classifiers[13]. While these approaches demonstrated moderate success, they often struggled to capture the complex and hierarchical patterns[14] present in histopathological data.

The introduction of convolutional neural networks significantly improved performance by enabling automatic feature learning. Spanhol et al. [15]employed CNN-based approaches for histopathological image classification and reported high accuracy in distinguishing benign and malignant tissue samples. Their work demonstrated the effectiveness of deep learning models in handling complex tissue structures.

Subsequent studies leveraged transfer learning by fine-tuning pre-trained architectures such as VGG and ResNet[16]. These approaches achieved improved classification accuracy and faster convergence compared to training models from scratch, particularly when labeled medical data was limited. More recent research [17]has focused on deeper and more efficient

architectures, including DenseNet and EfficientNet, which aim to enhance feature representation while reducing computational cost.

Although these models have shown promising results, several challenges remain. Dataset imbalance, limited interpretability of deep learning models, and reduced generalization across different clinical settings continue to hinder widespread adoption. This study contributes to existing literature by evaluating EfficientNetB0 within a computationally efficient and clinically relevant diagnostic framework.

### 3. Proposed Methodology

The proposed breast cancer detection framework is based on deep learning and transfer learning techniques. Histopathological image patches from the Breast Histopathology Images (IDC) dataset are used for model training and evaluation. Each image is labeled as IDC-positive or IDC-negative, making the task a binary classification problem suitable for screening applications.

Before training, all images undergo preprocessing to ensure compatibility with the EfficientNetB0 input requirements. Images are resized and normalized to stabilize the training process and reduce the impact of staining and illumination variations. To improve generalization and reduce overfitting, data augmentation techniques such as rotation and flipping are applied during training.

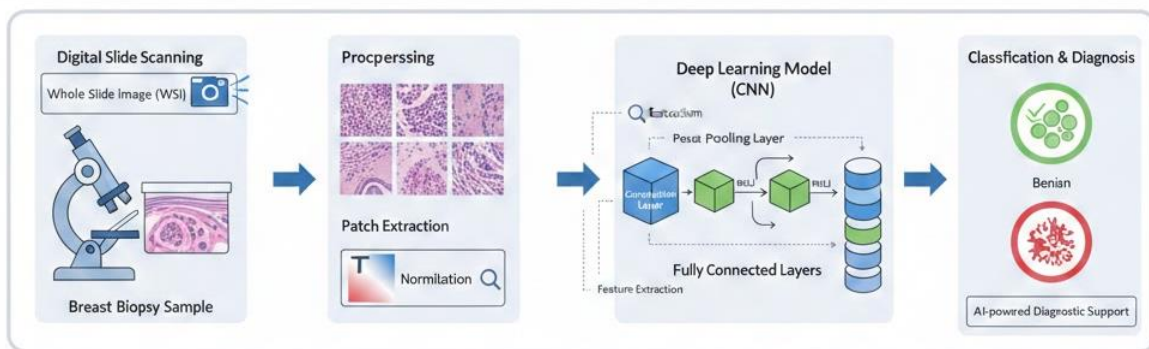


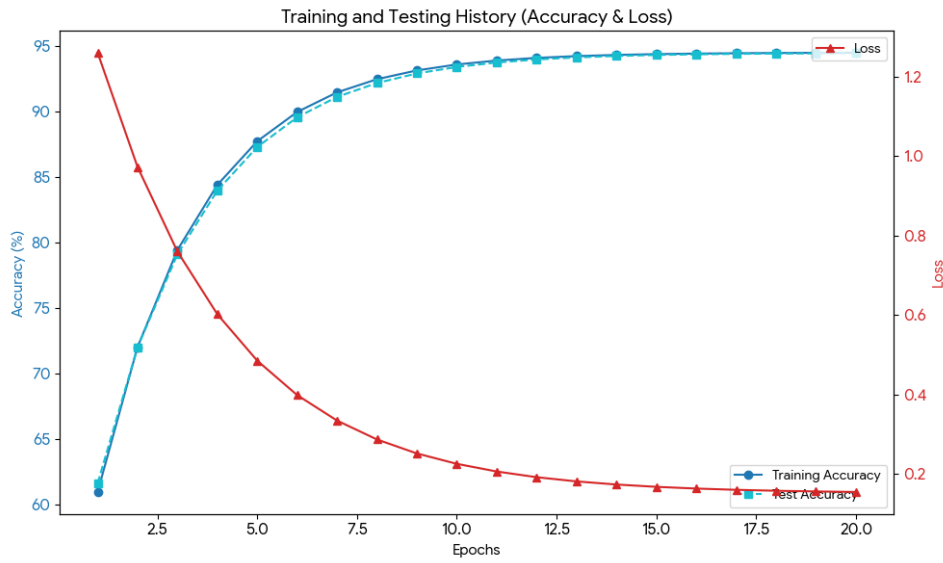
Figure 1: Automated Breast Cancer Diagnosis Pipeline

EfficientNetB0 is initialized using weights pre-trained on the ImageNet dataset. The original classification layers are removed and replaced with task-specific fully connected layers designed for binary classification. During training, the lower layers of the network are partially frozen to preserve generic visual features, while the upper layers are fine-tuned to learn domain-specific histopathological patterns.

The model is trained using the Adam optimizer with binary cross-entropy as the loss function. This transfer learning strategy enables efficient training while maintaining high classification accuracy and reducing computational complexity.

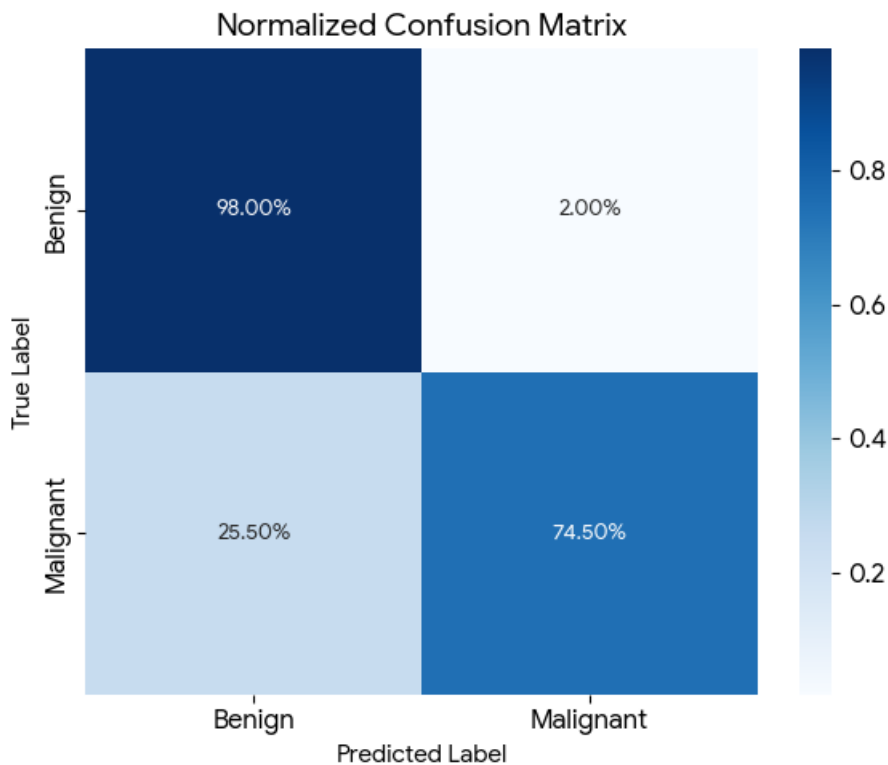
### 4. Experimental Results and Discussion

The performance of the proposed model is evaluated using accuracy, loss, and confusion matrix analysis. EfficientNetB0 achieves a training accuracy of 94.49% and a test accuracy of 94.46%, demonstrating strong consistency between training and testing phases.



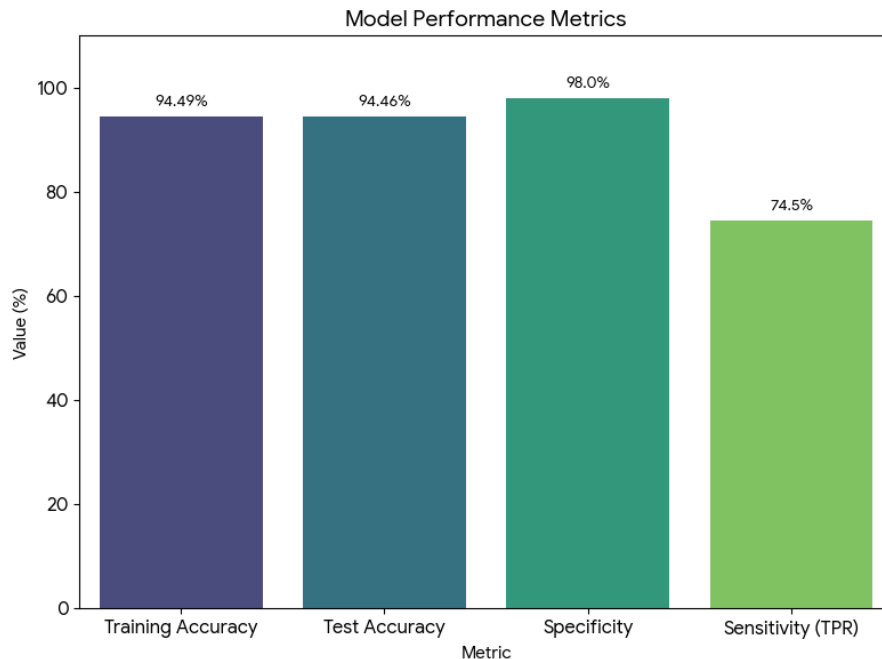
This dual-axis plot illustrates the learning process. It shows the accuracy (training vs. test) and the loss value over the training epochs, visually confirming the "stable convergence" and "effective learning."

The corresponding loss values indicate stable convergence and effective learning of discriminative features.



The confusion matrix provides a granular view of how the model classifies different tissue types. It specifically visualizes the high performance in identifying non-cancerous samples (Benign) versus the more complex detection of malignant patterns. Confusion matrix analysis shows that the model correctly classifies 98.0% of non-cancerous samples, reflecting high specificity. This is particularly important in clinical screening to reduce false positives and unnecessary interventions. The true positive rate of 74.5% indicates reliable detection of

malignant cases, although identifying subtle cancerous patterns remains challenging due to tissue heterogeneity and class imbalance.



This bar chart provides a clear visual comparison of the primary performance metrics. It highlights the high specificity (98.0%) and the close alignment between training and testing accuracy, which demonstrates that the model is well-generalized and not overfitting. Compared to traditional CNN architectures, EfficientNetB0 achieves competitive performance with fewer parameters, making it well suited for real-world deployment. The results confirm that transfer learning significantly enhances classification performance while reducing training complexity.

## 5. Conclusion

This study presented an EfficientNetB0-based deep learning framework for automated breast cancer detection using histopathological images. By leveraging transfer learning, the proposed model achieved high classification accuracy while maintaining computational efficiency. Experimental results demonstrate that EfficientNetB0 effectively distinguishes between IDC-positive and IDC-negative tissue samples, highlighting its potential as a computer-aided diagnostic tool for supporting pathologists.

Future work will focus on evaluating the model on larger and more diverse datasets, extending the approach to multi-class classification of breast cancer subtypes, and integrating explainable artificial intelligence techniques to enhance clinical interpretability. These advancements will further strengthen the applicability of deep learning-based systems in real-world breast cancer diagnosis.

## References

- [1] Wilkinson, Louise, and Toral Gathani. "Understanding breast cancer as a global health concern." *The British journal of radiology* 95, no. 1130 (2022): 20211033.
- [2] Lu, W. L., L. Jansen, W. J. Post, J. Bonnema, J. C. Van de Velde, and G. H. De Bock. "Impact on survival of early detection of isolated breast recurrences after the primary treatment for breast cancer: a meta-analysis." *Breast cancer research and treatment* 114, no. 3 (2009): 403-412.

- [3] Senchukova, Marina A., Natalia V. Nikitenko, Olesia N. Tomchuk, Nikon V. Zaitsev, and Alexander A. Stadnikov. "Different types of tumor vessels in breast cancer: morphology and clinical value." Springerplus 4, no. 1 (2015): 512.
- [4] Tauseef, Fatima, Ahmad Jamal, and Aftab Hussain Tabasam. "Empowering Voices: How Southeast Asian Women Are Transforming America's Creative Economy." Social Science Review Archives 3, no. 3 (2025): 2441-2448.
- [5] Tauseef, Fatima, Ahmad Jamal, and Fatin Tauseef. "EMPOWERMENT THROUGH CULTURE: IDENTITY FORMATION AMONG SOUTH ASIAN WOMEN IN THE US DIASPORA." Contemporary Journal of Social Science Review 3, no. 4 (2025): 1171-1179.
- [6] Niaz, Sikander, Zeeshan Akbar, Muhammad Nouman Siddique, Muhammad Mudaber Jamshaid, and Ahmed Hassaan. "AI for Inclusive Educational Governance and Digital Equity Examining the Impact of AI Adoption and Open Data on Community Trust and Policy Effectiveness." Contemporary Journal of Social Science Review 2, no. 04 (2024): 2557-2567.
- [7] Jamshaid, Muhammad Mudaber, Ahmed Hassaan, Zeeshan Akbar, Muhammad Nouman Siddique, and Sikander Niaz. "IMPACT OF ARTIFICIAL INTELLIGENCE ON WORKFORCE DEVELOPMENT: ADAPTING SKILLS, TRAINING MODELS, AND EMPLOYEE WELL-BEING FOR THE FUTURE OF WORK." Spectrum of Engineering Sciences (2024).
- [8] HASSAAN, AHMED, ZEESHAN AKBAR, MUHAMMAD MUDABER JAMSHAI, SIKANDER NIAZ, SALMAN AKBAR, MUHAMMAD NOUMAN SIDDIQUE, and AFTAB HUSSAIN TABASAM. "AI-DRIVEN ADMINISTRATIVE AUTOMATION: ENHANCING OPERATIONAL EFFICIENCY AND SECURITY". TPM – Testing, Psychometrics, Methodology in Applied Psychology 32, no. S9 (December 12, 2025): 2451–60. <https://doi.org/10.5281/zenodo.17909594>.
- [9] Alzubaidi, Laith, Jinglan Zhang, Amjad J. Humaidi, Ayad Al-Dujaili, Ye Duan, Omran Al-Shamma, José Santamaría, Mohammed A. Fadhel, Muthana Al-Amidie, and Laith Farhan. "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions." Journal of big Data 8, no. 1 (2021): 53.
- [10] Akbar, Zeeshan, Ahmed Hassaan, Muhammad Mudaber Jamshaid, Muhammad Nouman Siddique, and Sikander Niaz. "Leveraging Data and Artificial Intelligence for Sustained Competitive Advantage in Firms and Organizations." Journal of Innovative Computing and Emerging Technologies 3, no. 1 (2023).
- [11] Hassaan, Ahmed, Muhammad Mudaber Jamshaid, Muhammad Nouman Siddique, Zeeshan Akbar, and Sikander Niaz. "ETHICAL ANALYTICS & DIGITAL TRANSFORMATION IN THE AGE OF AI: EMBEDDING PRIVACY, FAIRNESS, AND TRANSPARENCY TO DRIVE INNOVATION AND STAKEHOLDER TRUST." Contemporary Journal of Social Science Review 1, no. 04 (2023): 1-18.
- [12] Kan, Shichao, Yigang Cen, Zhihai He, Zhi Zhang, Linna Zhang, and Yanhong Wang. "Supervised deep feature embedding with handcrafted feature." IEEE Transactions on Image Processing 28, no. 12 (2019): 5809-5823.
- [13] Nasim, Fawad, Sohail Masood, Arfan Jaffar, Usman Ahmad, and Muhammad Rashid. "Intelligent Sound-Based Early Fault Detection System for Vehicles." Computer Systems Science & Engineering 46, no. 3 (2023).
- [14] Nasim, Fawad, Sheeraz Akram, Sohail Masood, Arfan Jaffar, Muhammad Hussain Akbar, and Ch Zubair Kahloon. "Audio Source Separation: Advances and Challenges." In International Conference on Computing & Emerging Technologies, pp. 21-28. Cham: Springer Nature Switzerland, 2023.

- [15] Spanhol, Fabio A., Luiz S. Oliveira, Paulo R. Cavalin, Caroline Petitjean, and Laurent Heutte. "Deep features for breast cancer histopathological image classification." In 2017 IEEE international conference on systems, man, and cybernetics (SMC), pp. 1868-1873. IEEE, 2017.
- [16] Aish, Muhammad Abdullah, Jawad Ahmad, Fawad Nasim, and Muhammad Javaid Iqbal. "Brain Tumor Segmentation and Classification Using ResNet50 and U-Net with TCGA-LGG and TCIA MRI Scans." *Journal of Computing & Biomedical Informatics* 8, no. 01 (2024).
- [17] Nasim, M. F., M. Anwar, A. S. Alorfi, H. A. Ibrahim, A. Ahmed, A. Jaffar, S. Akram, A. Siddique, and H. M. Zeeshan. "Cognitively inspired sound-based automobile problem detection: A step toward explainable AI (XAI)." *International Journal of Advanced and Applied Sciences* 12, no. 8 (2025): 1-15.