AI Based Framework for Breast Cancer Detection for Histopathological Images

Ahmed M. Salaheldin[1*], Ahmed El Bialy[1], Rahma Saad[1]

ABSTRACT

Breast cancer continues to be one of the primary causes of death among women globally, and timely, accurate diagnosis is essential for enhancing survival rates. This research introduces a deep learning approach that utilizes the Inception-ResNet-V2 model to classify histopathological breast tissue images. The publicly available BreakHis dataset, encompassing of benign and malignant samples at multiple magnification levels, was used to evaluate the proposed method. Images were preprocessed through resizing and normalization before being fed into the model. For binary classification, the proposed model achieved an accuracy of 83.15%, with precision, recall, and F1-score of 0.9487, 0.7912, and 0.8628, respectively. In the multi-class scenario involving eight histological subtypes, the model achieved a testing accuracy of 94%, with macro and weighted F1-scores of 0.94, indicating consistent performance across all classes. The results demonstrate that the proposed approach effectively captures both low- and high-level features from histopathological images, offering a reliable tool for supporting breast cancer diagnosis. The study highlights the potential of deep learning models, particularly Inception-ResNet-V2, in enhancing diagnostic precision and reducing the burden on medical professionals.

Keywords: Breast Cancer Diagnosis, Histopathological slices Classification, Deep Learning, Computer-Aided Diagnosis (CAD)

I. INTRODUCTION

Cancer remains one of the leading causes of death worldwide, marked by the uncontrolled and accelerated growth of abnormal cells [1]. Among the various types of cancer, breast cancer is one of the most diagnosed globally and stands as the second leading cause of cancer-related mortality [2]. Currently, over 3.8 million women are living with breast cancer, underscoring its widespread prevalence—particularly among women in the United States. According to the American Cancer Society, approximately 1 in 8 women in the U.S. will be diagnosed with breast cancer at some point in their lives [3].

This identifies the urgent need for effective strategies in the detection and treatment of cancer. The complex structure of tumor masses and the presence of microcalcifications present significant challenges for radiologists in accurately diagnosing breast cancer. Consequently, early detection and accurate diagnosis are essential to improve treatment outcomes and increase survival rates among affected individuals [3], [4]. A variety of diagnostic methods support the early identification of breast abnormalities. In particular, pathology, when combined with Computer-Aided Diagnostic (CAD) systems, enables more precise assessments through the detailed analysis of tissue images.

[1] Systems and Biomedical Engineering Department, Higher Institute of Engineering, EL Shorouk Academy, Cairo, Egypt.

* Corresponding Author

Recent advancements have enabled pathologists to analyze a wide range of features in biopsy images, including texture, structural patterns, and morphological characteristics, all of which are crucial for accurate diagnosis. However, the manual process of scanning, focusing, zooming, and navigating across various magnification levels remains labor-intensive and time-consuming. This exhaustive approach not only demands considerable resources but also introduces the potential for human error, which can compromise the diagnosis of critical conditions like breast cancer. As a result, there is an ongoing pursuit to develop objective, quantifiable, and fully automated methods for image analysis [5].

To improve the precision and speed of breast cancer diagnosis, there has been an increasing move toward adopting deep learning technologies. These advanced methods support the development of diagnostic systems that minimize variability and improve the accuracy of disease detection. Deep learning has grasped particular importance in histopathology image classification due to its ability to automatically learn and extract relevant features. Neural network architectures such as Inception-ResNet-V2 have shown better performance than conventional techniques. However, these models typically require large datasets for successful training [8], [9]. At the same time, advancements in digital imaging have enabled the integration of machine learning and computer vision in pathological image analysis, further strengthening diagnostic effectiveness.

Tyagi et al. [6] presents a recent study highlights the effectiveness of Convolutional Neural Networks (CNNs) in enhancing the accuracy and efficiency of breast cancer diagnosis. Traditional manual inspection methods, though commonly used, are prone to human error and time constraints. The study employs a CNN-based Computer-Aided Diagnosis (CAD) system evaluated on the BreakHis histopathology dataset, achieving classification accuracies of 98% (70–30 split), 98% (80–20 split), and 99% (90–10 split).

Ain et al. [7] introduces a Genetic Programming (GP)-based feature learning method for breast cancer detection using histopathological images. Due to the high inter-class and intra-class variability in such images, the proposed method enhances classification by automatically selecting and combining effective image descriptors. Unlike traditional black-box models, this approach offers interpretable solutions that can aid medical professionals by highlighting key visual features.

Das et al. [8] proposes a model using Wavelet-Convolutional Neural Network (WCNN) to improve the detection of Invasive Ductal Carcinoma (IDC) in breast histopathology images. By integrating wavelet and convolutional filters within each layer, the model captures both frequency and spatial domain features, enabling effective multi-scale representation learning. Evaluated on a public dataset, the WCNN achieved a classification accuracy of 98.4%.

Addo et al. [9] presents BCHI-CovNet, a lightweight AI model designed for efficient and accurate classification of breast cancer histopathological images. Addressing limitations of traditional CNNs, the model introduces a multiscale depth-wise separable convolution to capture both low- and high-resolution patterns, alongside a pooling module that extracts rich second-order statistical features. Additionally, a multi-head self-attention mechanism enhances pixel diversity and long-range feature learning. Tested on the BreaKHis and BACH datasets, BCHI-CovNet achieved high classification accuracy across various magnification levels, reaching up to 99.38%.

The aim of the study is to implement an AI model for the detection of breast cancer from histopathological images which can be used as a decision support system to reduce human error and save time.

The rest of the paper is organized as follows: Section 2 describes the materials and methodology used in this work. Section 3 reports the main results, which are further examined in Section 4. Section 5 provides the conclusion and suggests possible directions for future research.

II. MATERIALS AND METHODS

This section provides a comprehensive overview of the materials, datasets, and methodologies employed in the study. It details the data acquisition process, preprocessing techniques, and the architecture of the proposed model. Additionally, it outlines the training configurations, evaluation metrics, and experimental setup used to assess the performance of the system. Each component is carefully selected to ensure accurate and reliable classification of breast cancer from histopathological images. The proposed pipeline in the study is presented in Fig. 1.

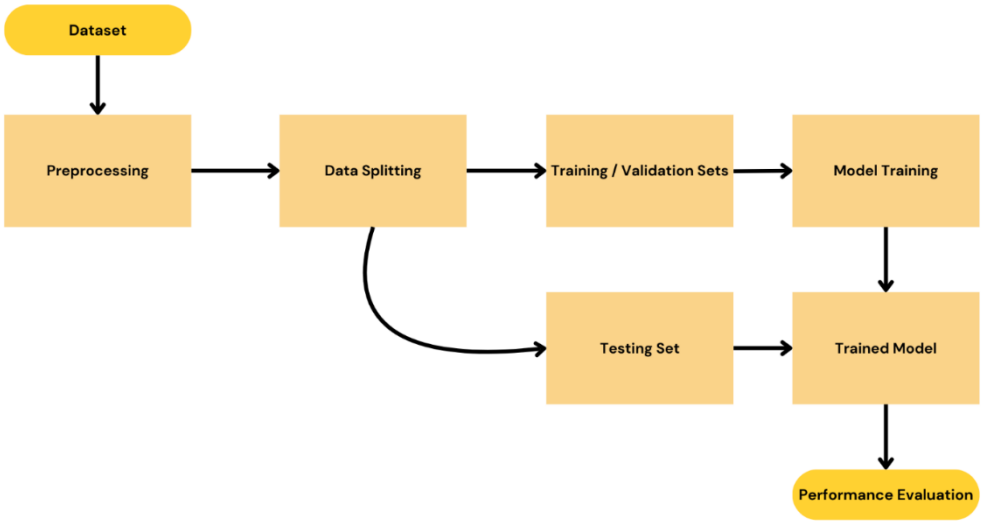
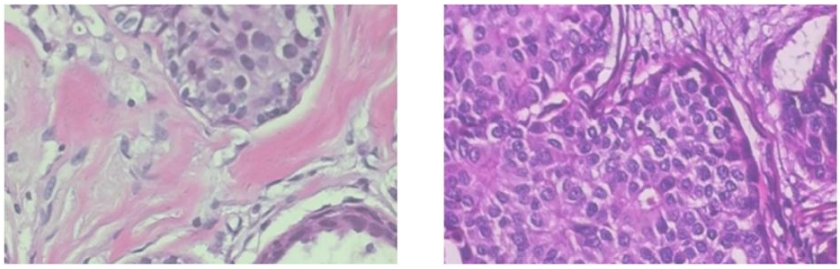


Figure 1 Block diagram for the proposed model

A. Dataset

The BreakHis histopathological image dataset [10] was employed to develop and evaluate the deep learning model for classifying images into benign and malignant categories. A representative sample from the dataset is shown in Fig. 2, while Table 1 presents the distribution of images across the two classes at various magnification levels.



a b

Figure 2 A sample of the utilized dataset a) Benign b) Malignant

Table 1 Distribution of the dataset for the target classes

Magnification Factor	Benign	Malignant	Total
40x	625	1,370	1,995
100x	644	1,437	2,081
200x	623	1,390	2,013
400x	588	1,232	1,820
Total	2,480	5,429	7,909

B. Preprocessing

To prepare the dataset for the training phase, a series of preprocessing steps were applied to ensure consistency and optimal performance of the deep learning model. Initially, all histopathological images were resized to match the input dimensions required by the network architecture, ensuring uniformity across the dataset. Subsequently, pixel values were normalized to a range between 0 and 1 by scaling the original 8-bit intensity values (0–255), a common practice that facilitates faster convergence and improved stability during training. These preprocessing steps are essential to enhance the model's ability to learn meaningful patterns from the input data.

C. Classification Model

This study utilized the Inception-ResNet-v2 architecture to classify histopathological breast cancer images into malignant and benign groups. To assess the model's ability to generalize, the dataset was split into 80% for training and 20% for testing. To ensure the robustness and reliability of the classification performance, 5-fold cross-validation was implemented during the training phase. In this approach, the training data is partitioned into five equally sized subsets D_1, D_2, D_3, D_4, D_5. For each fold i, the model is trained on $D \setminus D_i$ and validated on D_i, and the final performance is averaged over the five iterations:

$$Accuracy_{avg} = \frac{1}{k}\sum_{i=1}^{k} Accuracy_i \quad \text{where k} = 5 \qquad (1)$$

Inception-ResNet-v2, a hybrid architecture that combines the representational power of Inception modules with the training benefits of residual connections, is well-suited for learning complex patterns in medical images. The model utilizes convolutional layers with varying kernel sizes to extract multi-scale features and employs identity mappings to mitigate vanishing gradient issues. The final classification is achieved using a SoftMax activation function applied to the output logits:

$$\hat{y}_i = \frac{e^{z_i}}{\sum_{j=1}^{C} e^{z_i}}, i = 1, \ldots\ldots, C \qquad (2)$$

where $\hat{y}_i$ is the predicted probability for class is i, z_i the logit corresponding to class iii, and CCC is the number of output classes (in this case, 2: benign and malignant). This setup enables efficient and accurate classification while reducing the risk of overfitting, thanks to the cross-validation strategy and the architectural depth of Inception-ResNet-v2.

III. RESULTS

The proposed Inception-ResNet-V2-based deep learning approach was evaluated on the BreakHis dataset for both binary and multi-class classification tasks. In the binary classification setting, the model was assessed across four magnification levels. It demonstrated strong performance across all metrics, achieving an overall test accuracy of 83.15%, with a precision of

0.9487, recall of 0.7912, and F1-score of 0.8628. Specifically, for the benign class, the model achieved a precision of 0.68, recall of 0.91, and F1-score of 0.78, while for the malignant class, it reached a precision of 0.95, recall of 0.79, and F1-score of 0.86.

In the multi-class classification task, the model was trained for 20 epochs and achieved a validation accuracy of 93.42% with a final loss of 0.2640 and further attained a testing accuracy of 94%. The performance remained consistent across the eight histopathological subtypes, as reflected by the macro and weighted averages of 0.94 for precision, recall, and F1-score, indicating a balanced capability across all classes. Notably, the Adenosis class achieved a precision of 0.94, recall of 0.98, and F1-score of 0.96, while the Phyllodes Tumor class recorded the highest F1-score of 0.98. The Tubular Adenoma class also performed well with a recall of 0.99. Conversely, Ductal Carcinoma showed slightly lower performance with a recall of 0.83 and F1-score of 0.87, suggesting room for model enhancement in this category. Other classes, including Fibroadenoma, Lobular Carcinoma, Mucinous Carcinoma, and Papillary Carcinoma, maintained consistently high F1-scores ranging between 0.91 and 0.95, with Papillary Carcinoma achieving the highest precision of 0.98 and recall of 0.93.

IV. DISCUSSION

The results demonstrate the effectiveness of the Inception-ResNet-V2-based architecture in both binary and multi-class classification tasks for breast cancer histopathological images. In binary classification, the model achieved high overall performance, with particularly strong results in identifying malignant cases. The comparatively lower precision for benign samples suggests some misclassification, likely due to overlapping visual characteristics between benign and malignant tissues at certain magnification levels.

In the multi-class scenario, the model exhibited excellent generalization, evidenced by consistently high macro and weighted averages across all metrics. Classes such as Phyllodes Tumor, Adenosis, and Tubular Adenoma were detected with exceptional accuracy, indicating the model's robustness in capturing discriminative morphological features. However, the lower recall for Ductal Carcinoma indicates room for improvement, potentially through targeted data augmentation or class-specific fine-tuning strategies.

The use of 5-fold cross-validation, along with an 80–20 training-testing split, contributed to the reliability of the evaluation by minimizing overfitting and ensuring stable model behavior across varying data distributions. Additionally, the integration of deep feature extraction capabilities in Inception-ResNet-V2, combined with an optimized training strategy, played a vital role in achieving the reported performance levels.

Overall, the proposed framework demonstrates strong potential for practical application in computer-aided breast cancer diagnosis. Further improvements can be achieved by enriching the dataset, employing advanced augmentation techniques, or incorporating attention mechanisms to enhance the model’s focus on relevant image regions.

V. CONCLUSION

In this study, we proposed a deep learning-based approach utilizing the Inception-ResNet-V2 architecture for the classification of breast cancer histopathological images from the BreakHis dataset. The model was evaluated for both binary and multi-class classification tasks and demonstrated strong performance across all metrics, including accuracy, precision, recall, and F1-score. By leveraging an 80–20 train-test split and employing five-fold cross-validation, the model achieved robust generalization and consistent results across different magnification levels. For binary classification, the model effectively distinguished between benign and malignant samples, achieving an overall accuracy of 83.15%. In the multi-class setting, the model further demonstrated its strength by attaining a testing accuracy of 94% and balanced performance across all eight histological classes, with macro and weighted averages of 0.94 for all key metrics. These

findings affirm the model's capability to handle complex tissue patterns and class distributions, supporting its potential integration into computer-aided diagnostic systems. Future work may focus on enhancing performance for specific underperforming classes through targeted data augmentation, fine-tuning strategies, or incorporating attention mechanisms. Overall, the results validate the practical applicability of the proposed framework in aiding pathologists for early and accurate breast cancer diagnosis, ultimately contributing to improved clinical decision-making and patient outcomes.

REFERENCE

[1] S. D. Khalaf, M. M. Mahdi, and R. S. Mahmood, "THE CANCER: TYPES, THE MECHANISM OF CANER GROWTH AND DIAGNOSIS: A," World, vol. 3, no. 3, 2024.

[2] A. S. Bhat et al., "Cancer initiation and progression: a comprehensive review of carcinogenic substances, anti-cancer therapies, and regulatory frameworks," Asian J. Res. Biochem., vol. 14, no. 4, pp. 111–125, 2024.

[3] J. Ferlay et al., "Cancer statistics for the year 2020: An overview," Int. J. Cancer, vol. 149, no. 4, pp. 778–789, 2021, doi: 10.1002/ijc.33588.

[4] M. A. Aswathy and M. Jagannath, "An SVM approach towards breast cancer classification from H&E-stained histopathology images based on integrated features," Med. Biol. Eng. Comput., vol. 59, no. 9, pp. 1773–1783, 2021, doi: 10.1007/s11517-021-02403-0.

[5] X. Y. Liew, N. Hameed, and J. Clos, "A Review of Computer-Aided Expert Systems for Breast Cancer Diagnosis," Cancers, vol. 13, no. 11. 2021. doi: 10.3390/cancers13112764.

[6] S. Tyagi, S. Srivastava, B. C. Sahana, and I. S. Rajput, "A Deep Learning Framework for Automated Breast Cancer Detection Through Histopathological Image Analysis BT - Proceedings of Fifth International Conference on Computing, Communications, and Cyber-Security," P. J. S. Gonçalves, P. K. Singh, S. Tanwar, and G. Epiphaniou, Eds., Singapore: Springer Nature Singapore, 2025, pp. 321–333.

[7] Q. U. Ain, H. Al-Sahaf, B. Xue, and M. Zhang, "Genetic Programming for Malignancy Diagnosis From Breast Cancer Histopathological Images: A Feature Learning Approach," IEEE Trans. Emerg. Top. Comput. Intell., pp. 1–14, 2025, doi: 10.1109/TETCI.2024.3523769.

[8] A. Das and M. N. Mohanty, "Wavelet-Convolutional Neural Network: An Improved Deep Learning Model for Breast Cancer Detection from Histopathology Images: WCNN Breast Cancer," Arch. Breast Cancer, vol. 12, no. 1 SE-Original Article, pp. 73–84, Jan. 2025, doi: 10.32768/abc.202512173-84.

[9] D. Addo et al., "A hybrid lightweight breast cancer classification framework using the histopathological images," Biocybern. Biomed. Eng., vol. 44, no. 1, pp. 31–54, 2024, doi: https://doi.org/10.1016/j.bbe.2023.12.003.

[10] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "A Dataset for Breast Cancer Histopathological Image Classification," IEEE Trans. Biomed. Eng., vol. 63, no. 7, pp. 1455–1462, 2016, doi: 10.1109/TBME.2015.2496264.