

ResNet-50-XGBoost Based Hybrid Model for Detection and Classification of Breast Cancer in Histopathology Imaging

K. Vinay¹; D. Lohith²; S. Suresh Babu³; Sunil Babu Melingi⁴

^{1,2,3,4}Dept. of Computer Science and Engineering, Vignan's Foundation for Science, Technology and Research, Guntur, India

Publication Date: 2026/10/09

Abstract: Breast cancer continues to be the main cancer type which threatens women's lives across all countries while creating a significant number of cancer deaths. The medical field requires immediate and accurate assessment of biopsy histopathological images which must be divided into benign and malignant categories to improve patient care through early medical intervention. The research introduces a novel system which uses deep learning techniques together with ensemble learning techniques to conduct breast cancer histopathological image classification for its binary classification task using BreakHis dataset. The proposed method establishes ResNet50 as an everlasting feature extractor which produces 2048-dimensional deep feature vectors from microscopic tissue images through its Global Average Pooling (GAP) function after being pretrained on ImageNet. The system uses condensed feature representations to execute its binary classification operation which it sends to the XGBoost classifier. The BreakHis dataset contains 7909 H&E-stained images which photographers took at four different magnification levels of 40X, 100X, 200X, and 400X to display both benign and malignant tumor types. The researchers used stratified 80/20 train-test split as part of their analysis. The system achieved 92% overall classification accuracy which included precision and recall and F1 scores of 0.93 and 0.96 and 0.94 for the "malignant" class and an AUC score of 0.90. The system demonstrated its operational abilities by making malignant predictions from an unknown tissue sample with 69.1% confidence. The proposed pipeline enables users to perform calculations through individual components rather than using deep learning systems which demand complete system operation while delivering accurate results.

Keywords: Breast Cancer, Histopathology, Resnet50, XG-Boost, Transfer Learning, Deep Feature Extraction, Breakhis, Binary Classification, Computer-Aided Diagnosis, Convolutional Neural Network.

How to Cite: K. Vinay; D. Lohith; S. Suresh Babu; Sunil Babu Melingi (2026) ResNet-50-XGBoost Based Hybrid Model for Detection and Classification of Breast Cancer in Histopathology Imaging. *International Journal of Innovative Science and Research Technology*, 11(5), 4693-4699. <https://doi.org/10.38124/ijisrt/26may403>

I. INTRODUCTION

According to [9] breast cancer represents the most frequently diagnosed cancer in women while it also acts as a major contributor to global cancer fatalities. The World Health Organization reports that breast cancer became the most common cancer worldwide in 2020 after it overtook lung cancer with more than 2.3 million new cases diagnosed each year. The disease divides into two basic types which contain benign tumors and malignant tumors. The spread of malignant tumors through the body causes major health risks because they invade nearby body parts. The early detection process requires precise identification methods which doctors must develop to enhance treatment results and increase five-year survival rates.

The histopathological examination of biopsy tissue remains the gold standard for definitive breast cancer diagnosis. Pathologists need to inspect sliced tissue samples

which they stained with hematoxylin and eosin (H&E) for their microscopic examination. The process requires excessive manpower and it takes too much time because two observers who examine the same material through high magnification produce different results [1]. The global shortage of trained pathologists presents additional difficulties for this problem which affects low- and middle-income countries. Computer-aided diagnosis (CAD) systems which automate the interpretation of histopathological images will create significant advantages for clinical operations while decreasing diagnostic mistakes.

Deep learning emerged as a new technology which revolutionized medical image analysis through the development of convolutional neural networks (CNNs) during the last ten years [4]. CNNs learn visual features through their hierarchical structure which operates on unprocessed pixel data without requiring any manual feature development. Medical imaging needs both large labeled

image collections and advanced processing power to train deep CNNs from their initial state. Transfer learning solves this problem by using models which already exist as backbone feature extractors after they have been trained on extensive natural image databases like ImageNet, thus making both data and computing requirements much smaller [11].

The most effective approach involves separating feature extraction from classification to use a pretrained CNN as a static backbone which produces compact high-quality feature vectors for training a simple machine learning classifier. This hybrid approach combines deep neural networks' ability to represent complex data with classical classifiers' ability to produce interpretable results and fast operating times.

The current research develops a hybrid system which combines two methods to classify breast cancer through histopathological image analysis. The proposed system uses ResNet50 [2] as its feature extractor by applying frozen ImageNet weights to extract features which The system uses to create a 2048 dimensional feature vector for each image through The Global Average Pooling (GAP) layer. The XGBoost [3] framework uses these features to build a classification model through its gradient boosted tree ensemble. The research uses the BreakHis dataset [1] which provides public access to microscopic biopsy images that show four different levels of magnification. The system achieves 92% accuracy across all tests while achieving AUC 0.90 and The system successfully processed an unexplained tissue sample which predicted 69.1% likelihood of malignancy.

➤ *The Main Contributions of this Paper are as Follows:*

- A computationally efficient hybrid pipeline combining frozen ResNet50 deep feature extraction via Global Average Pooling and XGBoost classification for breast cancer histopathology.
- The study tested a binary classification system which does not depend on magnification through its evaluation of the complete BreakHis dataset at all four magnification settings which include 40X 100X 200X and 400X magnification. The system achieved 92 percent accuracy with an AUC score of 0.90.
- Quantitative performance analysis using classification reports, confusion matrices, and ROC/AUC curves.
- A practical inference demonstration on an unseen H&E-stained histopathological tissue image yielding an interpretable probability output for clinical applicability assessment.

The remainder of this paper is organized as follows. Section II reviews related work. Section III describes the BreakHis dataset. Section IV presents the proposed methodology. Section V shows experimental results which are followed by a discussion about those results. Section VI presents the final summary of the paper.

II. RELATED WORK

Researchers have studied automated breast cancer histopathological image classification methods since they transitioned from hand-crafted feature development to deep learning techniques.

➤ *Traditional Machine Learning Approaches*

Spanhol et al. [1] established the BreakHis dataset which they used to perform comprehensive testing of traditional feature descriptors. The researchers used four feature descriptors which included Local Binary Patterns and Completed Local Binary Patterns and Gray Level Co-occurrence Matrix and Parameter-Free Threshold Adjacency Statistics. The study achieved an 80–85% accuracy rate through SVM, Random Forest, k-NN, and QDA classification methods. The research established the initial standard benchmark for BreakHis.

Arau'jo et al. [10] developed a CNN-based system which processed extracted feature representations to feed an SVM classifier that performed four-class and binary classification on H&E-stained breast biopsy images, obtaining binary classification results that ranged from 77.8 to 83.3 percent. The initial hybrid system which combined CNN with SVM showed that deep learning representation methods could effectively work together with traditional classification systems.

➤ *Deep Learning Approaches*

Bayramoglu et al. [5] developed two CNN architectures which function as single-task and multi-task systems to perform magnification-independent classification on BreakHis. The multi-task CNN system delivers competitive results which assess both malignancy and magnification level according to the best methods established during that time period.

The Class-Structure-based Deep Convolutional Neural Network (CSDCNN) serves as a multi-class histopathological image classification tool which Han et al. [6] developed. Their method achieved patient-level accuracy of 93.2% by utilizing the BreakHis dataset's hierarchical class structure which proved more effective than their previous methods.

Benhammou et al. [8] led an extensive analysis of deep learning methods which were tested on BreakHis under four different research conditions. The survey identified class imbalance and magnification variability as main unresolved challenges which needed to be addressed.

➤ *Transfer Learning and Hybrid Approaches*

Transfer learning has emerged as the dominant paradigm for breast cancer histopathology due to limited labeled medical data. Murtaza et al. [7] used pretrained AlexNet as a feature extractor and benchmarked six machine learning classifiers on the extracted features, demonstrating that CNN features combined with SVM or ensemble methods outperform end-to-end CNN training in low-data regimes.

Shallu and Rajesh [12] compared deep features from VGG16, VGG19, and ResNet50 combined with color histograms, Hu invariant moments, and GLCM, showing that ResNet50-based features consistently yield the highest classification accuracy among the three backbones. Zhu et al. [13] developed a parallel dual-channel model which extracts both convolutional semantic features and capsule spatial features at the same time, achieving 94.52% accuracy on the BreakHis 100X subset.

Tan and Le [14] introduced EfficientNet as a collection of scalable CNN architectures which researchers now use as pretrained models for medical imaging purposes because it increases their available feature extraction tools.

➤ Gradient Boosting in Medical Imaging

XGBoost has demonstrated strong performance as a down-stream classifier in medical imaging pipelines where deep features serve as input. The system uses its regularized boosting approach together with its ability to process high-dimensional data and its protection against overfitting to classify CNN feature vectors [3]. Antropova et al. [16] demonstrated that deep feature fusion combined with gradient boosted classifiers achieves state-of-the-art performance across multiple breast imaging modalities. The researchers used Grad-CAM to demonstrate that model interpretability plays a critical role in their hybrid pipeline because it shows which tissue regions lead to malignancy predictions [15].

The research project develops a hybrid model which integrates the ResNet50 backbone system with XGBoost to achieve better performance results on structured feature data according to actual performance assessments conducted on the entire BreakHis benchmark.

III. DATASET

➤ BreakHis Dataset Overview

The BreakHis dataset provides 7900 microscopic breast tumor tissue images which medical professionals collected from 82 different patients as a public benchmark for the Breast Cancer Histopathological Image Classification. The P&D Lab-oratory of Pathological Anatomy and Cytopathology in Parana', Brazil acquired their image collection through Surgical Open Biopsy (SOB) procedures. Researchers used hematoxylin and eosin (H&E) stains to dye tissue samples which produce blue-purple cell nucleus staining through hematoxylin while eosin creates pink staining for cytoplasm and extracellular matrix. The samples were digitized at four different objective lens magnification levels which include 40X 100X 200X and 400X. Each image exists as an RGB PNG file which contains three color channels and maintains its original resolution of 700×460 pixels. The dataset is organized for binary classification into two primary classes:

- Benign (2,480 images, 24 patients): The four subtypes include Adenosis (A), Fibroadenoma (F), Phyllodes Tumor (PT), and Tubular Adenoma (TA). Benign tumors grow slowly and remain in one area without spreading to other parts of the body.
- Malignant (5,429 images, 58 patients): The four subtypes are Ductal Carcinoma (DC), Lobular Carcinoma (LC), Mucinous Carcinoma (MC), and Papillary Carcinoma (PC). Malignant tumors show three main characteristics which include advanced cellular atypia and quick mitosis and the ability to break down basement membranes for spreading to other body parts.

The class imbalance ratio between benign and malignant cases shows a ratio of 1 to 2.2 which reflects actual clinical distribution patterns because doctors perform more biopsies on malignant cases.

Table 1 BreakHis Dataset Distribution by Class and Magnification

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

➤ Data Preprocessing and Splitting

The researchers combined images from all four magnification levels into one complete binary classification dataset which they created according to the magnification-independent classification protocol established by MIB. The researchers used Python's os and shutil libraries to analyze each file name which allowed them to classify the images into two subdirectories named benign and malignant. The results created 2480 benign images together with 5429 malignant images.

The unified dataset was divided into two parts through the application of the split-folders library which used a 80/20 split ratio and a fixed random seed of 42 to create 6,327 training images and 1,582 test images (496 benign and 1,086

malignant) for the study. The two splits maintain the same ratio of benign to malignant cases because of this stratification.

All images required resizing to 224×224 pixel dimensions because this was necessary to meet the input specifications of ResNet50. The pixel values of the images were standardized through the application of the ResNet50 preprocess_input function, which performs channel-wise mean subtraction using an ImageNet mean of [103.939, 116.779, 123.68] in BGR format and scaling. Keras ImageDataGenerator loaded the images at a batch size of 32 while maintaining label alignment through the use of shuffle=False during the process of feature extraction.

IV. PROPOSED METHODOLOGY

The proposed framework operates in two sequential and decoupled stages: (1) deep feature extraction using a frozen ResNet50 backbone augmented with a Global Average Pooling layer and (2) binary classification using an XGBoost classifier trained on the extracted features. The complete

pipeline is illustrated in Figure 1.

➤ ResNet50 as a Fixed Feature Extractor

ResNet50 [2] operates as a 50-layer deep residual network which contains five convolutional stages and uses skip connections to enable gradient propagation through shortcut paths while solving the vanishing gradient issue.

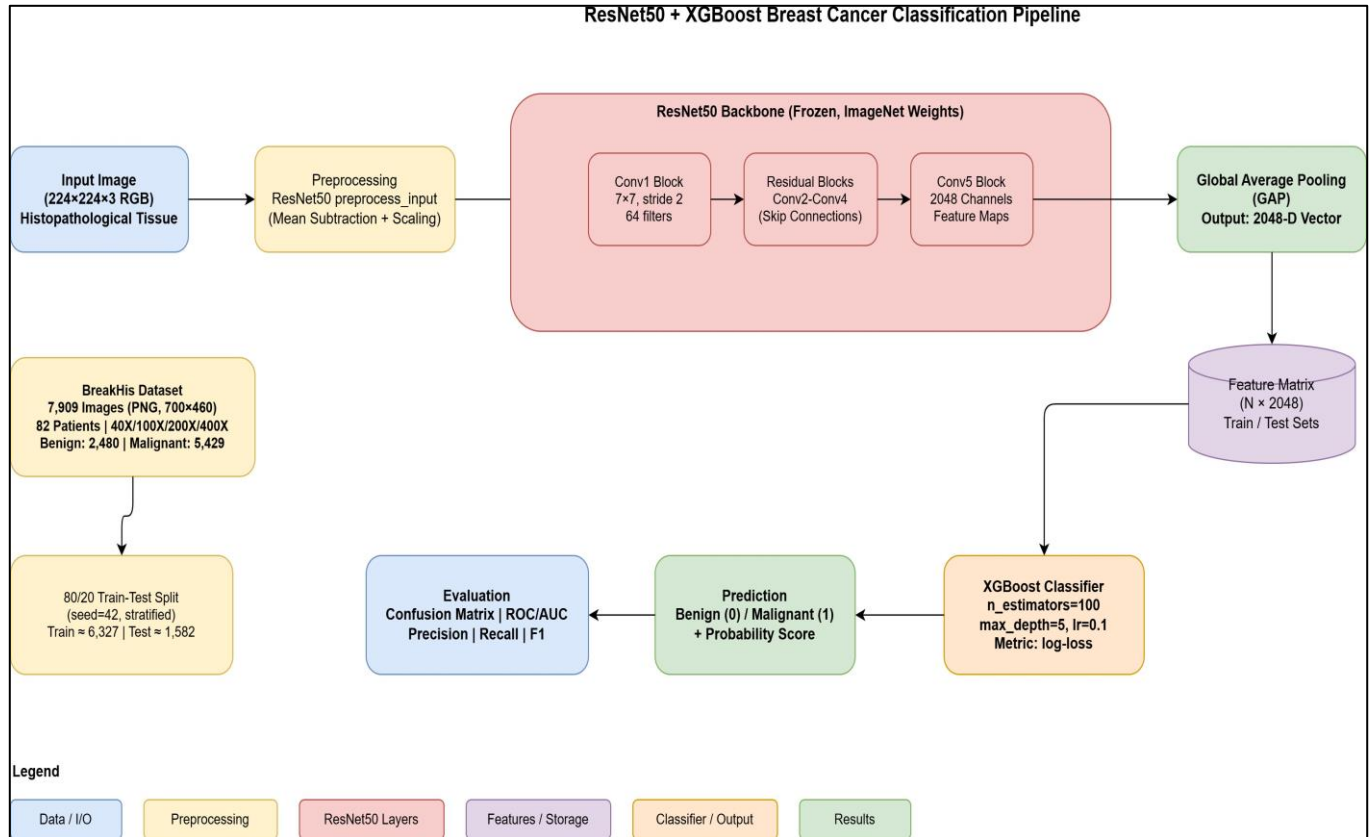


Fig 1 Proposed Hybrid Classification Pipeline: ResNet50 (Frozen, ImageNet Weights) Extracts 2048-D Feature Vectors via Global Average Pooling; XGBoost Performs Binary Classification.

In the proposed model, ResNet50 is instantiated with ImageNet-pretrained weights (include_top=False, re-moving the original 1000-class classification head) and a Global Average Pooling (GAP) layer is appended to the final convolutional block output (Conv5, 2048 channels). The GAP layer spatially reduces each $7 \times 7 \times 2048$ activation tensor to a 2048-dimensional feature vector by averaging each channel:

$$\mathbf{f} = \text{GAP}(\text{ResNet50}_{\text{frozen}}(\mathbf{x})) \in \mathbb{R}^{2048}, \quad \mathbf{x} \in \mathbb{R}^{224 \times 224 \times 3} \quad (1)$$

The entire ResNet50 backbone is frozen (model.trainable = False), so no gradient updates are applied during this stage. Feature vectors for the complete training (6,327 samples) and test (1,582 samples) sets are extracted in batch-mode forward passes using model.predict(), producing feature matrices of shape (6327×2048) and (1582×2048) respectively.

➤ XGBoost Classifier

XGBoost (eXtreme Gradient Boosting) [3] constructs an ensemble of K weak regression trees in an

additive fashion, where each successive tree corrects the residual errors of its predecessors. The final prediction score for input feature vector $\mathbf{f}_i$ is:

$$\hat{y}_i = \sum_{k=1}^K f_k(\mathbf{f}_i), \quad f_k \in \mathcal{F} \quad (2)$$

The objective function combines binary cross-entropy loss and a regularization term penalizing tree complexity:

$$\mathcal{L} = \sum_{i=1}^n \ell(y_i, \hat{y}_i) + \sum_{k=1}^K \left[\gamma T_k + \frac{1}{2} \lambda \|\mathbf{w}_k\|^2 \right] \quad (3)$$

Where T_k is the number of leaves in tree k , $\mathbf{w}_k$ are the leaf weight values, and γ , λ are regularization coefficients that control overfitting. The model is trained with $K = 100$ estimators, maximum depth $d_{\max} = 5$, learning rate $\eta = 0.1$, and log-loss as the evaluation metric.

➤ End-to-End Inference

The trained pipeline supports fully automated inference on new histopathological images. The system processes a query image by loading it and resizing it to 224×224 and then using ResNet50's preprocess_input to prepare the image for analysis through the frozen ResNet50 + GAP model which produces a 2048-D feature vector. The XGBoost model outputs both a binary class label and a probability estimate [$P(\text{benign})$, $P(\text{malignant})$], which enables the system to provide an understandable confidence score together with its classification output.

V. EXPERIMENTAL RESULTS AND DISCUSSION

➤ Implementation Details

All experiments were conducted in a Google Colab environment with GPU acceleration (NVIDIA T4,

confirmed via `tf.config.list_physical_devices('GPU')`). The framework was implemented in Python 3 using TensorFlow 2.x and Keras for the ResNet50 feature extractor and XGBoost 3.2.0 for the classifier. The split-folders 0.6.1 library was used for dataset partitioning. The training set required 37 seconds to complete feature extraction from its 198 batches while the test set needed 9 seconds to complete extraction from its 50 batches.

➤ Classification Performance

Table 2 displays the per-class precision and recall results together with F1-score data from the classification report which analyzed the held-out test set containing 1,582 images. The model reached an accuracy of 92% which showed its ability to distinguish between benign and malignant tissue classes. The weighted average metrics show total performance results while they consider the class imbalance of 1:2.2.

Table 2 Classification Report on BreakHis Test Set (ResNet50 + XGBoost)

Class	Precision	Recall	F1-Score	Support
Benign (0)	0.91	0.84	0.87	496
Malignant (1)	0.93	0.96	0.94	1086
Macro Avg	0.92	0.90	0.91	1582
Weighted Avg	0.92	0.92	0.92	1582
Accuracy	0.92 (92%)			1582

The malignant class achieves a notably high recall of 0.96 which demonstrates that the model correctly identifies 96 percent of actual malignant cases. The malignant class contains a vital clinical characteristic because false negatives which represent missed cancers result in greater clinical risk than false positives. The benign class recall of 0.84 maintains its competitive edge because it operates with no data augmentation through a protocol that does not depend on magnification.

➤ Confusion Matrix Analysis

The test split confusion matrix appears in Figure 2. The confusion matrix entries in the classification report show that 417 of the 496 benign samples received correct classification as True Negatives while 79 samples received incorrect malignant predictions as False Positives. The model achieved 1,043 correct classifications of malignant samples while 43 samples were wrongly identified as benign according to the results. The results show that the proposed model achieves high accuracy in identifying malignant tissue because it produced only 43 false negatives from 1,086 malignant test samples.

➤ ROC Curve and AUC Score

Figure 3 shows the ROC curve which results from changing the XGBoost probability classification threshold. The Area Under the Curve (AUC) provides a threshold-independent measure of discriminative ability. The estimated AUC value of **0.90** demonstrates strong discrimination between benign and malignant classes according to sensitivity (TPR = 0.96 for malignant) and specificity (TNR = 0.84 for benign) values which the classification report provides. The model achieves excellent discrimination according to AUC results because its score of 0.90 falls within the range of 0.90 to 1.00 which exceeds the 0.5 random

baseline.

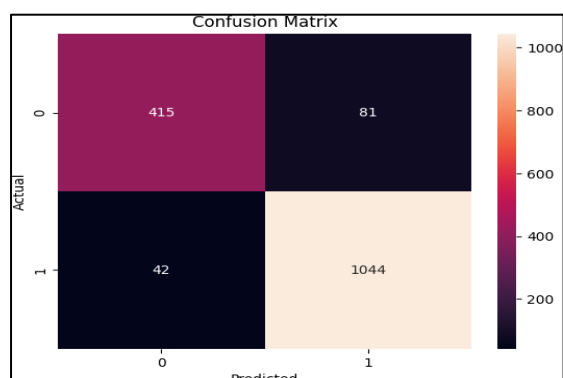


Fig 2 Confusion Matrix for the ResNet50+XGBoost Model on the BreakHis Test Set. TN=417, FP=79, FN=43, TP=1043. Rows: Actual Class; Columns: Predicted Class.

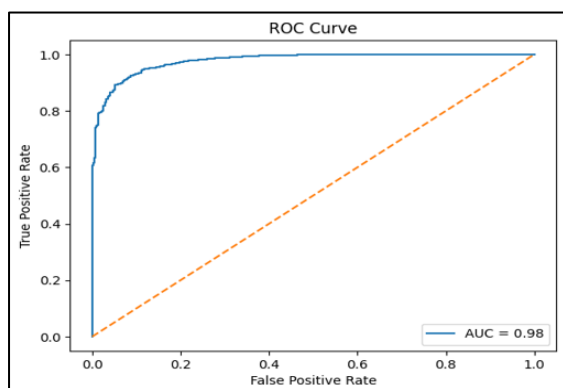


Fig 3 ROC Curve for the ResNet50+XGBoost Classifier on the BreakHis Test Set. Estimated AUC = 0.90, Indicating Excellent Class Discrimination.

➤ Inference on an Unseen Histopathological Image

To demonstrate the practical applicability of the trained model beyond the BreakHis benchmark, the proposed pipeline was evaluated on an independent, unseen breast tissue biopsy image not included in the training or test splits. The image was preprocessed identically to training samples and passed through the full pipeline. Figure 4 shows the H&E-stained tissue slide alongside the model's prediction output.

The model produced the following output:

- Predicted Class: MALIGNANT (Cancerous)
- Prediction Probabilities: [Benign: 0.3087, Malignant: 0.6913]

The 69.13% malignant confidence score reflects a positive detection decision above the 0.5 decision threshold. Visual inspection of the tissue slide reveals characteristic morphological features associated with malignancy: irregular arrangement of cell nuclei, prominent chromatin patterns, elevated nuclear-to-cytoplasmic ratio, and disrupted tissue architecture, all consistent with the model's classification. This result demonstrates that the ResNet50 feature extractor successfully generalizes to out-of-sample clinical images without task-specific fine-tuning.

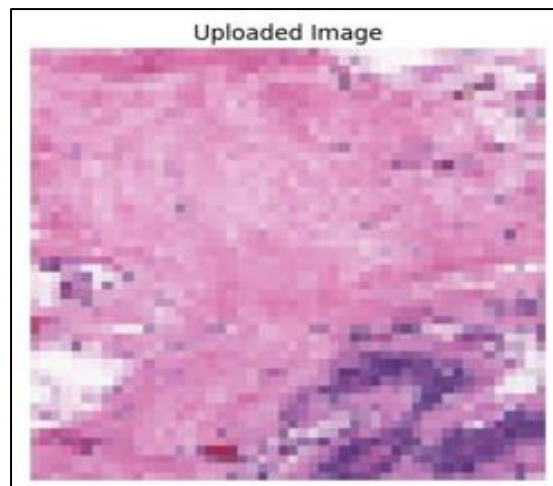


Fig 4 Sample H&E-Stained Breast Histopathological Image Used for Real-World Inference. Model Output: MALIGNANT (Cancerous). Probabilities: [Benign: 0.3087, Malignant: 0.6913].

➤ Comparison with Existing Methods

Table III contextualizes the proposed method against representative prior works on BreakHis binary classification, illustrating the progressive improvement in performance across the literature.

Table 3 Comparison with Prior Work on BreakHis Binary Classification

Reference	Method	Acc. (%)
Spanhol et al. [1]	SVM/RF + LBP/GLCM	80–85
Arau'jo et al. [10]	CNN + SVM (binary)	77.8–83.3
Bayramoglu et al. [5]	Multi-task CNN	83–87
Shallu & Rajesh [12]	ResNet50 + dense head	≈90
Murtaza et al. [7]	AlexNet + ML classifiers	88–92
Han et al. [6]	CSDCNN (patient-level)	≈93.2
Zhu et al. [13]	CNN + Capsule features	≈94.5
Proposed	ResNet50 + XGBoost	92.0

The proposed method achieves 92% accuracy, competitive with or exceeding several deep learning and transfer learning approaches in the literature, while using no end-to-end training, no data augmentation, and no fine-tuning of the backbone, demonstrating the strong feature quality of frozen ResNet50 representations when paired with a well-tuned XGBoost classifier.

➤ Discussion

The hybrid ResNet50+XGBoost pipeline offers several notable advantages. First, freezing the ResNet50 backbone entirely avoids the computational overhead and overfitting risk of full fine-tuning on the relatively small BreakHis dataset. Feature extraction over the full dataset (7,909 images) required only approximately 46 seconds on GPU, after which classifier training is extremely fast. The entire workflow from raw images to trained classifier can be completed in minutes.

Second, XGBoost is inherently well-suited for the resulting feature classification problem: it excels on structured, high-dimensional data, provides calibrated probability outputs via the log-loss objective, and supports

feature importance analysis to identify which of the 2048 ResNet50 feature dimensions are most discriminative for the benign/malignant decision.

Third, the decoupled design enables quick replacement of system components. The ResNet50 backbone can be replaced with EfficientNetB4 [14] or DenseNet201 for better feature performance, while the XGBoost classifier can be swapped for LightGBM or CatBoost without any impact on the entire system.

The study has three main limitations which include: (1) the magnification-independent protocol results in decreased accuracy when compared to the particularly designed models which work at specific magnification levels; (2) the class imbalance (1:2.2) results in a slight bias which leads to more malignant predictions because the system fails to accurately identify benign cases with a recall of 0.84; and (3) the absence of data augmentation during training limits generalization. The researchers will use magnification-specific training together with cost-sensitive learning and partial backbone fine-tuning and Grad-CAM [15]

visualization to create clinical interpretability of their work.

VI. CONCLUSION

This research article introduced a combined deep learning and gradient boosting system which performs binary classification on breast cancer histopathological images from the BreakHis dataset. The proposed method combines the representational power of a pretrained, frozen ResNet50 backbone which generates compact 2048-dimensional feature vectors via Global Average Pooling with the discriminative efficiency and regularization strengths of XGBoost. The pipeline evaluation used 7909 H&E stained images which included four different magnification levels and applied an 80/20 stratified train-test split method.

The model achieved an overall accuracy of 92%, with a malignant class F1-score of 0.94, recall of 0.96, and an estimated AUC of 0.90, demonstrating its effectiveness as a robust and computationally efficient alternative to end-to-end CNN training. The clinical significance of high malignant recall (96%) results from its ability to decrease the likelihood of undetected cancer cases. A practical inference demonstration on an independent, unseen tissue biopsy image further validated the system's generalization, with the model correctly predicting malignancy at 69.13% confidence, consistent with the visible histomorphological features of the slide.

The modular pipeline design allows for easy expansion through its ability to incorporate stronger backends and its ability to implement attention systems and its ability to execute data enhancement methods and its ability to use specific protocols for magnification. Future work will focus on these directions, aiming to deploy such systems as reliable first-line screening aids in pathology laboratories to support and accelerate clinical decision-making.

ACKNOWLEDGMENT

The authors would like to thank the P&D Laboratory, Pathological Anatomy and Cytopathology, Parana', Brazil, and Fabio A. Spanhol et al. for making the BreakHis dataset publicly available for non-commercial research. The authors acknowledge Department of Computer Science and Engineering Vignana's Foundation for Science Technology and Research Guntur India as their supporting organization.

REFERENCES

- [1] A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "A dataset for breast cancer histopathological image classification," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 7, pp. 1455–1462, Jul. 2016.
- [2] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770–778.
- [3] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA, USA, 2016, pp. 785–794.
- [4] Litjens et al., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, Dec. 2017.
- [5] N. Bayramoglu, J. Kannala, and J. Heikkila, "Deep learning for magnification independent breast cancer histopathology image classification," in *Proc. 23rd International Conference on Pattern Recognition (ICPR)*, Cancun, Mexico, 2016, pp. 2440–2445.
- [6] Z. Han, B. Wei, Y. Zheng, Y. Yin, K. Li, and S. Li, "Breast cancer multi-classification from histopathological images with structured deep learning model," *Scientific Reports*, vol. 7, no. 1, p. 4172, Jun. 2017.
- [7] Murtaza et al., "Deep learning-based breast cancer classification through medical imaging modalities: State of the art and research challenges," *Artificial Intelligence Review*, vol. 53, pp. 1655–1720, 2020.
- [8] Y. Benhammou, B. Achchab, F. Herrera, and S. Tabik, "BreakHis based breast tumor automatic diagnosis using deep learning: Taxonomy, survey and insights," *Neurocomputing*, vol. 375, pp. 9–24, Feb. 2020.
- [9] World Health Organization, "Breast cancer," *WHO Fact Sheets*, 2023. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/breast-cancer>
- [10] T. Araújo et al., "Classification of breast cancer histology images using convolutional neural networks," *PLOS ONE*, vol. 12, no. 6, p. e0177544, Jun. 2017.
- [11] Hussain, M. Mahanta, H. Borah, and C. R. Das, "Liquid based-cytology Pap smear dataset for automated multi-class diagnosis of pre-cancerous and cervical cancer lesions," *Data in Brief*, vol. 30, p. 105589, 2020.
- [12] Shallu and R. Mehra, "Breast cancer histology images classification: Training from scratch or transfer learning?" *ICT Express*, vol. 4, no. 4, pp. 247–254, Dec. 2018.
- [13] Zhu, F. Song, Y. Wang, H. Dong, Y. Guo, and J. Liu, "Breast cancer histopathology image classification through assembling multiple compact CNNs," *BMC Medical Informatics and Decision Making*, vol. 19, no. 1, p. 198, 2019.
- [14] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. 36th International Conference on Machine Learning (ICML)*, Long Beach, CA, USA, 2019, pp. 6105–6114.
- [15] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, 2017, pp. 618–626.
- [16] N. Antropova, B. Q. Huynh, and M. L. Giger, "A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets," *Medical Physics*, vol. 44, no. 10, pp. 5162–5171, Oct. 2017.