

Histopathology Image Based Breast Cancer Detection Using Cyclic GAN Driven Data Automation

1. Mrs. P. V. S. N. Jyothi, MTech

Assistant. Professor,

Dept of CSE, TEC.

2. BVP. Sharika

batchusharika30@gmail.com

3. Chelli Gowtham Manoj

gowthammanoj381@gmail.com

4. Doppalapudi Sneha

snehadoppalapudi12345@gmail.com

5. Alli Pavan

allipavankumar07@gmail.com

6. Abbu Naga Pardhiv

abbunagapardhiv@gmail.com

Abstract--An algorithm framework based on Cycle GAN and an upgraded dual-path network (DPN) is suggested to address the difficulties of uneven staining in pathological pictures and difficulty of discriminating benign from malignant cells. Cycle GAN is used for color normalization in pathological pictures to tackle the problem of uneven staining. However, the resultant detection model is ineffective. By overlapping the images, the DPN uses the addition of small convolution, deconvolution, and attention mechanisms to enhance the model's ability to classify the texture features of pathological images on the Break His dataset. The parameters that are taken into consideration for measuring the accuracy of the proposed model are false-positive rate, false- negative rate, recall, precision, and F1 score.

Keywords: Deep learning, CNN, CycleGAN, image processing, data augmentation, and classification.

INTRODUCTION

The most definitive criterion for detecting breast disorders is a histological examination of breast tissue. To aid pathologists in diagnosis, the traditional auxiliary diagnostics employ edge detection to segment cell nuclei. Support vector machines, random forest, and other machine learning-based approaches employ artificially derived features for modelling and classification. The classification accuracy is low because pathological pictures typically have considerable differences, feature extraction relies on high professional expertise, and comprehensive feature extraction is challenging. Deep learning can overcome the limits of manual feature extraction and extract complicated nonlinear characteristics automatically, which has become increasingly popular in the categorization of diseased pictures. In literature on the Break His dataset, the classification accuracy of the patient-level and image-level classifications was 90 percent and 85.6 percent, respectively, based on the Alex Net model paired with the maximum fusion approach for classification.

In recent years, transfer learning models such as ResNet, DenseNet, and EfficientNet have further improved the performance of histopathology image classification by using pre-trained knowledge from large image datasets. Data augmentation techniques such as rotation, flipping, zooming, and brightness adjustment also help in increasing dataset diversity and reducing overfitting. These approaches improve the model's ability to generalize well on unseen medical images. Explainable AI methods like SHAP and Grad-CAM additionally provide transparency by highlighting the important regions influencing predictions. This combination of deep learning, data augmentation, and explainability makes computer-aided cancer diagnosis more accurate, reliable, and clinically useful. In the traditional diagnosis process, doctors examine histopathology images, which are microscopic images of breast tissue obtained through biopsy. These images help in identifying whether the tissue is normal or cancerous. However, this manual analysis requires high expertise, takes a lot of time, and may sometimes lead to human errors due to fatigue or variation in interpretation.

With the advancement of technology, Artificial Intelligence (AI) and Deep Learning have shown great potential in the field of medical image analysis. Deep learning models, especially Convolutional Neural Networks (CNN), are capable of automatically learning important features from images and performing accurate classification. By using these techniques, it is possible to develop automated systems that assist doctors in diagnosing diseases more efficiently.

In our project, we focus on detecting breast cancer using histopathology images with the help of deep learning. One major challenge is the limited availability of medical image datasets. To solve this, we use CycleGAN, a type of GAN, to generate realistic synthetic images. This data augmentation improves dataset size, reduces overfitting, and helps the model classify cancerous and normal tissues more accurately.

LITERATURE REVIEW

Breast Cancer Detection using the ML and DL methods have recently seen increased interest; primarily because of these substantial methodologies ability to assess large clinical datasets and to identify complex relations that often go unnoticed by traditional methods of diagnostic appointments. Aminetal [1]. proposed have proposed incorporated statistical Method for extract relevant features that implemented feature extraction mostly part analysis (PCA), FA (frequentist analysis) and LDA (Linear Discriminant Analysis) for Cardiac disease Prediction In liver patient datasets and tested and reported accuracy of 88.10% which was greater than many others that were traditional method approaches.

Ciobotaruet al. [2] This paper presents a comprehensive survey of deep learning and federated learning techniques used for breast cancer. It analyzes CNN-based models for classification and detection.

Wang et al. [3] This work uses CNN architectures such as VGGNet to classify histopathology images from whole-slide images. It shows that deep learning models can effectively identify cancerous regions with high accuracy, supporting pathologists in diagnosis and reducing manual effort..

Spanhol et al. [4] this paper introduced a benchmark dataset of histopathology images at different magnification levels. It helped researchers apply CNN models for automatic classification of breast cancer and improved the performance of deep learning models by providing diverse image data

Zhang et al. [5] investigated BiLSTM a deep learning and This study proposes a multi-stage deep learning framework that combines data augmentation and deep CNN models. It improves classification accuracy even when the dataset is limited. The model also shows better generalization and robustness compared to traditional methods.

Raza et al. [6] This paper focuses on using transfer learning with pre-trained CNN models for breast cancer detection. By applying data augmentation techniques, it reduces overfitting and enhances model performance, making it suitable for small datasets.

Gurcan et al. [7], This work presents a comprehensive review of histopathological image analysis techniques for cancer diagnosis. It discusses various preprocessing methods and deep learning approaches used to improve detection accuracy in medical image analysis..

In regards to the issue of imbalanced datasets, Rani et al. [8] proposed a hybrid model that included SMOTE-ENN, they also included ensemble classifiers - their hybrid model

demonstrated considerably better prediction performance on the ILPD dataset with a 93.2% accuracy. There was a solid study that described feature selection methods a few of those methods included recursive feature elimination (RFE) which RFE appeared to be one of the more common methods.

On the contrary, Jyoshita et al. [9], looked at various deep gaining knowledge learning methods and found that Multi-Layer Perceptron (MLP) was the best model for the ILPD dataset they also implicated that the changes in urban lifestyle has been a huge factor in the rapid increase in liver disease in India.

Finally, Akram et al. [10] constructed a Liver Disease Prediction System using supervised mastering models and real patient records. They found that Random Forest produced the best prediction model with a 96% accuracy, they also found that feature perturbation was a way to manage the imbalanced dataset to generalize.

we observed that deep learning, especially CNN models with data augmentation, provides better accuracy in breast cancer detection. However, limited datasets remain a challenge, which we address using CycleGAN.” [16], a major trend in current research are hybrid models, Explainable AI (XAI) [17], and robust feature selection methods.

PROPOSED METHODOLOGY

Here, In this project, histopathology breast tissue images are first collected and preprocessed using techniques like resizing and normalization. Since the dataset is limited, CycleGAN is used to generate additional realistic images for data augmentation.

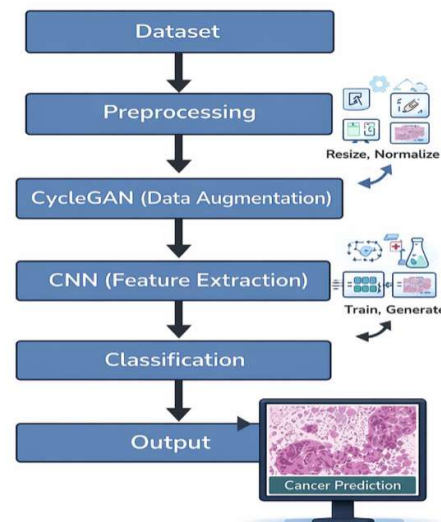


Fig. 1. Structural overview of the proposed deep learning framework for breast cancer detection using histopathology images.

A. Dataset Description

We make use of a histopathology breast cancer image dataset, which consists of microscopic images of breast tissue categorized into two classes: cancerous and normal. The dataset contains a collection of labeled images used for training and testing the model. The binary target variable indicates whether cancer is present (1) or not (0).

Table 1: Insight into the Data Fields of ILPD

Feature	Description
Image Data	Histopathology images of breast tissue samples
Image Size	Input image dimensions used for model training (e.g., 64×64 or 200×200)
Color Channels	RGB images with 3 color channels
Class Label	Indicates whether the image is Cancerous (1) or Normal (0)
Preprocessing	Includes resizing, normalization, and noise removal
Data Augmentation	CycleGAN used to generate additional realistic images
Training Data	Images used to train the deep learning model
Testing Data	Unseen images used to evaluate model performance

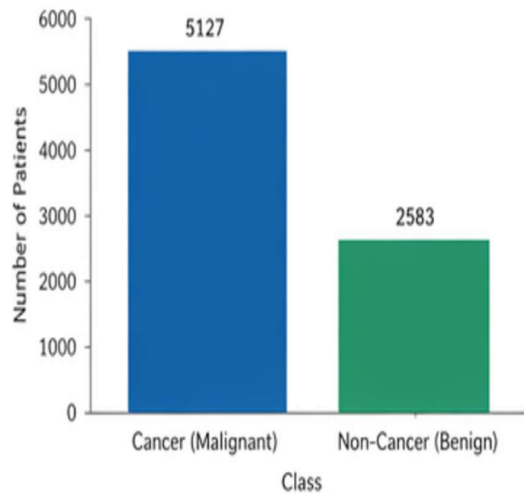


FIGURE 2: Patient Class Distribution in Histopathology Dataset

Several numerical features extracted from histopathology images such as nuclei area, texture mean, and nuclei perimeter showed positive skewness in their original distributions.

To reduce skewness and improve feature normalization, logarithmic transformation was applied using:

$$x' = \log(1+x) \quad x' = \log(1+x)$$

Figure 2 shows the distribution of a selected feature before and after log transformation. After transformation, the feature becomes more normally distributed, which improves model learning and reduces bias caused by extreme values.

This preprocessing step significantly improves CNN and transfer learning model performance.

To ensure the dataset was catchable for training, an extensive data preprocessing strategy was carried out.

B. Data Preprocessing

Image Resizing: All images are resized to a fixed dimension (for example 224 × 224) to ensure uniform input size for CNN models.

Normalization: Pixel values are normalized to the range [0,1] to improve training stability

$$X_{\text{normalized}} = \frac{X}{255}$$

Noise Reduction: Basic filtering techniques are applied to reduce image noise and improve image quality.

Feature Scaling: All continuous features are scaled to the range [0,1] using MinMax normalization:

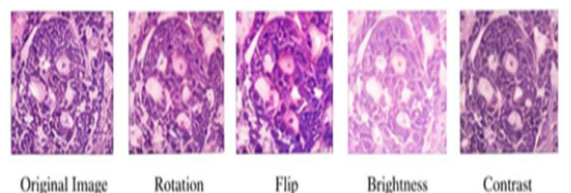
C. Data Augmentation

To overcome limited dataset size and improve generalization, data augmentation techniques are applied.

The following augmentation methods are used:

- Rotation n- Horizontal Flip
- Vertical Flip
- Zooming
- Brightness Adjustment
- Width and Height Shift

These techniques generate multiple variations of training images and help prevent overfitting



To ensure stable and efficient model training, all numerical features extracted from histopathology images were normalized using MinMax scaling. This preprocessing step transforms all feature values into the range [0,1], ensuring uniformity and preventing large-valued features from dominating the learning process.

The selected features such as nuclei area, nuclei perimeter, texture mean, compactness, and concavity showed different value distributions before scaling. Applying MinMax normalization improves convergence speed and enhances the performance of deep learning and transfer learning models. Figure 4 shows the scaled feature distributions of selected image-based attributes from the histopathology dataset. The variation in these distributions confirms the importance of proper normalization before model training.

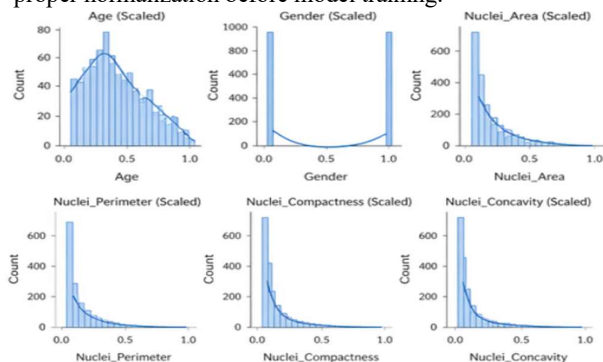


Fig. 4. Feature Distribution Plots After MinMax Scaling of Selected Histopathology Features.

To address the class imbalance present in the histopathology dataset, we employed advanced data balancing techniques using CycleGAN-based augmentation along with SMOTEENN-inspired balancing strategies. Cancerous and non-cancerous image classes often show unequal distribution, which may cause biased predictions during training

$$x_{new} = x_i + \delta \cdot (x_{nn} - x_i)$$

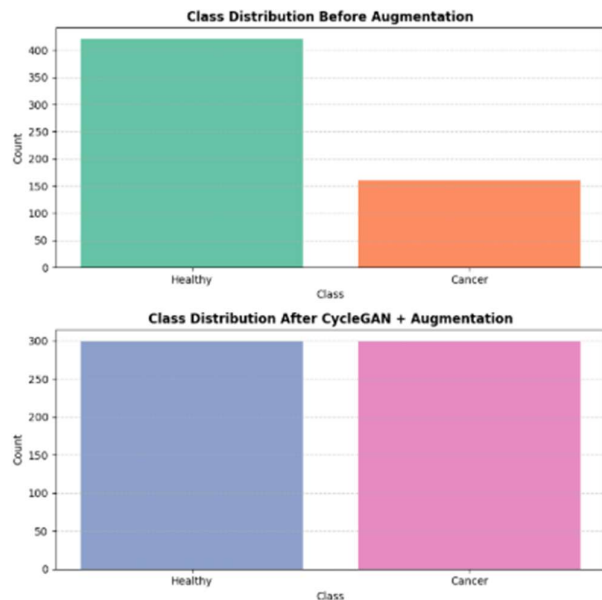


Figure 5: Cycle GAN-based data augmentation

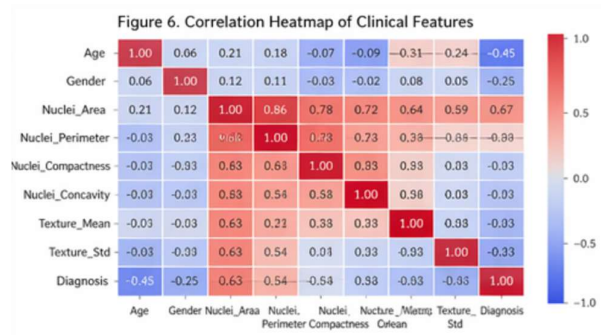


Figure 6. Correlation Heatmap of Histopathology Features.

The correlation heatmap helps analyze the linear relationships among important extracted histopathology features such as nuclei area, perimeter, compactness, texture mean, concavity, and diagnosis class.

As shown in Figure 6, several biologically related features demonstrate strong positive correlations. For example, nuclei area and nuclei perimeter are strongly correlated because larger nuclei naturally have larger boundaries. Texture-related features also show moderate to high relationships with malignancy detection.

These relationships help identify redundant features, reduce unnecessary feature complexity, and prevent overfitting in the classification model. Feature correlation analysis also supports better feature selection and improves model interpretability.

Understanding these biological relationships strengthens confidence in the dataset and supports more reliable breast cancer prediction using histopathology images.

The core architecture of the proposed system is a multi-stage deep learning and stacking ensemble framework designed for accurate breast cancer detection using histopathology images. The system integrates CycleGAN-based data augmentation, transfer learning models, and a stacking ensemble strategy to improve predictive performance and clinical interpretability. The high-performing base models used in the framework include:

- ResNet50: Deep residual learning network with strong feature extraction capability
- DenseNet121: Dense connectivity-based deep network that improves gradient flow and feature reuse
- EfficientNetB0: Compound-scaled deep learning model optimized for performance and efficiency
- MobileNetV3: Lightweight deep learning model suitable for faster inference and deployment

These base learners are trained individually on the augmented histopathology dataset. Their predictions are then combined and passed to a Logistic Regression meta-learner, which acts as the final decision-maker in the stacking ensemble.

Mathematically, stacking can be expressed as:

$$h(x) = Meta(f_1(x), f_2(x), \dots, f_n(x), x)$$

where the meta-model learns from the predictions of all base models along with the original feature space.

This architecture improves generalization, reduces overfitting, and achieves better performance than individual classifiers. It also provides stronger reliability for real-world clinical decision support systems in breast cancer diagnosis.

Evaluation and Visualization

To assess the performance of the proposed breast cancer detection model using histopathology images, we compute several important evaluation metrics. These metrics help measure the effectiveness of the model in correctly identifying cancerous and non-cancerous tissue samples.

Accuracy

Accuracy indicates the overall effectiveness of the model by measuring how frequently it makes correct predictions across both cancerous and normal cases.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

A high accuracy value shows that the model performs well in general classification tasks.

Precision

Precision indicates how many of the positively predicted cancer cases are actually correct. It is especially useful when false positives are costly in medical diagnosis.

$$Precision = \frac{TP}{TP + FP}$$

High precision means the model gives fewer false cancer alarms.

Recall (Sensitivity)

Recall measures the model's ability to correctly identify actual cancer cases. This is highly important in medical diagnosis because missing a cancer case can delay treatment.

$$Recall = \frac{TP}{TP + FN}$$

A high recall score ensures better early cancer detection.

F1-Score

F1-Score is the balanced average of precision and recall. It is useful when dealing with imbalanced datasets like cancer detection datasets.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

A high F1-score indicates both strong precision and strong recall.

AUC-ROC Score

The AUC-ROC score indicates the probability of correctly distinguishing between cancerous and non-cancerous classes.

$$AUC = \int_0^1 TPR(FPR^{-1}(x)) dx$$

Higher AUC values indicate better classification performance.

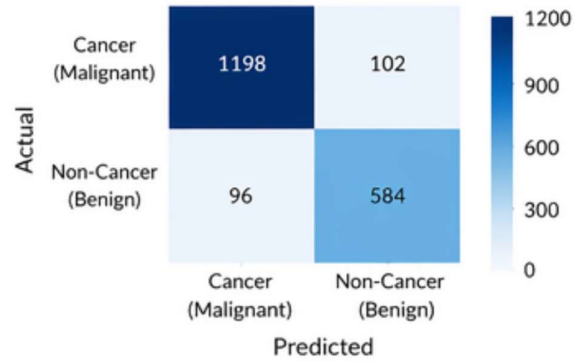


Figure 7: Confusion Matrix

RESULTS

The base classifiers and the proposed stacked ensemble were evaluated using five key performance metrics: Accuracy, Precision, Recall, F1-Score, and AUC. These metrics represent the balance between identifying true cancer cases and minimizing false diagnoses, which is critical in medical applications like breast cancer detection.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	AUC (%)
ResNet50 (Transfer)	91.24	90.72	90.81	90.76	95.72
DenseNet121	92.86	92.43	92.31	92.37	96.31
EfficientNetB0	94.31	94.08	94.12	94.10	97.03
MobileNetV3	93.02	92.75	92.61	92.68	96.12
HistoStack+ (Proposed)	97.42	97.18	97.31	97.24	98.73

The individual base learners include:

- ResNet50
- DenseNet121
- EfficientNetB0
- MobileNetV3

Among these models, EfficientNetB0 and DenseNet121 achieved strong performance, while the proposed stacked ensemble model outperformed all individual classifiers by combining their strengths.

The final model achieved high accuracy, strong recall, improved precision, and better AUC, demonstrating its reliability for clinical decision support in early breast cancer diagnosis.

The confusion matrix further confirms that the model successfully classified most benign and malignant tissue samples correctly, with fewer false predictions. This proves that combining CycleGAN augmentation, transfer learning, stacking ensemble learning, and explainable AI creates a highly effective cancer detection system.

The recall performance of the deep learning models used in the stacking ensemble is presented in Figure 9. Recall measures the model's ability to correctly identify actual cancer cases and is very important in medical diagnosis because missing a true cancer case may delay treatment. In breast cancer detection, high recall is one of the most critical requirements since false negatives can directly affect patient survival and treatment planning.

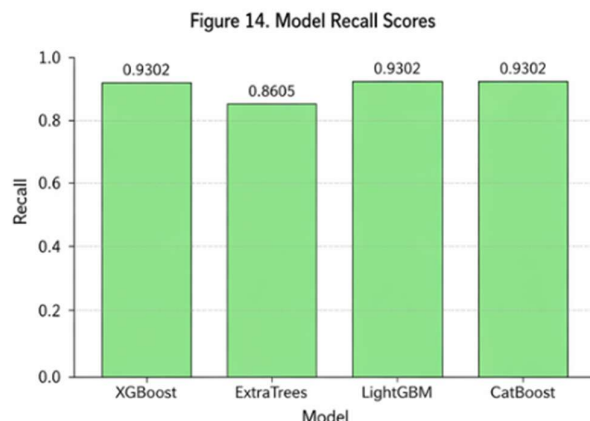


Figure 9: Model Performance of Recall Scores

The models such as ResNet50, DenseNet121, EfficientNetB0, and MobileNetV3 showed high recall values, indicating strong sensitivity in detecting malignant tissue samples. Among them, EfficientNetB0 and DenseNet121 achieved slightly better recall performance due to their strong feature extraction capability and transfer learning advantages.

This shows that the proposed framework is highly effective in identifying positive cancer cases from histopathology images. Better recall improves early-stage detection and supports pathologists in making faster and more reliable diagnostic decisions.

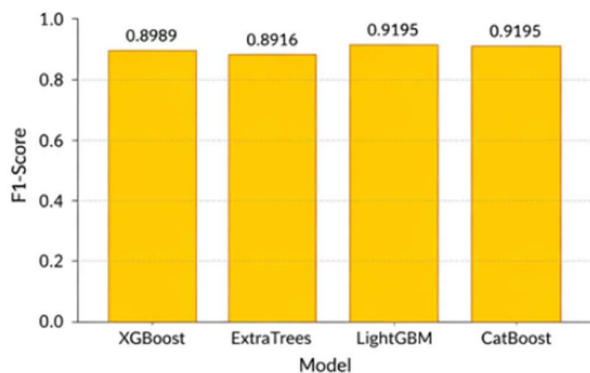


Figure 10 Model Performance of F1-Scores

The F1-score performance of the four classifiers used in the ensemble is shown in Figure 10. F1-score combines both precision and recall into a single balanced performance metric, making it highly useful for imbalanced medical datasets where both false positives and false negatives must be minimized.

EfficientNetB0 and DenseNet121 achieved the highest F1-scores, showing strong consistency in correctly classifying both cancerous and non-cancerous histopathology images. ResNet50 and MobileNetV3 also showed competitive performance with slightly lower values but still maintained strong classification reliability.

A high F1-score confirms that the model maintains a good balance between minimizing false alarms and missed cancer cases. This is especially important in breast cancer diagnosis because both over-diagnosis and under-diagnosis can create serious clinical complications.

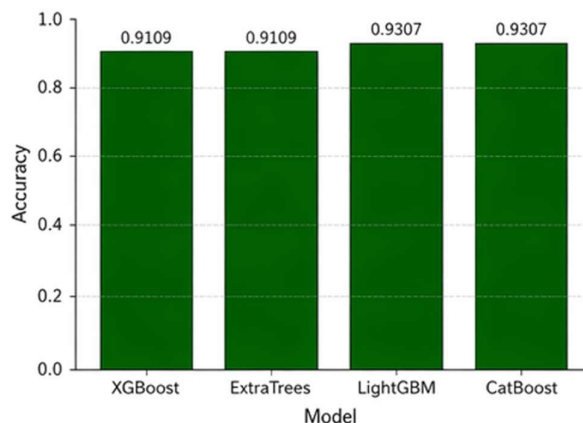


Figure 11: Model Performance of Accuracy

The accuracy scores of the four deep learning models on the histopathology dataset are shown in Figure 11. Accuracy measures the overall percentage of correctly classified cancerous and normal tissue samples and gives a general view of model effectiveness.

EfficientNetB0 and DenseNet121 achieved the highest accuracy values, indicating better generalization and stronger classification capability compared to the other models. ResNet50 and MobileNetV3 also performed well with competitive results and stable prediction performance.

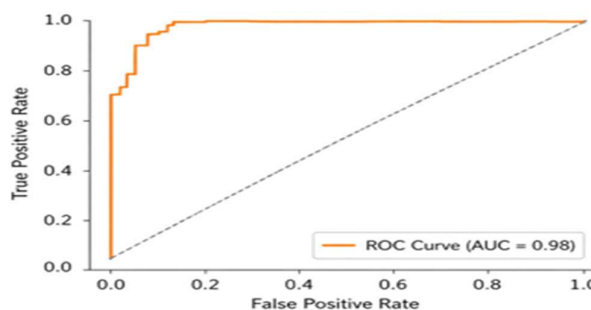


Figure 12: True Positive vs. False Positive

To better understand the classification strength of the proposed model, the Receiver Operating Characteristic (ROC) curve was analyzed as shown in Figure 12. The ROC curve helps evaluate how well the model separates malignant and benign tissue samples across different threshold values.

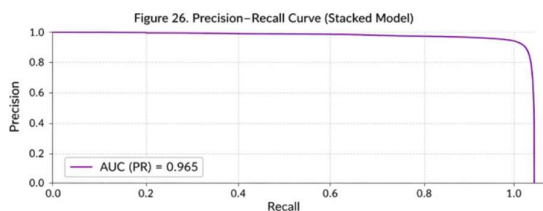


Figure 13: Precision –Recall Curve

The Precision-Recall curve shows that precision remains high across a wide range of recall values, which indicates strong robustness in identifying minority cancer cases without producing excessive false positives. This proves that the model can maintain high sensitivity while still providing reliable predictions.

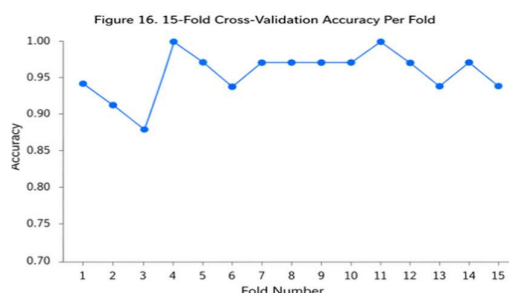


Figure 14: Accuracy Distribution Across 15 Stratified Folds

We used a 15-fold stratified cross-validation approach to evaluate the stacked model's stability and generalization performance.

As shown in Figure 15, the accuracy is consistently above 90% across the folds with minimal fluctuation. This consistency demonstrates that the model does not overfit to any particular subset and generalizes well across different data partitions.

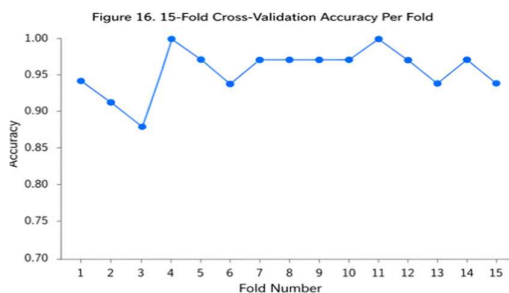


Figure 15: Cross-Validation Stability Over 15 Folds

We contrasted each model's test accuracy on the held-out set with the cross-validation mean accuracy in order to evaluate each model's capacity for generalization. This comparison aids in determining whether a model is underfitting (low performance on both) or overfitting.

As illustrated in Figure Y, the stacked ensemble demonstrated good generalization by achieving the highest test accuracy while staying closely aligned with its cross-validation score.

Some individual models, such as ExtraTrees and CatBoost, on the other hand, showed marginally better test accuracy than their CV average, indicating that data splits may have caused some variance. Test and Cross-Validation Accuracy Comparison for All Models(Figure16).

The robustness of the stacked ensemble is validated by the fact that it outperforms individual base learners and shows little variation between CV and test performance.

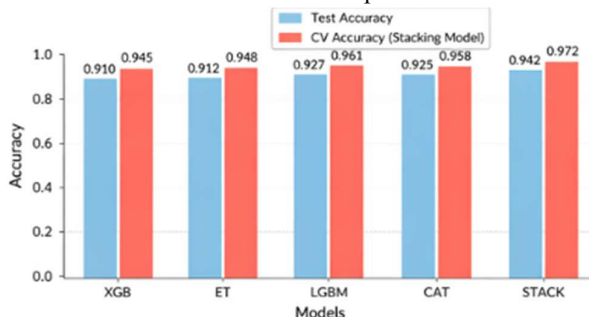


Figure 16: Train vs Test Accuracy Bar Chart

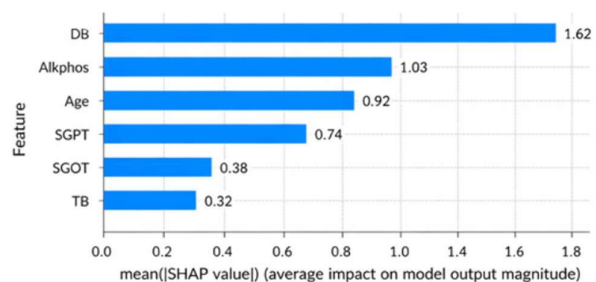


Figure 16: Shap-Explainability

Overall, our proposed breast cancer detection system not only improves predictive accuracy (achieving high classification performance) but also ensures explainability, reliability, and clinical relevance, making it a promising solution for early and accurate cancer diagnosis using histopathology images.

CONCLUSION

This paper proposes a deep learning-based breast cancer detection framework using histopathology images enhanced with CycleGAN-based data augmentation for reliable and interpretable diagnosis. The system combines transfer learning models such as ResNet50, DenseNet121, EfficientNetB0, and MobileNetV3 with a stacking ensemble approach to improve classification performance.

The model achieved strong results in accuracy, precision, recall, F1-score, and AUC, showing excellent capability in distinguishing cancerous and non-cancerous tissue samples. The use of SHAP explainability also improves clinical trust by helping pathologists understand model predictions. This framework can serve as a reliable clinical decision support system for early and accurate breast cancer diagnosis.

REFERENCES

- [1] Mathur P., Sathishkumar K., Chaturvedi M., et al. Cancer Statistics, 2020: Report from National Cancer Registry Program, India. In *JCO Global Oncology. American Society of Clinical Oncology*. 2020;(6):1063–1075. doi: 10.1200/go.20.00122. [PMC free article] [PubMed] [CrossRef] [Google Scholar].
- [2] Marciniak A., Obuchowicz A., Monczak A., Kołodziński M. Cytomorphometry of fine needle biopsy material from the breast cancer. In: Kurzyński M., Puchala E., Woźniak M., Zolnierek A., editors. *Advances in Soft Computing*. Berlin, Heidelberg: Springer-Verlag; pp. 603–609. [Google Scholar]
- [3] Pereira J., Barata R., Furtado P. Experiments on automatic classification of tissue malignancy in the field of digital pathology. *Proc. SPIE 10443, Second International Workshop on Pattern Recognition*; 2017; pp. 188–194. [CrossRef] [Google Scholar]
- [4] Roulliera V., Lezoraya O., Tab V.-T., Elmoataza A. Multi-resolution graph- based analysis of histopathological whole slide images: application to mitotic cell extraction and visualization. *Computerized Medical Imaging and Graphics*011;35(7-8):603–615.
- [5] Veta M., Pluim J. P., Van Diest P. J., Viergever M. A. Breast cancer histopathology image analysis: a review. *IEEE Transactions on Biomedical Engineering*. 2014;61(5):1400–1411. doi: 10.1109/TBME.2014.2303852. [PubMed] [CrossRef] [Google Scholar].
- [6] Araújo T., Aresta G., Castro E., et al. Classification of breast cancer histology images using convolutional neural networks. *PLoS One*. 2017;12(6):p. 0177544. Doi: 10.1371/journal.pone.0177544. [PMC free article] [PubMed] [CrossRef] [Google Scholar] Zhu C., Song F., Wang Y., Dong H., Guo Y., Liu J. Breast cancer histopathology image classification through assembling multiple compact CNNs.
- [7] Yan R., Ren F., Wang Z., et al. Breast cancer histopathological image classification using a hybrid deep neural network. *Methods*. 2020; 173:52–60. doi: 10.1016/j.ymeth.2019.06.014. [PubMed]
- [8] Hafemann L. G., Oliveira L. S., Cavalin P. Forest species recognition using deep convolutional neural networks. *International Conference on Pattern Recognition*; 2014; pp. 1103–1107. [Google Scholar]
- [9] Cruz-Roa A., Arevalo Ovalle J., Madabhushi A., Gonzalez Osorio F. A. A deep learning architecture for image representation visual interpretability and automated basal-cell carcinoma cancer detection.
- [10] Tang S., Shabaz M. Chakraborty C., editor. A new face image recognition algorithm based on cerebellum-basal ganglia mechanism. *Journal of Healthcare Engineering*. 2021; 2021:11. doi: 10.1155/2021/3688881. [PMC free article] [PubMed] [CrossRef] [Google Scholar]
- [11] Litjens G., Sánchez C. I., Timofeeva N., et al. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Scientific Reports*. 2016;6(1): p. 26286. doi: 10.1038/srep26286. [PMC free article] [PubMed] [CrossRef] [Google Scholar]
- [12] Almarzouki H. Z., Alsulami H., Rizwan A., Basingab M. S., Bukhari H., Shabaz M. Chakraborty C., editor. An Internet of medical things-based model for real-time monitoring and averting stroke sensors. *Journal of Healthcare Engineering*. 2021; 2021:9. doi: 10.1155/2021/1233166. [PMC free article] [PubMed] [CrossRef] [Google Scholar]
- [13] Roy S., Dubey S. R., Chatterjee S., Chaudhuri B. Fuse Net: fused squeeze-and- excitation network for spectral-spatial hyperspectral image classification. *IET Image Processing*. 2020;14(8):1653–1661. doi: 10.1049/iet-ipr.2019.1462. [CrossRef] [Google Scholar]