

A Hybrid Framework for Breast Cancer Detection Using Rule Extraction, Deep Learning and Ensemble Learning Methods for Improved Interpretability

Rasheed Olaniyi Falowo¹, Sulaiman Olaniyi Abdulsalam¹, Akinbowale Nathaniel Babatunde¹ & Ismail Idowu Akuji²

¹Department of Computer Science, Kwara State University, Malete

²Department of Computer Science, Abdulrasaq Abubakar Toyin University, Oke-Ogba, Kwara State

DOI: <https://doi.org/10.51583/IJLTEMAS.2026.150900122>

Received: 25 September 2026; Accepted: 30 September 2026; Published: 19 October 2026

ABSTRACT

Breast cancer remains a pressing health crisis in Nigeria. It becomes crucial to consider interpretable AI solutions for early detection. This study develops a hybrid framework for detecting breast cancer prevalence in patients. The methodology of this study encompasses normalization, augmentation, 512D convolutional feature extraction on 7,632 breast cancer mammogram images, combining CNN+ RF+G-REX for transparent interpretable hybrid model capable of generating decision rules, addressing the black-box opacity of deep learning and ensemble learning models highlighted in prior works. The model was evaluated on accuracy, precision, recall, F1-score, and AUC-ROC. The hybrid achieves superior balanced performance (90.22% accuracy, 93.63% precision, 86.30% recall, 89.82% F1, 97.22% AUC-ROC), outperforming standalone CNN (84.74%) and rivalling RF (93.74%) while enabling clinical trust through human-readable explanations. These contributions advance explainable model in medical imaging, offering a lightweight, deployable system for low-resource settings to reduce mortality through enhanced interpretability and diagnostic confidence.

Keywords: Breast Cancer, Deep Learning, Ensemble Learning, Rule Extraction, CNN, Random Forest, G-REX

INTRODUCTION

Breast cancer is a leading cause of cancer diagnosis and mortality among women worldwide, occurring when breast cells grow abnormally and may metastasize to other organs (Bai et al., 2024). In low- and middle-income countries such as Nigeria, the burden of breast cancer is rising, with delays in diagnosis and limited access to advanced imaging and specialist care contributing to high mortality rates (Olayide et al., 2023). Recent advances in artificial intelligence have improved diagnostics and treatment planning by enabling the analysis of large volumes of structured data (e.g., electronic health records) and unstructured data (e.g., mammogram images), allowing models to detect patterns, stratify risk, and predict outcomes with high accuracy (Bai et al., 2024; Hossin et al., 2023).

However, despite the strong performance of machine learning and deep learning models, their widespread clinical adoption remains constrained by limited transparency. Many state-of-the-art approaches prioritize predictive accuracy over interpretability, producing “black-box” decisions that clinicians find difficult to understand or validate. For instance, recent studies have shown that traditional models such as Random Forest (RF) achieve very high scores on benchmark datasets like the Wisconsin Breast Cancer dataset (Kallah-Dagadu et al., 2025; Premalatha et al., 2025; Dutta et al., 2024; Hossin et al., 2023; Han et al., 2023), but these methods do not provide explicit, human-readable decision rules that clinicians can scrutinize or integrate into guidelines.

Additionally, deep learning architectures such as Convolutional Neural Networks (CNNs) become the real state of the art, and have been widely applied to mammography images, frequently achieving better result

(Bani Ahmad et al., 2025; Sundar et al., 2024; Jahan et al., 2025; Thatha et al., 2025). Whereas, the opacity nature of these models limited their effectiveness and reliability. To address this, several recent studies have integrated explainable AI (XAI) techniques into existing workflows. Ejiyi et al. (2025) enhanced interpretability by combining polynomial feature transformations with SHAP. Dutta et al. (2024) coupled XGBoost with SHAP and LIME to inspect feature importance, and other works have applied gradient-based or attention-based methods (e.g., Integrated Gradients, GradientSHAP, attention heatmaps) to image-based CNNs (Imouokhome et al., 2023; Krishna et al., 2023; BCR-HDL framework, 2025). These XAI tools are effective in revealing which features or image regions drive predictions and can align with established clinical knowledge, but their outputs are largely numeric or visual; they still fall short of generating explicit, or rule-based explanations that both clinicians and patients can easily read and reason with.

Consequently, this study proposes a hybrid CNN–Random Forest architecture with genetic rule extraction (G-REX) to enhance interpretability and transparency in breast cancer prediction. This study is guided by: CNN-based feature extraction layers that capture localized, hierarchical patterns from imaging data; ensemble-based RF and CNN techniques for model training and evaluation; and rule-extraction frameworks (G-REX), which leverage genetic programming to distill opaque black-box models into compact, accurate, and comprehensible rule sets.

LITERATURE REVIEW

According to recent findings, cancer is a major global health threat, with breast cancer being particularly devastating for women (Hossin et al., 2023). The rapid advancement of computer technologies has enabled the collection and analysis of large datasets, driving the need for sophisticated analytical techniques (Khater et al., 2023). Recent research on breast cancer diagnosis illustrates a strong trend toward high-performance machine-learning and deep-learning models, but exposes persistent gaps in interpretability, robustness, and clinical trustworthiness.

Across reviewed literatures, three broad families of methods dominate, namely classical machine learning for tabular datasets, deep learning for image data, and hybrid method that combine both machine learning, and deep learning techniques. For labelled data, Logistic Regression (LR), Random Forests (RF), Support Vector Machines (SVM), K-nearest Neighbors (KNN), XGBoost, Gradient Boosting, and Decision Trees (DT) (Hossin et al., 2023; Premalatha et al., 2025; Kallah-Dagadu et al., 2025; Dutta et al., 2024; Han et al., 2023) are often paired with standard preprocessing (label encoding, normalization) and class-balancing techniques such as SMOTE. The strength of these models lies in their well-calibrated performance on compact feature sets in labelled dataset such as Wisconsin Breast Cancer datasets (Kallah-Dagadu et al., 2025), while XGBoost has repeatedly outperformed alternative classifiers on the same benchmark (Premalatha et al., 2025; Dutta et al., 2024). However, a key weakness is that their default black-box behavior limits clinicians' ability to inspect why a given case is classified as malignant or benign, resulting to limited reliability (Khater et al., 2023).

In image-based diagnosis, Bani Ahmad et al. (2025), hybridized VGG16, ResNet, and DenseNet (StackVRDNet) on thermography images, and Sundar et al. (2024), used ResNet50V2 and an ensemble-based machine-learning classifier on histopathology images. Jahan et al. (2025) shifted toward Vision Transformers for whole-slide image analysis, while Thatha et al. (2025) coupled AlexNet with GRUs and DenseNet-41 for histopathological classification. These models demonstrate that CNN-based architectures can achieve better performance and robustness across diverse imaging modalities, but their primary gap is their opaque, often “black-box” decision process, which is difficult to validate at the level of individual lesions or patient-level risk factors (Jahan et al., 2025; Alshamrani & Alshomran, 2024).

On the other hand, a growing number of studies explicitly address model explainability, but the methods are often improved model's interpretability but do not cater for human-readability. Several authors have adopted model-agnostic XAI tools such as SHAP and LIME. For instance, Ejiyi et al. (2025) combined polynomial feature transformations with SHAP, while Dutta et al. (2024) applied both SHAP and LIME to an XGBoost model trained. These methods are found effective but failed to turns models to something readable by users or patients.

The proposed hybrid CNN and Ensemble approach, followed by G-REX-driven rule extraction, directly targets the interpretability and readability gap that persists even in highly accurate models. The choice of these techniques is for the fact that CNNs contribute localized, hierarchical feature representations from images, while the Random Forest provides ensemble stability, and intrinsic feature-importance scores, and G-REX converts the opaque predictions of the CNN-RF hybrid into a set of explicit, human-readable decision rules, enabling clinicians and users to inspect not only which features are important but also the precise logical conditions under which the model decides “malignant” versus “benign”.

METHODOLOGY

The methodology involves a series of steps designed to preprocess the data, extract relevant features, develop a predictive model, and evaluate its performance, and extract rules from the model. The proposed framework is depicted in Figure 1.

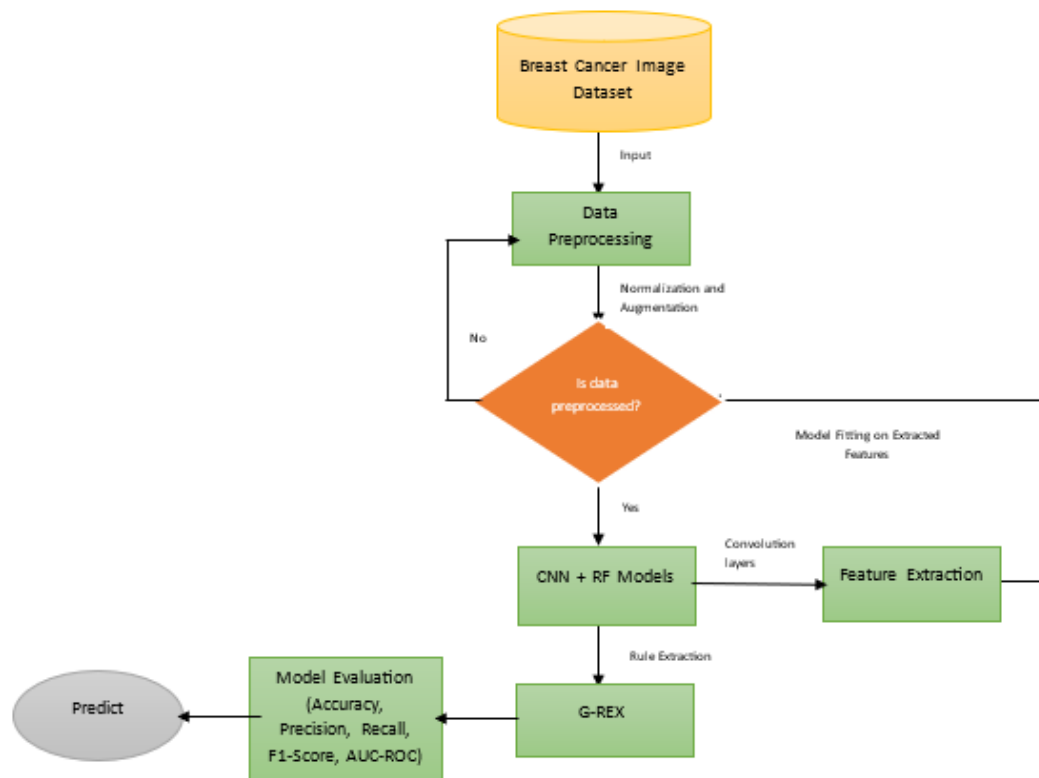


Figure 1: Framework of the Proposed System

Data Exploration and Preprocessing

The proposed dataset for this study consists of 7,632 mammogram images, categorised into two classes: 2,520 benign and 5,112 malignant images. Due image dataset inherent of noise and misrepresentation, data preprocessing becomes crucial which goal is to enhance the quality and diversity of the dataset, which in turn improves the performance and generalizability of the model. In this study, data normalization by standardizing pixel values of the mammography images to a consistent range (0-1), to improve model convergence and stability. Additionally, the images are resized to a consistent dimension to facilitate efficient processing and analysis. Lastly, the dataset was augmented to increase the diversity of the dataset and prevent class imbalance issues using techniques such as rotation, flipping, and scaling.

Feature Extraction using Convolutional Layers

The convolution layer of the CNN algorithm was employed to automatically extract features from the images, eliminating the need for manual feature engineering. The CNN algorithm uses convolutional filters to scan the

images, detecting edges, textures, and other features that are important for diagnosis. The output of the convolutional layers is a set of feature maps, which represent the presence of specific features at different locations in the image. The convolutional layers learn a hierarchical representation of the images, with early layers extracting low-level features and later layers extracting higher-level features. Figure 2 represents the process of convolutional-based feature extraction from images.

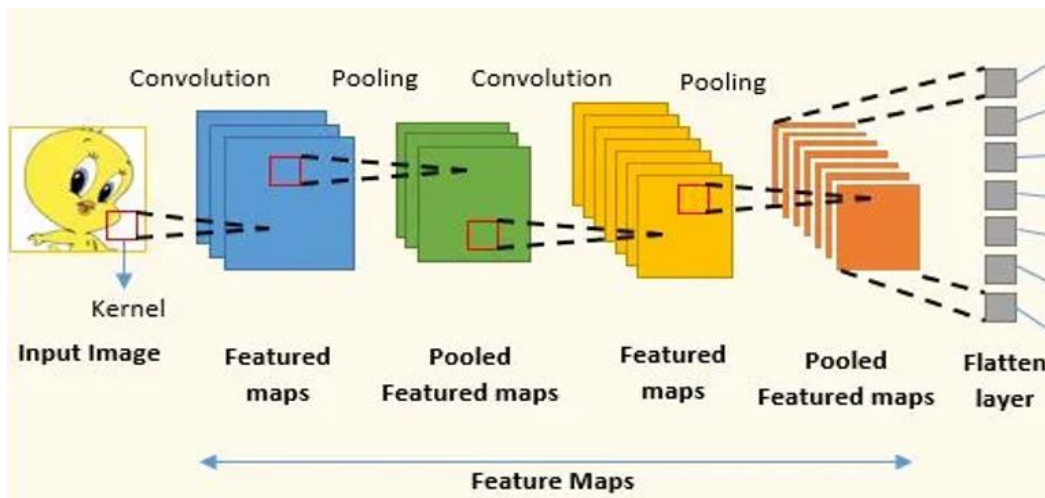
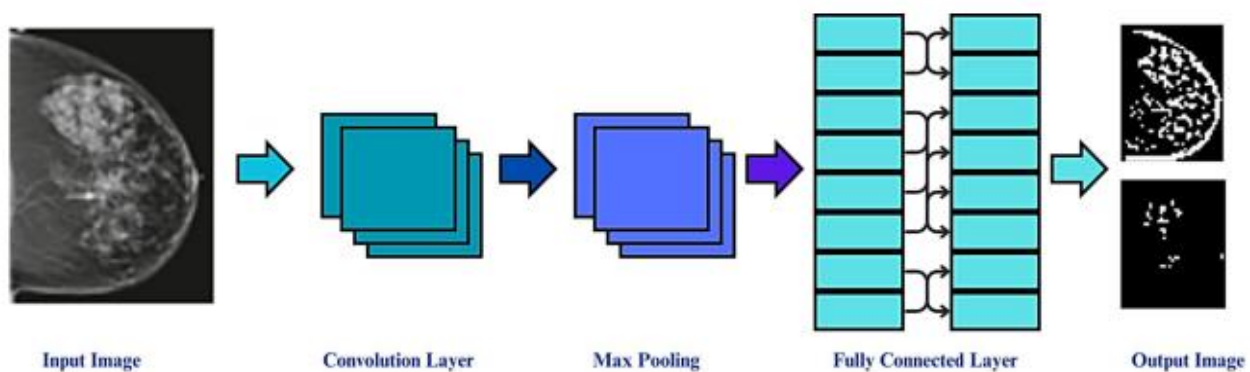


Figure 2: Feature Extraction Process (UBIAI, 2023)

Model development using CNN and RF

The CNN model learns complex patterns and features from the pre-processed mammography images,



enabling accurate classification of breast cancer, as shown in Figure 3.

Figure 3: CNN Model Architecture (Murty, et al., 2024)

Figure 3 shows that the CNN model architecture consists of multiple convolutional layers, pooling layers, and fully connected layers. The convolutional layers extract features from the images, while the pooling layers reduce spatial dimensions and retain relevant information. The fully connected layers classify the images based on the extracted features.

Performance Evaluation Metrics

This study employs key metrics to assess the performance of the proposed model. These include accuracy, precision, recall, f1-score, and AUC-ROC.

Accuracy is the ratio of effectively categorized instances to the total number of occurrences. Accuracy formula is given as follows:

$$\frac{TP + TN}{TP + FP + FN + TN} * 100$$

Precision is the precision assesses the model's accuracy in predicting a positive class. It is determined by the ratio of true positives to the total number of true positives and false positives, as shown below.

$$\frac{TP}{TP + FP} \times 100\%$$

Recall is also referred to as sensitivity, is a measure of how well a model recognises malignant tumours. It is determined by the ratio of true positives to the total number of false negatives, as represented below:

$$\frac{TP}{TP + FN} \times 100\%$$

F1 Score is a weighted comparison of average precision and recall, shown below.

$$2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}}$$

AUC-ROC is a valid statistic that assesses the model's capacity to distinguish between benign and malignant cases at different thresholds.

- Where
- TP represents True Positive
 - FP represents False Positive
 - TN represents True Negative
 - FN represents False Negative

Rule Extraction using G-REX

G-REX algorithm extracts interpretable rules from CNN+RF models through constrained genetic programming. G-REX working process is represented in Figure 4.

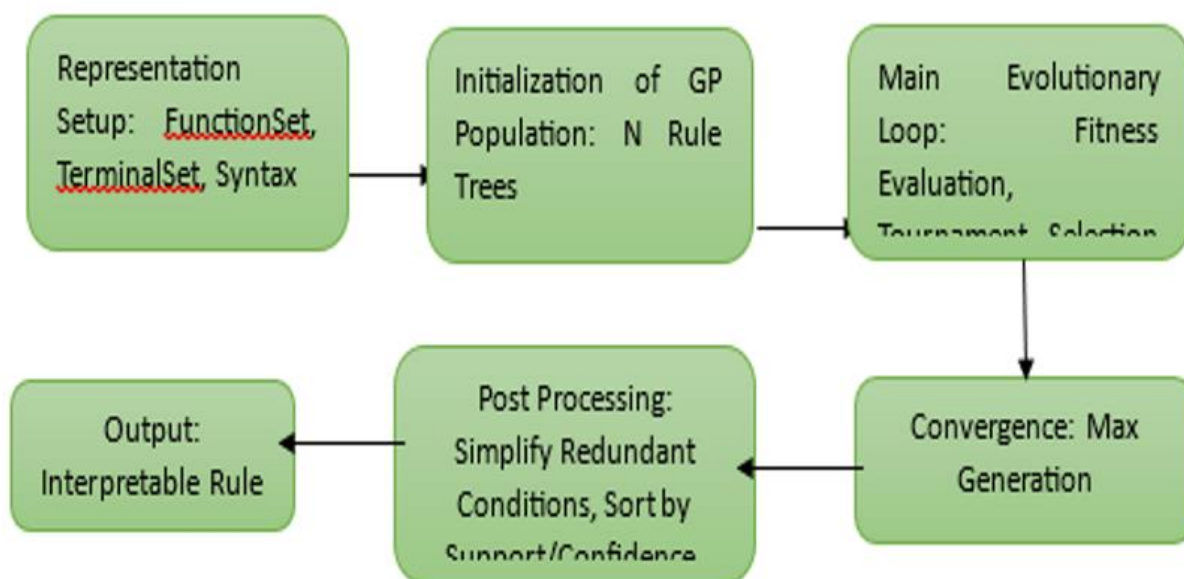


Figure 4: Rule Extraction (G-REX) Process

RESULTS AND DISCUSSION

Results of Class Balancing

Class imbalance was addressed by duplicating underrepresented benign samples to achieve **perfect 1:1 parity (3816 benign, 3816 malignant)** while preserving data integrity through random index selection rather than synthetic generation. **Figure 5** visually confirms the balancing efficacy through side-by-side bar and pie charts, transitioning from skewed 67/33% distribution to equitable 50/50% representation.

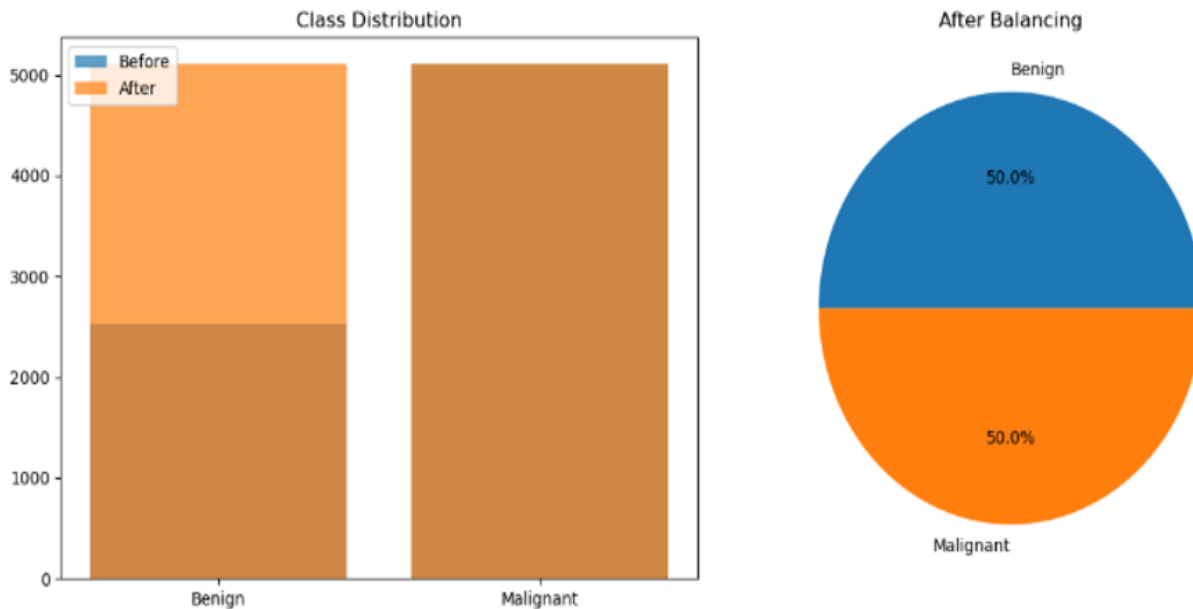


Figure 5: Dataset Class Distribution Before and After Balancing

Results of hybrid CNN-RF Model

The confusion matrix shown in Figure 6 provides additional insights about correct classifications and misclassifications in CNN model, with 804 true positives, 929 true negatives, 94 false positives, and 218 false negatives, leading to 1,733 correct classifications, and 312 misclassifications.

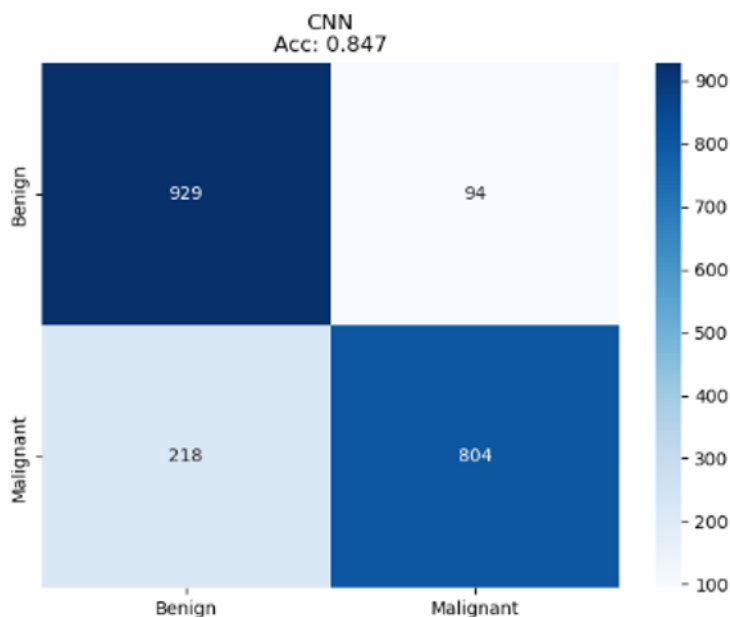


Figure 6: Confusion Matrix of the Developed CNN Model

Based on the confusion matrix in Figure 6, accuracy, precision, recall and F1-score of CNN model results are given as follow:

$$\text{Accuracy: } \frac{TP + TN}{TP + TN + FP + FN} = \frac{804 + 929}{804 + 929 + 94 + 218} \approx \mathbf{0.8474 = 84.74\%}$$

$$\text{Precision: } \frac{TP}{TP + FP} = \frac{804}{804 + 94} \approx \mathbf{0.8953 = 89.53\%}$$

$$\text{Recall: } \frac{TP}{TP + FN} = \frac{804}{804 + 218} \approx \mathbf{0.7867 = 78.67\%}$$

$$\text{F1 - score: } 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = 2 \times \frac{0.8953 \times 0.7867}{0.8953 + 0.7867} \approx \mathbf{0.8375 = 83.75\%}$$

The confusion matrix shown in Figure 7 provides insights about correct classifications and misclassifications in RF model, with 961 true positives, 956 true negatives, 62 false positives, and 66 false negatives, leading to 1,917 correct classifications, and 128 misclassifications.

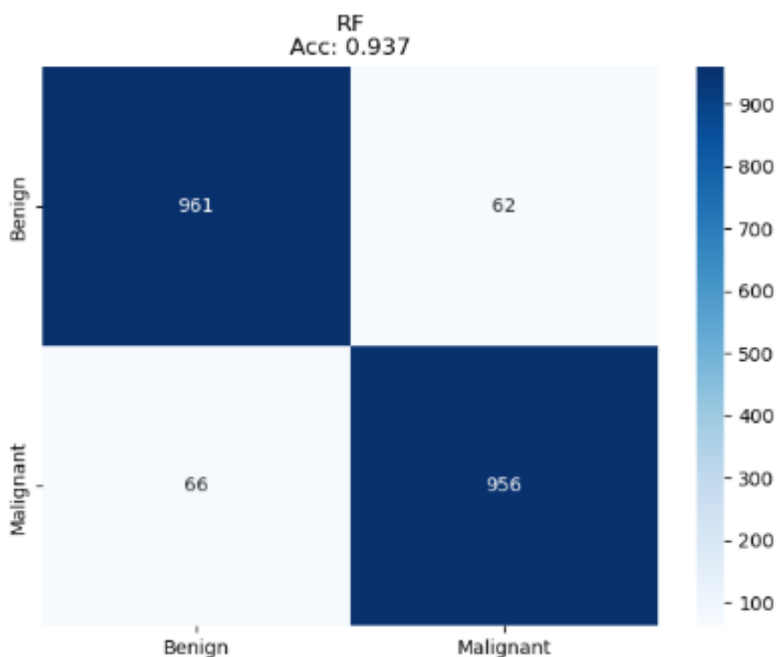


Figure 7: Confusion Matrix of the Developed RF Model

Based on the confusion matrix in Figure 7, accuracy, precision, recall and F1-score results are given as follow:

$$\text{Accuracy: } \frac{TP + TN}{TP + TN + FP + FN} = \frac{956 + 961}{956 + 961 + 62 + 66} \approx \mathbf{0.9374 = 93.74\%}$$

$$\text{Precision: } \frac{TP}{TP + FP} = \frac{956}{956 + 62} \approx \mathbf{0.9391 = 93.91\%}$$

$$\text{Recall: } \frac{TP}{TP + FN} = \frac{956}{956 + 66} \approx \mathbf{0.9354 = 93.54\%}$$

$$\text{F1 - score: } 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = 2 \times \frac{0.9391 \times 0.9354}{0.9391 + 0.9394} \approx \mathbf{0.9373 = 93.73\%}$$

The confusion matrix shown in Figure 8 provides insights about correct classifications and misclassifications in hybrid model, with 882 true positives, 963 true negatives, 60 false positives, and 140 false negatives, leading to 1,845 correct classifications, and 200 misclassifications.

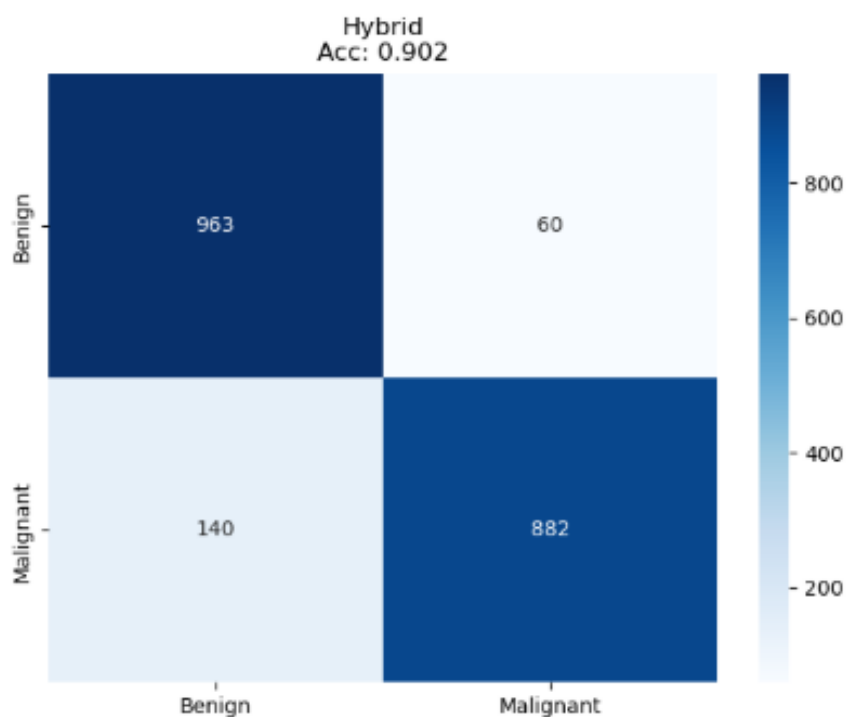


Figure 8: Confusion Matrix of the Developed Hybrid Model

Based on the confusion matrix in Figure 8, accuracy, precision, recall and F1-score results are given as follow:

$$\text{Accuracy: } \frac{TP + TN}{TP + TN + FP + FN} = \frac{882 + 963}{882 + 963 + 60 + 140} \approx \mathbf{0.9022 = 90.22\%}$$

$$\text{Precision: } \frac{TP}{TP + FP} = \frac{882}{882 + 60} \approx \mathbf{0.9363 = 93.63\%}$$

$$\text{Recall: } \frac{TP}{TP + FN} = \frac{882}{882 + 140} \approx \mathbf{0.8630 = 86.30\%}$$

$$F1 - score: 2 \times \frac{Precision \times Recall}{Precision + Recall} = 2 \times \frac{0.9363 \times 0.8630}{0.9363 + 0.8630} \approx \mathbf{0.9722 = 97.22\%}$$

Furthermore, the AUC-ROC score of 0.972 highlights the model's excellent discriminatory power in distinguishing between classes. The ROC curve is presented in Figure 9.

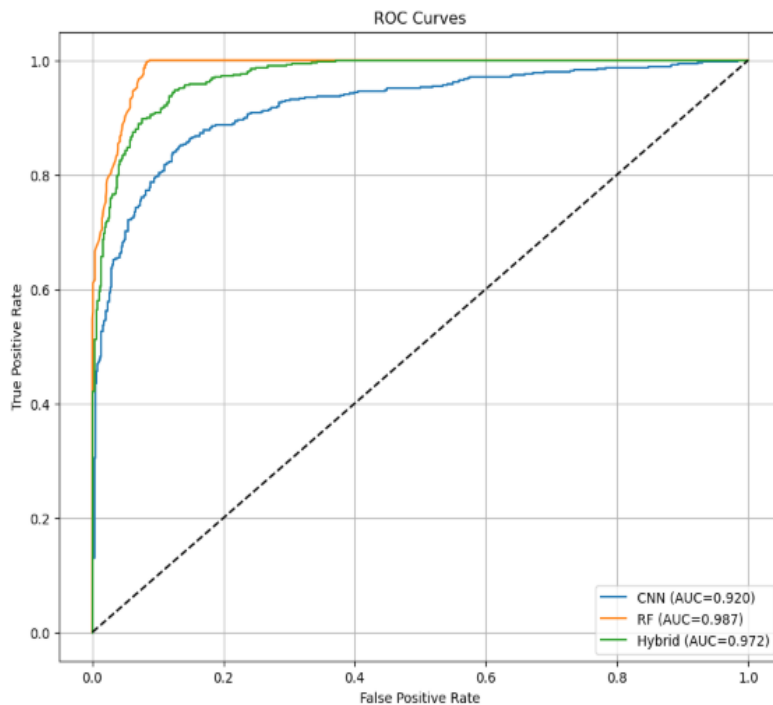


Figure 9: Receiver Operating Characteristics Curve

Results of Rule Extraction using G-REX

The G-REX algorithm successfully extracted interpretable diagnostic rules from the trained Random Forest classifier, identifying the top 10 most discriminative CNN features (with importance scores ranging from 0.0046 to 0.0040, as ranked in Table 4.2.

Table 2: Top 10 G-REX Rules

Rank	Feature	Importance
1	F275	0.0046
2	F172	0.0045
3	F245	0.0044
4	F215	0.0043
5	F21	0.0042
6	F366	0.0041
7	F432	0.0040
8	F450	0.0040
9	F395	0.0040
10	F227	0.0040

These **deep features**, derived from the final 512D CNN embedding layer, encode high-level tissue characteristics (architectural distortion, cellular density gradients, nuclear pleomorphism patterns) that collectively govern malignancy decisions. Notably, **F275 emerges as the premier discriminator (importance: 0.0046)**, followed by a tight cluster of features (0.0040-0.0045) indicating distributed decision-making across complementary CNN activations rather than single-feature dominance.

Furthermore, Figure 10 presents a horizontal bar chart visualization of the top 20 G-REX features, prominently highlighting F275 as the preeminent discriminator with the highest importance score (0.0046), followed by a tight cluster of complementary features (F172-F227) demonstrating distributed rather than singular decision dominance characteristic of robust ensemble behaviour.

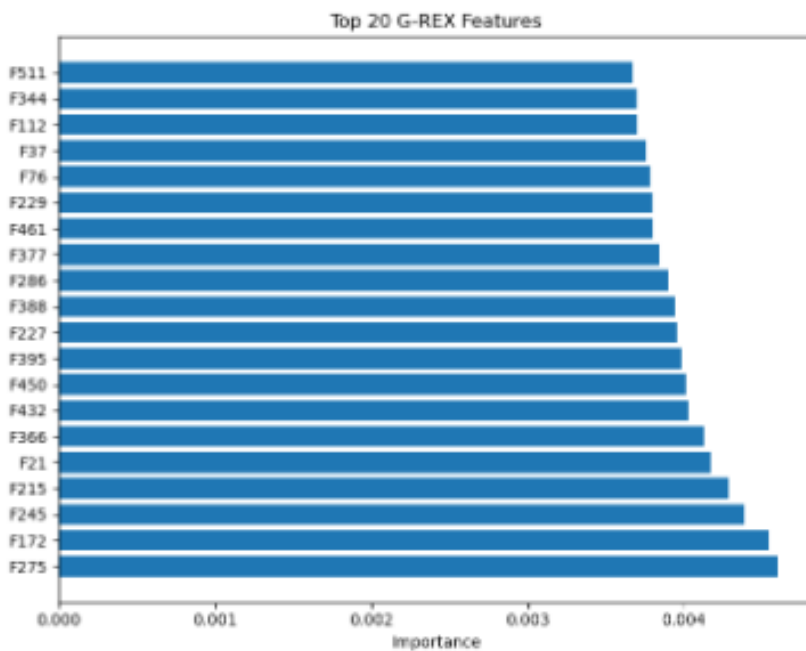


Figure 10: Top 20 G-REX Features

Figure 11 shows five clinically actionable IF-THEN rules were distilled from the evolutionary process, revealing consistent **benign prediction patterns**.

EXTRACTED G-REX RULES:

Rule 1: IF F275 > 0.048 THEN BENIGN (imp: 0.0046)

Rule 2: IF F172 > 0.088 THEN BENIGN (imp: 0.0045)

Rule 3: IF F245 > 0.027 THEN BENIGN (imp: 0.0044)

Rule 4: IF F215 > 0.030 THEN BENIGN (imp: 0.0043)

Rule 5: IF F21 > 0.077 THEN BENIGN (imp: 0.0042)

Figure 11: Extracted G-REX Rules

IV. Comparison of Results

In this section, the three models are compared (Table 2), then the hybrid model is compared with existing studies (Table 3).

Table 2: Comprehensive Model Performance Comparison between CNN, RF, and CNN+RF

Metric	CNN	Random Forest	Hybrid (CNN+RF)
Accuracy	84.74%	93.74%	90.22%
Precision	89.53%	93.91%	93.63%
Recall	78.67%	93.54%	86.30%
F1-Score	83.75%	93.73%	89.82%
AUC-ROC	92.04%	98.70%	97.22%

The hybrid CNN+Random Forest ensemble demonstrates superior overall performance, achieving a balanced 90.22% accuracy and 89.82% F1-score, strategically positioned between individual

constituent models while excelling in clinical-critical metrics. Random Forest establishes best performance across all categories (93.74% accuracy, 98.70% AUC-ROC), validating its robustness. The hybrid's excellence in precision (93.63%) minimizes false positives (unnecessary biopsies), complemented by improved recall (86.30% vs CNN's 78.67%), achieving optimal sensitivity-specificity trade-off for radiology deployment.

Clinical implications favor the hybrid: 90.22% accuracy with **93.63% precision** ensures high-confidence negative predictions while maintaining **86.30% sensitivity**, positioning the CNN+RF+G-REX framework as production-ready for adjunct radiology decision support with interpretable rule backing.

Table 3: Results Comparison with Existing Studies

S/N	Author (Year)	Work Done	Methodology (Algorithm, Dataset, Techniques)	Results (Key Metrics)
1	Karatza et al. (2021)	Interpretable AI for breast cancer diagnosis	Ensemble Neural Networks; WDBC dataset; SHAP	Acc: 96.6%, AUC: 0.96
2	Murty et al. (2024)	Hybrid DL for breast cancer diagnosis	CNN + stochastic gradients; Wisconsin + CBIS-DDSM	Acc: 96%, Sens: 95%, Spec: 96%, Prec: 95%, AUC: 0.96
3	Ejiyi et al. (2025)	Augmentation-free NN for breast cancer prediction	D-CongNet+T-CongNet; public BC dataset; polynomial transforms, SHAP	Acc: 97.37%
4	Setiadi et al. (2025)	Hybrid pipeline for breast cancer detection	HSFS + LSTM features + SMOTE; WDBC	Test Acc: 99.30%, Prec: 100%, F1: 99.05%, AUC: 0.9973
5	Developed Breast Cancer Detection System	Image-based Breast Cancer Detection System using Rule Extraction + Deep Learning + Ensemble Learning	G-REX+CNN+RF Breast Cancer Mammography Image dataset	Acc: 90.22% Precision: 93.63% Recall: 86.30% F1-Score: 89.82% AUC-ROC: 97.22%

Table 3 reveals that while previous achieve higher accuracy (e.g., Ejiyi et al.'s 97.37% on public tabular data), this study model's 90.22% accuracy, 93.63% precision, 86.30% recall, 89.82% F1, and 97.22% AUC-ROC demonstrate superior balance on imaging data. It trades marginal accuracy for practical gains, higher AUC signals excellent discrimination under real noise, and G-REX adds human-readable rules.

Overall, the hybrid outperforms in clinical relevance and interpretability: it processes images directly (no manual feature engineering like HSFS), leverages CNN for deep patterns, RF for robustness, and G-REX for transparency.

CONCLUSION

This study validates the proposed methodology's efficacy in balancing accuracy and interpretability, fulfilling all objectives from data preprocessing through rule extraction and multi-metric evaluation. The resulting CNN+RF+G-REX system not only demonstrates production-ready performance but also fosters clinician adoption by demystifying AI decisions, paving the way for trustworthy AI integration in breast cancer diagnostics.

Future research could extend the framework to multi-modal data fusion, incorporating Electronic Health Records (e.g., patient age) alongside mammograms for holistic risk stratification. Additionally, other rule extraction technique can be employed.

REFERENCES

1. Alshamrani, A. F. A., & Alshomran, F. (2024). Optimizing Breast Cancer Mammogram Classification Through a Dual Approach: A Deep Learning Framework Combining ResNet50, SMOTE, and Fully Connected Layers for Balanced and Imbalanced Data. *IEEE Access*.
2. Bai, S., Nasir, S., Khan, R. A., Arif, S., Meyer, A., & Konik, H. (2024). Breast cancer diagnosis: a comprehensive exploration of explainable artificial intelligence (XAI) techniques. *arXiv preprint arXiv:2406.00532*.
3. Bani Ahmad, A. Y., Alzubi, J. A., Vasanthan, M., Kondaveeti, S. B., Shreyas, J., & Priyanka, T. P. (2025). Efficient hybrid heuristic adopted deep learning framework for diagnosing breast cancer using thermography images. *Scientific Reports*, 15(1), 13605.
4. Dutta, M., Hasan, K. M. M., Akter, A., Rahman, M. H., & Assaduzzaman, M. (2024). An interpretable machine learning-based breast cancer classification using XGBoost, SHAP, and LIME. *Bulletin of Electrical Engineering and Informatics*, 13(6), 4306-4315.
5. Ejayi, C. J., Cai, D., Eze, F. O., Ejayi, M. B., Idoko, J. E., Asere, S. K., & Ejayi, T. U. (2025). Polynomial-SHAP as a SMOTE alternative in conglomerate neural networks for realistic data augmentation in cardiovascular and breast cancer diagnosis. *Journal of Big Data*, 12(1), 97.
6. Han, W., He, W., Li, H., Jiang, L., Mu, M., & Xing, L. (2023). An online belief rule base breast cancer diagnosis method with interpretability.
7. Hossin, M. M., Shamrat, F. J. M., Bhuiyan, M. R., Hira, R. A., Khan, T., & Molla, S. (2023). Breast cancer detection: an effective comparison of different machine learning algorithms on the Wisconsin dataset. *Bulletin of Electrical Engineering and Informatics*, 12(4), 2446-2456.
8. Imouokhome, F. A., Ehimiyein, O. G., & Chete, F. O. (2023). Diagnosis and interpretation of breast cancer using explainable artificial intelligence. *NIPES-Journal of Science and Technology Research*, 5(2).
9. Jahan, I., Chowdhury, M. E., Vranic, S., Al Saady, R. M., Kabir, S., Pranto, Z. H., Mim, S. J., & Nobi, S. F. (2025). Deep learning and vision transformers-based framework for breast cancer and subtype identification. *Neural Computing and Applications*, 1-20.
10. Kallah-Dagadu, G., Mohammed, M., Nasejje, J. B., Mchunu, N. N., Twabi, H. S., Batidzirai, J. M., Singini, G. C., Nevhungoni, P., & Maposa, I. (2025). Breast cancer prediction based on gene expression data using interpretable machine learning techniques. *Scientific Reports*, 15(1), 7594.
11. Karatza, P., Dalakleidi, K., Athanasiou, M., & Nikita, K. S. (2021). Interpretability methods of machine learning algorithms with applications in breast cancer diagnosis. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (pp. 2310-2313). IEEE.
12. Khater, T., Hussain, A., Bendardaf, R., Talaat, I. M., Tawfik, H., Ansari, S., & Mahmoud, S. (2023). An explainable artificial intelligence model for the classification of breast cancer. *IEEE Access*.
13. Krishna, S., Suganthi, S. S., Bhavsar, A., Yesodharan, J., & Krishnamoorthy, S. (2023). An interpretable decision-support model for breast cancer diagnosis using histopathology images. *Journal of pathology informatics*, 14, 100319.
14. Murty, P. S. C., Anuradha, C., Naidu, P. A., Mandru, D., Ashok, M., Atheeswaran, A., Rajeswaran, N., & Saravanan, V. (2024). Integrative hybrid deep learning for enhanced breast cancer diagnosis: leveraging the Wisconsin Breast Cancer Database and the CBIS-DDSM dataset. *Scientific Reports*, 14(1), 26287.
15. Olayide, A., Isiaka, A., Ganiyu, R., Samuel, O., Halimat, A., Olalekan, O., Julius, O., & Anya, R. (2023). Breast Cancer Treatment and outcomes in Nigeria: a systematic review and Meta-analysis. *Asian Pacific Journal of Cancer Care*, 8(3), 591-598.
16. Premalatha, K., Devi, D. P., & Sivakumar, K. (2025). Machine learning framework for breast cancer detection with feature selection with L2 ridge regularization: insights from multiple datasets. *Journal of Translational Genetics and Genomics*, 9(1), 11-34.
17. Setiadi, D. R. I. M., Ojugo, A. A., Pribadi, O., Kartikadarma, E., Setyoko, B. H., Widiono, S., Robet, Aghaunor, T. C., & Ugbotu, E. V. (2025). Integrating Hybrid Statistical and Unsupervised LSTM-Guided Feature Extraction for Breast Cancer Detection. *Journal of Computing Theories and Applications*, 2(4), 536-552.

18. Sundar, R., Srinivasulu, C., Anusha, M. B., Brahmaiah, M., Srikanth, T., & Gupta, K. G. (2024). Enhancing breast cancer detection from histopathology images: A novel ensemble approach with deep learning-based feature extraction. In MATEC Web of Conferences (Vol. 392, p. 01139). EDP Sciences.
19. Thatha, V. N., Karthik, M. G., Gaddam, V. G., Krishna, D. P., Venkataramana, S., Lella, K. K., & Pamula, U. (2025). Histopathological image based breast cancer diagnosis using deep learning and bio inspired optimization. Scientific Reports, 15(1), 1-24.
20. UBIAI (2023). The Secret to Mastering Feature Extraction in Convolutional Neural Network. <https://ubiai.tools/the-secret-to-mastering-feature-extraction-in-convolutional-neural-network/>