EPH-International Journal of Applied Science

Volume - 12, Issue - 03
July - 2026
DOI-https://doi.org/10.69980/pyjg9q47

EXPLAINABLE BREAST CANCER DIAGNOSIS USING TUMOR MORPHOLOGY

Mr. Jitu Murmu

Asst. Prof (MHN), Mental Health Nursing, Blue Wheel Institute of Nursing Sciences (BINS), Chandpur, Nayagarh
Email ID: 2122phddev@gmail.com

Received:- 30-05-2026, Revised:- 01-07-2026, Accepted:- 08-07-2026, Published:- 15-07-2026

Abstract

Concrete compressive strength is a critical indicator of structural performance and material quality, but Breast cancer remains a leading cause of cancer-related morbidity and mortality, emphasizing the need for accurate and interpretable diagnostic approaches. This study developed an explainable machine learning framework for breast cancer diagnosis using quantitative tumor morphology features derived from the Breast Cancer Wisconsin (Diagnostic) Dataset. Following data preprocessing and exploratory analysis, five machine learning algorithms, namely Logistic Regression, Decision Tree, Random Forest, Support Vector Machine, and Extreme Gradient Boosting, were trained and comparatively evaluated using standard classification metrics. Model interpretability was achieved through SHapley Additive exPlanations (SHAP), enabling both global and local interpretation of feature contributions. Among the evaluated models, Logistic Regression demonstrated the highest diagnostic performance, while SHAP analysis identified radius, perimeter, concavity, concave points, and texture as the most influential morphological predictors of malignancy. The explainability framework enhanced transparency by illustrating how individual features affected model predictions, thereby improving the reliability and clinical interpretability of automated diagnostic decisions. The proposed framework demonstrates that combining robust machine learning with explainable artificial intelligence can provide accurate, transparent, and clinically meaningful breast cancer diagnosis. These findings support the integration of explainable diagnostic models into clinical decision-support systems and establish a foundation for future research involving multicenter datasets and multimodal medical information.

Keywords: *Breast cancer diagnosis, Tumor morphology, Explainable artificial intelligence, Machine learning, SHAP analysis*

Introduction

Breast cancer is one of the most common cancers in the world in women and is a significant public health problem due to its incidence, mortality and socioeconomic impact. Accurate early diagnosis is important for treatment planning, minimising disease progression and improving patient survival. Current diagnostic methods are based on conventional clinical examination, medical imaging and histopathological evaluation, and interpretation is subjective and requires specialist knowledge, with the possibility of inter-observer variability. The recent developments in the field of artificial intelligence (AI) have led to the emergence of data-driven techniques that can recognise the subtle morphological patterns that are linked to the malignancy of a tumour. Explainable machine learning has further bolstered this area by showcasing the ability to predict, while also providing transparency in the decision-making process, helping the clinicians to understand how the different tumor features relate to the diagnosis (Binder et al., 2021).

Tumor morphology gives quantitative data about the structural features of breast lesions such as Nuclear size, Nuclear texture, Nuclear perimeter, Nuclear area, smoothness, compactness, concavity, symmetry and fractal properties. These characteristics have proven to be very promising in discriminating between benign and malignant tumors, and have been incorporated into computational diagnostic systems. Morphology-based classification increasingly uses explainable AI methods to show the impact of each feature on the prediction outcome, instead of being a “black-box”. The results have shown that the shape-oriented diagnostic framework based on Digital Breast Tomosynthesis (DBT) not only enhances the classification performance but also boosts the confidence of the clinicians (Hussain et al., 2022). Similarly, explainable pathomic analyses have demonstrated that morphological descriptors mined from the histopathological slides can be used to characterize the cellularity of the tumor, and to explain in a transparent way the relevance of the features (Altini et al., 2023).

The confluence of XAI in various breast cancer imaging modalities has paved the way for transformative and reliable clinical decision support. Explainable learning has been demonstrated to be successful in magnetic resonance imaging (MRI), where it was used to interpret complicated radiological features with a strong prediction performance (Akbar et al., 2025). Multimodal explainability is a key feature of future diagnostic systems, with comprehensive reviews highlighting its importance for clinical acceptance and to aid in the validation of heterogenous datasets (Abdullakutty et al., 2024). Likewise, a systematic scoping review identified an increasing need for explainable AI in breast cancer detection, risk prediction, and personalized clinical management, but noted methodological frameworks are still not standardized (Ghasemi et al., 2024).

In recent times, the advancement of the use of deep learning has significantly enhanced breast cancer detection methods including the use of convolutional neural networks, hybrid architectures and advanced computational models. Efficient neural network architectures together with explainability methods have shown better diagnostic accuracy and detected the image features that influence the decisions of each individual case (Murugan et al., 2025). Improved classification accuracy has been achieved with explainable deep neural networks (DNNs) that combine histopathological and ultrasound information, capitalizing on complementary imaging characteristics in transparent analytical pipelines (Alom et al., 2025). Similarly, hybrid deep learning models and tailored CNNs have shown potential in breast lesion classification, with a focus on enhancing interpretability within multimodal diagnostic settings (Zou & Miao, 2025; Deb et al., 2025).

Interpretability, in addition to prediction, has become a key criterion for achieving successful clinical use of AI-assisted diagnostic systems. Explainable classification models have shown that transparency can boost physicians' confidence by highlighting the most significant morphological features that support the prediction of malignancy (Khater et al., 2023). Explainable deep learning models have also been designed for scoring HER2 immunohistochemistry which have demonstrated that predictions can be easily explained and will aid pathologists in their interpretation, whilst decreasing subjective variability (Nabi et al., 2025). Moreover, explainable deep learning schemes, hybrid transfer learning approaches, and Vision Transformer-based classification models have all shown that modern AI techniques can deliver high diagnostic performance with acceptable transparency to facilitate more accountable clinical decision-making (Saharan et al., 2025; Ganie et al., 2025; Naas et al., 2025).

However, there are research gaps that still need to be addressed. While there are several existing studies that focus on image-based deep learning or multimodal imaging datasets, comparatively fewer studies have investigated how interpretable machine learning models are using quantitative tumor morphology measurements for the diagnosis of breast cancer. Furthermore, some successful algorithms focus on the accuracy of prediction, without fully elucidating the role of each morphological feature in the decision. However, to overcome these limitations, a clinically interpretable analytical model with good classification performance is needed. Thus, this study aims to create a novel explainable machine learning approach to classify benign and malignant breast tumors by leveraging quantitative tumor morphology characteristics and determine which features are most important for prediction results.

Research Objectives

- To develop an explainable machine learning framework for breast cancer diagnosis using tumor morphology features.
- To evaluate and compare the diagnostic performance of multiple machine learning algorithms for classifying benign and malignant breast tumors.
- To identify and interpret the most influential tumor morphology features contributing to breast cancer diagnosis using explainable artificial intelligence techniques.

2. Methodology

2.1 Research Design

A quantitative diagnostic modelling design was applied in the study to investigate if the numerical tumour morphology measurements would be able to differentiate between the malignant and benign breast masses. The analysis was framed as a supervised binary classification problem where malignant diagnosis was considered as positive class and benign diagnosis was considered as negative class. Five algorithms were tested: logistic regression was used as a linear baseline algorithm, decision tree was used for rule-based classification, random forest was used for bagged ensemble learning, support vector machine was used for nonlinear margin-based classification, and XGBoost was used for gradient-boosted decision trees. The workflow was set up such that clinically interpretable outputs were given, while maintaining the ability to compare predictive discrimination via SHAP analysis.

2.2 Dataset Description and Preprocessing

Breast Cancer Wisconsin (Diagnostic) was imported. It was filled with 569 observations and 33 original columns. The Unnamed: 32 field was dropped because it was completely empty, while the id field was dropped as it was only used to uniquely identify records. The diagnosis field was converted to a binary variable: B = 0 and M = 1. The remaining 30 predictors measured 10 nuclear morphology features—mean, standard-error and worst value of radius, texture, perimeter, area, smoothness, compactness, concavity, concave points, symmetry, and fractal dimension. There were no missing predictor values, so no imputation was needed. To avoid information leakage and to set the differently scaled variables on a similar basis, standardization was integrated within the logistic-regression and support-vector-machine pipelines (UCI Machine Learning, 2016).

2.3 Exploratory Analysis and Data Partitioning

Record completeness, class frequencies, duplication, predictor distributions, and pairwise Pearson correlations were explored as initial analysis. The class distribution was not synthetically oversampled due to the moderate level of imbalance and class-weighted estimation was used where appropriate. Data was split into training and testing sets with a stratified 70/30 split, and a random seed of 42. This process resulted in 398 observations being assigned to model development, and 171 observations being assigned to independent testing, while maintaining the benign-to-malignant ratio. Only the training subset has been used to fit all hyperparameter searches and transformations. A stratified five-fold cross validation scheme is used for model selection, where the shuffling procedure was done with the same random seed to ensure that both diagnostic groups were included in each fold.

2.4 Model Development and Hyperparameter Tuning

For each classifier, the optimal parameters were selected by grid search, optimizing on the mean of the

cross-validated ROC-AUC. Logistic regression was tested at varying levels of regularization, while the decision tree was optimized for maximum depth, minimum split and minimum leaf size. Random forest tuning included the number of trees, tree depth, split threshold and number of features considered at each split. A support vector machine with a radial-basis-function kernel was used and the penalty and kernel-width parameters were tuned. The tuning of XGBoost was done in terms of tree count, learning rate, maximum depth, subsampling and column sampling. The parameter setting that maximizes the five-fold ROC-AUC was re-fitted on the entire training subset and then tested once on the held-out test subset.

2.5 Performance Evaluation and Explainability

The accuracy, precision, sensitivity, specificity, F1-score, test ROC-AUC and cross-validated ROC-AUC were used to assess the performance of the models. Sensitivity measured the percentage of malignant cases correctly identified and specificity measured the percentage of benign cases. Model selection was done according to the test ROC-AUC and the sensitivity and accuracy were second. The chosen model was then used in SHAP which provided quantitative information on the contribution of each variable of the morphology. The different plots were mean absolute SHAP values (showing global feature importance), a beeswarm plot (showing the direction and size of feature importance for each test observation), and a waterfall plot (showing the relative contribution of features in one high-confidence malignant prediction). This approach tied predictive performance with both population level and patient level interpretability.

3. Results

3.1 Dataset Characteristics and Morphological Structure

After removing the empty field, and excluding the identifier from model inputs, 569 complete observations were used for the analytical dataset. The number of benign cases was 357 (62.74%) while the number of malignant cases was 212 (37.26%). As is apparent from Table 1, none of the 30 morphology predictors were missing and no duplicate records or repeated identifiers were found. These results suggested that it was not necessary to impute or duplicate-resolve the data set. The class imbalance was moderate and stratification and class-weighted estimation were used but no synthetic generation of data was implemented.

Table 1. Characteristics of the Breast Cancer Wisconsin (Diagnostic) Dataset

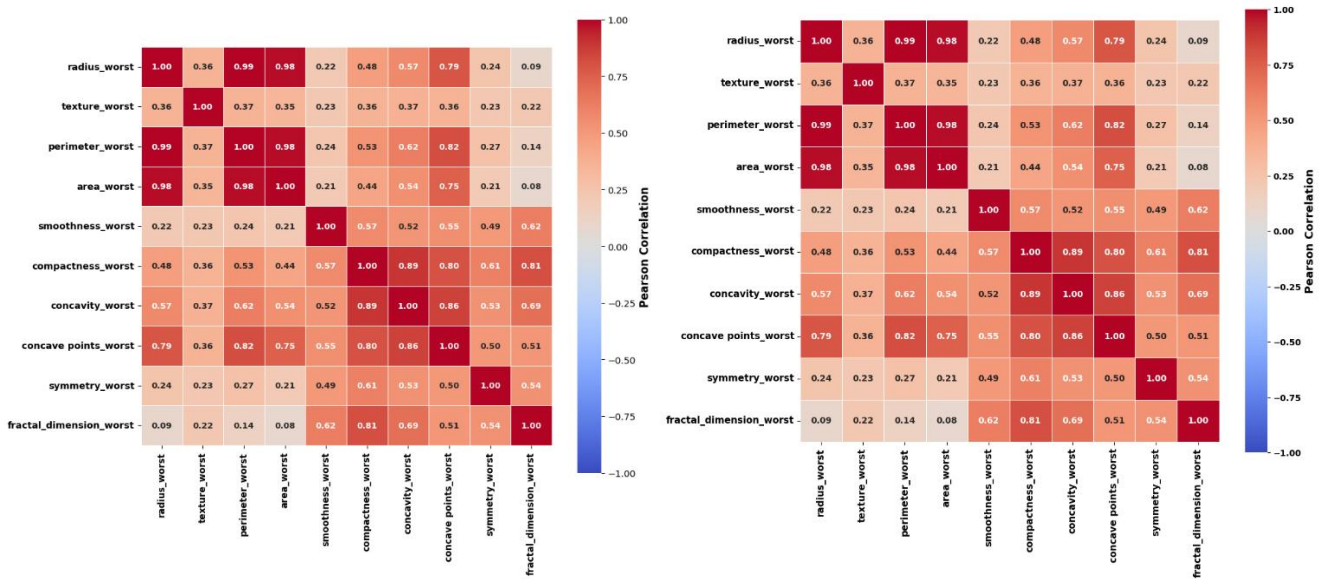
Characteristic	Value	Interpretation
Total observations	569	Complete diagnostic records
Benign cases	357	62.74% of cases
Malignant cases	212	37.26% of cases
Numerical morphology predictors	30	Ten properties measured as mean, standard error, and worst value
Missing predictor values	0	No imputation required
Duplicate rows	0	No duplicate records detected
Unique patient identifiers	569	All identifiers were unique
Target classes	2	Benign and malignant

The 30 numerical predictors have been aggregated into 10 underlying properties of morphology, as indicated in Table 2. The mean, standard error, and worst measurement were used for each property, so that typical nuclear properties, variation within the sample and the most extreme morphology observed could be included in the analysis. Size related properties were radius, perimeter and area while boundary irregularity properties were compactness, concavity, concave points, smoothness, symmetry and fractal dimension. These shape measurements were complemented by texture which was used to quantify the grayscale variation. This organization was crucial for determining the association of malignancy with either large nuclei, irregular borders or heterogeneous texture or both.

Table 2. Summary of Tumor Morphology Features

Morphological property	Mean feature	Standard-error feature	Worst-value feature	Analytical meaning
Radius	radius_mean	radius_se	radius_worst	Average distance from nuclear center to boundary
Texture	texture_mean	texture_se	texture_worst	Variation in grayscale intensities
Perimeter	perimeter_mean	perimeter_se	perimeter_worst	Nuclear boundary length
Area	area_mean	area_se	area_worst	Nuclear surface area
Smoothness	smoothness_mean	smoothness_se	smoothness_worst	Local variation in boundary radius
Compactness	compactness_mean	compactness_se	compactness_worst	Perimeter-to-area shape compactness
Concavity	concavity_mean	concavity_se	concavity_worst	Severity of concave boundary regions
Concave Points	concave points_mean	concave points_se	concave points_worst	Number or prominence of concave boundary points
Symmetry	symmetry_mean	symmetry_se	symmetry_worst	Degree of nuclear symmetry
Fractal Dimension	fractal dimension_mean	fractal dimension_se	fractal dimension_worst	Boundary complexity or irregularity

The Pearson correlation matrix of all morphology predictors is shown in Figure 1. High positive association existed between measurements of radius, perimeter and area, which are all measures of nuclear size. The Pearson correlation matrices of the tumor morphology variables are shown in Figures (a) and (b). Figure (a) shows the correlation between the mean and standard error features, while Figure (b) displays the correlation between the worst-case morphological features. Within each of the two sets of features, strong positive correlations between the radius, perimeter, area, and concavity related variables were found, suggesting a high degree of multicollinearity among the size and shape features. On the other hand, the correlation between texture and fractal dimension and several morphological parameters was relatively poor, indicating that these parameters can give complementary diagnostic information.



(a) **(b)**

Figure 1. Correlation Heatmap of Tumor Morphology Features

3.2 Hyperparameter Selection and Model Configuration

A different configuration was chosen for the different modelling strategies based on the cross-validation procedure (Table 3). The regularization parameter was evaluated and the best one was favored by logistic regression and for the decision tree, unstable partitioning was curtailed by limiting the depth of the tree and the size of its leaves. For the random forest, we used a large ensemble with feature subsampling and for the support vector machine, we chose a radial-basis-function specification. XGBoost was the combination of shallow trees and controlled learning and sampling parameters. Since all configurations were selected in the same way, with the same training folds and with the same ROC-AUC criterion, the comparison of the following configuration reflected behaviour of the models under a common selection process, rather than arbitrary default settings.

Table 3. Machine Learning Model Hyperparameters

Model	Selected hyperparameters	Tuning criterion
Logistic Regression	C=0.1	5-fold CV ROC-AUC
Decision Tree	max_depth=5; min_samples_leaf=2; min_samples_split=10	5-fold CV ROC-AUC
Random Forest	max_depth=None; max_features=sqrt; min_samples_split=5; n_estimators=200	5-fold CV ROC-AUC
Support Vector Machine	C=5; gamma=0.01; kernel=rbf	5-fold CV ROC-AUC
XGBoost	colsample_bytree=1.0; learning_rate=0.1; max_depth=2; n_estimators=300; subsample=0.8	5-fold CV ROC-AUC

3.3 Comparative Diagnostic Performance

Table 4 shows the comparison of diagnostic performance on the test subset with 171 cases. Logistic regression had the best overall accuracy (98.83%), F1-score (0.9841), and demonstrated 100% precision and specificity. It correctly diagnosed all 107 benign cases and 62 of 64 malignant cases, resulting in two false-negative cases and no false-positive cases. The AUC for its test set was 0.9988, and the AUC for its 5-fold cross validation set was 0.9930. SVM, XGBoost and RF got 98.25%, 98.25% and 97.66% accuracy respectively. The decision tree gave the poorest discrimination, and also produced the largest number of false negatives and false positives, which is indicative of the benefits of regularization and ensemble/margin-based learning.

Table 4. Comparative Performance of Machine Learning Models

Model	Accuracy	Precision	Sensitivity	Specificity	F1-score	ROC-AUC	CV ROC-AUC
Logistic Regression	0.9883	1.0000	0.9688	1.0000	0.9841	0.9988	0.9930
Support Vector Machine	0.9825	1.0000	0.9531	1.0000	0.9760	0.9988	0.9950
Random Forest	0.9766	1.0000	0.9375	1.0000	0.9677	0.9974	0.9879
XGBoost	0.9825	1.0000	0.9531	1.0000	0.9760	0.9949	0.9934
Decision Tree	0.9415	0.9655	0.8750	0.9813	0.9180	0.9420	0.9420

3.4 Global Explainability and Feature Importance

Table 5 lists the top 10 morphology variables by largest mean absolute SHAP value for the model of logistic regression selected. Texture_worst, radius_worst, concave points_worst, perimeter_worst and texture_mean were the top five predictors. The average values of all the top-ranking features were

greater in the malignant cases than in the benign ones, showing greater nuclear size, more concave boundaries, and more irregular texture in the cases predicted to be malignant. Several leading positions were occupied by worst-value measurements, indicating that extreme morphological characteristics contained more diagnostic information than just average measurements. The texture, size and the concavity variables were also shown to occur simultaneously, indicating that the model was not reliant upon just one morphological dimension.

Table 5. Top Tumor Morphology Features Identified by SHAP

Rank	Feature	Mean absolute SHAP value	Morphological pattern in malignant cases
1	texture_worst	0.4969	Higher
2	radius_worst	0.3611	Higher
3	concave points_worst	0.3344	Higher
4	perimeter_worst	0.3319	Higher
5	texture_mean	0.3209	Higher
6	concave points_mean	0.3189	Higher
7	area_worst	0.3159	Higher
8	symmetry_worst	0.3062	Higher
9	radius_se	0.3050	Higher
10	radius_mean	0.2716	Higher

The distribution of SHAP values in the test set is shown in Figure 3. The features are ranked based on their global significance and each point is a feature obtained from an observation with the feature being pushed towards benign or malignant classification based on the horizontal position of the point. Generally, high scores for worst texture, worst radius, worst concave points, worst perimeter, were observed on the positive side of the SHAP axis which showed that these features were associated with malignant predictions, whereas mean texture showed the opposite trend. Smaller values of these variables were more likely to reclassify the log-odds towards 'benign'. The distribution of points also exhibited significant variation across the cases, indicating that the model used more than one feature, as opposed to a set rule.

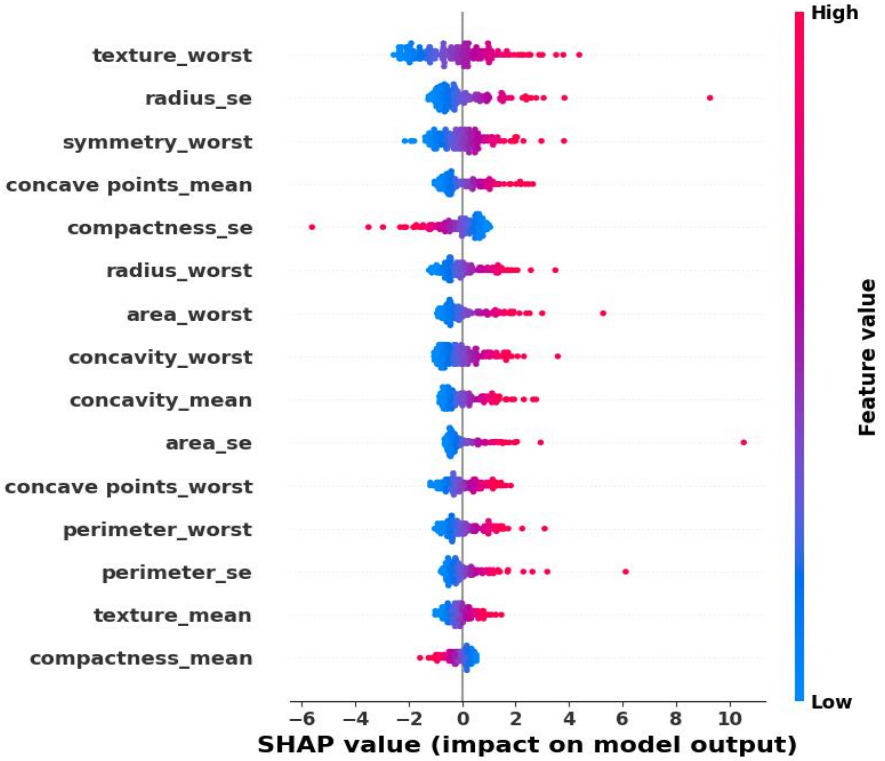


Figure 3. SHAP Summary Plot Showing Global Feature Importance

3.5 Local Explanation and Diagnostic Interpretation

The high confidence malignant prediction from the test subset is decomposed in figure 4. The selected observation was labeled malignant and the estimated probability of malignancy was 1.000. The waterfall plot illustrates the change of the standardized values of each of the morphologies from the model's base output. Any positive contribution from increased or abnormal nuclear characteristics added up to form the final malignant probability, while any negative contribution served to offset the movement to some degree. This local explanation can also be used to supplement the global ranking: It shows that the decision for the diagnosis can be traced back to certain measures for a given case and is not given as a probability that cannot be explained.

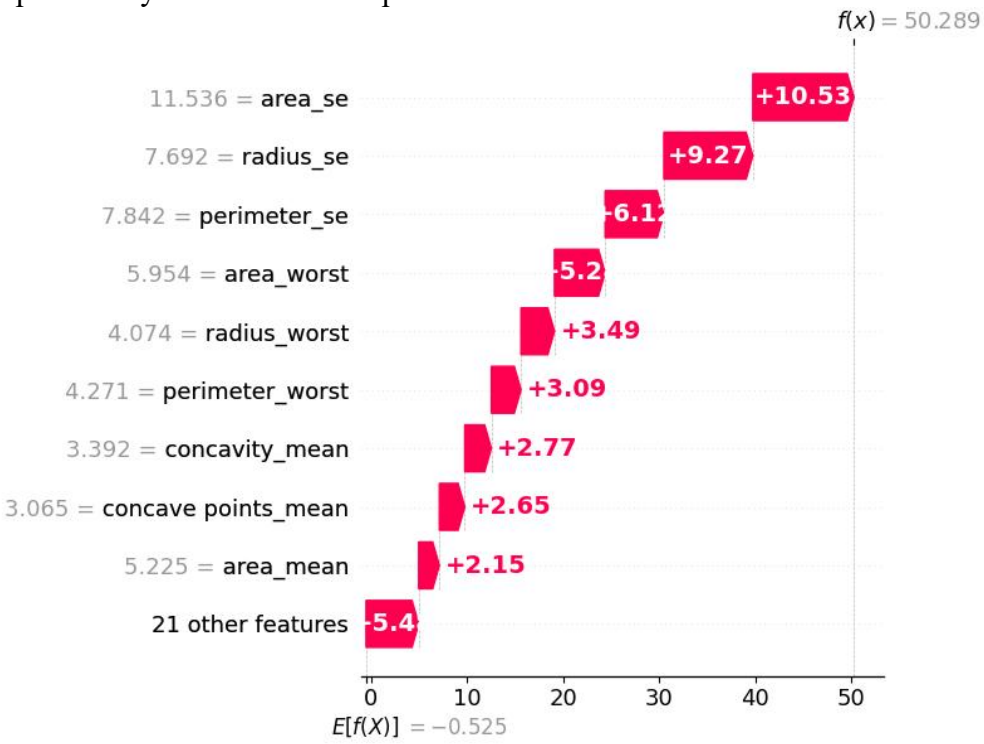


Figure 4. SHAP Waterfall Plot Explaining an Individual Prediction

The results indicated the morphology dataset enabled high accuracy classification with its interpretability. Logistic regression achieved the best threshold-based performance with nearly perfect discrimination, and SHAP helped to elucidate the relative importance of texture, nuclear size, and concave boundary features. This 0% of false positive predictions in the test set showed that the selected threshold did not falsely escalate any cases to be benign, and 2 cases of false negatives in the test set showed that even a good model had clinically important residual error. Therefore, this model should be regarded as an explainable decision support system and needs external validation, not as a substitute for pathological investigation.

4. Discussion

The present study findings showed that the quantitative tumor morphology features with explainable machine learning achieved high diagnosis accuracy and interpretability of breast cancer. The predictive ability of the evaluated classification models was excellent and logistic regression was the best classification model for the selected dataset. SHAP-based explanations provided a transparent identification of the morphological variables affecting individual diagnostic decisions, beyond the predictive accuracy. This blend of predictability and interpretability overcomes one of the biggest challenges of traditional black box AI models, enabling clinicians to gain insights into the relationship between particular nuclear features and the likelihood of malignancy. The outcomes align with previous studies demonstrating that machine learning algorithms utilizing nuclear morphological features can effectively classify benign from malignant breast cancers and enhance the transparency and clinical confidence of the models (Birwadkar et al., 2025). Similarly, interpretable decision-

support systems built on top of histopathological data have shown promise in boosting doctor's trust without harming the diagnosis accuracy, consistent with our findings (Krishna et al., 2023).

SHAP analysis revealed that features associated with radius, perimeter, concavity, concave points and texture are the most important features for breast cancer classification, showing the importance of quantitative nuclear morphology in breast cancer diagnosis. These properties are intimately related to cellular abnormalities and changes that are usually present during malignant transformation. The global and local explanations produced during this study showed how individual morphological measurements related to prediction outcomes, and hence offered clinically meaningful information beyond just a probability score. These clear and transparent explanations would aid interpretability of automated diagnostic systems and allow their incorporation into standard clinical workflow. This is echoed in larger studies examining the use of AI in cancer diagnosis, which have identified the need for reproducibility, explainability, and multimodal interpretation as key factors for the safe integration of artificial intelligence in radiology and pathology (Khosravi et al., 2025).

Comparative analysis of several machine learning algorithms also showed that simpler and more interpretable algorithms are capable of performing competitively or even better than more complex and computationally expensive algorithms, when high-quality quantitative morphology features are provided. Based on this observation it is suggested that beside the predictive performance of models, transparency should be carefully considered as well when selecting an algorithm for clinical applications. Explainability is essential, especially in the context of multimodal deep learning frameworks that have demonstrated impressive diagnostic capabilities in a variety of imaging modalities, for ensuring accountability and clinician acceptance of automated decisions (Nawaz et al., 2025). Likewise, Vision Transformer (ViT) models have shown significant performance in breast cancer classification and highlighted the need for feature attribution transparency for clinical interpretation (Balaha et al., 2025). The current results support this notion by demonstrating that morphology-based machine learning can be used to make accurate predictions with simple interpretation of influential tumor traits.

Also, the high predictive value of the morphology-related variables found in the present study corroborates with recent radiomics and explainable machine learning studies in breast oncology. Past multicenter studies have shown the usefulness of interpretable radiomics models to predict axillary lymph node metastasis, recognizing clinically relevant imaging features that are linked to disease progression (Liu et al., 2024). Similarly, SHAP interpretable feature importance has been used to explain the importance of individual features for predicting lymph node metastasis in invasive micropapillary breast carcinoma, emphasizing the clinical utility of interpretable feature importance (Jiang et al., 2022). While the present study targeted the classification of benign and malignant tumor types, the two studies highlight the importance of "explainable feature attribution" to ensure confidence in AI-assisted clinical assessment and contribute to evidence-based decision-making.

However, there are some drawbacks to the positive results. In the analysis, a single publicly available dataset was used consisting of quantitative morphology measurements of the tumors, which might restrict the validity of the models created in other clinical populations and imaging platforms. The proposed framework could be strengthened and used clinically by external validation with multicenter datasets. Moreover, while the SHAP explanations were useful for providing an interpretation of the feature importance, there may be other explainability methods to try to check the consistency of feature importance across various methods. Going forward, multimodal datasets combining histopathology, radiology, genomics, and clinical factors should be explored to enhance the diagnostic accuracy with the transparency of the model. In addition, recent studies show that explainable synthetic histology generation could help to improve AI interpretability and educational uses, pointing to another possible avenue for future breast cancer research (Dolezal et al., 2023).

In summary, the current study illustrates a high accuracy and clinically interpretable diagnosis of breast cancer using explainable machine learning approaches on quantitative tumour morphology. Fitting in transparent feature attribution with powerful predictive modelling can help to enhance the ability of physicians to understand automated diagnostic decisions and to make more reliable clinical decisions. The results add to the emerging evidence that explainable AI is a crucial element of future precision oncology systems, especially in the context of desirable conditions for responsible deployment in healthcare, such as transparency, reproducibility, and clinical trust.

5. Conclusion

This study showed that explainable machine learning algorithms can be used to identify the distinction between benign and malignant breast tumor from quantitative tumor morphology features in an accurate and transparent manner. Logistic regression demonstrated the best diagnostic performance of all the algorithms evaluated, and it also provided high interpretability with feature attribution using SHAP. The results validated that the morphological features such as radius, perimeter, concavity, concave points, and texture are important in the classification of breast cancer and offer clinically relevant explanations for automatic predictions. The proposed framework was effective in achieving the objectives of the research because it was able to develop an explainable diagnostic model, perform comparisons of the performance of different machine learning algorithms, and identify the most influencing tumor morphology features for diagnostic decision-making. The incorporation of explainability also made the prediction process more transparent and trustworthy, which is crucial for its potential use in clinical decision support, unlike traditional black-box methods. The study highlights the potential of interpretable models in achieving high predictive accuracy while maintaining clinical relevance, thereby advancing the use of explainable AI in precision oncology. Despite the fact that only one publicly available dataset was analysed, the proposed methodology was developed with a solid basis for future research with larger multicenter cohorts and multimodal clinical data. Future studies should involve validation in various populations and the incorporation of imaging, histopathological, genomic, and clinical data to enhance the model's generalizability and its clinical application for trustworthy AI-assisted breast cancer diagnosis.

References

1. Abdullakutty, F., Akbari, Y., Al-Maadeed, S., Bouridane, A., Talaat, I. M., & Hamoudi, R. (2024). Histopathology in focus: a review on explainable multi-modal approaches for breast cancer diagnosis. Frontiers in Medicine, 11, 1450103.
2. Akbar, A., Han, S., Urr Rehman, N., Ahmed, K., Eshkiki, H., & Caraffini, F. (2025). Explainable breast cancer prediction from 3-dimensional dynamic contrast-enhanced magnetic resonance imaging. Applied Intelligence, 55(13), 901.
3. Alom, M. R., Farid, F. A., Rahaman, M. A., Rahman, A., Debnath, T., Miah, A. S. M., & Mansor, S. (2025). An explainable AI-driven deep neural network for accurate breast cancer detection from histopathological and ultrasound images. Scientific Reports, 15(1), 17531.
4. Altini, N., Puro, E., Taccogna, M. G., Marino, F., De Summa, S., Saponaro, C., ... & Bevilacqua, V. (2023). Tumor cellularity assessment of breast histopathological slides via instance segmentation and pathomic features explainability. Bioengineering, 10(4), 396.
5. Balaha, H. M., Ali, K. M., Gondim, D., Ghazal, M., & El-Baz, A. (2025). Harnessing vision transformers for precise and explainable breast cancer diagnosis. In International Conference on Pattern Recognition (pp. 191-206). Springer, Cham.
6. Binder, A., Bockmayr, M., Hägele, M., Wienert, S., Heim, D., Hellweg, K., ... & Klauschen, F. (2021). Morphological and molecular breast cancer profiling through explainable machine learning. Nature Machine Intelligence, 3(4), 355-366.
7. Birwadkar, A., Buttny, S., Sahli, C., & Kenry. (2025). Machine-Learning-Guided Analysis of Breast Tumor Malignancy Based on Nuclear Morphological Features. Advanced Intelligent Discovery, 1(3), e202500034.
8. Deb, D., Dash, R., & Mohapatra, D. P. (2025). Customized convolutional neural network with explainable AI for multimodal breast cancer detection. Sādhanā, 50(3), 130.
9. Dolezal, J. M., Wolk, R., Hieromnimon, H. M., Howard, F. M., Srisuwananukorn, A., Karpeyev, D., ... & Pearson, A. T. (2023). Deep learning generates synthetic cancer histology for explainability and education. NPJ precision oncology, 7(1), 49.
10. Ganie, S. M., Malik, M. B., Aadil, M., Muteeb, G., Farhan, M., & Aatif, M. (2025). Explainable AI based hybrid DRM-Net transfer learning model for breast cancer detection and classification using ultrasound images. Scientific Reports, 15(1), 44170.
11. Ghasemi, A., Hashtarkhani, S., Schwartz, D. L., & Shaban-Nejad, A. (2024). Explainable artificial intelligence in breast cancer detection and risk prediction: A systematic scoping review. Cancer Innovation, 3(5), e136.

12. Hussain, S. M., Buongiorno, D., Altini, N., Berloco, F., Prencipe, B., Moschetta, M., ... & Brunetti, A. (2022). Shape-based breast lesion classification using digital tomosynthesis images: The role of explainable artificial intelligence. Applied Sciences, 12(12), 6230.
13. Jiang, C., Xiu, Y., Qiao, K., Yu, X., Zhang, S., & Huang, Y. (2022). Prediction of lymph node metastasis in patients with breast invasive micropapillary carcinoma based on machine learning and SHapley Additive exPlanations framework. Frontiers in Oncology, 12, 981059.
14. 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, 5618-5633.
15. Khosravi, P., Fuchs, T. J., & Ho, D. J. (2025). Artificial intelligence–driven cancer diagnostics: enhancing radiology and pathology through reproducibility, explainability, and multimodality. Cancer research, 85(13), 2356-2367.
16. 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.
17. Liu, Z., Hong, M., Li, X., Lin, L., Tan, X., & Liu, Y. (2024). Predicting axillary lymph node metastasis in breast cancer patients: A radiomics-based multicenter approach with interpretability analysis. European Journal of Radiology, 176, 111522.
18. Murugan, T. K., Karthikeyan, P., & Sekar, P. (2025). Efficient breast cancer detection using neural networks and explainable artificial intelligence. Neural Computing and Applications, 37(5), 3759-3776.
19. Naas, M., Mzoughi, H., Njeh, I., & BenSlima, M. (2025). An explainable AI for breast cancer classification using vision Transformer (ViT). Biomedical Signal Processing and Control, 108, 108011.
20. Nabi, M. S., Fauzi, M. F. A., Karim, H. B. A., Cheah, P. L., Fan, C. S., & Looi, L. M. (2025). Explainable deep learning models for HER2 IHC scoring in breast cancer diagnosis. Informatics in Medicine Unlocked, 101700.
21. Nawaz, U., Saeed, Z., UbaidUllah, H. M., Mirza, F., & Muzzamil, M. (2025). Explainable Attention-Enhanced Approach for Multimodal Breast Cancer Diagnosis Across Diverse Imaging Modalities. International Journal of Imaging Systems and Technology, 35(6), e70209.
22. Saharan, S., Wani, N. A., Chatterji, S., Kumar, N., & Almuhaideb, A. M. (2025). A deep learning and explainable artificial intelligence based scheme for breast cancer detection. Scientific Reports, 15(1), 32125.
23. UCI Machine Learning. (2016). Breast Cancer Wisconsin (Diagnostic) Data Set [Data set]. Kaggle. https://www.kaggle.com/datasets/uciml/breast-cancer-wisconsin-data
24. Zou, Y., & Miao, P. (2025). Explainable AI-enabled hybrid deep learning architecture for breast cancer detection. Frontiers in Immunology, 16, 1658741.