e-ISSN 2549-7472
Edumatic: Jurnal Pendidikan Informatika
Volume 10, Issue 2, **330-339.**
https://e-journal.hamzanwadi.ac.id/index.php/edumatic
DOI: 10.29408/edumatic.v10i2.34993

Explainable Transfer Learning for Breast Cancer Histopathology Classification Using Grad-CAM

Ade Fatahillah [1,*] , Fandy Setyo Utomo [1] , Taqwa Hariguna [1]

[1] Universitas Amikom Purwokerto, Indonesia
[*] Corresponding author: Ade Fatahillah, Universitas Amikom Purwokerto, Indonesia
✉ amikom@amikompurwokerto.ac.id

Received: 23 May 2026 | Revised: 1 June 2026 | Accepted: 21 June 2026 | Published: 9 July 2026

Abstract
Breast cancer remains one of the leading causes of cancer-related mortality among women worldwide, highlighting the need for diagnostic systems that are both accurate and interpretable. Although transfer learning has shown promising results in histopathological image classification, studies simultaneously examining predictive performance, statistical reliability, and interpretability remain limited. This study proposes an explainable transfer learning framework for breast cancer histopathology classification and investigates the relationship between classification performance and visual interpretability. Experiments were conducted using 2,013 histopathological images from the BreakHis dataset at 200× magnification. Three pretrained architectures, ResNet50, DenseNet121, and EfficientNetB0, were trained and evaluated under identical preprocessing, augmentation, and training settings. Performance was assessed using accuracy, precision, recall, F1-score, AUC, confidence intervals, McNemar testing, confusion matrix analysis, and Grad-CAM visualization. Results showed that DenseNet121 achieved the most balanced classification performance and the highest discriminative capability among the evaluated models. Statistical analysis confirmed significant performance differences, while Grad-CAM visualizations demonstrated more focused and diagnostically relevant activation regions. These findings suggest that models learning more discriminative histopathological representations tend to generate more meaningful visual explanations. The study emphasizes integrating predictive performance, statistical validation, and explainability to support reliable and transparent artificial intelligence systems for breast cancer diagnosis.

Keywords: breast cancer; explainable artificial intelligence; grad-cam; histopathology classification; transfer learning

To cite this article: Fatahillah, A., Utomo, F. S., & Hariguna, T. (2026). Explainable Transfer Learning for Breast Cancer Histopathology Classification Using Grad-CAM. *Edumatic: Jurnal Pendidikan Informatika*, *10*(2), 330–339. https://doi.org/10.29408/edumatic.v10i2.34993

INTRODUCTION

Breast cancer remains one of the leading causes of cancer-related morbidity and mortality among women worldwide. Early and accurate diagnosis is essential for improving treatment outcomes, reducing mortality, and increasing patient survival. Histopathological examination is widely regarded as the gold standard for breast cancer diagnosis because it provides detailed information about cellular morphology and tissue architecture. However, manual assessment of histopathological images is a complex and time-consuming process that requires substantial expertise. Variations in tissue morphology, staining quality, and cellular patterns may also contribute to diagnostic inconsistency and inter-observer variability, potentially affecting

clinical decision-making (Jiang et al., 2024; Sadeghi et al., 2024; Soliman et al., 2024; Ivanov et al., 2025).

Recent advances in artificial intelligence (AI), particularly deep learning, have created new opportunities for computer-aided diagnosis in medical imaging (Ananda et al., 2024; Anhar et al., 2026; Parameswara et al., 2026). Convolutional Neural Networks (CNNs) have demonstrated strong performance in various medical image analysis tasks, including breast cancer histopathology classification, by automatically learning discriminative features directly from image data. Furthermore, transfer learning has become a widely adopted strategy for overcoming the limited availability of annotated medical datasets by leveraging knowledge acquired from large-scale pretrained models. Consequently, architectures such as ResNet50, DenseNet121, and EfficientNetB0 have been extensively applied in breast cancer image classification and have reported promising predictive performance (Jiang et al., 2024; Stanescu et al., 2025; Priya et al., 2024).

High classification accuracy has been widely reported in medical image analysis studies; however, concerns regarding model transparency and interpretability continue to limit the practical adoption of AI in healthcare. Deep learning models are frequently criticized as "black-box" systems because their decision-making processes are often difficult to understand and verify. In clinical environments, healthcare professionals must be able to assess whether AI-generated predictions are supported by medically relevant evidence before incorporating them into diagnostic decision-making. Consequently, the concepts of Explainable Artificial Intelligence (XAI) and transparent AI systems have become increasingly important in healthcare research. According to contemporary healthcare AI frameworks, explainability represents a key component of transparent and interpretable AI systems because it enables users to understand the reasoning behind model predictions. By providing transparent explanations, XAI can enhance clinician confidence, facilitate model validation, and support responsible AI adoption in clinical practice. Therefore, explainability is not merely a technical feature but an important requirement for developing interpretable and clinically relevant AI systems. Among various XAI approaches, Gradient-weighted Class Activation Mapping (Grad-CAM) is widely used to generate visual explanations by highlighting image regions that contribute most strongly to model predictions. In breast cancer histopathology analysis, such visual explanations can help determine whether classification decisions are based on diagnostically meaningful tissue structures (Ghasemi et al., 2024; Talaat et al., 2024; Houssein et al., 2025).

Existing studies have demonstrated that transfer learning models can achieve high classification accuracy in breast cancer histopathological image analysis, while recent investigations have also highlighted the value of Grad-CAM for improving model transparency and interpretability (Naas et al., 2025; Ochoa-Ornelas et al., 2025; Stanescu et al., 2025). Thus, it is well established that transfer learning can provide effective predictive performance and that explainable AI techniques can enhance model transparency. However, several important issues remain insufficiently explored. Existing studies on breast cancer histopathology classification have predominantly emphasized improving classification accuracy, while relatively limited attention has been devoted to model interpretability and trustworthiness. Furthermore, many investigations evaluate only a single transfer learning architecture, restricting comprehensive comparisons of different models under identical experimental settings. In addition, studies that jointly assess predictive performance and visual interpretability across multiple transfer learning architectures using the BreakHis dataset at 200× magnification remain scarce. Consequently, the relationship between classification effectiveness and the quality of visual explanations has not been sufficiently explored, leaving uncertainty as to whether models achieving superior predictive performance also generate more meaningful and clinically relevant interpretability outcomes.

The relationship between predictive performance and visual interpretability in breast cancer histopathology classification remains insufficiently explored. Although high predictive accuracy is essential, clinical adoption also requires models that are interpretable and trustworthy. Therefore, evaluating how different transfer learning architectures balance classification performance and explainability is important for developing reliable and clinically applicable AI-based diagnostic systems.

Therefore, this study aims to compare the performance of three widely used transfer learning architectures, namely ResNet50, DenseNet121, and EfficientNetB0, for breast cancer histopathology classification using 2,013 images from the BreakHis dataset at 200× magnification. In addition to evaluating classification performance, this study employs Grad-CAM to assess model interpretability by examining the tissue regions that influence classification decisions. The scientific contribution of this study lies in its integrated evaluation of predictive performance and explainability within a unified experimental framework. Unlike previous studies that primarily emphasize accuracy or focus on a single architecture, this research investigates how different transfer learning models balance classification effectiveness and visual interpretability. The findings are expected to provide empirical evidence regarding the relationship between explainability and predictive performance while supporting the development of accurate, transparent, and interpretable AI systems for breast cancer diagnosis.

METHOD

This study employed a quantitative experimental approach to evaluate the classification performance, statistical reliability, and interpretability of three transfer learning architectures (ResNet50, DenseNet121, and EfficientNetB0) for breast cancer histopathology classification. The overall workflow is shown in Figure 1.

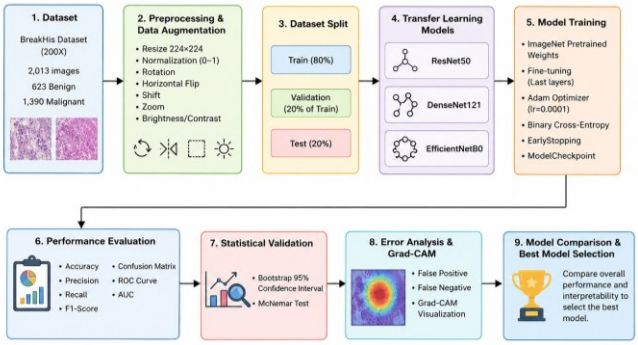


Figure 1. Workflow of breast cancer histopathology classification and grad-cam analysis

Figure 1 illustrates the main stages of the study, including dataset preparation, preprocessing and augmentation, transfer learning model development, performance evaluation, statistical validation, error analysis, and Grad-CAM interpretability assessment. The BreakHis breast cancer histopathological dataset obtained from Kaggle was used, consisting of 2,013 images captured at 200× magnification, including 623 benign and 1,390 malignant samples. This magnification level was selected because it provides a balance between cellular morphology and overall tissue architecture.

All images were resized to 224 × 224 pixels and normalized to the range of 0–1. Data augmentation was performed using TensorFlow ImageDataGenerator, including random rotation, horizontal flipping, width and height shifting, and zooming to improve model generalization. Class weights were applied during training to mitigate class imbalance. The dataset was partitioned at the image level because patient identifiers were unavailable in the public dataset. Images were divided into training and testing sets using an 80:20 ratio with

random_state = 42, while 20% of the training data was used for validation. Transfer learning was implemented using ImageNet-pretrained weights. Fine-tuning was performed by unfreezing the final layers of each architecture while preserving previously learned representations. All models employed the same classification head consisting of Global Average Pooling, Batch Normalization, a 256-neuron dense layer with ReLU activation, Dropout (0.4), and a sigmoid output layer.

Model training utilized the Adam optimizer with a learning rate of 0.0001, binary cross-entropy loss, a batch size of 16, and a maximum of 25 epochs. EarlyStopping and ModelCheckpoint were used to improve training stability and retain the best-performing model. Performance evaluation was conducted on the independent test set using accuracy, precision, recall, F1-score, confusion matrix analysis, classification reports, receiver operating characteristic (ROC) curves, and area under the curve (AUC). Statistical reliability was assessed through bootstrap resampling to estimate 95% confidence intervals for accuracy. McNemar's test was subsequently performed to determine whether performance differences between models were statistically significant. To provide a deeper understanding of model behavior, error distribution analysis was conducted by examining false-positive and false-negative predictions. Model interpretability was evaluated using Gradient-weighted Class Activation Mapping (Grad-CAM), which visualizes image regions contributing to classification decisions. Interpretability assessment was performed qualitatively by comparing activation localization, concentration, and pathological relevance across architectures. The final model selection was based on classification performance, statistical validation, error analysis, and Grad-CAM interpretability.

RESULTS AND DISCUSSION

Results

The experiments were conducted using the BreakHis breast cancer histopathology dataset at 200× magnification consisting of 2,013 images, including 623 benign images and 1,390 malignant images. Three transfer learning architectures, namely ResNet50, DenseNet121, and EfficientNetB0, were evaluated under identical preprocessing, augmentation, training, and testing conditions. Model performance was evaluated using Accuracy, Precision, Recall, F1-score, Area Under the Curve (AUC), confidence interval estimation, McNemar statistical testing, confusion matrix analysis, error distribution analysis, and Grad-CAM visualization.

Table 1. Performance comparison of transfer learning models

Model	Accuracy	Precision	Recall	F1-Score	AUC
ResNet50	0.7692	0.8745	0.7770	0.8229	0.8459
DenseNet121	0.9057	0.9444	0.9173	0.9307	0.9598
EfficientNetB0	0.6898	0.6898	1.0000	0.8164	0.6238

Table 2. Bootstrap accuracy confidence intervals

Model	Mean Accuracy	95% CI Lower	95% CI Upper
DenseNet121	0.9051	0.8734	0.9330
ResNet50	0.7685	0.7270	0.8089
EfficientNetB0	0.6895	0.6452	0.7345

The results indicate that DenseNet121 achieved the highest overall classification performance among the evaluated architectures. DenseNet121 obtained an accuracy of 90.57%, precision of 94.44%, recall of 91.73%, F1-score of 93.07%, and AUC of 0.9598. ResNet50 achieved moderate performance, whereas EfficientNetB0 produced the lowest overall accuracy

despite obtaining perfect recall. These findings suggest that DenseNet121 provided the most balanced classification capability for distinguishing benign and malignant breast cancer histopathology images. To assess the reliability and stability of model performance, bootstrap resampling was performed to estimate 95% confidence intervals of classification accuracy

The confidence interval estimates demonstrate that DenseNet121 consistently achieved higher classification performance than the competing architectures. The relatively narrow confidence interval further indicates stable model behavior and reduced variability across bootstrap samples. To determine whether the observed performance differences were statistically significant, McNemar's test was performed using model prediction outcomes on the independent testing dataset.

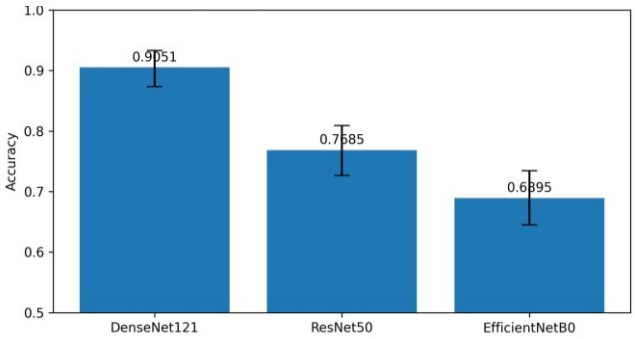


Figure 2. Accuracy comparison with 95% confidence intervals

Figure 2 visually summarizes the mean classification accuracy and associated uncertainty for each model. DenseNet121 demonstrated the highest mean accuracy, with its confidence interval remaining consistently above those of ResNet50 and EfficientNetB0. The relatively narrow confidence interval suggests stable model behavior across bootstrap samples. In contrast, ResNet50 and EfficientNetB0 showed lower accuracy values and wider uncertainty ranges. These findings reinforce the observation that DenseNet121 provides superior and more reliable classification performance under the experimental conditions used in this study.

Table 3. McNemar statistical test result

Comparison	n01	n10	Statistic	p-value	Effect Size
DenseNet121 vs EfficientNetB0	110	23	55.609	<0.001	0.654
DenseNet121 vs ResNet50	76	21	30.0619	<0.001	0.567
EfficientNetB0 vs ResNet50	62	94	6.1603	0.0131	0.205

The McNemar test revealed statistically significant differences among all evaluated model pairs. DenseNet121 significantly outperformed both EfficientNetB0 and ResNet50 ($p < 0.001$), indicating that the observed performance improvements were unlikely to be caused by random variation. The corresponding effect sizes were moderate to large, suggesting meaningful practical differences in classification behavior. In contrast, although EfficientNetB0 and ResNet50 also differed significantly ($p = 0.0131$), the smaller effect size indicates a comparatively weaker practical distinction between the two architectures. These findings provide statistical evidence that the superiority of DenseNet121 is both statistically significant and practically meaningful.

Receiver Operating Characteristic (ROC) analysis was performed to evaluate model discrimination ability across different classification thresholds. DenseNet121 achieved the

highest AUC value of 0.9598, indicating excellent discriminative capability between benign and malignant samples. ResNet50 achieved an AUC of 0.8459, whereas EfficientNetB0 achieved the lowest AUC value of 0.6238. The ROC curves further confirm the superiority of DenseNet121, as its curve consistently remained above those of the competing architectures across most operating thresholds. To better understand model behavior, classification errors were examined using confusion matrix statistics.

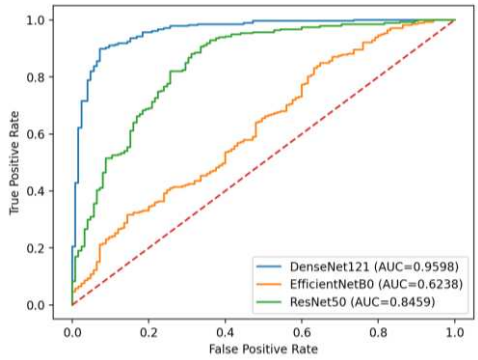
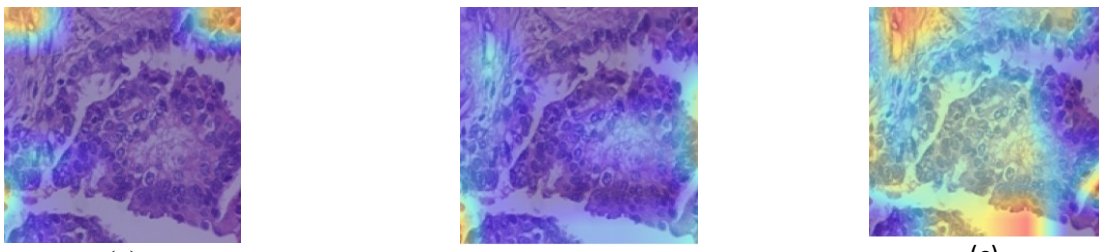


Figure 3. ROC curve comparison of densenet121, resnet50 and efficientnetb0

Table 4. Confusion matrix summary

Model	True Positive	True Negative	False Positive	False Negative
DenseNet121	255	110	15	23
ResNet50	277	18	107	1
EfficientNetB0	278	0	125	0

The error distribution analysis revealed substantial differences in prediction behavior among the evaluated architectures. DenseNet121 produced the lowest number of overall classification errors, indicating strong discriminative capability for identifying both benign and malignant tissue patterns. ResNet50 generated a considerably larger number of false-positive predictions, suggesting a tendency to classify uncertain samples as malignant. EfficientNetB0 exhibited the most extreme prediction bias, classifying nearly all testing samples as malignant and consequently failing to correctly identify benign cases.



(a) (b) (c)

Figure 4. (a) Comparative Grad-CAM visualizations of DenseNet121, (b) ResNet50, and (c) EfficientNetB0

Grad-CAM visualization was applied to investigate the regions contributing most strongly to model predictions. Figure 4 presents representative heatmaps generated by the evaluated architectures. DenseNet121 produced concentrated activation regions focused on diagnostically relevant tissue structures. ResNet50 generated broader activation patterns distributed across larger image regions, while EfficientNetB0 exhibited more dispersed activation maps with weaker localization. These visual results demonstrate distinct attention

behaviors among the architectures and provide additional evidence supporting the superior classification performance achieved by DenseNet121. Although Grad-CAM visualizations revealed clear differences in attention localization among the evaluated architectures, quantitative explainability evaluation was not performed because the BreakHis dataset does not provide expert-annotated lesion masks or pixel-level pathological annotations. Consequently, metrics such as Intersection over Union (IoU), Localization Accuracy, Pointing Game Accuracy, and Energy-based Pointing Game could not be objectively calculated. Future studies utilizing expert-annotated datasets may enable quantitative validation of explanation quality.

Discussion

The findings demonstrated substantial differences in predictive performance and interpretability among the evaluated transfer learning architectures. DenseNet121 consistently outperformed ResNet50 and EfficientNetB0 across classification metrics, statistical validation, and Grad-CAM analysis, indicating superior representation learning under identical experimental conditions. The superior performance of DenseNet121 may be attributed to its dense connectivity mechanism, which facilitates feature reuse, enhances gradient flow, and preserves discriminative image representations across network layers (Bundea et al., 2024; Mezei et al., 2024). These characteristics are particularly advantageous for histopathological image analysis, where diagnostically relevant information is distributed across multiple spatial and morphological scales. Breast cancer diagnosis requires the simultaneous assessment of subtle cellular features, such as nuclear atypia and chromatin abnormalities, as well as higher-order tissue structures, including glandular organization and stromal architecture. The dense connectivity mechanism enables DenseNet121 to integrate low-level texture information with high-level semantic representations more effectively, thereby capturing the complex hierarchical patterns that characterize histopathological images. This characteristic likely contributed to its superior classification performance compared with ResNet50 and EfficientNetB0. Furthermore, the Grad-CAM visualizations revealed that DenseNet121 generated more focused activation regions on diagnostically relevant tissue structures, suggesting that improved predictive performance was accompanied by enhanced interpretability. These findings indicate that the ability to learn more discriminative histopathological representations may contribute not only to higher classification accuracy but also to more meaningful visual explanations.

A notable finding of this study is the observed relationship between predictive performance and interpretability. DenseNet121 achieved the strongest classification performance while producing the most localized and diagnostically plausible Grad-CAM activation regions. In contrast, ResNet50 and EfficientNetB0 generated broader and less focused activation patterns. These observations suggest that models learning more discriminative feature representations may also provide more coherent visual explanations, consistent with previous studies indicating that improved feature localization can enhance the interpretability of deep learning models in medical imaging applications (Bhati et al., 2024; Talaat et al., 2024). The observed relationship between predictive performance and Grad-CAM localization further supports the view that explainability and predictive effectiveness should be considered complementary dimensions when evaluating artificial intelligence systems for healthcare applications (Rosenbacke et al., 2024). EfficientNetB0 demonstrated the weakest performance despite its success in general computer vision applications. The model exhibited a strong tendency to predict the malignant class, resulting in high sensitivity but poor specificity and overall classification performance. This behavior may be associated with the domain gap between natural images and histopathological images, as well as the class imbalance present in the BreakHis dataset. Grad-CAM visualizations further indicated less focused activation

patterns, suggesting reduced capability to learn discriminative benign tissue representations. These findings indicate that architectural efficiency alone does not guarantee superior performance in computational pathology applications. From a theoretical perspective, the findings suggest that predictive performance and explainability may be closely related through the quality of learned feature representations. Models that capture more discriminative histopathological patterns appear more likely to generate diagnostically meaningful visual explanations, indicating that representation learning quality may influence both classification effectiveness and interpretability (Bhati et al., 2024; Klauschen et al., 2024).

The present findings are generally consistent with previous studies reporting the effectiveness of DenseNet-based architectures in medical image analysis and histopathological classification (Klauschen et al., 2024; Mezei et al., 2024; Ochoa-Ornelas et al., 2025). Nevertheless, direct comparison across studies should be interpreted cautiously because classification outcomes are influenced by dataset characteristics, magnification levels, preprocessing procedures, augmentation strategies, and experimental protocols. Unlike many previous studies that primarily focus on predictive performance, the present work employed a comprehensive evaluation framework incorporating confidence interval estimation, ROC analysis, statistical significance testing, error distribution analysis, and Grad-CAM interpretability assessment. The statistical validation results confirmed that the observed superiority of DenseNet121 was unlikely to be attributable to sampling variation alone.

A quantitative comparison with previous studies further supports the present findings. Rafiq et al. (2025) reported a binary classification accuracy of 98.50% and an AUC of 0.98 using a DenseNet121-based model on the BreakHis dataset. In the present study, DenseNet121 achieved an accuracy of 90.57% and an AUC of 0.9598. The lower performance observed in this study may be associated with differences in experimental design and evaluation procedures. Patient-level evaluation protocols provide a more realistic assessment of model generalization than image-level splitting. Nevertheless, both studies consistently demonstrate the effectiveness of DenseNet-based architectures for breast cancer histopathology classification (Stanescu et al., 2025).

This study contributes both theoretically and methodologically to explainable medical image analysis. From a theoretical perspective, the findings support the proposition that dense feature connectivity facilitates more discriminative histopathological representations and may improve explanation quality. From a methodological perspective, the study introduces a comprehensive evaluation framework integrating classification metrics, confidence intervals, statistical significance testing, ROC analysis, error analysis, and Grad-CAM visualization within a unified experimental pipeline. The findings also have important clinical implications. The combination of strong classification performance and focused visual explanations suggests that DenseNet121 may provide greater transparency in computer-assisted pathology systems. Although such systems are not intended to replace pathologists, they may assist in prioritizing suspicious tissue regions, supporting diagnostic decision-making, and improving workflow efficiency. Furthermore, visual explanations may facilitate model verification and increase confidence in artificial intelligence-assisted diagnosis by providing evidence regarding the tissue regions contributing to classification decisions

Several limitations should be acknowledged. First, the study utilized only the BreakHis dataset at 200× magnification, which may limit generalizability to other datasets and imaging conditions. Second, image-level partitioning was performed because patient identifiers were unavailable, preventing patient-wise validation. Third, the investigation focused exclusively on binary classification and did not evaluate breast cancer subtypes. Fourth, Grad-CAM evaluation remained qualitative because expert-annotated lesion masks were unavailable in the dataset. Consequently, objective explainability metrics such as Intersection over Union (IoU), Localization Accuracy, Pointing Game Accuracy, and Energy-based Pointing Game could not

be computed. Therefore, the Grad-CAM results should be interpreted as qualitative visual evidence supporting model behavior rather than quantitative measurements of explanation quality. Future research should perform external validation using multi-institutional datasets and patient-wise evaluation protocols. The availability of expert-annotated lesion masks would enable quantitative explainability assessment using IoU, Localization Accuracy, Pointing Game Accuracy, and Energy-based Pointing Game metrics. Further investigations may also explore transformer-based architectures, foundation models, self-supervised learning strategies, and multimodal approaches integrating histopathological, molecular, and clinical information to improve the reliability, transparency, and clinical applicability of artificial intelligence systems for breast cancer diagnosis.

CONCLUSION

This study demonstrates that evaluating predictive performance together with explainability provides a more comprehensive assessment of transfer learning models for breast cancer histopathology classification. Among the evaluated architectures, DenseNet121 consistently achieved the best overall performance, combining superior classification accuracy with more focused and diagnostically meaningful Grad-CAM visualizations, indicating its stronger ability to learn discriminative histopathological representations. These findings highlight that predictive effectiveness and interpretability should be considered complementary criteria when developing trustworthy artificial intelligence systems for computational pathology. The proposed evaluation framework contributes to the advancement of transparent AI by integrating performance metrics, statistical validation, and visual explainability within a unified assessment pipeline. Future studies should validate these findings using multi-center datasets, patient-level evaluation protocols, and quantitative explainability metrics to further establish the robustness and clinical applicability of explainable deep learning models.

REFERENCES

Ananda, I. K., Fanani, A. Z., Setiawan, D., & Wicaksono, D. F. (2024). Penerapan Random Oversampling dan Algoritma Boosting untuk Memprediksi Kualitas Buah Jeruk. *Edumatic: Jurnal Pendidikan Informatika*, *8*(1), 282–289. https://doi.org/10.29408/edumatic.v8i1.25836

Anhar, A., Wajidi, F., & Insani, C. N. (2026). Two-Stage Transfer Learning with EfficientNetB0 for Four-Class Banana Ripeness Classification. *Edumatic: Jurnal Pendidikan Informatika*, *10*(2), 300-309. https://doi.org/10.29408/edumatic.v10i1.34588

Bhati, D., Neha, F., & Amiruzzaman, M. (2024). A Survey on Explainable Artificial Intelligence (XAI) Techniques for Visualizing Deep Learning Models in Medical Imaging. *Journal of Imaging*, *10*(10), 239. https://doi.org/10.3390/jimaging10100239

Bundea, M., & Danciu, G. M. (2024). Pneumonia Image Classification Using DenseNet Architecture. *Information*, *15*(10), 611. https://doi.org/10.3390/info15100611

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. https://doi.org/10.1002/hsr2.70017

Houssein, E. H., Gamal, A. M., Younis, E. M. G., & Mohamed, E. (2025). Explainable artificial intelligence for medical imaging systems using deep learning: a comprehensive review. *Cluster Computing*, *28*(7), 469. https://doi.org/10.1007/s10586-025-05281-5

Ivanov, V., Khalid, U., Gurung, J., Dimov, R., Chonov, V., Uchikov, P., Kostov, G., & Ivanov, S. (2025). Use of AI Histopathology in Breast Cancer Diagnosis. *Medicina*, *61*(10), 1878. https://doi.org/10.3390/medicina61101878

Jiang, B., Bao, L., He, S., Chen, X., Jin, Z., & Ye, Y. (2024). Deep learning applications in breast cancer histopathological imaging: diagnosis, treatment, and prognosis. *Breast*

Cancer Research, *26*(1), 137. https://doi.org/10.1186/s13058-024-01895-6

Klauschen, F., Dippel, J., Keyl, P., Jurmeister, P., Bockmayr, M., Mock, A., Buchstab, O., Alber, M., Ruff, L., Montavon, G., & Müller, K.-R. (2024). Toward Explainable Artificial Intelligence for Precision Pathology. *Annual Review of Pathology: Mechanisms of Disease*, *19*(1), 541–570. https://doi.org/10.1146/annurev-pathmechdis-051222-113147

Mezei, T., Kolcsár, M., Joó, A., & Gurzu, S. (2024). Image Analysis in Histopathology and Cytopathology: From Early Days to Current Perspectives. *Journal of Imaging*, *10*(10), 252. https://doi.org/10.3390/jimaging10100252

Naas, M., Benyettou, A., Chougard, H., & Hlou, L. (2025). An explainable artificial intelligence framework for breast cancer classification using vision transformers and explainability techniques. *Biomedical Signal Processing and Control*, *99*, 106984. https://doi.org/10.1016/j.bspc.2025.108011

Ochoa-Ornelas, R., Gudiño-Ochoa, A., García-Rodríguez, J. A., & Uribe-Toscano, S. (2025). A robust transfer learning approach with histopathological images for lung and colon cancer detection using EfficientNetB3. *Healthcare Analytics*, *7*, 100391. https://doi.org/10.1016/j.health.2025.100391

Parameswara, D. A. D., Berlilana, B., & Saputro, R. E. (2026). Utilitarian vs Human-Centered AI Acceptance: Explaining Students' Adoption of ChatGPT in Higher Education. *Edumatic: Jurnal Pendidikan Informatika*, *10*(1), 190–199. https://doi.org/10.29408/edumatic.v10i1.34218

Priya C V, L., V G, B., B R, V., & Ramachandran, S. (2024). Deep learning approaches for breast cancer detection in histopathology images: A review. *Cancer Biomarkers*, *40*(1), 1–25. https://doi.org/10.3233/CBM-230251

Rafiq, A., Jaffar, A., Latif, G., Masood, S., & Abdelhamid, S. E. (2025). Enhanced Multi-Class Breast Cancer Classification from Whole-Slide Histopathology Images Using a Proposed Deep Learning Model. *Diagnostics*, *15*(5), 582. https://doi.org/10.3390/diagnostics15050582

Rosenbacke, R., Melhus, Å., McKee, M., & Stuckler, D. (2024). How Explainable Artificial Intelligence Can Increase or Decrease Clinicians' Trust in AI Applications in Health Care: Systematic Review. *JMIR AI*, *3*, e53207. https://doi.org/10.2196/53207

Sadeghi, Z., Alizadehsani, R., CIFCI, M. A., Kausar, S., Rehman, R., Mahanta, P., Bora, P. K., Almasri, A., Alkhawaldeh, R. S., Hussain, S., Alatas, B., Shoeibi, A., Moosaei, H., Hladík, M., Nahavandi, S., & Pardalos, P. M. (2024). A review of Explainable Artificial Intelligence in healthcare. *Computers and Electrical Engineering*, *118*, 109370. https://doi.org/10.1016/j.compeleceng.2024.109370

Soliman, A., Li, Z., & Parwani, A. V. (2024). Artificial intelligence's impact on breast cancer pathology: a literature review. *Diagnostic Pathology*, *19*(1), 38. https://doi.org/10.1186/s13000-024-01453-w

Stanescu, L., & Stoica-Spahiu, C. (2025). Clinically Oriented Evaluation of Transfer Learning Strategies for Cross-Site Breast Cancer Histopathology Classification. *Applied Sciences*, *15*(23), 12819. https://doi.org/10.3390/app152312819

Talaat, F. M., Gamel, S. A., El-Balka, R. M., Shehata, M., & ZainEldin, H. (2024). Grad-CAM Enabled Breast Cancer Classification with a 3D Inception-ResNet V2: Empowering Radiologists with Explainable Insights. *Cancers*, *16*(21), 3668. https://doi.org/10.3390/cancers16213668