ADBA
Computers and Electronics in Medicine
Cutting-Edge Scientific Solutions

e-ISSN: 3023-8609
RESEARCH ARTICLE
Vol.3 / No.1 / 2026 / pp.54-59
https://doi.org/10.69882/adba.cem.2026017

Towards Robust CAD Systems for Digital Pathology: Evaluating Transformer-Based Backbones for Breast Cancer Classification

Seref Koyuncu *,1, **Yigitcan Cakmak** *,2 **and Ishak Pacal** *,β,3

*Department of Computer Engineering, Faculty of Engineering, Igdir Universty, 76000, Igdir, Türkiye, βDepartment of Electronics and Information Technologies, Faculty of Architecture and Engineering, Nakhchivan State University, AZ 7012, Nakhchivan, Azerbaijan.

ABSTRACT Breast cancer is one of the greatest global health burdens today and demands accurate diagnosis because of the vast histological variety. CNN-based systems had been the dominant technology in Digital Pathology, but with their inability to create a global representation has allowed other technologies such as Vision Transformers to compete. This paper evaluate the performance of three different transformer-based backbone architectures (DeiT Base, Swin Base, and ViT Base) for classifying breast histopathological images into eight granular classes using the BreaKHis database. To facilitate this comparison, we utilize transfer learning and distinct data augmentation methods. Each architecture was fine-tuned to classify four benign and four malignant subtypes with a minimum reported accuracy of 94%, with Swin Base performing more optimally than either of the other two approaches, obtaining highest reported accuracy of 0.9511 and an F1 score of 0.9434. The unique design and shifted windowing processes of Swin Base have allowed this architecture to capture detailed nuclear information as well as the larger context regarding breast cancers, to an extent greater than the other two architectures. Additionally, we provide an in-depth study of confusion matrices in conjunction with high classification accuracy, even when dealing with minor morphological overlap, to further support their claim regarding the ability of Swin Base and the remaining transformer architectures to successfully differentiate between histologically similar classes.

KEYWORDS
Breast cancer
Histopathology images
Multi-class classification
Computer-aided diagnosis
Digital pathology

INTRODUCTION

Breast cancer remains one of the most formidable challenges in global healthcare, consistently ranking as a leading cause of oncological mortality among women worldwide (Getu *et al.* 2025; Chikkala *et al.* 2025). The complexity of the disease, characterized by its high histological heterogeneity, demands diagnostic precision that is both rapid and highly accurate. In recent years, the field of pathology has undergone a significant paradigm shift with the advent of Digital Pathology (DP) and Whole Slide Imaging (WSI) (Karthiga *et al.* 2025; Hayat *et al.* 2024). This transition from traditional glass slides to high-resolution digital patches has paved the way for Computer-Aided Diagnosis (CAD) systems to assist pathologists in navigating the immense workload and reducing the inherent subjectivity of manual assessments (Murphy and Singh 2024; Logu and Thangaraj 2024; Asha 2025).

Deep learning and convolutional neural networks (CNNs) are transforming the way Computer-aided design (CAD) systems operate (Singh and Kaswan 2024; Alshehri 2025). Early CAD systems relied very heavily on the handcrafted features created by DLT and modified by the CAD software (Abdulaal *et al.* 2024b; Behzadpour *et al.* 2024). CNV models are designed specifically to recognize and represent locality patterns at a small scale (in a local-neighborhood context) within an image, constraining their ability to relate and connect long-range or distant regional features to each other and to map across an image with respect to the overall context (Ramamoorthy *et al.* 2024; Ukwuoma *et al.* 2025; Jackson *et al.* 2025). This limitation has led researchers to the Vision Transform is (ViT) architecture, in which all patches of an image processed sequentially to incorporate their information into a complete image representation (Kansal *et al.* 2025; Chitta *et al.* 2025; Jakkaladiki and Maly 2024). However, before ViTs can be routinely used in clinical practice, the full measure of robustness, computational efficiency, and classification accuracy across the full spectrum of breast cancer subtypes must be assessed and validated (Akshaya *et al.* 2024; Abdulaal *et al.* 2024a; Simonyan *et al.* 2024).

The primary objective of this study is to evaluate the efficacy of transformer-based backbones, specifically DeiT Base, Swin Base, and ViT Base, in the multi-class classification of breast histopathology images. Unlike binary classification tasks that merely distinguish between benign and malignant tissue, our work tackles an 8-class challenge involving specific subtypes: Adenosis, Ductal Carcinoma, Fibroadenoma, Lobular Carcinoma, Mucinous Carcinoma, Papillary Carcinoma, Phyllodes Tumor, and Tubular Adenoma. This granular approach is essential for providing clinicians

Manuscript received: 1 September 2025,
Revised: 28 November 2025,
Accepted: 30 November 2025.

[1]serefkoyuncu77@gmail.com
[2]ygtcncakmak@gmail.com(**Corresponding author**)
[3]ishakpacal@igdir.edu.tr

with the detailed diagnostic insights necessary for personalized treatment planning.

RELATED WORKS

The clinical management of breast cancer has been fundamentally reshaped by the transition to DP and the subsequent integration of CAD systems. Historically, pathologists relied on manual microscopic examination, a process prone to inter-observer variability and high cognitive load. Recent advancements in DL have mitigated these challenges by enabling the automated extraction of high-level features that capture the complex morphological heterogeneity of breast tissue. Current research in this domain focuses on enhancing diagnostic reliability through hybrid architectures, ensemble strategies, and novel feature descriptors, particularly when dealing with the granular multi-class classification of histological subtypes.

Ogundokun *et al.* (2024) proposed a robust hybrid DL framework that integrates CNN with Artificial Neural Networks (ANN) specifically for 400x magnification images. By utilizing a dual-pathway approach for feature processing, their model achieves exceptional diagnostic precision, reaching a near-perfect accuracy of 99.94%, which significantly reduces the margin for diagnostic error in high-resolution histopathology. Gul (2025) introduced an innovative textural analysis method termed Quad Star Local Binary Pattern (QS-LBP) paired with a customized 20-layer CNN architecture. This combination effectively encodes fine-grained tissue textures, allowing the system to outperform existing methodologies in distinguishing subtle morphological differences between benign and malignant growths, boasting a peak accuracy of 98.27

Rajaram *et al.* (2024) conducted a systematic evaluation of various ResNet architectures to determine their effectiveness in classifying BreaKHis dataset samples. Their comparative study highlighted that deeper residual networks, when combined with transfer learning, provide a highly scalable solution for multi-class classification, successfully capturing the intrinsic hierarchical patterns of cancerous cell structures. Balasubramanian *et al.* (2024) developed an ensemble learning strategy that fuses the outputs of VGG16, ResNet34, and ResNet50 models to tackle both cancer subtyping and invasiveness. This multi-model approach increases the overall stability of the CAD system and minimizes the risk of misclassification by leveraging a diversified feature set from multiple architectural families.

This work on the performance benchmarking of the BreaKHis dataset through the application of SVM on the textural features as proposed by Thakur *et al.* (2025) sets the standard for the BreaKHis dataset. This work on the benchmarking of performance through the experimental work on patients according to the patient-centric benchmarking methodology of this work on the rigorous benchmarking of patients helps to prevent data leakage. Aldakhil *et al.* (2025) provided a comprehensive review and performance assessment of transfer learning-based architectures, specifically ResNet18, Inception-V3, and ShuffleNet. Their findings emphasize that fine-tuned pre-trained models offer a computationally efficient path toward autonomous breast cancer detection, achieving high accuracy rates while significantly reducing the training time and data requirements typically associated with training from scratch.

MATERIALS AND METHODS

Dataset and Data Preprocessing

The BreaKHis (Breast Cancer Histopathological Image Classification) dataset (Dataset 2025), which is widely used as a benchmark in the field of digital pathology (DP), was used to conduct experiments and evaluate the proposed transformer-based architectures. This dataset contains a complete set of all types of Normal Breast Tissue (eight types), and there are two distinct classes of breast tissue; four (four types) benign, and four (four types) malignant. The dataset consists of 1995 high resolution images. The training and test sets were created by utilizing a structured data split, where 70% of the datasets is used to train the models (1393 samples), 15% is used for validation (295 samples), and 15% is used for independent testing (307 samples) in order to ensure that each model is trained appropriately and evaluated fairly. The exact breakdown of the samples across the eight types of breast tissue is provided in Table 1 and depicts the inherent imbalances found in clinical datasets.

The visual complexity and histological diversity of the samples are demonstrated in Figure 1, where some sample images of the eight histopathological classes are shown. These sample images demonstrate the level of cellular architecture similarity between the different classes, an issue that has often been known to make the task of pathological evaluation challenging when performed on a slide by slide basis. To eventually preprocess the visual data in a form conducive to the transformers chosen in this work as their base models, namely DeiT, Swin, and ViT, some image processing operations such as resizing and normalization have been performed on the sample images. Moreover, in order to help the transformers attain the required level of generalizability and avoid the problems of a small number of samples in classes such as Adenosis, a data augmentation process has been followed.

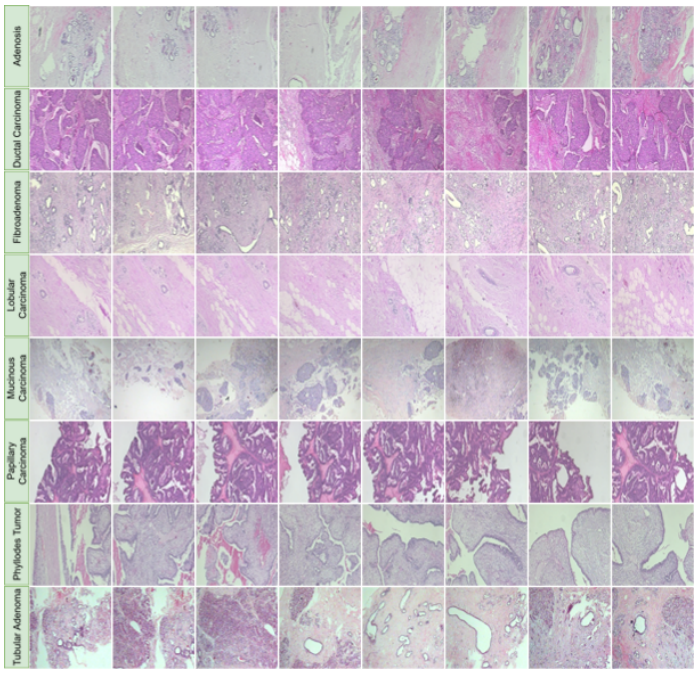

Figure 1 Sample images from the breast histopathology classes in the dataset.

Vision Transformers (ViTs)

The advent of ViTs marks the beginning of an important shift in the paradigm of medical image processing, diverging from the local receptive field requirement imposed by conventional CNNs. Although CNNs excel in local texture description, their efficacy in describing global dependencies in the image is inherently

Table 1 Dataset Distribution by Class.

Class	Original Dataset	Train Set (70%)	Validation Set (15%)	Test Set (15%)
Adenosis (ADE)	114	79	17	18
Ductal Carcinoma (DC)	864	604	129	131
Fibroadenoma (FIB)	253	177	37	39
Lobular Carcinoma (LC)	156	109	23	24
Mucinous Carcinoma (MC)	205	143	30	32
Papillary Carcinoma (PC)	145	101	21	23
Phyllodes Tumor (PT)	109	76	16	17
Tubular Adenoma (TA)	149	104	22	23
Total	1995	1393	295	307

limited, making it difficult to capture the essence of complicated histopathological patterns without considering the whole image in the process. This is where ViTs accentuate the advantage with their patch-based division of the image into a series of discrete segments and self-attention techniques, which help the model extract correlations even in the far-off regions of cells right from the beginning, opposed to CNNs, which tend to concentrate on local cell morphologies without emphasis on global architecture in digital pathology, where the difference in tissue might be quite subtle and relied on entirely on the global pattern.

In this study, we evaluate three distinct transformer-based architectures, ViT Base (Dosovitskiy *et al.* 2025), DeiT Base (Touvron *et al.* 2025), and Swin Base (Liu *et al.* 2025), to determine their effectiveness in classifying eight specific breast cancer subtypes. While the standard ViT Base provides a robust foundation for learning global representations, the Data-efficient Image Transformer (DeiT Base) is specifically designed to perform effectively on smaller datasets through a teacher-student distillation strategy, making it highly relevant for specialized medical imaging tasks where data can be scarce. Furthermore, we investigate the Swin Transformer (Swin Base), which introduces a hierarchical structure and shifted windowing scheme. This approach allows the model to process images at multiple scales, effectively capturing both fine-grained nuclear details and broader tissue patterns, which is vital for the granular 8-class classification challenge involving diverse benign and malignant subtypes.

Transfer Learning and Data Augmentation Strategy

The size and variability of the BreaKHis dataset and the depth of the deep transformer models make deep learning converge inefficiently. To counter these issues, we employed the concept of transfer learning. As ViTs are huge and demand the availability of vast amounts of data to train and converge properly, doing so on the relatively small number of images in the histopathology domain would be inefficient and would indulge the model into the problem of overfitting. To overcome this problem, we made use of the PyTorch Image Models library to load the pre-trained models of DeiT Base, Swin Base, and ViT Base pre-trained models on the ImageNet-1k dataset. The use of these models enables them to focus and work on the high-level visual representations right from the start, which would be further tuned to focus on the complexities of the eight sub-types of breast cancer.

As a way of compensating for the natural imbalance between classes in the dataset and to ultimately strengthen our model's performance, we created a highly effective method of augmenting our images through the use of a complete set of augmentation strategies provided by the timm library. Examples of these strategies include random resizing, random cropping, and random horizontal mirroring. The application of these augmentation techniques in the field of digital pathology represents a novel approach because they replicate the variations that occur naturally in a tissue's orientation and staining intensity during slide processing. By exposing our transformer models to the diverse range of visual challenges that arise from these augmentation methods during training, our models were trained to learn the underlying pattern of abnormality rather than simply learning specific signs of pathology through visual memorization (Wang *et al.* 2024; Mumuni *et al.* 2024).

Performance Evaluation Metrics

In this paper, we have conducted a rigorous quantification of the diagnostic efficacy of the transformer-based backbones that were studied using a comprehensive set of performance evaluation metrics derived from the confusion matrix. Overall, classification performance was primarily conducted with Accuracy, representing the ratio of correctly identified benign and malignant samples out of all total predictions defined in Equation (1). Given the clinical need to minimize both false positives and false negatives in breast cancer screening, Precision and Recall (Sensitivity) were calculated. Precision, defined in Equation (2), reflects the model's ability not to label a negative sample as positive. On the other hand, Recall, defined in Equation (3), characterizes the ability to detect all positive examples within the dataset. Since there are severe class imbalances within our histological subtypes, we present the F1-Score as the single, balanced metric that considers the possible trade-offs between those two measures, which is the harmonic mean of Precision and Recall, and defined in Equation (4). The virtues of such multi-faceted performance metrics ensure a robust evaluation of the models' reliability in high-stakes computer-aided diagnosis.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

Table 2 Comparison of Performance Metrics and Computational Complexity of the Models.

Models	Accuracy	Precision	Recall	F1-score	Params	GFLOPs
DeiT Base	0.9446	0.9466	0.9192	0.9308	85.80M	336.955
Swin Base	0.9511	0.9548	0.9332	0.9434	86.75M	303.375
ViT Base	0.9414	0.9353	0.9319	0.9323	85.80M	33.6955

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

$$\text{F1-score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

RESULTS AND DISCUSSION

The experimental results obtained in this study underscore the remarkable potential of ViTs in navigating the complex histological landscape of breast cancer. Our comprehensive evaluation across eight distinct tissue subtypes reveals that transformer-based architectures can effectively learn the subtle, high-dimensional features required for precise computer-aided diagnosis. As summarized in Table 2, all three evaluated models, DeiT Base, Swin Base, and ViT Base, exhibited robust performance, with accuracy scores consistently exceeding 94%. This high baseline suggests that the self-attention mechanism is inherently well-suited for capturing the structural variations present in BreaKHis histopathology images.

The Swin Base model outperformed all other backbones tested in this experiment, reaching a peak accuracy of 0.9511 and an F1-score of 0.9434. The Swin architecture has a hierarchical design along with a shifted windowing method; this allows the Swin model to handle tissue slides at multiple scales. The ability to work at various scales is critical when detecting small nuclear atypia and more general architectural distortions. The second and third most successful backbones in accuracy were the DeiT Base model with an accuracy of 0.9446 and the ViT Base model with an accuracy of 0.9414. Despite being the least computationally intensive backbone (GFLOPs of 33.69), the ViT Base model still demonstrated a high recall value of 0.9319, indicating that standard transformer-based models can also have excellent detection sensitivity to malignant cells. However, the Swin Base model's high levels of precision (0.9548) and recall (0.9332) make it the most suitable choice for clinical use, where the prevention of false positives and false negatives is critical.

The diagnostic reliability of our leading machine learning (ML) model can be evaluated using the confusion matrix in Figure 2. Based on the confusion matrix, the Swin Base model demonstrated considerable accuracy for high-frequency immortalized (cancerous) classes, such as Ductal Carcinoma (DC), which correctly classified 128 out of 131 cases of Ductal Carcinoma (DC). Additionally, the Swin Base model exhibited strong predictive ability for Fibroadenomas (FIB); there was only one case in which Fibroadenomas (FIB) were incorrectly classified. The Swin Base model encountered minor confusion regarding histologically similar subtypes: four out of four Lobular Carcinomas (LC) were misclassified as Ductal Carcinomas (DC). As can be expected from a biological standpoint, Ductal Carcinomas (DC) and Lobular Carcinomas (LC) are both malignant carcinomas that may show overlap in morphology, depending on their magnification level. Additionally, the Swin Base model had slightly more difficulty with Tubular Adenomas (TA) than with Fibroadenomas (Fib), and there were occasions in which Tubular Adenomas (TA) were misclassified as Fibroadenomas (Fib). Nevertheless, the relatively high diagonal values for all classes in the Swin Base model (Figure 2) indicate that it has successfully generalized to the considerable histologic heterogeneity present in the dataset.

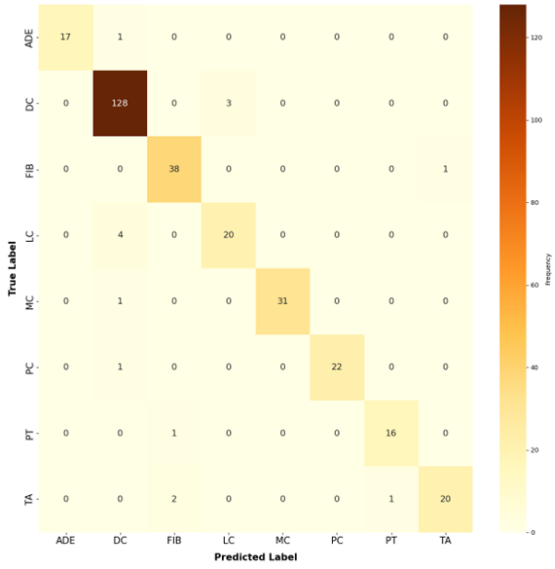

Figure 2 Confusion Matrix of the Swin Base Model.

The findings of this study indicate that transitioning away from utilizing convolutional filters to employing hierarchical transformer systems can enhance the overall experience one has with viewing digital slides. In particular, the excellent performance of the Swin Transformer shows the value of being able to extract features at multiple scales in digital pathology. Thus, capturing long-range spatial relationships between various features in conjunction with maintaining local texture analysis allows for the development of a strong base for future CAD systems that are designed to help pathologists throughout their highly complicated workflows related to diagnosing patients.

CONCLUSION

This study has systematically assessed how effective ViTs are as a new model for multi-class classification of breast cancer histopathology. Results indicate that the ability of the self-attention mechanism to manage multiple structural complexities of digital slides allows for a broader understanding of tissue architecture than the limited local receptive field of Classic CNNs. Of the models tested, the Swin Base was determined to provide the most

reliable performance, achieving an ideal balance of precision and recall, while also successfully addressing the inherent imbalances among the eight class categories present in the BreaKHis dataset. Transfer learning combined with extensive data augmentation was critical to mitigate the impact of infrequent impressions typically seen in this type of medical imaging task and ensure the models created general representations of pathology and did not rely on representations of image artifacts. Mismatches did occur within histologically similar subtypes (e.g., Lobular and Ductal Carcinoma), but the very high diagonal values achieved within the performance evaluation indicate the strong capacity of hierarchical transformers to differentiate between the eight tissue categories.

Ethical standard

The authors have no relevant financial or non-financial interests to disclose.

Availability of data and material

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

LITERATURE CITED

Abdulaal, A. H., M. Valizadeh, M. C. Amirani, and A. F. S. Shah, 2024a A self-learning deep neural network for classification of breast histopathological images. Biomedical Signal Processing and Control **87**: 105418.

Abdulaal, A. H., R. A. Yassin, M. Valizadeh, A. H. Abdulwahhab, A. M. Jasim, *et al.*, 2024b Cutting-edge cnn approaches for breast histopathological classification: The impact of spatial attention mechanisms. ShodhAI: Journal of Artificial Intelligence **1**: 109–130–109–130.

Akshaya, K., A. Bhan, and S. Pathan, 2024 Deep learning based classification approach to improve breast cancer screening using histopathological images. 2024 International Conference on Advances in Modern Age Technologies for Health and Engineering Science, AMATHE 2024 .

Aldakhil, L. A., H. F. Alhasson, S. S. Alharbi, R. U. Khan, and A. M. Qamar, 2025 Image-based breast cancer histopathology classification and diagnosis using deep learning approaches. Applied Computational Intelligence and Soft Computing **2025**: 7011984.

Alshehri, H. M., 2025 Brenet: Attention-enhanced multi-scale cnn framework for breast cancer classification in histopathological images. IEEE Access .

Asha, N., 2025 Breast cancer data augmentation with detection using cnn model in deep learning. Lecture Notes in Electrical Engineering **1246 LNEE**: 325–339.

Balasubramanian, A. A., S. M. A. Al-Heejawi, A. Singh, A. Breggia, B. Ahmad, *et al.*, 2024 Ensemble deep learning-based image classification for breast cancer subtype and invasiveness diagnosis from whole slide image histopathology. Cancers **16**.

Behzadpour, M., B. L. Ortiz, E. Azizi, and K. Wu, 2024 Breast tumor classification using efficientnet deep learning model .

Chikkala, R. B., C. Anuradha, P. S. C. Murty, S. Rajeswari, N. Rajeswaran, *et al.*, 2025 Enhancing breast cancer diagnosis with bidirectional recurrent neural networks: A novel approach for histopathological image multi-classification. IEEE Access **13**: 41682–41707.

Chitta, S., S. Sharma, and V. K. Yandrapalli, 2025 Hybrid deep learning model for enhanced breast cancer diagnosis using histopathological images. Procedia Computer Science **260**: 245–251.

Dataset, 2025 Breakhis - breast cancer histopathological dataset.

Dosovitskiy, A., L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, *et al.*, 2025 An image is worth 16x16 words: Transformers for image recognition at scale .

Getu, M. A., C. Lu, Y. Liu, A. Mehmood, Z. Iqbal, *et al.*, 2025 Bcnet: A novel deep learning model for enhanced breast cancer classification using histopathological images .

Gul, M., 2025 A novel local binary patterns-based approach and proposed cnn model to diagnose breast cancer by analyzing histopathology images. IEEE Access **13**: 39610–39620.

Hayat, M., N. Ahmad, A. Nasir, and Z. A. Tariq, 2024 Hybrid deep learning efficientnetv2 and vision transformer (effnetv2-vit) model for breast cancer histopathological image classification. IEEE Access **12**: 184119–184131.

Jackson, J., L. E. Jackson, C. C. Ukwuoma, M. D. Kissi, A. Oluwasanmi, *et al.*, 2025 A patch-based deep learning framework with 5-b network for breast cancer multi-classification using histopathological images. Engineering Applications of Artificial Intelligence **148**: 110439.

Jakkaladiki, S. P. and F. Maly, 2024 Integrating hybrid transfer learning with attention-enhanced deep learning models to improve breast cancer diagnosis. PeerJ Computer Science **10**: e1850.

Kansal, K., S. Kumar, and K. Kansal, 2025 Advances in deep learning techniques for breast cancer classification: A comprehensive review. Archives of Computational Methods in Engineering pp. 1–36.

Karthiga, R., K. Narasimhan, N.Raju, and R. Amirtharajan, 2025 Automatic approach for breast cancer detection based on deep belief network using histopathology images. Multimedia Tools and Applications **84**: 4733–4750.

Liu, Z., Y. Lin, Y. Cao, H. Hu, Y. Wei, *et al.*, 2025 Swin transformer: Hierarchical vision transformer using shifted windows .

Logu, K. and S. J. J. Thangaraj, 2024 Progressing breast cancer assessment: Precise tumor categorization through densenet201-based deep learning. TQCEBT 2024 - 2nd IEEE International Conference on Trends in Quantum Computing and Emerging Business Technologies 2024 .

Mumuni, A., F. Mumuni, and N. K. Gerrar, 2024 A survey of synthetic data augmentation methods in machine vision. Machine Intelligence Research 2024 21:5 **21**: 831–869.

Murphy, G. and R. Singh, 2024 Comparative analysis and ensemble enhancement of leading cnn architectures for breast cancer classification .

Ogundokun, R. O., A. R. T. Abdullahi, A. R. Adenike, C. Awoniyi, H. B. Akande, *et al.*, 2024 Hybrid deep learning for breast cancer diagnosis: Evaluating cnn and ann on breakhis_v1_400x. In *International Conference on Science, Engineering and Business for Driving Sustainable Development Goals, SEB4SDG 2024*, Institute of Electrical and Electronics Engineers Inc.

Rajaram, G., V. Rajinikanth, and B. S. Latha, 2024 Classification of breast histology into benign and malignant using resnet-variants: A study with breakhis database. In *2024 9th International Conference on Applying New Technology in Green Buildings, ATiGB 2024*, pp. 479–483, Institute of Electrical and Electronics Engineers Inc.

Ramamoorthy, P., B. R. R. Reddy, S. S. Askar, and M. Abouhawwash, 2024 Histopathology-based breast cancer prediction using deep learning methods for healthcare applications. Frontiers in Oncology **14**: 1300997.

Simonyan, E. O., J. A. Badejo, and J. S. Weijin, 2024 Histopathological breast cancer classification using cnn. Materials Today: Proceedings **105**: 268–275.

Singh, A. and K. S. Kaswan, 2024 Empirical analysis on breast cancer datasets with machine learning. Proceedings - International Conference on Computing, Power, and Communication Technologies, IC2PCT 2024 pp. 223–227.

Thakur, N., S. Shrivastava, S. Shukla, and M. Gyanchandani, 2025 Patient-centric multilabel classification with svm on breakhis dataset. In *2025 3rd International Conference on Data Science and Network Security (ICDSNS)*, pp. 1–7, IEEE.

Touvron, H., M. Cord, M. Douze, F. Massa, A. Sablayrolles, *et al.*, 2025 Training data-efficient image transformers & distillation through attention .

Ukwuoma, C. C., D. Cai, E. O. Eziefuna, A. Oluwasanmi, S. F. Abdi, *et al.*, 2025 Enhancing histopathological medical image classification for early cancer diagnosis using deep learning and explainable ai – lime & shap. Biomedical Signal Processing and Control **100**: 107014.

Wang, Z., P. Wang, K. Liu, P. Wang, Y. Fu, *et al.*, 2024 A comprehensive survey on data augmentation .

How to cite this article: Koyuncu, S., Cakmak, Y. and Pacal, I. Towards Robust CAD Systems for Digital Pathology: Evaluating Transformer-Based Backbones for Breast Cancer Classification. *Computers and Electronics in Medicine*, 3(1), 54-59, 2026.