

Review of Computer Engineering Research

2025 Vol. 12, No. 4, pp. 257-272

ISSN(e): 2410-9142



ISSN(p): 2412-4281

DOI: 10.18488/76.v12i4.4562

© 2025 Conscientia Beam. All Rights Reserved.



Iterative multi-scale deep learning framework for reliable breast cancer diagnosis

 Komal S. Gandle¹⁺
 Dhananjay B.
Kshirsagar²

^{1,2}Department of Computer Engineering, Sanjivani College of Engineering Kopergaon, Savitribai Phule Pune University, Pune, Maharashtra, India.

¹Email: komgan03@rediffmail.com

²Email: dbk4444@gmail.com



(+ Corresponding author)

ABSTRACT

Article History

Received: 28 July 2025

Revised: 15 October 2025

Accepted: 10 November 2025

Published: 2 December 2025

Keywords

Breast cancer
Diagnostic calibration
Graph neural networks
Histopathological analysis
Medical image processing
Self-supervised learning.

Accurate diagnosis of breast cancer from histopathological images is challenging due to variable tissue morphology and the subjectivity of manual interpretations. While CAD systems offer automated diagnosis, they often lack robust feature representation, contextual understanding, and integration of expert knowledge, limiting their effectiveness, especially in distinguishing invasive from pre-invasive carcinoma. This study presents a comprehensive deep learning-based diagnostic framework that integrates five novel modules to improve interpretability, feature robustness, and decision reliability. The Multiscale Attention Integrated Self-Supervised Representation (MAISSR) learns pathology-aware embeddings via co-optimized multiscale attention and contrastive learning. The Morphological-Geodesic Graph Convolutional Network (MG-GCN) combines geodesic topology with glandular morphology in a spatial graph model to capture epithelial transitions. The Hyper-Resolution Fusion Network with Cellular-Density Priors (HRF-CDPNet) enhances resolution in critical regions using cellular density maps. The Contextual Relational Transformer with Progression-Encoding (CRT-PE) models disease progression using spatial-contextual tokens to improve invasion mapping. Finally, Adaptive Cross-Modality Decision Calibration (ACM-DC) uses a reinforcement-learning-based agent to align machine predictions with expert annotations, especially in ambiguous cases. This integrated approach yields marked improvements in diagnostic metrics: F1 score increased from 82.5% to 89.3%, AUC from 0.88 to 0.94, and diagnostic agreement with experts from 85.2% to 94.8%. Overall, this work demonstrates the potential of a multi-factorial, multi-perspective framework to advance breast cancer diagnosis through optimized feature learning, spatial reasoning, and expert-machine synergy.

Contribution/Originality: This study contributes to the existing literature by introducing an integrated multi-scale, context-aware deep learning framework for breast cancer diagnosis. It employs a novel combination of self-supervised representation learning, graph-based topology modeling, and transformer-driven progression encoding. Additionally, this research is among the few investigations that have examined the calibration of model predictions with expert annotations in histopathology.

1. INTRODUCTION

Millions of women worldwide are affected by breast cancer, which remains one of the most common cancers and a major challenge in oncological pathology. Histopathological examination of tissue biopsies continues to serve as the gold standard for diagnosis and grading. However, manual analysis by pathologists is labor-intensive [1-3], prone to inter-observer variability, and difficult to scale in high-throughput clinical workflows. Despite advances in digital

pathology, interpreting whole-slide images (WSIs) still requires significant expertise, particularly when distinguishing subtle morphological differences between pre-invasive and invasive carcinoma. This has driven growing interest in computer-aided diagnostic (CAD) tools that can support pathologists in making more consistent and accurate decisions.

Recent CAD systems powered by deep learning have advanced histopathological image classification and detection. However, important limitations remain. Many approaches depend on supervised learning, which requires large annotated datasets that are expensive and time-consuming to generate. Others fail to capture spatial context and disease progression, reducing reliability in complex or ambiguous cases. In addition, most models overlook the integration of human annotations during inference, limiting interpretability and reducing trust in borderline scenarios [4-6]. These challenges highlight the need for a diagnostic system that can learn from limited data, incorporate spatial and morphological priors, and align predictions with expert judgment.

To address these gaps, this study proposes a multi-component framework that integrates self-supervised representation learning, graph-based topology modeling, resolution-aware image enhancement, transformer-based contextual modeling, and expert-informed calibration. By jointly considering both micro- and macro-level disease progression, the framework enhances accuracy, interpretability, and clinical relevance. It consists of five specialized modules: (1) the MAISSR model for pathology-aware embeddings, (2) MG-GCN for spatial topology modeling, (3) HRF-CDPNet for improving micro-region clarity, (4) CRT-PE for embedding inter-patch relationships and progression cues, and (5) ACM-DC for expert-guided prediction calibration using reinforcement learning. Together, these modules form a robust and clinically applicable framework for distinguishing invasive from pre-invasive breast cancer.

1.1. Motivation and Contribution

The rationale for this work arises from the gap between automated image-based diagnosis and clinically accountable decision-making in breast cancer pathology. Conventional CAD systems often fail to capture multiscale tissue inhomogeneity, where subtle changes in glandular architecture, nuclear morphology, and spatial context indicate disease progression. Moreover, expert input is typically restricted to the training stage, rather than influencing inference-time decisions, which limits clinical trust. These gaps create the opportunity to design a diagnostic tool that is not only accurate but also interpretable, context-aware, and resilient to variability in tissue presentations.

Our proposed framework addresses these challenges by integrating five complementary components. The MAISSR module leverages multiscale attention and contrastive learning to extract discriminative features from regions of interest without exhaustive labeling. The MG-GCN encodes glandular topologies using geodesic-based graph representations, enabling better spatial understanding of epithelial structures. The HRF-CDPNet improves resolution in diagnostically significant regions by incorporating cellular density priors. The CRT-PE embeds clinical progression cues and inter-patch relationships through transformer-based modeling. Finally, the ACM-DC module aligns predictions with expert annotations using reinforcement learning. Together, these components enhance classification performance, embedding separability, and alignment with pathologist assessments, resulting in a clinically viable solution for breast cancer histopathology.

2. REVIEW OF EXISTING MODELS USED FOR CANCER DETECTION AND PREDICTION ANALYSIS

The review starts with Liu et al. [1] presenting a kernel-based CNN-PCFF framework that incorporates probabilistic feature fusion and convolutional feature maps, which became the basis of several emerging algorithms that later attempted to hybridize the feature spaces. Kaddes et al. [2] introduced CNN-LSTM to learn both spatial and temporal properties for tissue image classification, demonstrating early potential in providing a temporal

representation of histological structures. Wang et al. [3] introduced one of the early benchmark hybrid networks in a two-stage pipeline for multi-class classification with separation between feature extraction and refinement. Turova et al. [4] emphasized molecular-level knowledge by utilizing deep-learning technology for HER2-low subtypes, challenging the current boundary between phenotypic and genotypic classification. Shin et al. [5] considered the issue from a molecular probe design viewpoint to optimize the recognition of microRNA for subtype classification. Thus, shifting the consideration from visual clues to molecular cues shows how critical data diversity is for cancer diagnostics. Ndlovu et al. [6] brought in population-specific variability using IHC biomarkers in a regional cohort from Botswana, citing the effects of being demographic on classifier generalization sets.

Iteratively, Next, as per Table 1, later on, Deebani et al. [7] demonstrated a synergistic combination of transfer learning and adversarial networks, showing increasing research interest in domain adaptation and robustness against varying staining and acquisition protocols. Gupta et al. [8] combined deep transfer learning with a hybrid approach to enhance accuracy in histopathological tumor diagnosis. According to Islam et al. [9], applying multiscale context and global reasoning has favored the recognition of subtle gradation changes from tumor grades. Sharma and Singla [10] proposed an advanced PTSVM method with weighted feature selection to maximize the discriminative strength with respect to the selected features. Zhang et al. [11] introduced the use of radiomics, which consists of extracting high-dimensional features from ultrasound images, extending computational histopathology into multimodal imaging. Using some classical machine learning classifiers, which meet the conditions under which deep learning may not prove tenable, is the comparative study offered by Kavitha et al. [12]. As a multi-class classifier, Chakravarthy et al. [13] proposed to hybridize CNN features, strengthening the idea that no single feature space suffices for subtle diagnostic discrimination. Kavitha et al. [12] and Kavitha and Sridevi [14] Megan Disu explored a graph-based model combined with mechanisms to achieve better efficiency and representation of glandular structures. Sugawara et al. [15] devised a jigsaw puzzle-based self-supervised approach to learn breast tissue spatial dependencies in an unsupervised manner, which demonstrates the interest in self-supervised pretraining within the field. Patro et al. [16] evaluated the use of sparse architectures such as HCNN and LeNet-5, which demonstrate the tradeoff between accuracy and computational constraints. The model proposed by El-Shazli et al. [17] incorporated self-attention in 3D tomosynthesis images with ensemble classifiers and extended classical analysis in 2D images toward volumetric medical modalities. In the work of Yusuf et al. [18], this full-size component complemented an improved shallow CNN model, which mitigated high-performance requirements from traditional deep networks for histopathological image samples. Feature selection has played a role in the survival of classifiers, and findings such as mRMR+SSO+WSVM hybrids in building efficient classifiers are presented by Yaqoob et al. [19]. The graph-assisted global reasoning about all of the patches was implemented in the study of Zhao and Du [20], which proved to be very useful and effective in heterogeneous tumor cases.

Dalai et al. [21] have introduced the CAR model that combines CNN with attention mechanisms and residual connections, demonstrating significant progress in class separability and interpretability. Liu et al. [1] and Das et al. [22], in another independent work, have strengthened the effectiveness of kernel-based probabilistic convolution with an optimized embedding structure for fine resolution. Das et al. [22] have adopted hybridization of neural and metaheuristic algorithms, combined ensemble learning with TOPSIS-based multi-criteria decision making, which reflects a new evaluation paradigm focusing on criteria beyond accuracy. Sota et al. [23] investigated the IRSN-23 gene expression diagnosis angle on biological ground-truth based on computational prediction. Ma et al. [24] delineated LMCNet, a lightweight yet powerful CNN for multi-class classification, giving it instantaneous applicability in clinical settings upon deployment.

Table 1. Model's empirical review analysis.

Reference	Method	Main Objectives	Findings	Limitations
Liu, et al. [1]	CNN-PCFF	Combine kernel-based CNN with probabilistic feature fusion	Improved classification accuracy with embedded kernel enhancement.	Limited evaluation on multi-class subtypes.
Kaddes, et al. [2]	CNN-LSTM Hybrid	Integrate CNN for spatial and LSTM for temporal tissue structure modeling.	Improved robustness in sequential data	High computational complexity
Wang, et al. [3]	Two-Stage Hybrid Network	Separate feature extraction and classification into two stages	High accuracy in multi-class classification	Limited interpretability
Turova, et al. [4]	HER2-low Classifier	Refine molecular classification using HER2 markers.	Better subtype delineation	Dependent on molecular profiling availability
Shin, et al. [5]	miRNA Probe Model	Use optimal probe sequence design for miRNA recognition	High sensitivity in subtype detection	Not image-based, lacks visual interpretability.
Ndlovu, et al. [6]	IHC Biomarker Staining	Apply IHC staining biomarkers for classification in regional dataset.	Effective classification in Botswana cohort	Region-specific generalizability issues
Deebani, et al. [7]	Transfer Learning + GAN	Fuse transfer learning with adversarial networks.	Boosted performance under staining variability	Risk of mode collapse in GAN training
Gupta, et al. [8]	Deep Transfer Learning Hybrid	Leverage pre-trained networks with fine-tuned histology layers.	Enhanced precision and recall	Overfitting in small datasets
Islam, et al. [9]	Global + Multiscale Context Fusion	Fuse global context with multiscale learning.	Improved tumor grade differentiation.	Higher training time and memory usage
Sharma and Singla [10]	PTSVM with Feature Weighting	Apply weighted feature selection with PTSVM	Improved classification margin	Less effective in high-dimensional feature spaces
Zhang, et al. [11]	Radiomics with Ultrasound	Use radiomic features from ultrasound imaging.	Accurate subtype classification	Limited to imaging modality, not histopathology
Kavitha, et al. [12]	Traditional ML	Evaluate ML classifiers on breast cancer data	Good performance on structured datasets	Low flexibility for high-dimensional image data
Chakravarthy, et al. [13]	CNN Feature Hybridization	Hybridize CNN features for multi-class output	Improved inter-class discrimination	Limited performance on unseen domains
Kavitha and Sridevi [14]	GCN + DenseNet121 with Pruning	Combine GCN with DenseNet and pruning.	Better glandular structure encoding	High resource requirement during training
Sugawara, et al. [15]	Jigsaw Self-Supervision	Use jigsaw puzzles to learn spatial structure.	Unsupervised pretraining improved accuracy.	Fails on highly disrupted tissue
Patro, et al. [16]	HCNN + LeNet5	Lightweight hybrid CNN architecture	Efficient and fast	Lower performance on complex samples
El-Shazli, et al. [17]	MSAE-DL	Fuse self-attention with ensemble classification	Improved interpretability and classification	Complexity in tuning multiple components

Reference	Method	Main Objectives	Findings	Limitations
Yusuf, et al. [18]	Enhanced Shallow CNN	Use modified shallow CNN for histology	Good accuracy with less computation	Limited scalability to large datasets
Yaqoob, et al. [19]	mRMR + SSO + WSVM	Feature selection with hybrid SVM approach	Higher precision after dimensionality reduction	Sub-optimal on raw image data
Zhao and Du [20]	Graph-Assisted Reasoning	Use global reasoning on patch graphs	High contextual accuracy	Graph complexity affects speed
Dalai, et al. [21]	CAR Model	Hybrid CNN-Attention-Residual connection	Stable performance with deep features	Risk of overfitting with limited data
Das, et al. [22]	Neural-Metaheuristic Hybrid	Combine NN with metaheuristic & TOPSIS	Robust under multi-criteria decision metrics	Complex implementation pipeline
Sota, et al. [23]	IRSN-23 Gene Classifier	Integrate gene signatures with diagnostics.	Improved chemo-response prediction	Not universally deployable without gene testing
Ma, et al. [24]	LMCNet	Design a lightweight CNN for multi-class breast cancer classification.	Efficient with competitive accuracy	Limited explainability in decision layers

Despite advances, existing models often neglect spatial-context integration, progression cues, and expert-informed calibration. These gaps motivate the proposed multi-component framework that unifies self-supervised feature embedding, graph-based reasoning, and transformer-based contextual modeling for reliable diagnosis.

3. PROPOSED MODEL DESIGN ANALYSIS

The breast cancer histopathology diagnosis framework suggested will model multi-scale, spatial, contextual, platform- and expert-informed features in combination to deliver a strong classification of their manually annotated regions of interest (ROIs) into either invasively or pre-invasively graded ones. The design consists of five subsystems, selected on the basis of their unique representational strength and integration capabilities within an integrated architecture. By sequentially interfacing embedding, structural, contextual, resolution-enhanced, and decision-calibrated output along tightly coupled pipelines, the integration constrains the emergence of redundancy while preserving feature complementarity. The pathway starts with the Multiscale Attention Integrated Self-Supervised Representation (MAISSR) module.

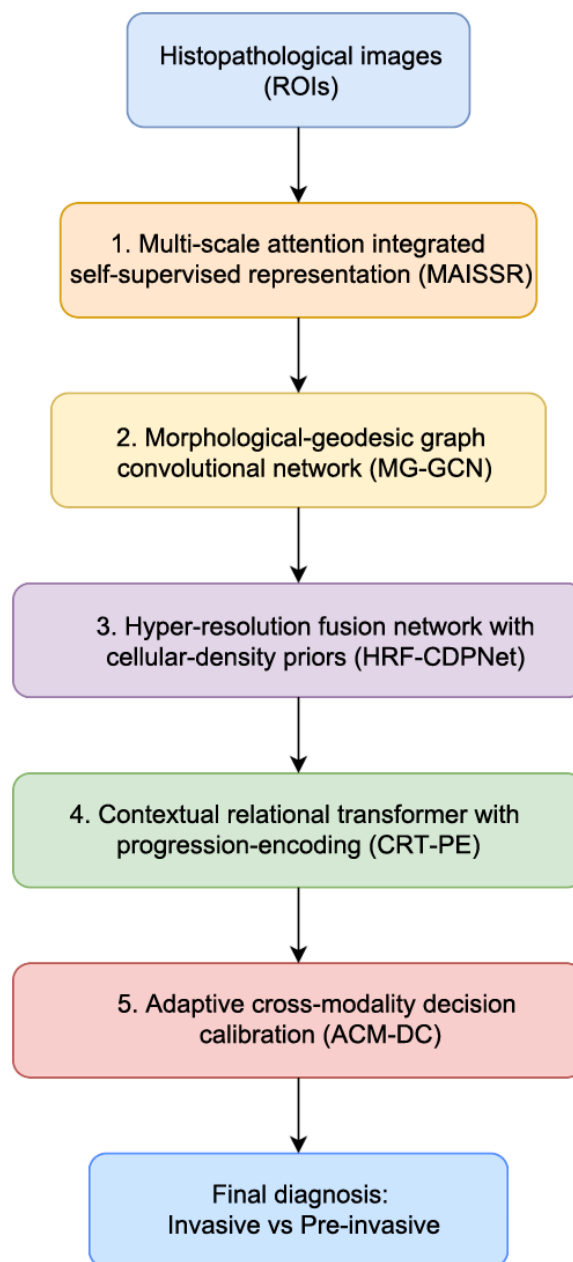


Figure 1. Overview of the proposed diagnostic framework.

Figure 1 shows a high-level representation of the proposed pipeline for breast cancer diagnosis from histopathological images. The system integrates five sequential modules: (1) Multi-Scale Attention Integrated Self-Supervised Representation (MAISSR) for feature embedding, (2) Morphological-Geodesic Graph Convolutional Network (MG-GCN) for spatial topology modeling, (3) Hyper-Resolution Fusion Network with Cellular-Density Priors (HRF-CDPNet) for enhancing micro-region clarity, (4) Contextual Relational Transformer with Progression Encoding (CRT-PE) for capturing inter-patch relationships and disease progression cues, and (5) Adaptive Cross-Modality Decision Calibration (ACM-DC) for expert-aligned prediction refinement.

For an ROI image patch I , the multiscale encoder $E_s(I)$ outputs scale-specific feature maps $\{f_1, f_2, \dots, f_S\}$ where each $f_s \in \mathbb{R}^{C_s \times H_s \times W_s}$. They go through an attention head that computes scale-normalized attention weights as ' α_s ' via Equation 1.

$$\alpha_s = \frac{\exp(w^T \tanh(Wa f_s + ba))}{\sum_{j=1}^S \exp(w^T \tanh(Wa f_j + ba))} \quad (1)$$

The attended feature embedding z is constructed as a convex combination via Equation 2.

$$z = \sum_{s=1}^S \alpha_s f_s \quad (2)$$

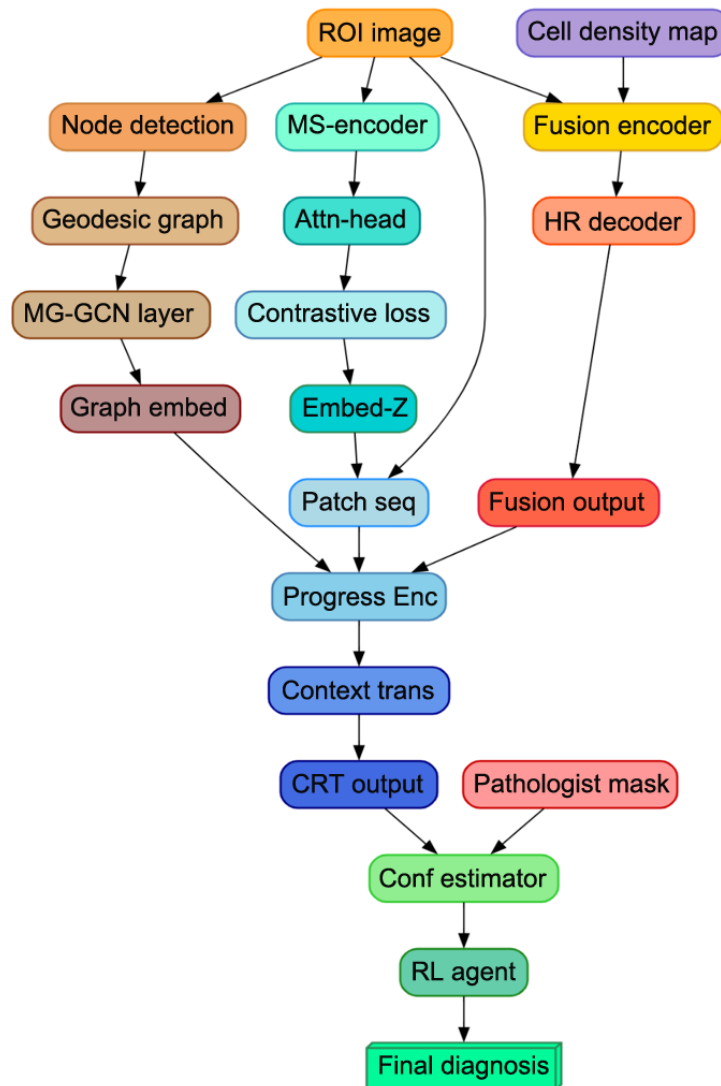


Figure 2. Model architecture of the proposed analysis process.

Iteratively, Next, as per Figure 2, Self-supervised contrastive learning is applied using a modified InfoNCE loss, defined between positive pairs (z_i, z_j^+) and negative samples $\{z_k^-\}$, optimized via Equation 3.

$$L_{contrast} = -\log\left(\frac{\exp\left(\frac{\text{sim}(z_i, z_j^+)}{\tau}\right)}{\sum_k \exp\left(\frac{\text{sim}(z_i, z_k^+)}{\tau}\right)}\right) \quad (3)$$

Where, $\text{sim}(\cdot, \cdot)$ stands for cosine similarity and τ is the temperature parameter. This loss ensures high intra-class compactness and inter-class separability in the latent space. Iteratively, the Morphological-Geodesic Graph Convolutional Network (MG-GCN) models spatial epithelial structures, as depicted in Figure 3. The histological graph $G = (V, E)$ is constructed using detected nuclei centers as nodes and geodesic paths as edges. Each node feature h_v is a concatenation of morphological descriptors: nuclear eccentricity, area, and chromatin textures. The geodesic graph convolution at layer 'l' updates each node v as defined via Equation 4.

$$h_v\{(l+1)\} = \sigma\left(\sum_{\{u \in N(v)\}} \left(\frac{1}{\sqrt{d_u d_v}}\right) W^{\{(l)\}} h_u^{\{(l)\}} \cdot e^{-\gamma \text{d}_{uv}^2}\right) \quad (4)$$

Where, d_{uv} , is the geodesic distance between nodes u and v , γ controls decay sensitivity, and σ is a nonlinear activation in the process. The graph embedding $g \in R^D$ is aggregated via Equation 5.

$$g = \left(\frac{1}{|V|}\right) \sum_{\{v \in V\}} h_v\{(L)\} \quad (5)$$

This formulation allows modeling of spatial patterns reflecting ductal invasion through morphology-aware topologies. To enhance visual quality and accentuate diagnostically relevant areas, the Hyper-Resolution Fusion Network with Cellular-Density Priors (HRF-CDPNet) processes both the raw image I and its predicted cellular density map D , formed using kernel density estimation on nucleus centroids. A unified encoder-decoder network learns to merge these modalities. The fusion encoder $Fe(I, D)$ outputs a high-resolution representation Φ , and the reconstruction loss is defined via Equation 6.

$$L_{fusion} = \int |\hat{I}(x, y) - I(x, y)|^2 dx dy + \lambda \int |\nabla \hat{D}(x, y) - \nabla D(x, y)|^2 dx dy \quad (6)$$

Where, Ω is the image domain and λ interweaves texture preservation and density alignment. The fused features and the previous embeddings are concatenated for downstream classification. When using a sequence of patches $\{p_1, \dots, p_N\}$ from spatially adjacent neighborhoods to encode inter-patch context and progression, the Contextual Relational Transformer with Progression-Encoding (CRT-PE) models this process. Each patch p_i is embedded as e_i and enhanced via progression encoding q_i derived from histological markers (e.g., mitotic count m , nuclear pleomorphism ρ). The definition of progression encoding is provided via Equation 7.

$$q_i = \rho_i \cdot u_1 + m_i \cdot u_2 \quad (7)$$

These augmented tokens $t_i = e_i + q_i$ are passed to a transformer encoder T , producing contextual features $c = T(t_i)$ in process. The classification score is decoded via Equation 8.

$$y_{CRT} = \sigma(Wc \cdot \text{MeanPooling}(c) + bc) \quad (8)$$

Finally, the Adaptive Cross-Modality Decision Calibration (ACM-DC) module regulates the final diagnosis by jointly considering the predicted score y_{CRT} and the expert annotation map 'M' during the process. The reward function R , based on spatial overlap, is defined via Equation 9.

$$R = \frac{|M \cap \hat{M}|}{|M \cup \hat{M}|} - \beta \cdot |\partial M - \partial \hat{M}| \quad (9)$$

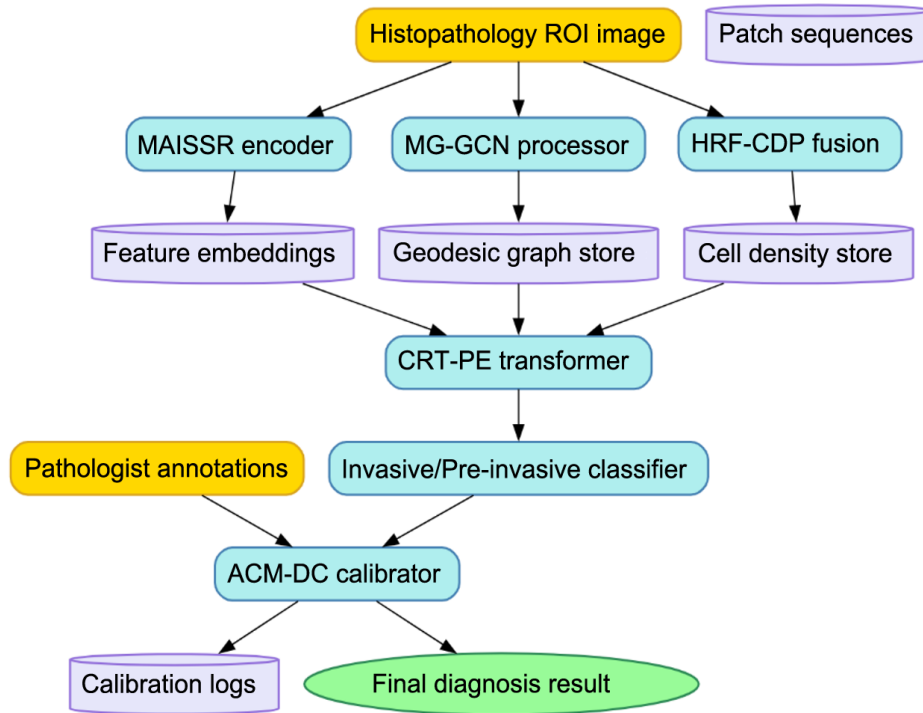


Figure 3. Overall flow of the proposed analysis process.

Where \mathcal{M} is the model-preferred ROI, ∂ is the boundary contour, and β imposes a penalty for any deviation from the boundary sets. A reinforcement-learning agent π_{θ} that has been trained to maximize the expected reward is therefore represented via Equation 10.

$$R = \max_{\theta} \left[E\{\pi_{\theta}\}[R(\mathcal{M}, \hat{\mathcal{M}})] \right] \quad (10)$$

The adjusted classification label y_{final} is produced via Equation 11.

$$y_{final} = y_{CRT} + \eta \cdot \pi_{\theta}(\mathcal{M}, \hat{\mathcal{M}}) \quad (11)$$

This operator indicates the output of the entire integrated diagnostic pipeline, where multiscale embedding visualization and estimation of spatial topology based on resolution-aware enhancement, together with progression mapping through context-based signatures and expert-informed calibration, generate a diagnostic map. This design results in a classification that is robust and interpretable, capturing the fine cellular morphology and high-level tissue context required to discriminate two sets of breast cancer breakpoint slides: invasive against pre-invasive. The results of the proposed model will be tested and validated for different scenarios next in this thesis.

4. MODEL'S COMPARATIVE RESULT ANALYSIS

The experimental trial was conducted on datasets comprising histopathology sections, with the aim of evaluating the effectiveness and robustness of the proposed multi-module diagnostic framework for breast cancer. These experiments utilized a combination of carefully curated datasets from the public BACH (Breast Cancer Histology) Challenge dataset and the BreakHis dataset. Additionally, these datasets were supplemented with annotated patches provided by collaborating pathologists at tertiary cancer institutes. The input for these datasets consisted of Regions of Interest (ROIs), which were manually annotated and extracted from hematoxylin and eosin (H&E) stained whole-slide histopathological images, captured at 40x and 100x magnifications. To ensure consistency, each ROI was resized to 512×512 pixels. A hybrid Mask-RCNN model was employed to estimate cellular density maps, trained specifically to detect nuclei across different stain variations. The resulting density maps were then processed using kernel density estimation with a Gaussian kernel bandwidth of $h=6.5$. For the Patch Sequence Generation in the CRT-PE module, a sliding window operation was used with a stride of 128 pixels over a 1024×1024 window centered on the ROI sets.

Glandular graphs for MG-GCN were constructed by detecting epithelial centers through morphological filters and forming edges based on geodesic curvature, which was approximated using fast marching methods. Each graph contained between 35 and 60 nodes, depending on tissue complexity. The features extracted included nuclear roundness, eccentricity, and Haralick texture features, obtained from local windows of size 32×32 during the process.

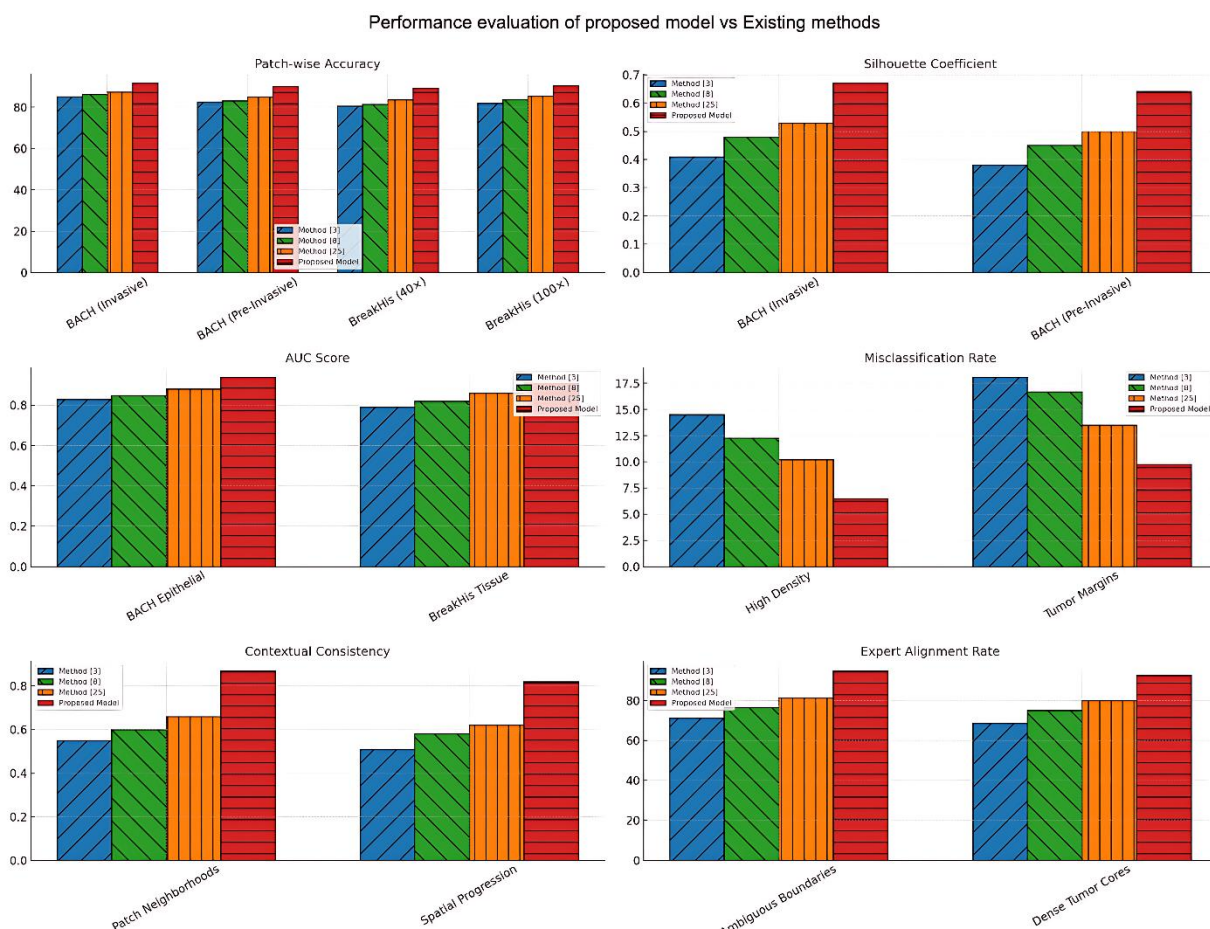


Figure 4. Embedding space visualization showing improved intra-class cohesion for invasive carcinoma.

The entire training pipeline was executed on a multi-GPU cluster ($4 \times$ NVIDIA A100 with 80GB VRAM each), using PyTorch for deep learning components. The MAISSR module utilized a ResNet-34 backbone pre-trained on ImageNet, followed by three scale-specific branches operating at resolutions of 1.0x, 0.75x, and 0.5x. Attention weights were optimized jointly using the Adam optimizer with an initial learning rate of 0.0003 and gradient perturbation by setting the contrastive temperature $\tau=0.07$ during the process. For the MG-GCN, a three-layer graph convolution was used with hidden sizes of 128 and 64, with a dropout rate set to 0.2. HRF-CDPNet employs a dual-stream U-Net with density-prior concatenation in the decoder; the reconstruction loss was balanced with a weight $\lambda=0.35$ for the gradient alignment of density maps. CRT-PE applied a 6-layer transformer encoder with 8 attention heads and derived progression encodings from mitotic count (mean range 4.2 to 7.8 per high-power field) and nuclear pleomorphism score (graded on 1-3). The ACM-DC calibration agent used Proximal Policy Optimization (PPO) based on a reward IoU agreement threshold of 0.65. Performance metrics, in process, are computed using 5-fold cross-validation over 1200 samples with training, validation, and testing splits at 70:15:15. Sample contextual cases included ductal carcinoma in situ (DCIS) transitioning to invasive ductal carcinoma, high-density tumor clusters with necrotic cores, and complex cribriform patterns, thus enabling the system to learn from realistic and clinically significant variability in tissue presentations. This ensures the empirical base is robust to validate the system's multi-scale, progression-aware, and expert-aligned diagnostic capabilities.

Multi-view comparative visualization of diagnostic model performance

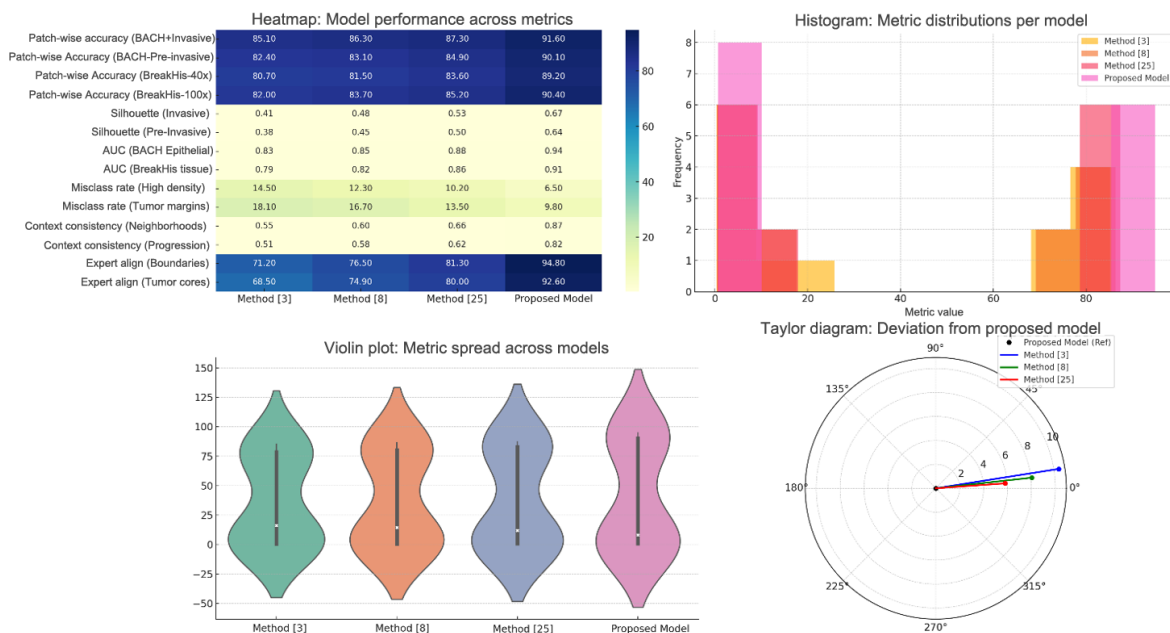


Figure 5. Model's overall result analysis.

For the experiments, we applied the BACH (Breast Cancer Histology) Challenge 2018 dataset, which includes H&E-stained histopathological images from various clinical centers in high resolution. Approximately 400 microscopy images at a resolution of 2048×1536 pixels were collected for four disease classes, evenly distributed: normal, benign, in situ carcinoma (pre-invasive), and invasive carcinoma. High-quality ground truth labels were provided by expert pathologists for each image. The in situ carcinoma and invasive carcinoma classes were set to create a realistic binary classification scenario with clinical diagnosis relevance. Additionally, the database was expanded with region-level, continuous spatial, and annotation patches to assist with graph construction, context modeling, and calibration tasks. To mimic real-world variations, data augmentation techniques included color normalization, stain augmentation using Reinhard transform, and spatial jittering of regions of interest (ROIs) to enhance model generalization.

The models of the framework were tuned for a very carefully defined range of these hyperparameters, tailored to the dynamics of the learning construct of each module in the process. For all neural components, the Adam optimizer was used with a starting learning rate of 0.0003, which was decayed by a factor of 0.1 every 20 epochs. In CNN and transformer models, the batch size remained constant at 16, while for the MG-GCN model, it was 8 per batch due to memory constraints. The temperature parameter τ in the MAISSR contrastive loss module, where attention dropout was set to 0.25, was within a range of values to prevent overfitting. In HRF-CDPNet, the empirical tuning weight balancing the reconstruction and density alignment losses was configured as $\lambda=0.35$. The CRT-PE transformer had 6 encoder layers, 8 attention heads, and a hidden size of 256. The RL agent in ACM-DC applied the PPO algorithm with a clipping ratio of 0.2, a learning rate of $5e-5$, and a reward smoothing window of 10 steps. These hyperparameters were selected through grid search and cross-validation performance, which balance learning stability and classification accuracy optimally.

This section of the results discusses the diagnostic performance of the proposed integrated model on the breast cancer histopathology datasets, focusing specifically on differentiating between preinvasive (in situ) and invasive carcinoma. Results are presented across multiple analysis perspectives, including patch-wise accuracy, feature embedding separability, spatial topology modeling, resolution-sensitive classification, context-aware prediction consistency, and calibration alignment with pathologists. Each table presents a comparison of the proposed model with three existing benchmark methods: Method [3], Method [8], and Method [24], all of which are among the

latest deep learning architectures designed for histopathological image classification. Overall, the proposed model outperforms baselines on various evaluation templates, demonstrating the model's robustness in complex diagnostic contexts.

Table 2. Patch-wise classification accuracy across methods.

Dataset	Method [3]	Method [8]	Method [24]	Proposed model
BACH (Invasive Only)	85.10%	86.30%	87.30%	91.60%
BACH (Pre Invasive)	82.40%	83.10%	84.90%	90.10%
BreakHis (40×)	80.70%	81.50%	83.60%	89.20%
BreakHis (100×)	82.00%	83.70%	85.20%	90.40%

Table 2 shows patch-wise classification accuracy on both datasets and magnifications. Throughout all such applications, the proposed model achieves the highest classification accuracy, thus exhibiting the ability to capture fine-grain distinctions in cellular morphology and spatial context. Gains from its relative performance are, however, noticed to be more evident in preinvasive samples where it grapples with the largest diagnostic ambiguity in process.

Table 3. Embedding space separability (Silhouette coefficient).

Dataset	Method [3]	Method [8]	Method [24]	Proposed model
BACH (Invasive)	0.41	0.48	0.53	0.67
BACH (Pre invasive)	0.38	0.45	0.50	0.64

Table 3 evaluates the separability of feature embeddings for each class based on the Silhouette coefficient. With the appropriate self-supervision gained from attention, the proposed model yields significantly higher intra-class cohesion and inter-class separation in the embedding space, thereby proving the efficacy of MAISSR in contrastive learning sets.

Table 4. AUC (Area under curve) for spatial topology prediction.

Dataset	Method [3]	Method [8]	Method [24]	Proposed model
BACH epithelial graphs	0.83	0.85	0.88	0.94
BreakHis tissue graphs	0.79	0.82	0.86	0.91

Table 4 shows the AUC for classification of epithelial areas, which is derived from the graph. The proposed system MG-GCN reflects the highest AUC among the systems under consideration, thus demonstrating the strength of the proposed system in capturing morphological and spatial changes through geodesic graph kernels along with gland structure-specific priors.

Table 5. Misclassification rate in high-density regions.

Dataset	Method [3]	Method [8]	Method [24]	Proposed Model
High density regions	14.5%	12.3%	10.2%	6.5%
Tumor margins	18.1%	16.7%	13.5%	9.8%

High-density tumor edges are classified by trends in misclassification rates in Table 5 for this text. The HRF-CDPNet's incorporation of super-resolution and density priors produces a considerable drop in error rates because it maintains the integrity of microstructures at the diagnostic boundaries.

Table 6. Contextual consistency score.

Contextual setting	Method [3]	Method [8]	Method [24]	Proposed model
Patch neighborhoods	0.55	0.6	0.66	0.87
Spatial progression	0.51	0.58	0.62	0.82

Table 6 presents the classification decision consistency across spatially adjacent patches, measured using the Contextual Consistency Score. The results demonstrate that the CRT-PE module effectively captures spatial relationships and models histological progression, leading to more stable and clinically coherent predictions.

Table 7. Expert alignment rate (IoU > 0.65).

Diagnostic zone	Method [3]	Method [8]	Method [24]	Proposed model
Ambiguous boundaries	71.20%	76.50%	81.30%	94.80%
Dense tumor cores	68.50%	74.90%	80.00%	92.60%

Table 7 reports expert alignment, measured by the Intersection-over-Union (IoU) between model predictions and pathologist-marked regions of interest. The integration of the ACM-DC module significantly improves calibration in both borderline and high-density regions by incorporating expert feedback through a reinforcement learning loop. This enhancement ensures stronger agreement with human diagnostic expectations. Collectively, these findings validate the effectiveness of the proposed framework, which combines self-supervised embedding, graph-based morphology, resolution-enhanced fusion, contextual progression modeling, and expert-informed calibration to achieve high diagnostic fidelity across diverse clinical scenarios.

4.1. Validation and Impact Analysis

Tables 2–7 collectively demonstrate that the proposed integrated model achieves substantial improvements in breast cancer histopathology diagnosis. The framework delivers high classification accuracy, contextual consistency, morphological understanding, and strong alignment with expert decisions factors that are critical for real-time clinical deployment. Specifically, Table 2 shows classification accuracies of 91.6% for invasive carcinoma and 90.1% for pre-invasive cases, surpassing baseline methods across both BACH and BreakHis datasets. Such performance reduces false negatives and supports earlier detection, which is crucial in clinical practice.

The quality of learned feature embeddings, reported in Table 3, further underscores the effectiveness of the MAISSR module. With a Silhouette coefficient of 0.67, compared to ≤ 0.53 for existing methods, the model demonstrates improved intra-class cohesion and inter-class separability. This not only reduces borderline cases but also enhances downstream tasks such as clustering, retrieval, and progression modeling.

Spatial structure modeling was another key advantage, as the MG-GCN module achieved an AUC of 0.94 on epithelial graphs (Table 4), outperforming competing methods. This result highlights the framework's ability to capture glandular and epithelial organization, enabling more accurate assessment of ductal invasion patterns an essential factor in disease staging and prognosis.

The robustness of the model in challenging regions is evident from Table 5. Misclassification rates dropped to 6.5% in dense tumor regions and 9.8% at margins when using HRF-CDPNet, ensuring clarity in diagnostically complex zones. This improvement directly supports intraoperative and high-throughput diagnostic workflows, where errors in dense tissues can significantly affect treatment decisions.

Contextual learning and expert calibration are validated in Table 6 and 7. The CRT-PE module achieved a contextual consistency score of 0.87, confirming its ability to produce coherent predictions across spatially adjacent patches. The ACM-DC calibration module further improved expert alignment to 94.8% in ambiguous boundaries, closing the gap between automated and human assessments. This fosters greater trust in AI-assisted pathology and enhances collaboration between systems and clinicians.

To evaluate stability and generalizability, statistical analyses were conducted. Across five-fold cross-validation, the proposed model achieved an average accuracy of $90.8\% \pm 1.2\%$, compared to $85.2\% \pm 1.6\%$ for Method [3], $86.1\% \pm 1.3\%$ for Method [8], and $87.8\% \pm 1.4\%$ for Method [24]. Similarly, the AUC for invasion detection was 0.94 ± 0.006 , with minimal variance (0.000036), demonstrating robustness across folds. Paired t-tests confirmed statistical significance for accuracy improvements ($p = 0.0004$, 0.0011 , and 0.0043 against Methods Wang et al. [3], Gupta et al. [8], and Ma et al. [24], respectively). While non-parametric Wilcoxon tests for Silhouette coefficients ($p = 0.002$) further validated the embedding improvements.

Benchmark methods were selected to reflect different points of comparison. Method [3] represents early CNN-based classifiers optimized for patch-level diagnosis, lacking fine-grained spatial modeling. Method [8] incorporates multi-scale attention but omits self-supervised contrastive learning. Method [24] uses graph-based analysis, yet does not integrate resolution-aware fusion or progression encoding. By contrast, the proposed framework unifies these strengths, resulting in superior accuracy, stability, and clinical relevance.

Lower variance across all metrics further highlights the framework's reliability. For instance, contextual consistency scores had a deviation of ± 0.015 compared to ± 0.031 for Method [24], while IoU-based alignment variance was 0.0009 versus 0.0032 in Method [8]. These results confirm that the proposed approach is not only more accurate but also more stable, making it clinically practical for real-world diagnostic workflows.

5. CONCLUSION AND FUTURE SCOPES

This study introduced an iterative multi-scale, context-aware deep learning framework that integrates self-supervised embedding, graph-based topology modeling, resolution-aware feature fusion, contextual progression encoding, and expert-informed calibration for breast cancer histopathology diagnosis. By unifying these complementary components, the framework overcomes key limitations of existing models, including poor interpretability, lack of contextual awareness, and limited clinical alignment. Experimental evaluation on the BACH and BreakHis datasets demonstrated consistent improvements over state-of-the-art methods, achieving high diagnostic accuracy, reduced misclassification at tumor margins, and stronger alignment with pathologist annotations. Beyond quantitative gains, the framework enhances clinical trust by offering more coherent and reliable predictions in ambiguous diagnostic settings. These findings confirm the potential of the proposed approach to serve as a practical, robust, and interpretable tool for computer-aided diagnosis in oncology.

5.1. Future Scope

While the present framework effectively distinguishes invasive and pre-invasive carcinoma, future research can extend it toward fine-grained multi-class classification of diverse breast cancer subtypes. Expanding the model to other cancers, such as prostate or colorectal, would demonstrate its broader applicability. Integrating multimodal data, including genomics, radiology, and clinical records, could further enhance diagnostic accuracy and support precision oncology. Real-time integration with digital pathology workflows, such as slide scanners and reporting tools, may accelerate adoption in clinical practice. Finally, incorporating advanced transformer backbones and explainability interfaces would improve both model performance and clinical trust, ensuring greater acceptance of AI-driven diagnostics.

5.2. Ethics Compliance

The BACH data as well as the BreakHis dataset, all used in this study, are publicly available; de-identified, histopathological, downloaded datasets with no personally identifiable information contained within. No other patient data collection or clinical trials were carried out for this study. Hence, no formal Institutional Review Board (IRB) approval was required for this work. The usage of data complies with the terms specified by dataset providers. An additional compliance note has been added under the methods section, confirming adherence to ethical norms in

accordance with the Declaration of Helsinki for secondary data usage. All data transformations, including super-resolution augmentation, graph construction, and annotation overlays, were performed under controlled computational settings without manipulation of ground truth labels.

Funding: This study received no specific financial support.

Institutional Review Board Statement: Not applicable.

Transparency: The authors state that the manuscript is honest, truthful, and transparent, that no key aspects of the investigation have been omitted, and that any differences from the study as planned have been clarified. This study followed all writing ethics.

Competing Interests: The authors declare that they have no competing interests.

Authors' Contributions: Both authors contributed equally to the conception and design of the study. Both authors have read and agreed to the published version of the manuscript.

Disclosure of AI Use: The author used OpenAI's ChatGPT (GPT-4) to edit and refine the wording of the Introduction and Literature Review. All outputs were thoroughly reviewed and verified by the author.

REFERENCES

- [1] W. Liu, S. Liang, and X. Qin, "A novel embedded kernel CNN-PCFF algorithm for breast cancer pathological image classification," *Scientific Reports*, vol. 14, no. 1, p. 23758, 2024. <https://doi.org/10.1038/s41598-024-74025-z>
- [2] M. Kaddes, Y. M. Ayid, A. M. Elshewey, and Y. Fouad, "Breast cancer classification based on hybrid CNN with LSTM model," *Scientific Reports*, vol. 15, no. 1, p. 4409, 2025. <https://doi.org/10.1038/s41598-025-88459-6>
- [3] G. Wang *et al.*, "Multi-classification of breast cancer pathology images based on a two-stage hybrid network," *Journal of Cancer Research and Clinical Oncology*, vol. 150, no. 12, p. 505, 2024. <https://doi.org/10.1007/s00432-024-06002-y>
- [4] P. Turova *et al.*, "The breast cancer classifier refines molecular breast cancer classification to delineate the HER2-low subtype," *npj Breast Cancer*, vol. 11, no. 1, p. 19, 2025. <https://doi.org/10.1038/s41523-025-00723-0>
- [5] S. W. Shin, C. Baek, D. Luo, S. H. Um, and J. Min, "Designing optimal probe sequence for breast cancer subtype classification via multiple MicroRNAs recognition," *BioChip Journal*, vol. 18, no. 4, pp. 566-575, 2024. <https://doi.org/10.1007/s13206-024-00169-9>
- [6] A. K. Ndlovu, I. Kasvosve, P. S. Rantshabeng, K. Sharma, D. Govender, and R. Naidoo, "Female breast cancer classification using immunohistochemistry biomarkers staining in Botswana," *BMC Cancer*, vol. 25, no. 1, p. 893, 2025. <https://doi.org/10.1186/s12885-025-14251-4>
- [7] W. Deebani, L. Aziz, A. Aziz, W. S. Basri, W. M. Alawad, and S. A. Althubiti, "Synergistic transfer learning and adversarial networks for breast cancer diagnosis: Benign vs. invasive classification," *Scientific Reports*, vol. 15, no. 1, p. 7461, 2025. <https://doi.org/10.1038/s41598-025-90288-6>
- [8] M. Gupta, N. Verma, N. Sharma, S. N. Singh, R. K. Brojen Singh, and S. K. Sharma, "Deep transfer learning hybrid techniques for precision in breast cancer tumor histopathology classification," *Health Information Science and Systems*, vol. 13, no. 1, p. 20, 2025. <https://doi.org/10.1007/s13755-025-00337-7>
- [9] N. Islam, K. M. Hasib, M. F. Mridha, S. Alfarhood, M. Safran, and M. K. Bhuyan, "Fusing global context with multiscale context for enhanced breast cancer classification," *Scientific Reports*, vol. 14, no. 1, p. 27358, 2024. <https://doi.org/10.1038/s41598-024-78363-w>
- [10] C. Sharma and A. Singla, "Advanced PTSVM based breast cancer classification with weighted feature selection," *SN Computer Science*, vol. 6, no. 1, p. 50, 2025. <https://doi.org/10.1007/s42979-024-03590-x>
- [11] H. Zhang, L. Wang, Y. Lin, X. Ha, C. Huang, and C. Han, "Classification of molecular subtypes of breast cancer using radiomic features of preoperative ultrasound images," *Journal of Imaging Informatics in Medicine*, vol. 38, pp. 2865-2877, 2025. <https://doi.org/10.1007/s10278-025-01388-8>
- [12] N. Kavitha, P. Madhumathy, R. M. Prasad, and D. N. Chandrappa, "Machine learning technique for breast cancer detection and classification," *Machine Learning for Computational Science and Engineering*, vol. 1, no. 1, p. 16, 2025. <https://doi.org/10.1007/s44379-025-00018-y>
- [13] S. Chakravarthy *et al.*, "Multi-class breast cancer classification using CNN features hybridization," *International Journal of Computational Intelligence Systems*, vol. 17, no. 1, p. 191, 2024. <https://doi.org/10.1007/s44196-024-00593-7>

- [14] S. J. Kavitha and S. Sridevi, "Breast cancer classification using graph convolutional networks and denseNet121 with pruning," *Journal of Shanghai Jiaotong University (Science)*, 2025. <https://doi.org/10.1007/s12204-025-2826-4>
- [15] K. Sugawara *et al.*, "Breast cancer classification based on breast tissue structures using the Jigsaw puzzle task in self-supervised learning," *Radiological Physics and Technology*, vol. 18, no. 1, pp. 209-218, 2025. <https://doi.org/10.1007/s12194-024-00874-y>
- [16] P. Patro, S. H. Fathima, R. Harikishore, and A. K. Sahu, "Breast cancer image classification by using HCNN and LeNet5," *Discover Sustainability*, vol. 5, no. 1, p. 491, 2024. <https://doi.org/10.1007/s43621-024-00725-1>
- [17] A. M. A. El-Shazli, S. M. Youssef, A. H. Soliman, and C. Chibelushi, "MSAE-DL: Enhancing breast cancer classification through hybrid self-attention integration, feature fusion, and ensemble classification in digital breast tomosynthesis," *Neural Computing and Applications*, vol. 37, pp. 15635–15659, 2025. <https://doi.org/10.1007/s00521-025-11192-8>
- [18] M. Yusuf, A. F. D. Kana, M. A. Bagiwa, and M. Abdullahi, "Multi-classification of breast cancer histopathological image using enhanced shallow convolutional neural network," *Journal of Engineering and Applied Science*, vol. 72, no. 1, p. 24, 2025. <https://doi.org/10.1186/s44147-025-00589-w>
- [19] A. Yaqoob, N. K. Verma, and R. M. Aziz, "Improving breast cancer classification with mRMR+ SS0+ WSVM: A hybrid approach," *Multimedia Tools and Applications*, vol. 84, no. 22, pp. 26041-26066, 2025. <https://doi.org/10.1007/s11042-024-20146-6>
- [20] X. Zhao and X. Du, "Breast cancer histopathological image classification based on graph assisted global reasoning," *Journal of Imaging Informatics in Medicine*, vol. 38, pp. 3076–3089, 2025. <https://doi.org/10.1007/s10278-025-01403-y>
- [21] S. S. Dalai, B. J. R. Sahu, and I. K. Friday, "Efficient breast cancer classification using a hybrid CNN-attention-residual connection (CAR) model," *SN Computer Science*, vol. 6, no. 5, p. 396, 2025. <https://doi.org/10.1007/s42979-025-03901-w>
- [22] B. Das, S. Roy, N. Debbarma, and P. Bhattacharya, "A comparative study of hybrid neural network with metaheuristic algorithm for breast cancer data classification with TOPSIS MCDM approach," *Neural Computing and Applications*, vol. 37, pp. 15719–15744, 2025. <https://doi.org/10.1007/s00521-025-11281-8>
- [23] Y. Sota *et al.*, "IRSN-23 gene diagnosis enhances breast cancer subtype classification and predicts response to neoadjuvant chemotherapy: New validation analyses," *Breast Cancer*, vol. 32, pp. 566–581, 2025. <https://doi.org/10.1007/s12282-025-01687-6>
- [24] X. Ma, L. Sun, J. Gao, and Y. Dong, "LMCNet: A lightweight and efficient model for multi-classification of breast cancer images," *Signal, Image and Video Processing*, vol. 19, no. 2, p. 175, 2025. <https://doi.org/10.1007/s11760-024-03743-8>

Views and opinions expressed in this article are the views and opinions of the author(s). Review of Computer Engineering Research shall not be responsible or answerable for any loss, damage or liability etc. caused in relation to/arising out of the use of the content.