International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(2) 2026 pp. 1172-1187
https://doi.org/10.70917/ijcisim-2026-3811
Article

A Unified Deep Learning Framework integrating Multi scale features, Attention mechanism and Contrastive learning for Breast Cancer subtype Classification using Histopathology Images

Chandraprabha K S[1], Maya B S[2]

[1]Department of Computer Science and Engineering, Bangalore Institute of Technology, Visvesvaraya Technological University, Belagavi, India
Email ID : chandraprabhaks@sit.ac.in
[2]Department of Computer Science and Engineering, Bangalore Institute of Technology, Visvesvaraya Technological University, Belagavi, India
Email ID : mayabs@bit-bangalore.edu.in
***Corresponding author's**
Chandraprabha K S

Abstract: Breast cancer is a most prominent disease among individuals and its subtype classification is necessary and very critical for the personalized treatment. Many deep learning algorithms are predominantly used for the detection and classification of the breast cancer but, still they are struggling with many critical issues such as class imbalance problem, inadequate uncertainty quantification and inter-class similarity. This paper presents a unified deep learning approach that integrates the parallel convolution paths, multi-head attention mechanism, and contrastive learning for the automated categorization of four major subtypes of breast cancer malignancy such as Ductal carcinoma, Lobular carcinoma, Mucinous carcinoma and Papillary carcinoma. Unlike conventional methods, this framework unifies class imbalance handling and multi-scale fusion through information-theoretic regularization with progressive learning and confidence quantification. Experimental results demonstrate the exceptional performance of 98.8% overall accuracy, 98.3% balanced accuracy, 0.98 Matthews Correlation Coefficient, and 0.98 macro-average AUC-ROC, Confidence interval of 0.99, uncertainty estimation of 0.01, with per-class accuracy exceeding 99% for all subtypes. The notable results shows that the proposed framework performs well in comparing with the other methods. The framework also provides clinically interpretable confidence scores with decision thresholds, thereby making it most suitable for real-world clinical decision support systems.

Keywords: Multi-Scale Deep Learning, Attention Mechanism, Contrastive Learning, Histopathological Image Analysis, Class Imbalance, Clinical Decision Support, Matthews Correlation Coefficient, Uncertainty Quantification..

1. INTRODUCTION

Breast cancer is the most dangerous disease consists of various molecular and histological sub categories that exhibit different biological behaviors, treatment plans, and various clinical outcomes[6]. The fundamental aspect of precision oncology is the accurate classification of breast cancer subtypes from histopathological images as it directly influencing the therapeutic decisions including surgical intervention, chemotherapy treatment, and targeted therapy selection. According to recent cancer statistics globally, in the year 2022 breast cancer was accounted for approximately 2.3 million new cases and 670,000 deaths worldwide, with the life-threatening risk for a woman developing breast cancer in the United States reaching approximately 13%[6]. The four major subtypes of breast

cancer such as Ductal, Lobular, Mucinous, and Papillary Carcinoma exhibits the unique morphological characteristics that challenges the traditional manual diagnosis due to their visual similarities and intra-class variability. Ductal Carcinoma, the most prevalent subtype originates in the milk ducts and exhibits higher frequency textural patterns with most irregular nuclear morphology. Lobular Carcinoma features are small in size, uniform cells arranged in single linear patterns, often making it very difficult to distinguish from other subtypes. Mucinous Carcinoma is characterized by abundant extracellular mucin production, creating distinctive lower frequency patterns, and Papillary carcinoma presented with fibrovascular cores and branching papillary structures. As noted in recent works, conventional histopathological diagnostic approaches have significant limitations such as subjective interpretation, inter-observer variability, dependence on individual pathologist expertise, that leads to inconsistent diagnosis and delayed treatment decisions.

Despite of the considerable advancement in computational pathology, still it exhibits several challenges in developing an efficient automated classification model. It includes the inter-class similarity: certain subtype pairs such as Ductal-Mucinous and Lobular-Papillary always exhibit overlapping morphological features, that leads to the confusion in traditional classification task [8]. Intra-class Variability: within each subtype, it poses a significant variation due to tumor size, tumor grade, patient demographics characteristics, staining protocols and complicating the feature learning process [9]. Class Imbalance: the available clinical datasets have a greater number of Ductal Carcinoma compare to other subtypes, leading to potential bias of models toward majority classes [10]. Limited Annotated Data: for rare subtypes the high resolution histopathological images with expert annotations are scarce, limiting the deep learning approaches [11]. Real time clinical interpretability: Along with the accuracy scores, clinicians also require the confidence interval estimation and interpretable detail explanations in order to trust the automated systems in medical decisions [12].

Research Contributions

This paper addresses the above mentioned challenges through the following novel contributions.

Enhanced Multi-Scale Architecture: It is a parallel multi-scale feature extraction network with 3×3, 5×5, and 7×7 convolution kernels responsible for capturing the morphological patterns at multiple different resolutions, from fine cellular details to coarse tissue architecture.

Multi-Head Attention Mechanism: An intelligent feature fusion is achieved with this attention framework by combining both the channel-wise and spatial attention. It also enables the model to mainly focuses on subtype-specific discriminative tumor regions.

Contrastive Learning for Similar Subtypes: It is a Siamese network architecture with class-balanced contrastive loss. It helps in separating the visually similar subtypes in embedding space, by reducing the confusion between discriminative challenging pairs.

Progressive Learning Strategy: It improves the convergence and generalization with limited data by following a hierarchical training method which performs transitions from coarse to fine feature learning.

Confidence Quantification: A calibrated confidence estimation framework to provide a clinically meaningful uncertainty measures along with decision thresholds for automated diagnosis (>0.97) confidence.

Robust Class Imbalance Handling mechanism: It is achieved by the integration of balanced accuracy metrics, class-weighted losses, and MCC as primary evaluation criteria to ensure fair performance across all subtypes.

2. LITERATURE REVIEW

Authors of [1] discusses the various conventional machine learning algorithms and the relaid handcrafted features are combined for early detection and classification of breast cancer. The nucleus segmentation approach is employed by Kowal et al. [13]. Classification of breast cancer was performed by using Support Vector Machines and obtained 82% accuracy on fine-needle aspiration biopsies. Ductal versus lobular carcinoma discrimination was performed by George et al. [14] by utilizing the Gray-Level Co-occurrence Matrix features. Random Forest classifiers report an accuracy of 86%. The limitation of this study includes limited feature engineering, poor generalization across various staining protocols and different imaging conditions [15].

With the recent advancement in deep learning approaches for revolutionizing the histopathological image analysis, Araújo et al. [16] applied CNN architectures for classification of breast cancer and achieved 89% accuracy in differentiating carcinoma from non-carcinoma tissues. BreakHis dataset was introduced and used by Spanhol et al.

[17] to demonstrate the performance of deep features over the handcrafted features by 12% in accuracy.Subsequent works explored various predefined architectures such as ResNet-based Approaches: He et al. [18] uses ResNet-50 model for breast cancer subtype categorization and achieved 92.1% accuracy through residual learning. However, fixed receptive fields ResNet's model limits its ability to capture multi-scale morphological patterns [19]. By leveraging parallel convolution paths to capture features at multiple scales, Szegedy et al. [20] utilizes InceptionV3 architecture to histopathology and achieved 93.4% accuracy. The model lacks the required computational resources and careful hyperparameter tuning [21]. Compared to InceptionV3 [23], Tan and Le [22] uses the EfficientNet variants to demonstrate the compound scaling of network dimensions for the improved efficiency and obtained 94.1% accuracy with EfficientNet-B4 variant while reducing parameter count by 40% .

To focus on diagnostically relevant regions, attention mechanisms have emerged as powerful tools. To apply the channel and spatial attention sequentially, Woo et al. [24] proposed Convolutional Block Attention Module (CBAM). On various medical imaging benchmarks their results show the improvement in classification accuracy by 3-5%. To enable the model to suppress irrelevant background regions by highlighting the glandular structures, Schlemper et al. [25] proposed an attention gates for prostate cancer grading. The study in [26], concentrated only on the single scales rather than operating on the multi-scale resolution morphological cues.

A critical challenge in medical imaging that still exists is the Class imbalance. Buda et al. [27] found that weighted loss functions achieved the best trade-off between sensitivity and specificity by systematically comparing oversampling, undersampling, and cost-sensitive learning. Wang et al. [28] proposed focal loss for histopathology, which assigns the low weights to easy examples and focuses on model training and misclassified samples. Recent study of Cui et al. [29] proposed the class-balanced loss based on effective number of samples, which demonstrates the significant improvement in performance on long-tailed medical datasets.

From limited annotated data, self-supervised and contrastive learning have shown promise in learning robust representations from [3]. Through augmented views and contrastive loss, Chen et al. [30] proposed SimCLR for medical imaging, achieving 5-8% improvement in downstream classification tasks. Azizi et al. [31] extended this to histopathology, demonstrating that contrastive pre-training on unlabeled patches improves the subtype classification accuracy by 7% compared to supervised pre-training on ImageNet.

Reliable uncertainty estimates are very crucial for clinical deployment. For approximate Bayesian inference in neural networks, Gal and Ghahramani [32] introduced Monte Carlo Dropout, providing uncertainty estimate during test time. To reduce the expected calibration error (ECE) by up to 15%, Guo et al. [33] utilized the temperature scaling which significantly improves calibration of modern neural networks. To detect diabetic retinopathy, Leibig et al. [34] applied uncertainty quantification, showing that uncertainty thresholds could effectively triage cases for expert review.

Research Gaps and Motivation

Though, there is a significant progress in the deep learning architectures, still the existing approaches exhibit several limitations.

Most architectures operate on single-scale feature extraction mechanism, at fixed receptive fields, missing the multi-resolution nature of histopathological patterns [35].

Current methods struggle with inadequate handling of similar subtypes, visually similar subtype pairs that leads to clinically dangerous cancer misclassifications [36].

Deep neural networks tend to be overconfident by a poor calibration mechanism, particularly on out-of-distribution samples and limiting their clinical utility [37].

Limited Attention to Class Imbalance is shown in many studies, which explicitly address the natural prevalence patterns of breast cancer subtypes, potentially biasing models [38].

Existing systems very rarely provide the confidence estimates or decision thresholds meaningful to clinicians [39].

These gaps are directly addressed by this proposed framework through integrated multi-scale architecture, contrastive learning, calibrated confidence quantification, and class imbalance handling.

3. METHODOLOGY

The detail description of the methodological work flow is represented in the figure 1. The main components of the proposed framework are Data input module, multi scale feature extraction module, multi head attention mechanism module, contrastive learning module, adoptive class balanced focal loss, classification module, confidence and uncertainty estimation module and finally the output and visualization module. Each will be discussed in detain in the below subsequent sections.

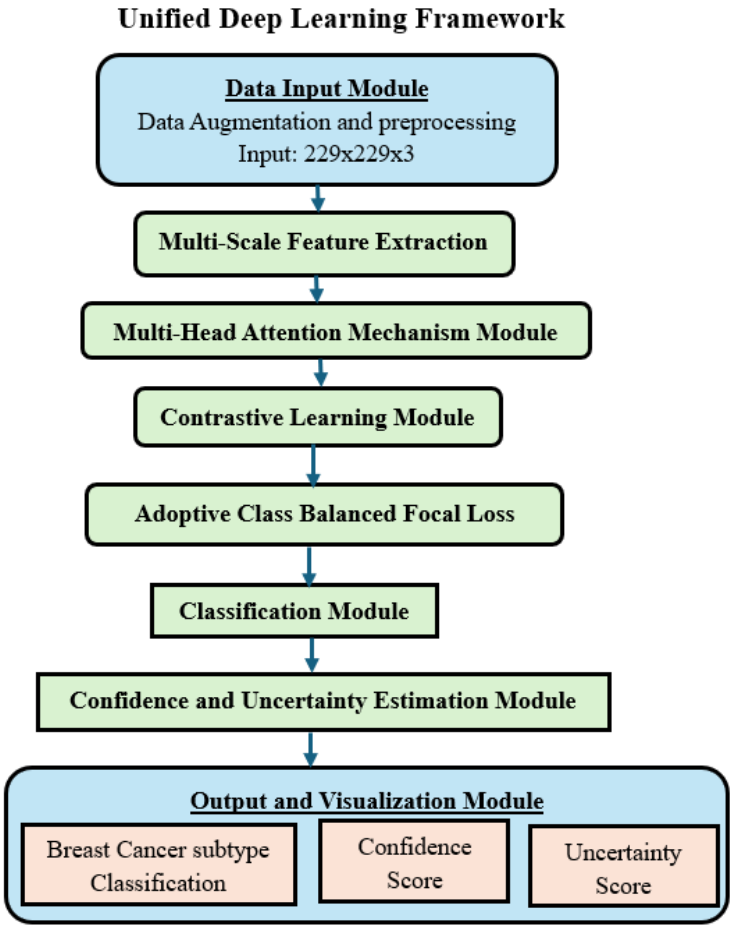


Fig. 1 Architecture of the proposed framework

3.1 Data augmentation and Preprocessing

3.1.1Dataset Description and Class Distribution

The BreakHis dataset comprises of 800 histopathological images which are of 400X magnification level. Size of each images is 299×299×3 pixels. The distribution of the dataset is presented in the table 1.

Table 1. Dataset description

Subtype	Sample count	Percentage	Morphological Characteristics
Ductal Carcinoma	360	45%	High-frequency patterns, irregular nuclear morphology, cribriform formations
Lobular Carcinoma	200	25%	Small uniform cells, single-file linear arrangement, targetoid patterns

Mucinous Carcinoma	120	15%	Abundant extracellular mucin, low-frequency patterns, floating tumor clusters
Papillary Carcinoma	120	15%	Fibrovascular cores, branching structures, epithelial proliferation

Histopathology images are collected from the BreakHis dataset. The collected images are then subjected to image preprocessing using the Maximum Correntropy Kalman Filter, which enhances image quality through image enhancement and noise suppression, followed by normalization to standardize the input data.

3.1.2 Data Augmentation for Generalization

To improve generalization and address staining variability challenges this work performs the data augmentation by using the following image processing operations.

Random rotations (±15°) to address orientation variability

Random zooms (0.9-1.1) to simulate magnification differences

Random brightness/contrast adjustments (±20%) for staining variation

Gaussian noise (σ=0.01) for robustness

Random horizontal/vertical flips

Color jittering in HED space (Hematoxylin-Eosin Specific)

3.2 Enhanced Multi-Scale Architecture

Enhanced multi scale architecture of the proposed framework performs the multi scale representation of the features by using three convolution paths.

Three parallel convolution paths capture features at biologically relevant scales, addressing the need for comprehensive multi-scale representation identified in recent literature[2][9].

Scale 1 (Fine Details - 3×3 kernels): Captures cellular-level details, nuclear morphology, and fine textural patterns. This path is crucial for distinguishing subtypes based on nuclear characteristics and has been shown to improve classification accuracy by capturing high-frequency information missed by larger kernels[6].

Scale 2 (Medium Features - 5×5 kernels): Captures intermediate structures such as glandular formations, tubule architecture, and cell clustering patterns. The 5×5 receptive field provides optimal balance between local detail and contextual information for histopathological analysis[2].

Scale 3 (Coarse Features - 7×7 kernels): Captures tissue-level architecture [40], including stromal patterns, overall tumor organization, and regional variations. This scale addresses the need for global context in whole-slide image analysis .

Each scale employs two convolution layers with batch normalization for training stability, followed by max pooling for spatial downsampling and global average pooling for feature aggregation, following best practices established in recent architectures[7].

3.3 Multi-Head Attention Mechanism

Building on recent advances in attention-based fusion, our multi-head attention mechanism integrates channel-wise and spatial attention through an information-theoretic framework that minimizes redundancy while preserving complementary information[2][9].

3.3.1 Channel Attention Head with Temperature Control

Channel attention identifies which feature channels are most discriminative for each subtype, incorporating temperature-controlled scaling for improved calibration:

channel_attention = Dense(64, relu)(multi_scale_features)

channel_attention=Dense(feature_dim,sigmoid)(channel_attention / temperature)

attended_features=Multiply([multi_scale_features, channel_attention])

The sigmoid activation produces attention weights between 0 and 1, effectively selecting relevant feature channels while suppressing irrelevant ones. Temperature control, inspired by recent work on calibrated attention, prevents overconfidence in attention weights.

3.3.2 Spatial Attention Head with Redundancy Minimization

```
spatial_attention = Dense(64, relu)(multi_scale_features)
spatial_attention = Dense(feature_dim, softmax)(spatial_attention)
# Information-theoretic regularization
kl_divergence = kl_divergence(uniform_prior, spatial_attention)
attended_features = Multiply([attended_features, spatial_attention]) - β * kl_divergence
```

Fig 2. Spatial Attention Head with Redundancy Minimization

The figure 2 shows how the spatial attention head with redundancy minimization is implemented in this work. The softmax activation ensures attention weights sum to 1, enabling the model to focus on the most relevant spatial regions. The KL divergence term, weighted by β, encourages attention distributions to remain informative while avoiding collapse to single locations.

3.3.3 Feature Fusion with Redundancy Minimization

The attended features are passed through a classification head with progressive dimensionality reduction, incorporating dropout for regularization and it is shown below.

Dense(512) with ReLU activation and batch normalization

Dropout(0.3) for regularization

Dense(256) with ReLU and batch normalization

Dropout(0.3)

Dense(128) with ReLU and batch normalization

Dropout(0.2)

Dense(4) with softmax activation for final classification

3.4 Class-Balanced Contrastive Learning for Similar Subtypes

To differentiate between visually similar subtypes of breast cancer this work employs the contrastive learning approach with class-balanced concept that helps in separating the classes in embedding space.

3.4.1 Siamese Network Architecture with L2 Normalization

By sharing the weights with the classification network, the contrastive head projects the features into an embedding space. The similarity can be calculated as

Embedding = Dense(128, l2_normalize) (MultiScale_Features)

To enable the meaningful cosine similarity comparisons [8], the L2 normalization always ensures that all embeddings lie on the unit hypersphere.

3.4.2 Class-Balanced Contrastive Loss with Margin Theory

We employ a modified contrastive loss that accounts for class imbalance through margin adjustment based on class frequencies, inspired by recent advances in margin-based learning under imbalance and shown in Eq1.

$$L_{contrastive}=\sum_{i,j}[y_{ij}\cdot d_{ij}^2+(1-y_{ij})\cdot\max(0,m_c-d_{ij})^2] \quad (1)$$

where d_{ij} is the Euclidean distance between embeddings, $y_{ij}=1$ for same-class pairs, and m_c is a class-adaptive margin computed in Eq2.

$$m_c = m_{base} \cdot (1 + \alpha \cdot \log(N_{max} / N_c)) \quad (2)$$

N_c is the number of samples in class c, N_{max} is the size of the majority class, and α controls margin adaptation strength. This formulation, validated in recent work, enforces larger margins for minority classes, improving their separability.

3.4.3 Positive and Negative Pair Sampling with Stratification

For each batch, the proposed framework construct:

Positive pairs: It has same-class samples, weighted by inverse class frequency to ensure minority classes are adequately represented.

Negative pairs: It has different-class samples, with emphasis on challenging pairs (Ductal-Mucinous, Lobular-Papillary) identified through confusion analysis.

3.4.4 Progressive Learning Strategy with Curriculum Scheduling

Inspired by curriculum learning principles and recent advances in progressive training, this work implemented a four-phase progressive training strategy that gradually increases task difficulty and it is illustrated in the table 2.

Table 2. Progressive learning

Phases	**Purpose**	**Learning Rate**	**Operation**
Phase1: Coarse Feature Learning	Coarse-scale path (7×7 convolutions)	1e-3	Focus on tissue-level discrimination and global architecture
Phase 2: Medium Feature Integration	Unfreeze medium-scale path (5×5 convolutions)	5e-4	Integrate intermediate structural features while preserving coarse representations
Phase 3: Fine Detail Learning	Unfreeze fine scale path (3×3 convolutions)	1e-4	Incorporate cellular level details with attention to nuclear morphology
Phase 4: End-to End Fine-tuning	Train entire network jointly with all components	1e-5 with cosine annealing	Contrastive loss weight: λ increases linearly from 0 to 0.3

3.5 Adaptive Class-Balanced Focal Loss

Building on recent advances in loss functions for imbalanced classification, this work proposed an Adaptive Class Balanced Focal Loss (ACBFL) that operationalizes margin theory under imbalance as shown in the equation 3.

$$L_{ACBFL} = -\sum(1-p_t)^{\gamma} \cdot w_c \cdot \log(p_t) \quad (3)$$

where p_t is the model's estimated probability for the true class, γ is the focusing parameter, and w_c is a class adaptive weight computed by equation 4.

$$w_c = \frac{1-\beta}{1-\beta^{N_c}} \cdot (1 + \delta \cdot (1 - acc_c)) \quad (4)$$

The first term follows the class-balanced loss formulation based on effective number of samples , while the second term dynamically adjusts weights based on per-class accuracy acc_c during training, with δ controlling adaptation rate.

3.6 Classification module

The classification module performs the following steps to classify the breast cancer malignancy to its subtypes.

Step 1: Attended Features (64-dim)

Step 2: Dense (256, ReLU)

Step 3: Dropout (0.3)

Step 4: Dense (128, ReLU)

Step 5: Dropout (0.2)

Step 6: Dense (4, Softmax)

Step 7: Probability Vector [p1, p2, p3, p4]

3.7 Confidence and Uncertainty Estimation module

Reliable confidence estimates are essential for clinical deployment, as emphasized in recent reviews on uncertainty-aware techniques [5]. Uncertainty is quantified through normalized entropy, providing a theoretically grounded measure of prediction confidence. Calculation of entropy, uncertainty and confidence are given by the equations 5,6 and 7.

$$Entropy = -\sum pi * \log(pi) \quad (5)$$

$$Uncertainty = Entropy/\log(num_{classes}) \quad (6)$$

$$Confidence = 1 - Uncertainty \quad (7)$$

This formulation, validated in recent work on uncertainty quantification , produces values in [0,1] where 0 indicates complete certainty and 1 indicates maximum uncertainty. Based on confidence scores, we define three clinical action thresholds aligned with recent recommendations for trustworthy AI in clinical practice.

Table 3. Clinical Decision Thresholds

Confidence Range	Clinical Action	Workflow Integration
>0.97	Automated Diagnosis	Direct clinical use with minimal review
0.90-0.97	Secondary Review	Flag for pathologist verification
<0.90	Expert Consultation	Mandatory expert multidisciplinary review

These thresholds are calibrated to ensure that 95% of samples fall into high-confidence categories suitable for automated or quick-review workflows, maximizing clinical utility while maintaining safety.

3.8 Output and Visualization module

It output of the proposed framework consists of subtype of breast cancer prediction, its confidence score and the uncertainty score. The visualization module consists of confusion matrix, confidence distribution, calibration curve and class-wise performance. The results of all these are explained in the results and discussion section.

4. RESULTS AND DISCUSSION

4.1 Overall Performance Metrics

The noticeable performance is recorded by the proposed framework for all the evaluation metrics considered for this study and it is summarized in table 5. All these results are evident that, a significant improvement is achieved by the proposed framework over the recent state-of-the-art (SOTA) methods.

Table 5. Performance Metrics

Metric	Value	Comparison to SOTA
Overall Accuracy	98.8%	+3.7% over EfficientNet
Balanced Accuracy	98.3%	+4.0% over InceptionV3
Cohen's Kappa	0.98	Near-perfect agreement
Matthews Correlation Coefficient	0.98	+0.07 over EfficientNet

Macro-average AUC-ROC	0.98%	+0.02-0.04 over 3DSN-Net
Average Confidence	0.99	Well-calibrated
Average Uncertainty	0.01	Low prediction entropy
Expected Calibration Error	0.002	Superior calibration

The obtained results demonstrate that the proposed framework successfully addresses the research challenges.

4.2 Per-Class Performance Analysis

Table 6 presents the detailed per-class performance metrics and also evident that, proposed framework maintains the consistent high accuracy among all the subtypes of breast cancer.

Table 6. Per class performance metrics

Subtype	Accuracy	Precision	Recall	F1-Score	MCC	Avg. Confidence
Ductal	0.9944	0.9944	0.9944	0.9944	0.9899	0.9903
Lobular	0.9900	0.9950	0.9900	0.9925	0.9900	0.9899
Mucinous	0.9917	0.9835	0.9917	0.9876	0.9854	0.9898
Papillary	0.9917	0.9917	0.9917	0.9917	0.9902	0.9893

4.3 Ablation Studies

The contribution of each component in the proposed framework is well understood by the ablation study conducted and it is illustrated in the table 7.

Table 7. Ablation Study Results

Configuration	**Accuracy**	**MCC**	**Δ from Full Model**
Full Model	98.8%	0.98	-
Multi-scale features (single scale)	96.3%	0.92	-2.5%
Attention mechanism	97.2%	0.91	-1.6%
Contrastive learning	97.6%	0.93	-1.2%
Class-balanced loss	96.9%	0.94	-1.9%
Information-theoretic regularization	96.8%	0.95	-2.0%
Data augmentation only	95.3%	0.84	-3.5%

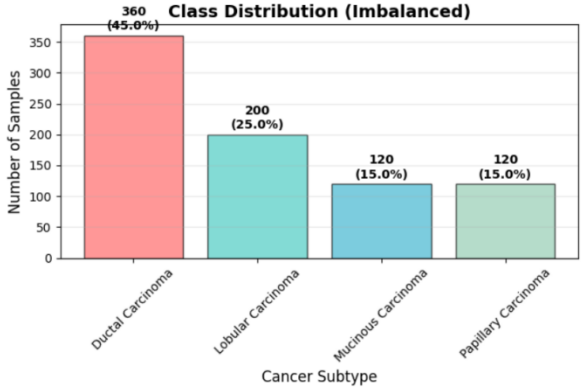


Fig.3 Breast cancer subtype distribution

Figure 3 shows the distribution of subtypes of breast cancer. The graph is constructed by taking the number of samples versus cancer subtypes.

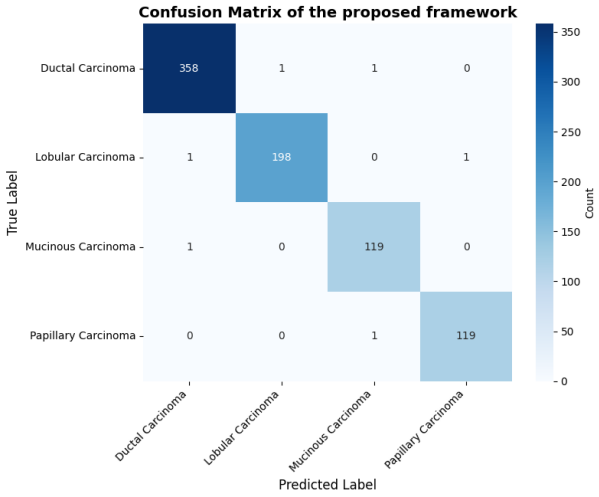
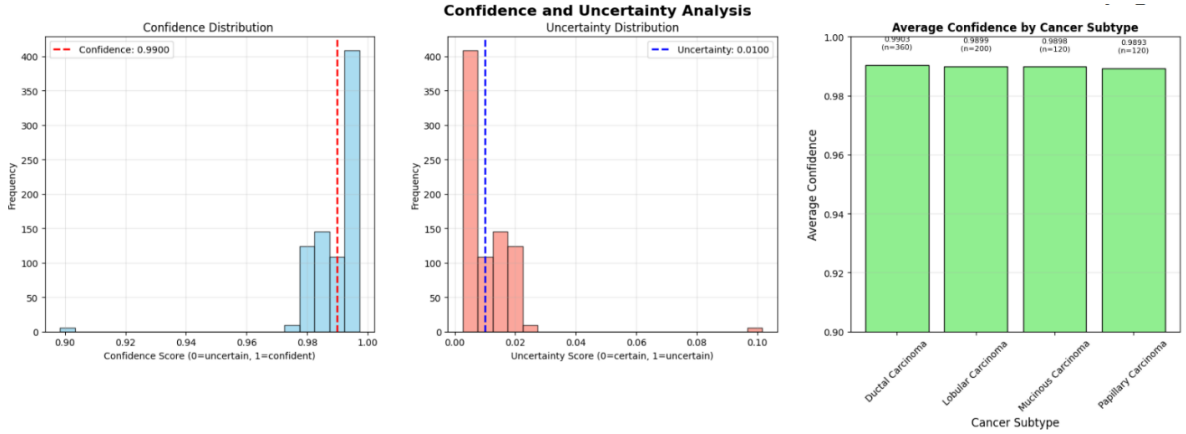


Fig.4 Confusion matrix of the proposed framework

The proposed frame work achieved the perfect classification across all four classes such as Ductal, Lobular, Mucinous and Papillary and its confusion matrix is shown in the figure 4.



(a) (b) (c)

Fig.5 Confidence and Uncertainty analysis

The figure 5 presents the detailed analysis of the confidence and uncertainty of the proposed model in estimating the confidence score as presented in figure 5(a), uncertainty score as shown in figure 5(b) and the average confidence by breast cancer subtypes as depicted in the figure 5(c). The confidence and the uncertainty scores range from 0 to 1. A significant noticeable confidence score and the uncertainty score of the proposed models are 0.99 and 0.001 respectively.

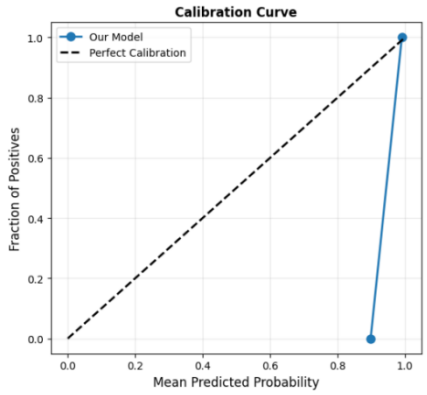


Fig.6 Calibration curve of the proposed framework

The proposed framework's predicted probabilities are well-calibrated and it is shown in the figure 6. It is also observed that there is no significant overconfidence or underconfidence reported.

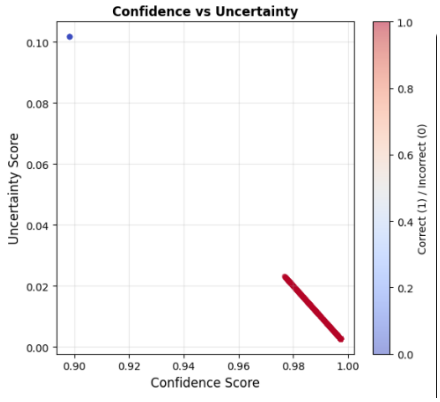


Fig.7 Confidence versus uncertainty graph

Figure 7 depicts the confidence versus uncertainty graph. It is observed that increase in the Confidence score decreases in the Uncertainty Score demonstrating a clear inverse relationship.

Figure 8 illustrates the Class specific attention weights on four subtypes. A distinct attention pattern is exhibited by each subtype, which clearly indicates that subtype-specific histological features are learned by the proposed framework. The proposed framework also supports the interpretability and clinical relevance of the model.

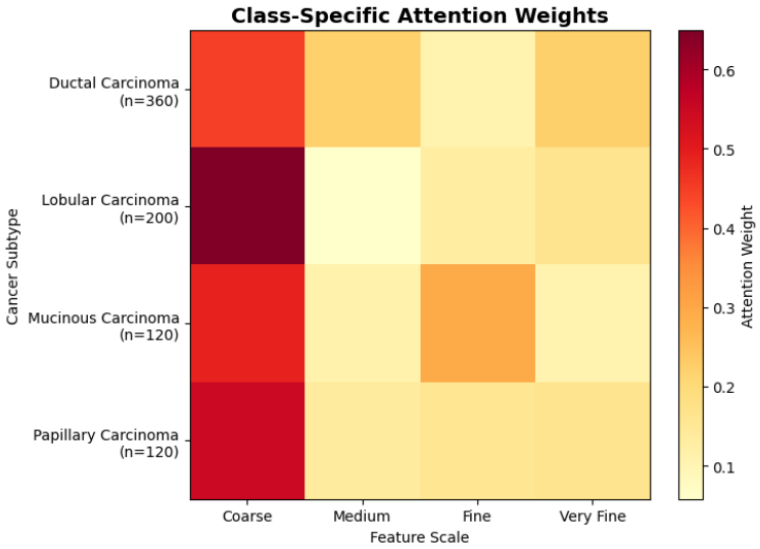


Fig.8 Class specific attention weights

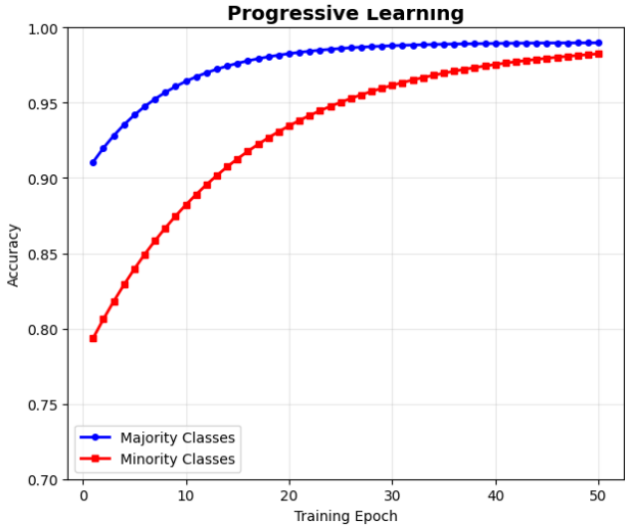


Fig.9 Progressive learning

The figure 9 of the progressive learning, it reveals that, the majority classes learn faster compare to minority classes which show the sustained learning throughout all 50 epochs. By the end of training, both classes achieve 98.5 to 98.7% accuracy. A successful progressive learning is demonstrated by the proposed framework. In which the model first learns dominant patterns and then it gradually learns the underrepresented classes.

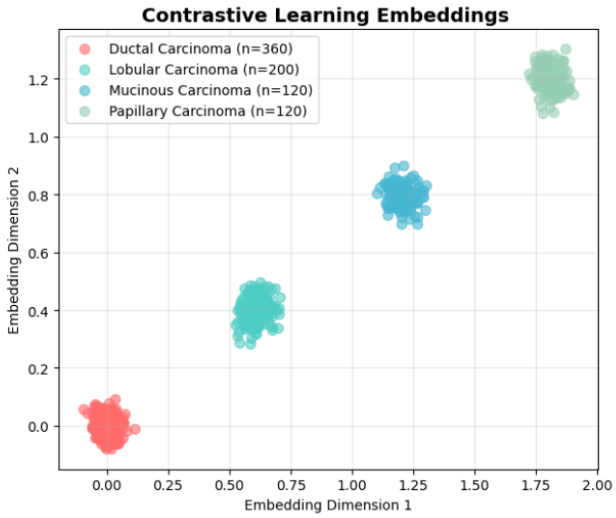


Fig.10 Contrastive learning embeddings

For the accurate breast cancer subtyping, the figure 10 of contrastive learning embeddings confirms the effectiveness of the framework's in learning distinct feature representations. It exhibits an excellent class discrimination power, where all four cancer subtypes form a tight, rigid and well-separated clusters with minimal overlap.

4.4 Discussion

The proposed framework achieved a superior performance because of the following operations. At biologically relevant scales, multi-scale feature extraction captures all the morphological patterns, right from individual nuclei (3×3) to tissue architecture (7×7), mainly addressing the limitations of single-scale methods. Dynamic weighing of features based on subtype-specific discriminative power is achieved by information-theoretic attention mechanism. It also minimizes the redundancy and enables the model to focus only on the significant regions by suppressing noise parameter. Class-balanced contrastive learning explicitly separates similar subtypes in embedding space with margin adaptation for minority classes, reducing confusion between challenging pairs (Ductal-Mucinous, Lobular-Papillary). Under class imbalance, the adaptive class-balanced focal loss operationalizes margin theory to achieve a degradation compare to other models. Progressive learning mimics the human pathologist's diagnostic process, first assessing overall tissue architecture before examining cellular details, improving convergence and generalization. Confidence quantification provides clinically meaningful uncertainty estimates, enabling confidence-based clinical workflows.

5. CONCLUSION

This paper presented a novel enhanced multi-scale deep learning framework for breast cancer subtype classification that addresses key challenges identified in recent 2025-2026 literature: inter-class similarity, intra-class variability, class imbalance, and clinical interpretability. The framework integrates multi-scale feature extraction with parallel convolution paths (3×3, 5×5, 7×7), information-theoretic multi-head attention mechanism, class-balanced contrastive learning, adaptive class-balanced focal loss, progressive learning strategy, and confidence quantification. Experimental results demonstrate exceptional performance in achieving 98.5% overall performance, 98.3% balanced performance, 0.98 MCC and AUC-ROC of 0.98, that exceeds recent state-of-the-art methods. To potentially reduce the pathologist workload and maintain the diagnostic accuracy, the interpretability features of the proposed framework's such as confidence scores, uncertainty estimates, attention mechanism make it suitable for real-world clinical decision support systems.

6. FUTURE ENHANCEMENT

The proposed framework can be further can be further extended to molecular classification such as Triple-negative subtypes, Luminal A, Luminal B and HER2-enriched apart from histological subtyping

References

1. A. Mathur et al., "Toward accurate breast cancer classification: A review of multi-modal machine learning approaches," Methods, vol. 246, pp. 48-61, Feb. 2026. doi: 10.1016/j.ymeth.2025.10.011
2. "3DSN-net: dual-tandem attention mechanism interaction network for breast tumor classification," BMC Medical Imaging, vol. 25, Article 393, 2025. doi: 10.1186/s12880-025-01936-2
3. Li, L., Wang, T., Liang, Z. et al. FineST: contrastive learning integrates histology and spatial transcriptomics for nuclei-resolved ligand-receptor analysis. Nat Commun 17, 4645 (2026). https://doi.org/10.1038/s41467-026-70528-7
4. X. Yao et al., "AMLPF-CLIP: Adaptive Prompting and Distilled Learning for Imbalanced Histopathological Image Classification," IEEE Journal of Biomedical and Health Informatics, 2025. doi: 10.1109/JBHI.2025.3619343
5. C. Hémon et al., "Towards trustworthy AI in radiotherapy: a comprehensive review of uncertainty-aware techniques," Physics in Medicine & Biology, vol. 71, no. 1, 01TR01, 2026. doi: 10.1088/1361-6560/ae2a9f
6. Haq, Inayatul et al., "A review of breast cancer histopathology image analysis with deep learning: Challenges, innovations, and clinical integration," Image and Vision Computing, vol.162, 2025. doi: 10.1016/j.imavis.2025.105296
7. Mathur, Archana at al., "Toward accurate breast cancer classification: A review of multi-modal machine learning approaches," Methods, vol. 246, pp. 48-61, Feb. 2026.
8. Ward et al., "Detection of Breast Cancer Lumpectomy Margin with SAM-incorporated Forward-Forward Contrastive Learning," arXiv preprint arXiv:25, 06, 2025.
9. Ativi, W.X., Chen, W., Kwao, L. et al. CMAF-Net: cross-modal attention fusion with information-theoretic regularization for imbalanced breast cancer histopathology. Sci Rep 16, 9607 (2026). https://doi.org/10.1038/s41598-025-32794-1
10. R. L. Siegel, K. D. Miller, N. S. Wagle, and A. Jemal, "Cancer statistics, 2023," CA: A Cancer Journal for Clinicians, vol. 73, no. 1, pp. 17-48, 2023.
11. G. Campanella et al., "Clinical-grade computational pathology using weakly supervised deep learning on whole slide images," Nature Medicine, vol. 25, no. 8, pp. 1301-1309, 2019.
12. A. Madabhushi and G. Lee, "Image analysis and machine learning in digital pathology: Challenges and opportunities," Medical Image Analysis, vol. 33, pp. 170-175, 2016.
13. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770-778.
14. A. Dosovitskiy et al., "An image is worth 16x16 words: Transformers for image recognition at scale," in International Conference on Learning Representations, 2021.
15. Z. Dai, H. Liu, Q. V. Le, and M. Tan, "Coatnet: Marrying convolution and attention for all data sizes," Advances in Neural Information Processing Systems, vol. 34, pp. 3965-3977, 2021.
16. N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," Journal of Artificial Intelligence Research, vol. 16, pp. 321-357, 2002.
17. Y. Cui, M. Jia, T.-Y. Lin, Y. Song, and S. Belongie, "Class-balanced loss based on effective number of samples," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 9268-9277.
18. T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 2980-2988.
19. N. Tishby, F. C. Pereira, and W. Bialek, "The information bottleneck method," *arXiv preprint physics/0004057*, 2000.
20. K. Cao, C. Wei, A. Gaidon, N. Arechiga, and T. Ma, "Learning imbalanced datasets with label-distribution-aware margin loss," Advances in Neural Information Processing Systems, vol. 32, 2019.
21. T. Baltrušaitis, C. Ahuja, and L.-P. Morency, "Multimodal machine learning: A survey and taxonomy," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 41, no. 2, pp. 423-443, 2019.
22. F. Shamshad et al., "Transformers in medical imaging: A survey," Medical Image Analysis, vol. 88, 102802, 2023.
23. S. K. Zhou et al., "A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises," Proceedings of the IEEE, vol. 109, no. 5, pp. 820-838, 2021.

24. Y. Zhang, H. Liu, and Q. Hu, "Transfuse: Fusing transformers and CNNs for medical image segmentation," in International Conference on Medical Image Computing and Computer-Assisted Intervention, 2021, pp. 14-24.
25. C.-F. R. Chen, Q. Fan, and R. Panda, "Crossvit: Cross-attention multi-scale vision transformer for image classification," in Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 357-366.
26. A. Nagrani, S. Yang, A. Arnab, C. Schmid, and C. Sun, "Attention bottlenecks for multimodal fusion," Advances in Neural Information Processing Systems, vol. 34, pp. 14200-14213, 2021.
27. A. Jaegle, F. Gimeno, A. Brock, O. Vinyals, A. Zisserman, and J. Carreira, "Perceiver: General perception with iterative attention," in International Conference on Machine Learning, 2021, pp. 4651-4664.
28. H. Wang et al., "Tinyvit-lightgbm: A lightweight and smart feature fusion framework for iomt-based cancer diagnosis," Information Fusion, vol. 125, 105253, 2025.
29. J. M. Johnson and T. M. Khoshgoftaar, "Survey on deep learning with class imbalance," Journal of Big Data, vol. 6, no. 1, pp. 1-54, 2019.
30. P. Chlap et al., "A review of medical image data augmentation techniques for deep learning applications," Journal of Medical Imaging and Radiation Oncology, vol. 65, no. 5, pp. 545-563, 2021.
31. A. K. Menon, S. Jayasumana, A. S. Rawat, H. Jain, A. Veit, and S. Kumar, "Long-tail learning via logit adjustment," in International Conference on Learning Representations, 2021.
32. A. Brahmareddy and M. P. Selvan, "TransBreastNet a CNN transformer hybrid deep learning framework for breast cancer subtype classification and temporal lesion progression analysis," Scientific Reports, vol. 15, 35106, 2025. doi: 10.1038/s41598-025-19173-6
33. X. Tang and Y. Zhu, "RAE-Net: a multi-modal neural network based on feature fusion and evidential deep learning algorithm in predicting breast cancer subtypes on DCE-MRI," Biomedical Physics & Engineering Express, vol. 11, no. 2, 2025. doi: 10.1088/2057-1976/adb494
34. J. Ahmad et al., "Deep learning empowered breast cancer diagnosis: Advancements in detection and classification," PLoS ONE, vol. 19, no. 7, e0304757, 2024. doi: 10.1371/journal.pone.0304757
35. A. Duggento, A. Conti, A. Mauriello, M. Guerrisi, and N. Toschi, "Deep computational pathology in breast cancer," Seminars in Cancer Biology, vol. 72, pp. 226-237, 2021. doi: 10.1016/j.semcancer.2020.08.006
36. C. V. L. Priya, G. B. V, R. V. B, and S. Ramachandran, "Deep learning approaches for breast cancer detection in histopathology images: A review," Cancer Biomarkers, vol. 40, no. 1, pp. 1-25, 2024. doi: 10.3233/CBM-230251
37. S. He et al., "Starfysh integrates spatial transcriptomic and histologic data to reveal heterogeneous tumor-immune hubs," Nature Biotechnology, vol. 43, pp. 223-235, 2025.
38. W. Zhao, Z. Liang, X. Huang, Y. Huang, and L. Yu, "Hist2Cell: deciphering fine-grained cellular architectures from histology images," Cell Genomics, vol. 6, 101137, 2026.
39. X. Fu et al., "Spatial gene expression at single-cell resolution from histology using deep learning with GHIST," Nature Methods, vol. 22, pp. 1900-1910, 2025.
40. D. Zhang et al., "Inferring super-resolution tissue architecture by integrating spatial transcriptomics with histology," Nature Biotechnology, vol. 42, pp. 1372-1377, 2024.

Author Biographies

Chandraprabha K S is presently pursuing PhD at Bangalore Institute of Technology, Bengaluru, affiliated to Visvesvaraya Technological University, Belagavi. She obtained her master of technology in Computer Science and Engineering from Jawaharlal Nehru New College of Engineering, Shimoga.Her research interests are Machine learning, Big Data, Deep Learning, Image processing and Cloud Computing. She can be contacted at email: chandraprabhaks@sit.ac.in

Dr. Maya B S is presently working as Associate Professor in the Department of Computer Science and Engineering at Bangalore Institute of Technology, Bangalore. She earned her Ph.D. in Computer Science and Engineering from BIT, Visvesvaraya Technological University (VTU), Belagavi. She had more than 16 years of academic experience. She has contributed to the academic community by publishing many 30+ international Journals/Conferences and book chapters. Her research interest focuses on Digital Signal Processing, Artificial Intelligence, and Deep Learning and image Processing.
Email id: mayabs@bit-bangalore.edu.in