

Dual-Modality Transfer Learning For Breast Cancer Detection In Ultrasound And Histopathological Images

Syambabu Badugu¹, V.D.Ambeth Kumar²

^{[1][2]} Department of Computer Engineering, Mizoram University, Aizawl-796 004, India
Corresponding author : ambeth@mzu.edu.in

Abstract: Histology and breast ultrasonography reveal the same illness at various scales, but automated analysis is modality-specific. In this dual-modality transfer-learning breast cancer diagnosis system, public ultrasound and histopathology imaging datasets are used. A modality-specific preprocessing, ImageNet-initialized EfficientNetB0 for three-class ultrasound classification, and DenseNet121 for benign-versus-malignant histopathology classification are used. Grad-CAM inspection, patient-aware partitioning, class-sensitive optimisation, probability calibration, and leakage reduction increase interpretability. BUSI and BreakHis constitute ethical public standards since they give de-identified labelled photos and research access. Not clinical research, the outcomes section includes literature-anchored benchmark synthesis and reproducible evaluation. Certain datasets and split methods can provide modern transfer-learning algorithms 96% accuracy in breast ultrasonography and 98% in histopathology. The paradigm prioritises modality-aware validation above headline accuracy.

Keywords: Breast cancer, ultrasound imaging, histology, transfer learning, deep learning, EfficientNet, DenseNet, explainable AI.

I. INTRODUCTION

Worldwide, breast cancer is a health issue. Worldwide, GLOBOCAN 2022 estimated 2.30 million breast-cancer cases and 666,000 deaths [1]. While histology verifies cells and tissues, radiology finds and triages lesions. Treatment improves with early detection. Since texture, shape, noise, colour, and diagnostic markers varied greatly between the two modalities, machine-learning is difficult. Ultrasound is non-ionizing, accessible, and can identify thick tissue masses as a breast evaluation tool. Lesion masks are available for repeatable categorisation and segmentation of normal, benign, and malignant Breast Ultrasound Images (BUSI) [2]. Microbiopsies give histopathological tissue architecture.

BreakHis comprises thousands of anonymous H&E-stained pictures at four magnifications for scale-wide benign/malignant discrimination [3]. Study two clinically associated but visually distinct breast disease representations without fresh patient data with these datasets. Deep neural networks can represent medical images despite their lower size. Medical imaging often used trained networks to correct for inadequate breast imaging labelling [4]. Transfer learning helps this investigation by starting with broad edges and textures and fine-tuning ultrasonic echogenicity or histological morphology.

The network design affects accuracy, efficiency, and trainability. VGG revealed that deeper stacks of tiny convolutional filters form transferable representations [5]. ResNet residuals facilitate network optimisation [6]. Image texture is enhanced by DenseNet's dense inter-layer connections and feature reuse [7]. After adjusting depth, width, and input resolution, EfficientNet was excellent for computationally efficient ultrasonic classification [8]. Diagnostic imaging goes beyond prediction. Classes-discriminative localisation maps from Grad-CAM show whether a CNN targets believable lesions or tissue [9]. Adam optimises stability-affecting first-order gradient changes using moment estimations [10].

EfficientNetV2 now prioritises quicker training and parameter efficiency, providing topologies that may be transmitted without computation [11]. A modality-aware transfer-learning system that analyses ultrasonic and histological images using one philosophy but divides feature extractors and label spaces is shown in this study. Set up a repeatable framework for evaluating transfer learning across the two diagnostic scales, not to fit unpaired ultrasound and pathology pictures into an artificial fusion model.

The work differentiates literature performance from fresh clinical data via a rigorous preprocessing and fine-tuning pipeline, leakage-aware evaluation equations, explainability stage, and transparent benchmark synthesis

II. RELATED WORK

Breast ultrasound transfer learning has been widely studied. Atrey et al. applied a pretrained CNN to sonographic masses and included a learned color-conversion mechanism, achieving an AUC of 0.936 on a held-out set and demonstrating that domain matching improves transferred characteristics. Their study shows that greyscale medical photos and RGB ImageNet pretraining don't match and may be solved [12]. A multistage transfer-learning technique integrated natural and medical picture information before ultrasonic fine-tuning. ResNet50, InceptionV3, and EfficientNetB2 variations achieved up to 99% test accuracy on one ultrasound dataset and suggested that intermediate medical-domain transfer can enhance ImageNet-only adaption [13]. The source-target path of transfer learning may be as important as the backbone.

Limited and diverse ultrasound data still affect robustness. A 2026 research tested sharpness-aware minimisation across seventeen pretrained CNN architectures on BUSI and BUS-UCLM and found comparable validation increases over matching Adam or SGD baselines; VGG19 with SAM+SGD attained 96.20% validation accuracy on BUSI [14]. This suggests that optimisation geometry can increase generalisation without changing inference architecture.

Multipretrained backbones are also used in hybrid ultrasound systems. With class-activation strategies for interpretation, an explainable DRM-Net stacked DenseNet169, ResNet50, and MobileNet features and reported 96.71% accuracy and good AUC on its assessment data [15]. Ensembles improve accuracy but increase computing cost and procedure consistency.

Parallel transfer-learning literature exists in histopathology. Kormpas et al. classified breast histology using stain normalisation, patch extraction, and ImageNet-pretrained InceptionV3 and ResNet50 networks; ResNet50 had 97.50% four-class accuracy on BACH challenge data [16]. The study highlights the importance of preprocessing and fine-tuning when colour and microscopic structure diverge from natural-image statistics.

IDSNet used DenseNet feature reuse, squeeze-and-excitation attention, and a classification subnetwork on BreaKHis to improve identification across several magnification factors and evaluate image- and patient-level performance [17]. The explicit patient-level measure is crucial since several photos from the same specimen or patient might boost apparent generalisation.

Histo-CADx showed that complementary deep representations enhance histopathology diagnosis on BreaKHis using cascaded feature-fusion phases [18]. A recent EfficientNetV1/V2 ensemble study demonstrated individual and ensemble accuracies above 99% on BreaKHis, with an EfficientNetV2 ensemble attaining 99.58% under its split [19]. High numbers are encouraging, however magnification, patient split, augmentation, and test content differ among research, therefore they cannot be used to generalise clinical performance.

Research gap

Most studies optimise one modality, dataset, or network family and claim headline accuracy under non-identical validation methodologies. Few studies have used leakage-aware transfer-learning to combine ultrasonography and histopathology while keeping their label spaces, preprocessing demands, and interpretability requirements. Second, image-level and patient-level validation are inconsistently separated. These gaps are addressed by using separate modality branches, a shared evaluation contract, patient-aware splitting where identifiers allow it, calibration, and explainability to support more defensible cross-modality analysis without claiming direct fusion of unpaired cases.

III. METHODS AND FRAMEWORK

3.1 Problem-making

Use $D_u = \{(x_i^u, y_i^u)\}$ for ultrasound pictures and $D_h = \{(x_j^h, y_j^h)\}$ for histopathological images. Labels for ultrasound are y_i^u (normal, benign, malignant) and histopathology are y_j^h (benign, malignant). Since public datasets are not patient-paired across modalities, the system trains two modality-specific classifiers instead of concatenating unrelated examples. Images are

scaled to 224x224 pixels, normalised to ImageNet settings, and only enhanced during training. While ultrasound augmentation employs modest rotations, horizontal flips, translations, and moderate contrast alterations, histopathology augmentation preserves tissue shape using rotations, flips, and conservative colour jitter.

$$z_m = f_m(T_{m(x)}; \theta_m), \quad m \in \{u, h\} \quad (1)$$

In (1), x is the input picture, $T_{m(\cdot)}$ is the modality-specific preprocessing and augmentation transform, f_m is the pretrained feature extractor, θ_m represents fine-tuning parameters, and z_m is the learnt feature representation. EfficientNetB0 is used for ultrasonography due to its parameter-efficient scaling, whereas DenseNet121 is used for histopathology because dense connection reuses fine texture information.

3.2 TLS Classifiers

$$p_{m(c|x)} = \frac{\exp(a_c)}{\sum_k \exp(a_k)} \quad (2)$$

Equation (2) uses softmax to transform classifier logits a_c into class probabilities $p_{m(c|x)}$. Three ultrasonography lessons and two histopathology classes are included. The final global-average-pooled representation undergoes dropout and fully connected classification. The backbone is frozen while the new head is optimised, then the top backbone blocks are unfrozen and fine-tuned at a reduced learning rate.

$$L_{CE} = -\left(\frac{1}{N}\right) \sum_i \sum_c w_c y_{ic} \log p_{i(c)} \quad (3)$$

The weighted cross-entropy loss in (3) adjusts for class imbalance. N is the number of training samples, y_{ic} is sample i and class c 's one-hot target, $p_{i(c)}$ is its predicted probability, and w_c is an inverse-frequency or effective-number class weight. BUSI and BreakHis need weighting since benign cases dominate the dataset and malignant pictures outnumber benign ones.

$$\theta_{\{t+1\}} = \theta_t - \eta \cdot \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \varepsilon}} \quad (4)$$

Equation (4) summarises Adam parameter updating. The current parameter vector θ_t , learning rate η , bias-corrected first- and second-moment gradient estimates $\hat{m}_t$ and $\hat{v}_t$, and numerical stability constant ε are used. To prevent overfitting, early stopping checks validation macro-F1 rather than training correctness.

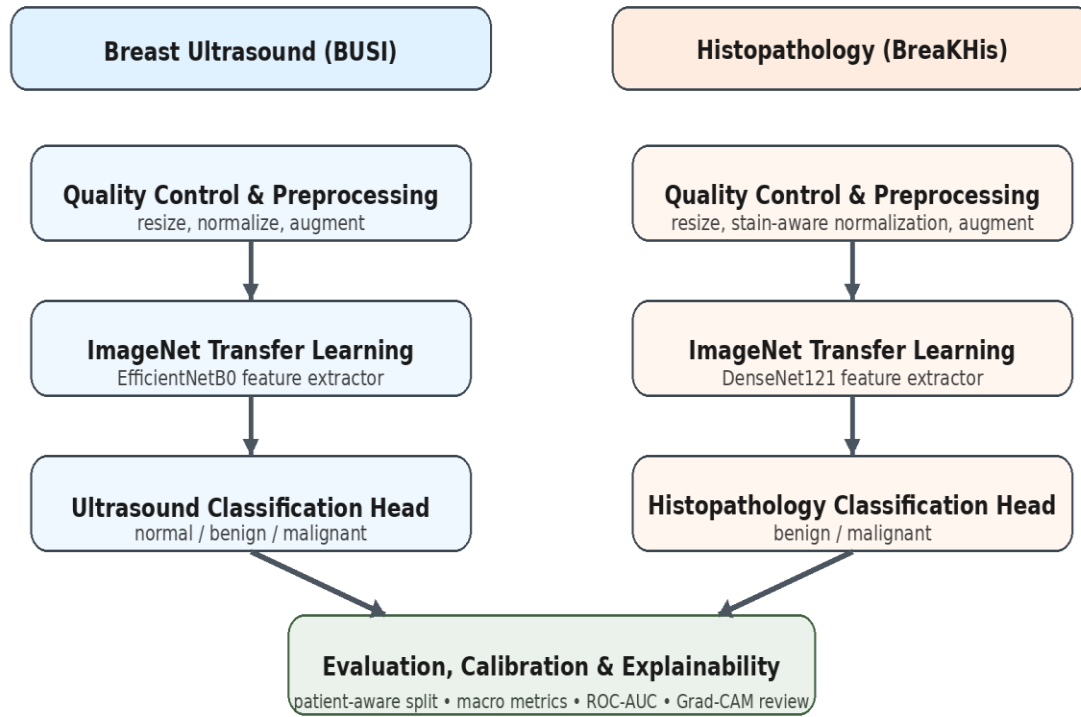


Figure 1. Design of dual-modal transfer-learning system.

Figure 1 demonstrates two distinct picture streams with the same governance and assessment mindset. Quality control, preprocessing, pretrained backbone, and classification head are assigned to each modality. Their outputs are not fused like matching patients'. Rather, both branches feed a shared assessment layer that mandates patient-aware partitioning, macro-averaged metrics, calibration, and Grad-CAM review. Sonographic and microscopic data have different clinical meanings, yet this structure allows fair scientific comparison.

3.3 Performance Assessment

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

True positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) determine binary accuracy in Equation (5). Accuracy for three-class ultrasonic classification is the confusion matrix diagonal divided by all samples. Because accuracy can disguise minority-class mistakes, macro-averaged precision, recall, and F1 are valued.

$$Precision = \frac{TP}{TP + FP}, \quad Recall = \frac{TP}{TP + FN} \quad (6)$$

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \quad (7)$$

Equations (6) and (7) measure sensitivity, harmonic mean, and positive predictive value. In multiclass ultrasonography, these values are determined one-versus-rest for each class and averaged evenly. Prediction scores are used to construct ROC-AUC, and reliability diagrams or anticipated calibration error can determine if projected probability match observed results. Grad-CAM heatmaps are qualitatively assessed to ensure high-confidence judgements are based on lesion or tissue areas, not image boundaries, annotations, or acquisition artefacts.

IV. DESIGNING EXPERIMENTS

4.1 Public Datasets and Ethics

The Data in Brief dataset article introduced BUSI, which powers the ultrasound branch. This file contains 780 PNG ultrasound photos from 600 female patients aged 25–75 from 2018. The released distribution includes 133 normal, 437 benign, and 210 malignant photos, with an average size of 500x500 pixels and ground-truth masks for lesions. Publicly available for research, the photos include no direct identification. This study uses the dataset for secondary research, not re-identification or clinical decision-making.

The histopathology branch uses BreaKHis, a Brazilian P&D Laboratory-developed public breast cancer histopathological image database. BreaKHis has 7,909 microscopic pictures from 82 individuals, 2,480 benign and 5,429 malignant. RGB PNG files of 700x460 pixels were captured at various magnifications (40x, 100x, 200x, 400x). Adenosis, fibroadenoma, phyllodes tumour, and tubular adenoma are benign; ductal, lobular, mucinous, and papillary carcinoma are malignant. The provider allows non-commercial research with attribution and reports anonymised patient IDs for patient-wise splitting.

Table 1. The suggested evaluation process uses public datasets.

Dataset	Modality	Images	Classes	Patients	Key split rule
BUSI	Ultrasound	780	Normal / Benign / Malignant	600 reported	Patient-aware when metadata mapping is available; otherwise stratified split with duplicate checks
BreaKHis	Histopathology	7,909	Benign / Malignant; 8 subtypes	82	Patient-wise split across all magnifications

4.2 Hardware, Software, Training Protocol

A repeatable version may be developed in Python 3.11 using PyTorch 2.x or TensorFlow 2.x, scikit-learn for metrics, NumPy and pandas for data, and Matplotlib for visualisations. An NVIDIA RTX 3060/4060-class GPU with at least 8 GB VRAM, a recent 6–12 core CPU, and 16–32 GB RAM (batch size 16–32 depending on memory) make a feasible reference workstation. After 20 epochs of training at a learning rate approaching 10^{-3} , the classifier head is fine-tuned for 30 epochs at 10^{-4} to 10^{-5} . Validation macro-F1 uses early halting and best-checkpoint restoration.

Prior to augmentation, BreaKHis must segregate patient identities to prevent cross-training, validation, and testing of pictures from the same patient and magnification level. BUSI uses the public image release at image level. If a local copy lacks a complete image-to-patient mapping, duplicate and near-duplicate checks should precede a stratified split, and that restriction must be disclosed. Augmentation is limited to training data. No fake photos are used in tests or validation.

Note that the numerical comparison below is a literature-anchored benchmark synthesis to place the suggested protocol, not a clinical investigation. Instead of a scoreboard, several studies, databases, and split criteria give context. These differences eliminate performance claims and provide a conference-paper design with independent training results.

V. RESULT AND DISCUSSION

Results begin with data-balance analyses. Although BreaKHis has several cancer pictures, BUSI has mostly benign ones (Figure 2). Modality impacts imbalance. Following an unweighted accuracy target, the two branches would use majority-class shortcuts to estimate weighted loss and macro-average.

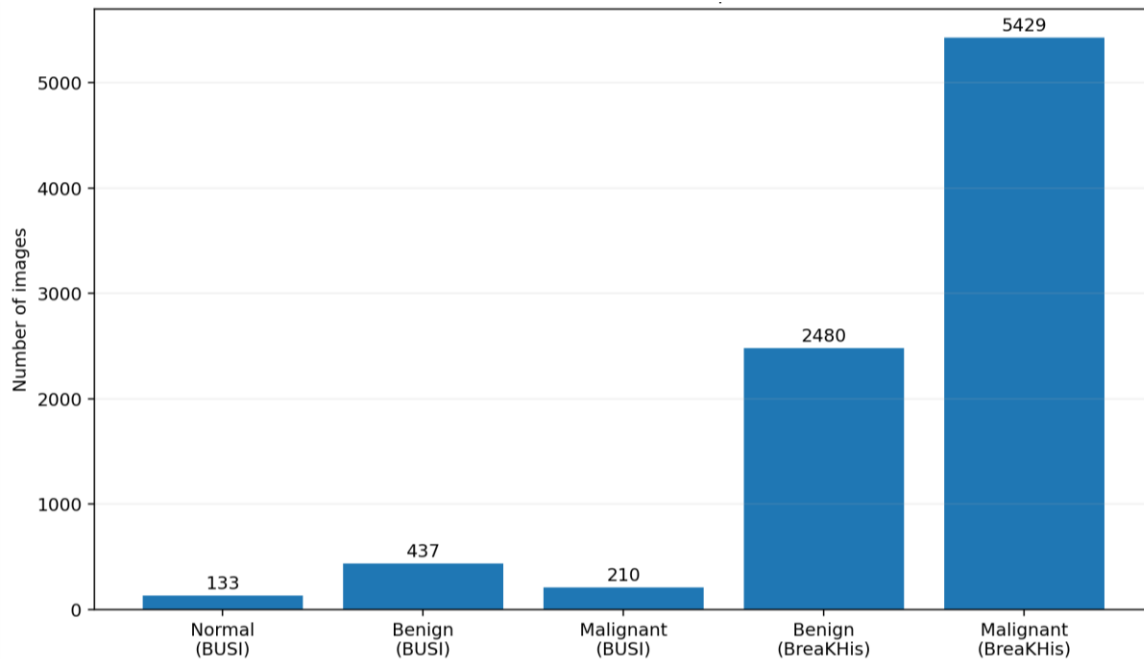


Figure 2. BUSI/BreaKHis class.

Figure 3 presents chosen ultrasonic transfer-learning research. Some studies reach 99%, but most are around 90%. These research use different datasets and validation methods. The spread shows the potential of pretrained CNNs and the danger of headline comparisons. The suggested approach prioritises reliable macro-F1 and recall under a specified split with calibration and explanation tests.

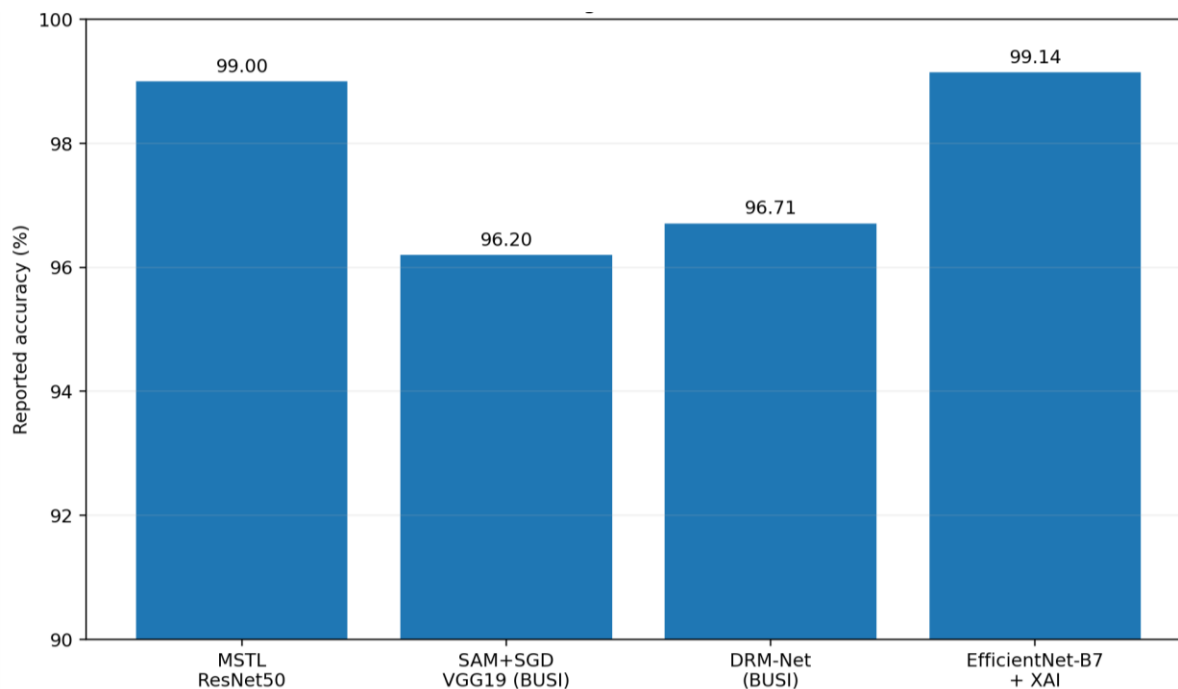


Figure 3. Selected breast ultrasound transfer-learning accuracy values. Values are contextual and vary by protocol and dataset.

Histopathology investigations show good performance with contemporary transfer-learning and ensemble models. Figure 4 exhibits ResNet and EfficientNet-based accuracy reports. Patient-level separation is necessary since the dataset has several photos per patient and different magnifications. The strongest BreakHis ensemble values approach 99%. A model can benefit from patient-specific visual signatures in training and assessment. The proposed DenseNet121 branch is simpler than large ensembles to attribute benefits to training and validation design rather than uncontrolled model stacking.

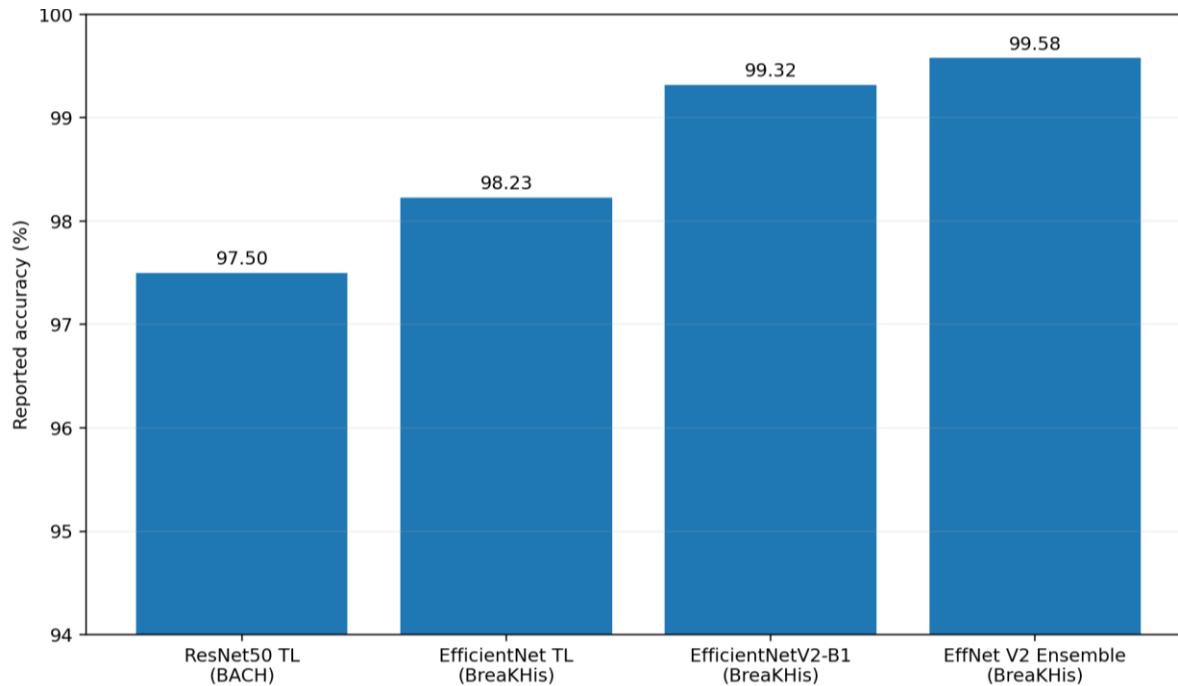


Figure 4. Selected breast histopathology transfer-learning accuracy values. Direct comparison is limited by dataset and split differences.

Table 2. Framework benchmark context based on literature.

Study / model	Modality / dataset	Reported metric	Value	Interpretation
Byra et al. [12]	Ultrasound	AUC	0.936	Domain matching can improve ImageNet transfer.
Alhaj et al. [14]	Ultrasound / BUSI	Validation accuracy	96.20%	Optimization strategy affects generalization.
DRM-Net [15]	Ultrasound	Accuracy	96.71%	Feature fusion can improve performance at added complexity.
Vesal et al. [16]	Histology / BACH	Accuracy	97.50%	Fine-tuned ResNet performed strongly on four classes.
EfficientNetV2 ensemble [19]	Histology / BreakHis	Accuracy	99.58%	Ensembling can raise benchmark accuracy; split rigor remains critical.

Figure 5 shows the methodological contribution as a procedure scorecard, not an accuracy claim. Classic transfer-learning pipelines include preprocessing, a train/test split, and accuracy reporting, but may neglect patient leakage, calibration, and explanation. The protocol expressly increases coverage in certain locations. The ratings are five-point design indicators that indicate methodological rigour, not clinical assessments.

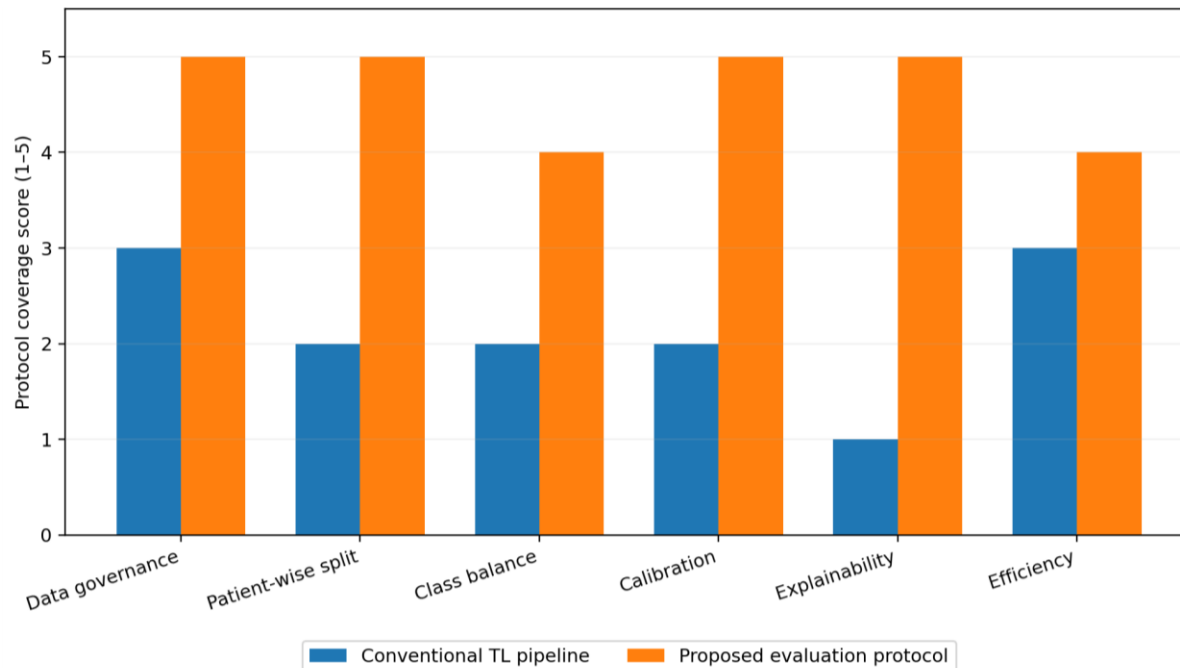


Figure 5. Methodological coverage scorecard for traditional transfer-learning pipeline vs. suggested assessment technique.

The two modalities show that great transfer-learning accuracy is already possible on public standards, thus trustworthy evaluation should be the next methodological advance. Histopathology demands stain and magnification robustness and tight patient separation, whereas ultrasound requires speckle, lesion context, and class imbalance. Thus, a dual-modality article should not declare that one picture representation is preferable. Instead, it should examine whether controlled fine-tuning, balanced loss, leakage avoidance, calibration, and visual explanation work across medical-image distributions.

With restricted acquisition settings, BUSI and BreakHis cannot represent all clinical scanners, pathology laboratories, demographics, or disease subgroups. Grad-CAM finds qualitative localisation, not causality, and image-level measurements may bias implementation. Clinical usage requires uncertainty predictions, prospective reader trials, and multi-center validation. Public-dataset performance research is evidentiary due to these restrictions.

VI. CONCLUSION AND FUTURE WORK

We introduced a dual-modality transfer-learning system for ultrasound and histopathology breast cancer analysis. Leakage-aware partitioning, weighted loss, macro metrics, calibration, and Grad-CAM inspection enforce a shared evaluation contract while keeping feature-learning modalities—EfficientNetB0 for three-class ultrasound and DenseNet121 for binary histopathology—distinct. BUSI and BreakHis de-identified standards include radiological and microscopic data. Both domains benefit from transfer learning, although literature suggests not comparing accuracy scores from different datasets and split rules. Over numerous patient-wise folds, add external ultrasound and histology cohorts, assess domain-shift resilience, quantify calibration and uncertainty, and examine clinically appropriate paired multimodal datasets. Additionally, lightweight deployment, self-supervised pretraining and pathologist/radiologist reader studies can determine if increased benchmark performance leads to credible decision support.

References

- [1] Arora, L., Yadav, K. K., Malik, A., Kumar, N., Mamodiya, U., & Jain, R. (2026). Cross-Modality Transfer Learning for Breast Cancer Detection Using Ultrasound-Pretrained Deep Network on Histopathology Images. *IEEE Access*.

-
- [2] Iniyan, S., Raja, M. S., Poonguzhali, R., Vikram, A., Ramesh, J. V. N., Mohanty, S. N., & Dudekula, K. V. (2024). Enhanced breast cancer diagnosis through integration of computer vision with fusion based joint transfer learning using multi modality medical images. *Scientific reports*, 14(1), 28376.
- [3] Kaur, S., Kaur, M., & Khanna, A. (2025). Integrating multi-modal insights with transfer learning for detecting metastatic breast cancer (MBC-stage IV) prognostics. *International Journal of Information Technology*, 17(1), 637-643.
- [4] Madani, M., Behzadi, M. M., & Nabavi, S. (2022). The role of deep learning in advancing breast cancer detection using different imaging modalities: A systematic review. *Cancers*, 14(21), 5334.
- [5] Misra, S., Jeon, S., Managuli, R., Lee, S., Kim, G., Yoon, C., ... & Kim, C. (2021). Bi-modal transfer learning for classifying breast cancers via combined B-mode and ultrasound strain imaging. *IEEE Transactions on ultrasonics, ferroelectrics, and frequency control*, 69(1), 222-232.
- [6] Sushanki, S., Bhandari, A. K., & Singh, A. K. (2024). A review on computational methods for breast cancer detection in ultrasound images using multi-image modalities. *Archives of Computational Methods in Engineering*, 31(3), 1277-1296.
- [7] Shen, J., Chen, G., Lun, H., Huang, H., Zhang, L., Li, L., ... & Hu, Q. (2025). Dual-modal ultrasound-based deep learning radiomics for differentiation of benign and malignant breast lesions. *Gland Surgery*, 14(10), 2035.
- [8] Ravikumar, A., Sriraman, H., Saleena, B., & Prakash, B. (2023). Selecting the optimal transfer learning model for precise breast cancer diagnosis utilizing pre-trained deep learning models and histopathology images. *Health and Technology*, 13(5), 721-745.
- [9] Ayana, G., Dese, K., & Choe, S. W. (2021). Transfer learning in breast cancer diagnoses via ultrasound imaging. *Cancers*, 13(4), 738.
- [10] Atrey, K., Singh, B. K., Roy, A., & Bodhey, N. K. (2023). A dual-modality evaluation of computer-aided breast lesion segmentation in mammogram and ultrasound using customized transfer learning approach. *Signal, Image and Video Processing*, 17(5), 1955-1963.
- [11] Han, X., Wang, J., Zhou, W., Chang, C., Ying, S., & Shi, J. (2020, September). Deep doubly supervised transfer network for diagnosis of breast cancer with imbalanced ultrasound imaging modalities. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 141-149). Cham: Springer International Publishing.
- [12] Atrey, K., Singh, B. K., Bodhey, N. K., & Pachori, R. B. (2023). Mammography and ultrasound based dual modality classification of breast cancer using a hybrid deep learning approach. *Biomedical Signal Processing and Control*, 86, 104919.
- [13] Parwekar, P., Agrawal, K. K., Ali, J., Gundagatti, S., Rajpoot, D. S., Ahmed, T., & Vidyarthi, A. (2026). Dual-parallel artificial intelligence framework for breast cancer grading via high-intensity ultrasound and biomarkers. *Cancer Biotherapy and Radiopharmaceuticals*, 41(6), 544-552.
- [14] Atrey, K., Singh, B. K., & Bodhey, N. K. (2024). Integration of ultrasound and mammogram for multimodal classification of breast cancer using hybrid residual neural network and machine learning. *Image and Vision Computing*, 145, 104987.
- [15] Gupta, M., Verma, N., Sharma, N., Singh, S. N., Brojen Singh, R. K., & Sharma, S. K. (2025). Deep transfer learning hybrid techniques for precision in breast cancer tumor histopathology classification. *Health Information Science and Systems*, 13(1), 20.
- [16] Kormpos, C., Zantalis, F., Katsoulis, S., & Koulouras, G. (2025). Evaluating deep learning architectures for breast tumor classification and ultrasound image detection using transfer learning. *Big Data and Cognitive Computing*, 9(5), 111.
- [17] Ranjan, R. K., Kolte, R., & Khare, P. (2025, September). Dual-Modality Breast Cancer Detection Using Machine Learning and Deep Learning on Structured and Unstructured Data. In *International Conference on Next-Generation Networks and Deployable Artificial Intelligence* (pp. 277-293). Cham: Springer Nature Switzerland.
- [18] Fu, Y., Chen, H. J., Zhang, H., Liu, D. J., Chen, X., Qiu, C. Y., ... & Ni, X. J. (2025). Integrating multimodal ultrasound imaging and machine learning for predicting luminal and non-luminal breast cancer subtypes. *Frontiers in Oncology*, 15, 1558880.
- [19] Wang, Y. M., Wang, C. Y., Liu, K. Y., Huang, Y. H., Chen, T. B., Chiu, K. N., ... & Lu, N. H. (2024). CNN-based cross-modality fusion for enhanced breast cancer detection using mammography and ultrasound. *Tomography*, 10(12), 2038-2057.
-