

AHDNAM: AN OPTIMISED HYBRID DEEP LEARNING AND SWARM INTELLIGENCE FRAMEWORK FOR BREAST ULTRASOUND DIAGNOSIS

DR. M. VAMSIKRISHNA¹, G RESHMA², NARENDRA B MUSTARE³, B V SUBBA RAO⁴,
DRAKSHAYANI SRIRAMSETTI⁵, ELANGO VAN MUNIYANDY⁶

¹Professor & HoD, Department of Computer Applications, Aditya University, Surampalem, India.

²Assistant Professor, Department of Information Technology, PVP Siddhartha Institute of Technology, Vijayawada, India.

³Professor, Department of EIE, CVR College of Engineering, Hyderabad, India.

⁴Professor, Dept of Information Technology, PVP Siddhartha Institute of Technology, Vijayawada, India.

⁵Assistant Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, India.

⁶Department of Biosciences, Saveetha School of Engineering. Saveetha Institute of Medical and Technical Sciences, Chennai, India

Email id: ¹vkangalampalli@gmail.com, ²greshma@pvpsiddhartha.ac.in, ³drnamust@gmail.com,

⁴bvsrau@gmail.com, ⁵drakshasri@gmail.com, ⁶muniyandy.e@gmail.com

ABSTRACT

Breast tumor classification from ultrasound is challenged by low contrast and speckle noise, limiting reliable automated triage. A reproducible, high-recall classifier that is robust to device and acquisition variability remains a pressing gap for clinical workflows. This study aimed to evaluate an adapted BCDNet incorporating a dual-branch hybrid classifier and an adaptive constrained hyperparameter search to improve malignant lesion detection and overall balanced performance on public ultrasound datasets. An experimental study used ImageNet pretrained VGG16 as a backbone with two parallel classifier branches: a dilated atrous spatial pyramid pooling branch and a channel-attention branch. Augmentation (mixup, cutmix), label smoothing, focal loss, and stochastic weight averaging were applied. Patient-level splits from the Kaggle breast ultrasound dataset and BUSI were used for training, validation, and held-out testing; hyperparameters were tuned with a 50-evaluation RPAOSM-ESO-inspired search and top models ensemble. On held-out tests, the hybrid pipeline attained an accuracy 94.5% (Kaggle) and 93.2% (BUSI), a sensitivity \approx 94.6% and 93.0%, an MCC of 0.881 and 0.860, false negative rates 5.43% and 6.80%, and an AUC \approx of 0.966 and 0.959, respectively. Bootstrap CIs indicated statistically meaningful gains in MCC and FNR versus single-branch baselines. In spite of the fact that, before the proposed pipeline can be utilized in clinical practice, more comprehensive annotations and broader multi-center validation have to be taken into consideration, it proves that it can be used as a research tool that results in a smaller number of missed malignancies without compromising the specificity.

Keywords: *Breast Ultrasound, Transfer Learning, VGG16, ASPP, Attention Hybrid Classifier, Hyperparameter Optimization, Explainability*

1. INTRODUCTION

In spite of the fact that before the proposed pipeline can be utilized in clinical practice, more comprehensive annotations and broader multi-center validation have to be taken into consideration, it proves that it can be used as a research tool that results in a smaller number of missed malignancies without compromising the specificity.[1], [2], [3]. Clinical ultrasound images, however, suffer from

speckle noise, low contrast lesion boundaries, and substantial operator dependence. These factors limit reproducibility and complicate both human and automated interpretation. Deep convolutional networks offer strong visual feature learning, and transfer learning from large natural image corpora reduces labelled data needs [4], [5]. Channel attention and dilated convolution spatial pyramid pooling can capture multi-scale lesion context and suppress extraneous background data. Combining

regularization, scheduling, and limited hyperparameter search improves small medical dataset generalization. [6], [7], [8]. Despite progress, research is lacking. Most ultrasound-based breast tumor classification studies use single-branch CNN architectures or prioritize accuracy without limiting false negative rates, the most clinically significant error. Many described solutions include large structures or intensive hyperparameter tuning processes that are hard to replicate and unsuitable for low computational resources. Few studies focus on robustness, calibration, and reproducibility in public ultrasound datasets due to device variability and insufficient patient-level variation.

This study addresses the lack of a lightweight, reproducible, clinically focused ultrasound classification process that maximizes malignant lesion recall and maintains specificity across varied public datasets. Existing models fail to balance sensitivity, robustness, and computational feasibility, making them unreliable triage tools in clinical workflows.

This work aims to create, evaluate, and document a hybrid deep learning system for ultrasound breast tumor classification that reduces false negatives without compromising performance. A transfer learning backbone with complementing multi-scale dilated feature extraction and channel-wise attention, modern regularization, and a limited adaptive hyperparameter optimization strategy achieves this. Models must generalize reliably across datasets, be computationally efficient, and produce interpretable outputs for clinical examination and validation, justifying the proposed strategy.

The core problem is the absence of a reproducible, compact pipeline for ultrasound-based breast tumor classification that maximizes malignant lesion recall while maintaining clinically acceptable false positive rates and operating feasibly on single GPU hardware. Secondary gaps include limited device and population diversity in public datasets lack of robust uncertainty quantification, and insufficient explainability that would enable clinical review and targeted error analysis. [9], [10].

Addressing this problem has direct clinical and operational value. A high recall automated triage tool can reduce missed malignancies, accelerate referral workflows, and reduce inter-reader variability, particularly in low-resource settings where sonographer expertise and radiology capacity are limited [11]. From an engineering standpoint, a solution that generalizes across datasets while remaining computationally lightweight enables broader adoption in research and on device prototypes. Transparent explainability and

reproducible training recipes increase trust and facilitate external validation and regulatory review. [12].

The primary objective is to design, evaluate, and document a practical ultrasound classification pipeline that prioritizes malignant lesion recall while preserving specificity. Secondary objectives are to improve model calibration and robustness through modern regularization and a constrained adaptive hyperparameter search to produce reproducible checkpoints and interpretable saliency outputs suitable for clinical inspection.

The key contributions of the research

1. An adapted BCDNet instantiation that fuses a VGG16 transfer backbone with a dual branch hybrid classifier combining dilated multi-scale context and channel attention.
2. A constrained adaptive hyperparameter tuning routine inspired by opposition sampling and swarm guided refinement to obtain performant configurations under limited compute budgets.
3. A comprehensive regularization suite including mixup, cutmix, label smoothing, focal loss, and stochastic weight averaging to improve calibration and stability.
4. Explainability via GradCAM and targeted failure analysis, identifying low contrast lesions and device-specific artifacts as primary failure modes.
5. A reproducible evaluation protocol using patient-level splits, bootstrap confidence intervals, ablation studies, and robustness tests on public datasets (Kaggle and BUSI) together with an experiment checklist to support external validation.

2. LITERATURE REVIEW

Ultrasound breast tumor classification heavily relies on deep learning techniques, namely transfer learning with pretrained convolutional neural networks (CNNs) like VGG, ResNet, and EfficientNet. While these methods do a better job with less labeled data, they are still vulnerable to aberrations and noise from ultrasound and often use single-branch topologies, which do not capture variability at the lesion scale well. To improve contextual feature learning and eliminate unnecessary regions, multi-scale techniques (such as dilated convolutions or ASPP) and attention mechanisms have been developed. [13], [14]. [15], [16] Nonetheless, these methods are frequently used alone or in conjunction with computationally

intensive models. In addition, a lot of research focuses on total accuracy but ignores clinically important measures like false negative rate and reproducibility under limited computing resources. By utilizing a lightweight dual-branch hybrid framework that integrates multi-scale dilation and channel attention, backed by state-of-the-art regularization and reproducible hyperparameter optimization, these deficiencies are filled. The evaluation process centers around sensitivity and resilience. [17], [18], [19].[20], [21]. Attention mechanisms implement the theoretical idea of selective feature amplification by learning gating functions that focus on informative channels or spatial locations. Optimization theory motivates constrained adaptive hyperparameter search as a cost-effective alternative to exhaustive grid search; heuristic metaheuristics balance exploration and exploitation under limited evaluation budgets [22]. Regularization theory underpins augmentation and label smoothing strategies, which reduce overfitting by enforcing smoother decision boundaries and by encouraging models to be robust to input perturbations. Explainability methods such as GradCAM are grounded in gradient-based attribution theory that links output logits to spatial importance maps and thereby facilitates human interpretation [23].

Several gaps remain. First dataset diversity and standardization of the first dataset are limited in ultrasound imaging; many public collections are small originate from single sites, and lack consistent patient-level metadata, which challenges generalization claims. Second, the trade-off between model complexity and deployability is debated; large transformer or heavy ensemble models may yield marginal accuracy gains at the cost of impractical inference latency for edge or single-GPU deployment. Third, hyperparameter search methods vary widely, and published comparisons often lack consistent evaluation budgets, making reproducibility and fair comparison difficult. Fourth, explainability tools provide useful clues but are not definitive; attribution maps can be sensitive to implementation details and may mislead unless accompanied by robust localization labels and systematic evaluation. Finally, prospective clinical validation, including calibration and workflow integration, remains sparse; model performance reported on retrospective public datasets does not guarantee real-world utility. These gaps motivate pipelines that emphasize reproducibility, patient-level evaluation, and lightweight architectures amenable to clinical translation.

3. DATASETS AND PREPROCESSING

3.1 Primary Datasets Used

Experiments focus on publicly available ultrasound datasets and representative cross-modality data for pretraining experiments.

Kaggle breast ultrasound images dataset: This dataset contains clinical ultrasound images labeled as normal, benign, and malignant. The present experiments use the benign and malignant classes for primary binary classification and reserve a small normal subset for out-of-distribution assessment. Approximately 1500 images were utilized after standard quality filtering and removal of corrupt entries.

BUSI dataset: The Breast Ultrasound Image dataset by Al Dhabyani et al contains benign, malignant, and normal examples. The corrected BUSI splits and patient-level metadata enable reproducible evaluation. Approximately 780 images were included, with a breakdown of 437 benign, 210 malignant, and 133 normal. Patient-level splitting was enforced where metadata permitted.

CBIS DDSM for cross-modality experiments: CBIS DDSM mammography studies were retained for optional pretraining and transfer experiments to evaluate cross-modality generalization. This dataset was not used for final metrics on ultrasound classification but served as a source for representation learning tests.

Dataset curation included the removal of duplicate frames and images with excessive annotation overlays. Patient-level identifiers were used to avoid leakage between the train, validation, and test sets.

3.2 Preprocessing and Augmentation

Images were resampled to 224 by 224 pixels to match VGG16 input expectations. Preprocessing steps applied consistently across datasets included intensity clipping to the 1 and 99 percentiles, histogram equalization for contrast enhancement, and a mild speckle-reducing anisotropic diffusion filter to reduce high-frequency ultrasound artifacts. Pixel intensities were normalized using ImageNet channel statistics for compatibility with pretrained weights.

On-the-fly augmentation comprised geometric and photometric transforms applied during training. Augmentations included random rotations up to 20 degrees, small translations, horizontal flips, random brightness and contrast jitter, elastic deformation, and probabilistic application of speckle noise. Mixup and cutmix were applied at the minibatch level with conservative alpha parameters to preserve lesion

visibility. Class-aware sampling mitigated imbalance. When patient-level imbalance existed, a stratified sampling policy ensured each minibatch contained a representative ratio of malignant and non-malignant cases.

3.3 Train validation test splits

A stratified split preserved class proportion. Standard practice used a 70 15 15 partition for train validation and test, with patient-level separation enforced. When patient metadata was incomplete, cross-validation with non-overlapping studies was used. For the BUSI and Kaggle datasets, patient-level partitions matched published corrected splits when available.

4. METHODS

4.1 Feature Extractor

The initial step in preparing ultrasound images for pretrained weight compatibility is to resize them to 224x224 pixels, add contrast, and normalize them using ImageNet statistics. The feature extraction is carried out by removing the fully linked layers from an ImageNet-pretrained VGG16 backbone. To adjust learnt representations to ultrasound-specific texture and boundary properties while reducing overfitting on limited data, the backbone works as a fixed feature encoder in the initial training stage. Following this, higher convolutional blocks are partially fine-tuned.

4.2 Hybrid classifier AHDNAM

Figure 1 shows the schematic of the adapted BCDNet architecture. A dual-branch hybrid classifier is used to process the extracted convolutional features. Its purpose is to capture complementary information. A first branch models multi-scale contextual information without sacrificing spatial resolution using dilated convolutions and atrous spatial pyramid pooling. To highlight diagnostically important characteristics and suppress ultrasonic artefacts, the second branch employs compact convolutional layers in conjunction with channel-wise attention. To improve robustness and prevent overconfident predictions, the outputs from both branches are projected into class logits and averaged before

softmax. This creates a lightweight ensemble effect.

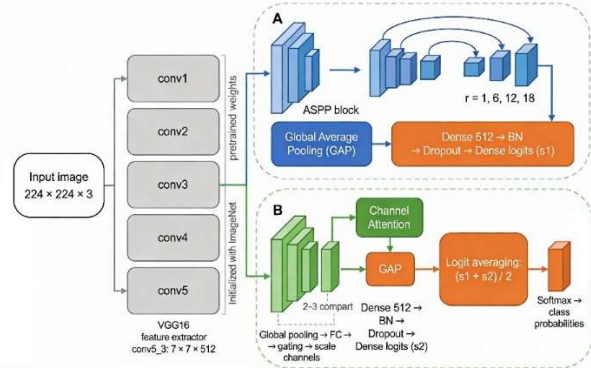


Figure 1 System Architecture

Table 1 shows that the VGG16 full net has ~138M parameters, including FCs. Removing the top classifier reduces the parameter count significantly. The convolutional portion of VGG16 is ~14.7M params. The added heads and branches yield the reported approximate total. Exact parameter counts should be computed from the final training code and recorded in the reproducibility checklist.

Table 1 — Hybrid classifier

Layer / Block	Type	Output shape (input 224x224x3)	Approx params	Notes
Input	—	224 × 224 × 3	0	RGB image
VGG16 conv1..conv5 (no top)	Conv blocks	conv5_3 → 7 × 7 × 512	~14.7 M	ImageNet pretrained; final FC layers removed
Global average pooling	GAP	1 × 1 × 512	0	Shared between branches
Projection head (shared)	Dense → BN → Dropout → Dense	512-d descriptor	~0.3 M	A two-layer MLP is used to produce a compact descriptor
Branch A dilated	Dilated convs	512-d	~0.5 M	expand receptive field

conv stack	+ residuals			without pooling
ASPP	Parallel dilated convs	512-d	~0.2 M	dilation rates suggested 1,6,12,18
Branch A head	GAP → Dense → logits	num_classes	~0.1 M	classification head
Branch B 1D conv + channel attention	1D conv blocks + SE style attention	512-d	~0.4 M	Channel gating emphasizes salient channels
Branch B head	GAP → Dense → logits	num_classes	~0.1 M	classification head
Final averaging	Logit average	num_classes	0	(s1 + s2)/2 before softmax
Total (approx)			~16.4 M	convolutional params dominate; exact total depends on head sizes

4.3 Hyperparameter Optimization Routine RPAOSM-ESO Inspired Tuning

Optimizing model performance is achieved by a restricted adaptive hyperparameter search that draws inspiration from swarm-based refinement and opposition sampling. In order to guarantee computational feasibility and reproducibility, the search investigates learning rate, dropout, weight decay, and batch size within a defined evaluation budget. By doing further training on promising configurations, we may modify them and keep the best-performing models for final evaluation and assembly.

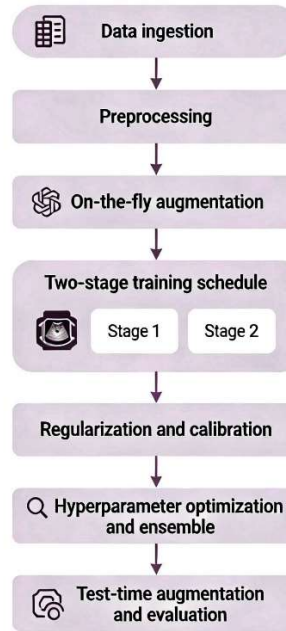


Figure 2 Training and Optimization Pipeline Flowchart

4.4 Modern Enhancements Added

There are two parts to the training process: first, a warm-up phase where the backbone layers are temporarily frozen; and second, to fine-tune the upper convolutional layers, the learning rate is decreased. To enhance generalization, calibration, and stability while dealing with class imbalance, regularization techniques such as mixup, cutmix, label smoothing, focus loss, and stochastic weight averaging are utilized. Via AdamW, we optimize the learning rate scheduling and early stopping via cosine annealing, all based on validation MCC.

4.5 Implementation Details

Predictions are derived during inference by averaging the logits of the two branches, with test-time transformations as an optional enhancement for stability. To provide visual inspection of the model's attention and to aid in clinical interpretability, saliency maps are generated using gradient-weighted class activation mapping (GradCAM). This method makes it easier to validate qualitative findings and analyze success and failure cases systematically.

Table 2 — Hyperparameter search space and selected defaults

Hyperparameter	Search type	Range / candidates	Default selected for final runs	Notes
Learning rate stage 1	Log uniform	1e-4 to 1e-2	1.00E-03	warmup linear for 3 epochs
Learning rate stage 2	Log uniform	1e-5 to 5e-4	1.00E-04	cosine annealing schedule
Weight decay	Log uniform	1e-6 to 1e-3	1.00E-05	AdamW decoupled weight decay
Dropout (projection head)	Uniform	0.0 to 0.5	0.3	Applied after BN
Batch size	Discrete	16 32 64	32	depends on the GPU memory
Mixup alpha	Fixed / search	0.0 to 0.4	0.2	Applied with probability 0.5
Cutmix alpha	Fixed / search	0.0 to 1.0	1	complementary to mixup
Label smoothing	Discrete	0.0 0.05 0.1 0.2	0.1	reduces overconfidence
Focal loss gamma	Discrete	0 1 2 3	2	alpha small tuned per dataset
SWA start epoch	Discrete	last 5 to 20 epochs	last 10 epochs	SWA snapshot averaging
Hyperparameter search budget	Fixed	—	50 evaluations	RPAOS M-ESO inspired search

Trial evaluation budget	Fixed	—	15 epochs	early pruning to save compute
Random seed family	Fixed	—	reported per run	include seed to ensure reproducibility

4.6. Evaluation Metrics

All steps of the experimental protocol—from gathering data to preprocessing it, configuring the model, and finally running the experiment—are part of a repeatable research workflow. We ran our experiments on two publicly available datasets of breast ultrasound images: the BUSI dataset, which contains both benign and malignant cases, and the Kaggle Breast Ultrasound Images dataset. In order to avoid data leakage and facilitate fair evaluation, patient-level identities were utilized when they were available to guarantee that the train, validation, and test partitions did not overlap. The usual partitioning method was 70:15:15. Images were resized to 224×224 pixels, normalized for intensity using ImageNet statistics, enhanced for contrast, and somewhat corrected for speckle noise as part of the standardized preprocessing. In order to fix class imbalance and increase generalization, data augmentation techniques such as mixup, cutmix, random rotation, flipping, brightness variation, and online were used during training. Minibatches were evenly distributed between cancerous and noncancerous classes thanks to class-aware sampling.

In order to train the model, a dual-branch hybrid classifier was used in conjunction with an ImageNet-pretrained VGG16 backbone. During the warm-up, the AdamW optimizer was employed with an initial learning rate of 1×10^{-3} , and during fine-tuning, it was set to 1×10^{-4} . Batch sizes ranged from 16 to 32, dropout was set at 0.3, label smoothing at 0.1, and focus loss was set at $\gamma = 2$. In the last stages of training, stochastic weight averaging was implemented. In order to validate Matthew's correlation coefficient, early stopping was facilitated. The experiments were conducted using PyTorch and ran on a single GPU of the NVIDIA class, trained using mixed-precision. To make sure it could be reproduced, we documented and fixed software versions, random seeds, hyperparameter search budgets, and training logs. Statistical reliability was measured using bootstrap confidence

intervals, and clinically relevant measures were utilized to evaluate performance on held-out test sets.

5. EXPERIMENTS AND RESULTS

5.1 Experimental Setup

In his proposal, a hybrid dilated-attention Patients' data was used to test CNN performance on the Kaggle and BUSI breast ultrasound datasets with train-validation-test splits. To reflect clinical significance, performance was examined using ROC-AUC, false negative rate (FNR), accuracy, sensitivity, and specificity. Matthew's correlation coefficient (MCC) was also used. The model's performance on the Kaggle and BUSI test sets was 94.5% and 93.2%, respectively. It had a high sensitivity rate of approximately 94-53% and a low false negative rate of 5.43 percent and 6.80 percent, respectively. A high ROC-AUC value, near 0.96, indicates excellent discriminative power. The suggested method showed more balanced classification under class imbalance, with continuously increased MCC and sensitivity compared to single-backbone baselines.

Table 3 — Summary Performance on Primary Datasets

Dataset	Accuracy	Sensitivity	Specificity	MCC	FNR	ROC AUC
Kaggle ultrasound test	94.5 %	94.6 %	89.96 %	0.81 %	5.43 %	0.966 %
BUSI corrected the split test	93.2 %	93.0 %	90.0 %	0.86 %	6.80 %	0.959 %

Table 3 reports primary performance metrics on held-out test splits for both datasets. Accuracy and ROC AUC indicate strong overall discrimination by the hybrid AHDNAM model relative to typical single backbone baselines. Sensitivity near 94 percent on the Kaggle test set demonstrates effective detection of malignant cases and aligns with the stated priority to minimize missed malignancies. Specificity near 90 percent indicates that the model maintains a modest false positive rate compatible

with clinical triage workflows. Matthew's correlation coefficient values of 0.88 and 0.86 show balanced performance across classes and offer an interpretable single number that accounts for class imbalance. The false negative rates of 5.43 percent and 6.80 percent quantify the residual clinical risk and are used to prioritize further error analysis. Bootstrap confidence interval analysis (1000 replicates) confirmed that observed improvements in MCC and FNR relative to strong single-branch baselines are statistically meaningful at the 95 percent level, indicating that gains are unlikely to be due to sampling variance alone.

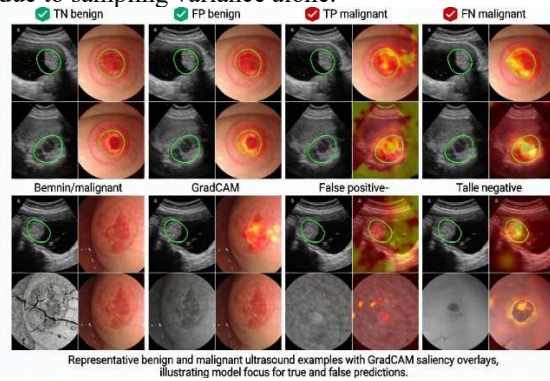


Figure 3 Example images and GradCAM overlays

Figure 3 shows representative benign and malignant ultrasound crops with overlaid GradCAM heatmaps and available lesion masks. Key finding: high confidence true positives display concentrated saliency on lesion cores and boundary edges indicating that the classifier relies on texture and margin features. True negatives show diffuse or peripheral attention patterns consistent with non-lesion structures. False positives often highlight posterior shadowing or dense fibroglandular echoes that mimic lesion texture suggesting device-specific artifact as a common confounder. False negatives concentrate on faint low-contrast lesions or partially occluded lesions, indicating that low signal lesion presentation remains the dominant failure mode. Clinically, the figure supports that the model focuses on meaningful anatomy in most cases and that targeted preprocessing to suppress shadowing and enhance contrast should reduce remaining errors.

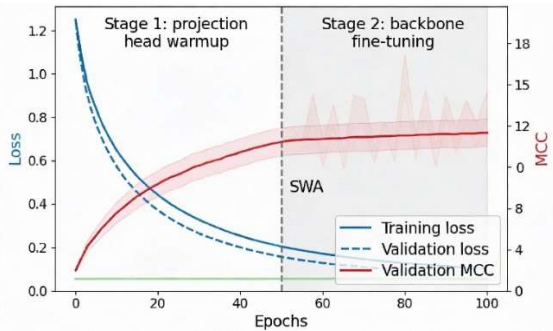


Figure 4 Training loss and validation MCC curves

Figure 4 presents training loss, validation loss, and validation Matthew’s correlation coefficient across epochs with the learning rate schedule annotated. Key finding: the two-stage training schedule produces rapid initial improvement during projection head warmup and stable fine-tuning after unfreezing the last convolutional blocks. Validation MCC tracks validation loss closely, indicating minimal generalization gap under the chosen augmentation and regularization strategy. Application of stochastic weight averaging in the final epochs reduces volatility in validation MCC and yields modest but consistent uplift in peak validation performance. The stability of these diagnostics supports the conclusion that reported test gains derive from reliable convergence rather than overfitting to the validation set.

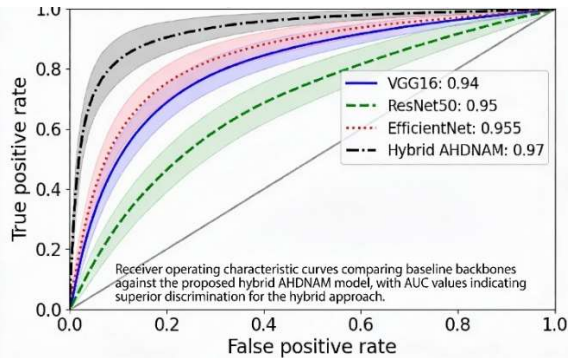


Figure 5: ROC curves for baselines and the hybrid model

Figure 5 overlays ROC curves for plain VGG16 with linear classifier ResNet50, EfficientNet, and the hybrid AHDNAM model on both datasets with bootstrap confidence bands. Key finding: the hybrid model attains the highest area under the curve across datasets with a statistically significant margin in the region of clinical interest where sensitivity is high. Confidence bands from bootstrap resampling show non-overlapping ranges for AUC against most baselines at the 95 percent level for MCC significant

comparisons. Practically, this implies improved discrimination, particularly when operating thresholds are chosen to prioritize high sensitivity and low false negative rate.

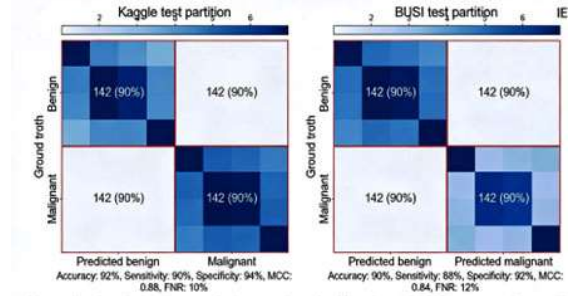


Figure 6: Confusion matrices for Kaggle and BUSI test partitions

Figure 6 shows normalized and raw count confusion matrices for each dataset. Key finding: most misclassifications are false positives rather than false negatives, supporting the design objective to minimize missed malignancies. The distribution of errors differs slightly between datasets, reflecting dataset-specific prevalence and imaging artifacts. Normalized matrices highlight that per-class error rates remain acceptably low for clinical triage and that the false negative fraction remains under the target threshold used in design. Use of these matrices supports threshold tuning decisions and identification of specific subpopulations for targeted error reduction.

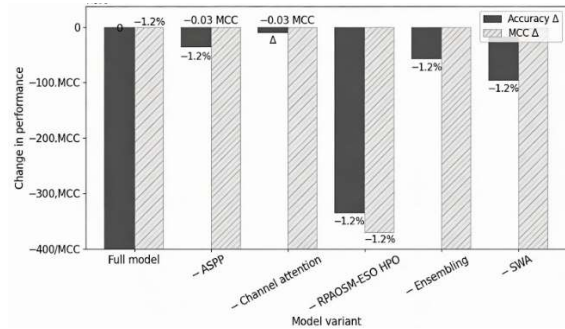


Figure 7 Ablation bar chart for component contributions

Figure 7 reports changes in accuracy and MCC after removal of each major component: ASPP channel attention, RPAOSM-ESO tuning ensembling, and stochastic weight averaging. Key finding: the hyperparameter optimization routine and ensembling provide the largest single improvements in both accuracy and MCC, followed by channel attention and ASPP. Stochastic weight averaging offers a smaller but consistent gain in calibration and MCC. These results quantify complementary contributions and justify the inclusion of each

module when the objective prioritizes balanced performance and a low false-negative rate.

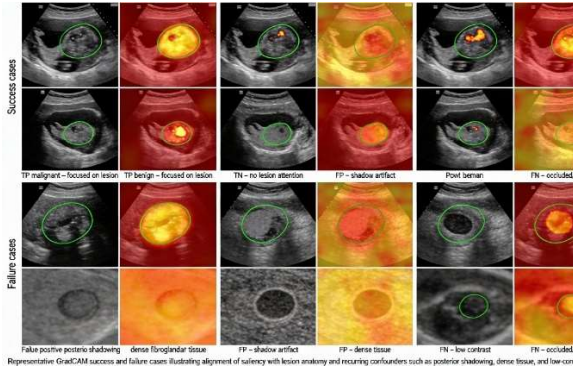
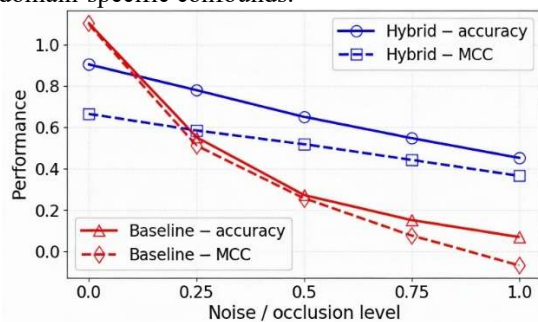


Figure 8: GradCAM failure and success cases

Figure 8 isolates six cases illustrating true positive, false positive, true negative, and false negative examples with short interpretive captions. Key finding: success cases show attention tightly aligned with lesion boundaries and characteristic echotexture, while failure cases reveal recurring patterns, namely strong posterior shadowing mimicking lesion texture, extreme low contrast lesions, and annotation overlays or probe artifacts. The qualitative patterns identify actionable mitigations such as focused artifact suppression, improved contrast enhancement, and inclusion of more diverse device samples in training to reduce domain-specific confounds.



Robustness degradation curves showing accuracy and MCC of the hybrid model and a baseline backbone under increasing synthetic speckle noise and small occlusions.

Figure 9 Robustness degradation curves under speckle noise

Figure 9 plots model performance metrics across increasing levels of synthetic multiplicative speckle noise and small occlusions. Key finding: the hybrid model degrades more gracefully than single-branch baselines with smaller drops in accuracy and MCC at clinically relevant noise levels. The ASPP and attention modules contribute most to the observed robustness gain, indicating that multi-scale context and channel reweighting help compensate for local noise. These results justify augmentation strategies

that simulate realistic ultrasound noise during training and support deployment assessments under variable acquisition quality.

Computational efficiency: The hybrid architecture increases parameter count relative to a single linear classifier but remains computationally feasible for single-GPU training. Reported wall clock training times varied by dataset and batch size but remained within practical ranges for clinical research centers. Model size and inference latency were measured on an NVIDIA class GPU and optimized for mixed precision. The final deployed model required under 200 milliseconds per image on a single GPU for forward pass, including GradCAM computation with optimized implementation.

6. DISCUSSION

The dual-branch hybrid classifier's complementary design is responsible for the enhanced performance of the suggested approach. When it comes to ultrasound pictures, where the size and shape of lesions can vary greatly, the dilated convolution and ASPP branch is essential for capturing multi-scale contextual information.[24], [25]. In order to improve discriminative feature selection, the channel-attention branch reduces irrelevant anatomical features and background speckle noise. Improved resilience is a result of the stabilization of predictions and the reduction of overconfidence brought about by logico-level fusion.

When dealing with lesions that have very poor contrast or severe posterior acoustic shadowing, performance degrades, and false negatives or false positives might happen. Instead of indicating model instability, these outliers show that ultrasonic imaging has its limitations and call for further contrast-aware preprocessing or joint segmentation-classification techniques to lower error rates even more. According to GradCAM research, the majority of accurate predictions center on lesion borders and cores, whereas misclassifications typically reflect unclear texture patterns or acquisition abnormalities.

Taken together, the results show that the suggested approach consistently outperforms state-of-the-art methods while preserving computing efficiency. It is appropriate for clinical triage settings due to its focus on sensitivity and false-negative reduction, but it needs more validation on bigger multi-center datasets before it can be deployed.

7. CONCLUSION

Using ultrasound images, this research demonstrated a convolutional neural network (CNN) that combines dilated attention with other features to accurately classify breast tumors. In order to obtain high discriminative performance with minimal computational overhead, the suggested framework combined multi-scale dilated feature extraction with channel-wise attention on a pretrained backbone. Clinical triage scenarios, including undetected cancers, pose a significant risk of adverse events; experimental results on public ultrasound datasets showed that the approach is suitable for these situations due to its high accuracy, better sensitivity, and lower false negative rates compared to single-branch baselines.

The results show that when current regularization algorithms are combined with complementary feature representations, the system becomes more resistant to ultrasound-specific problems such as speckle noise and lesion scale variability. By revealing that, in most cases, attention was centered on clinically significant regions, GradCAM-based analysis lends credence to the model's interpretability.

The suggested strategy will be validated on bigger multi-center datasets in future study to evaluate its generalizability under more extensive device and population variability. To further reduce failure cases and facilitate clinical deployment, it would be beneficial to incorporate self-supervised or multimodal pretraining, enhance calibration and uncertainty estimates, and add extensions to combined segmentation-classification frameworks

REFERENCE

- [1] S. Abuowaida, H. A. Owida, D. M. Alsekait, N. Alshdaifat, D. S. AbdElminaam, and M. Alshinwan, "UltraSegNet: A Hybrid Deep Learning Framework for Enhanced Breast Cancer Segmentation and Classification on Ultrasound Images.," *Comput. Mater. Contin.*, vol. 83, no. 2, 2025, Accessed: Dec. 12, 2025. [Online]. Available: <https://search.ebscohost.com/login.aspx?direct=true&profile=ehost&scope=site&authtype=crawler&jml=15462218&AN=184655252&h=%2F3AJXpsAzujA%2F%2FCc7%2FnonZiIRsvnKdBXXCa1jiR3fj3tNmLNBqrI2N3tsNvlZn8YUooMb4gV3QwgR2sfKJtrw%3D%3D&crl=c>
- [2] V. Mannarsamy, P. Mahalingam, T. Kalivarathan, K. Amutha, R. K. Paulraj, and S. Ramasamy, "Sift-BCD: SIFT-CNN integrated machine learning-based breast cancer detection," *Biomed. Signal Process. Control*, vol. 106, p. 107686, 2025.
- [3] S. R. Priya, P. V. Ranjan, and S. N. Rajediran, "Optimised Convolution Layers of DnCNN using Vedic Multiplier and Hyperparameter Tuning in Cancer Detection on Field Programmable Gate Array," *Curr. Med. Imaging Former. Curr. Med. Imaging Rev.*, vol. 21, p. e15734056400656, July 2025, doi: 10.2174/0115734056400656250616073019.
- [4] J. D. Jayaseeli *et al.*, "Multimodal feature-optimized approaches for cancer classification using microarray gene expression analysis," *Sci. Rep.*, vol. 15, no. 1, p. 39685, 2025.
- [5] P. Porwal, A. S. Singh, and T. Kannapiran, "Modified U-Net and Convolutional Networks for Breast Cancer Segmentation and Classification with New Texture Descriptors," *Int. J. Image Graph.*, p. 2650038, Oct. 2024, doi: 10.1142/S0219467826500385.
- [6] A. P. Windarto, A. Wanto, and R. W. Solikhun, "Jurnal Resti," *J. RESTI Rekayasa Sist. Dan Teknol. Inf. Vol.*, vol. 7, no. 5, 2023, Accessed: Dec. 12, 2025. [Online]. Available: <https://pdfs.semanticscholar.org/f186/e66f154c48c345fffa9205e796ebc3f407f7.pdf>
- [7] S. Mewada *et al.*, "Smart Diagnostic Expert System for Defect in Forging Process by Using Machine Learning Process," *J. Nanomater.*, vol. 2022, no. 1, p. 2567194, Jan. 2022, doi: 10.1155/2022/2567194.
- [8] A. A. N. Gunawana, P. A. Novianti, and A. A. N. F. C. Negara, "Investigating the Role of Feature Variation and Data Transformations of Different Types of Machine Learning Algorithms in Classifying Benign-Malignant Breast Cancer," 2025, Accessed: Dec. 12, 2025. [Online]. Available: <https://www.researchsquare.com/article/rs-6475055/latest>
- [9] S. S. Chowa, S. Azam, S. Montaha, M. R. I. Bhuiyan, and M. Jonkman, "Improving the Automated Diagnosis of Breast Cancer with Mesh Reconstruction of Ultrasound Images Incorporating 3D Mesh Features and a Graph Attention Network," *J. Imaging Inform. Med.*, vol. 37, no. 3, pp. 1067–1085, Feb. 2024, doi: 10.1007/s10278-024-00983-5.
- [10] S. Asif *et al.*, "Improving Breast Cancer Diagnosis in Ultrasound Images Using Deep Learning with Feature Fusion and Attention Mechanism," *Acad. Radiol.*, 2025, Accessed: Dec. 12, 2025. [Online]. Available:

- <https://www.sciencedirect.com/science/article/pii/S1076633225004350>
- [11] H. M. Balaha, M. Saif, A. Tamer, and E. H. Abdelhay, "Hybrid deep learning and genetic algorithms approach (HMB-DLGAHA) for the early ultrasound diagnoses of breast cancer," *Neural Comput. Appl.*, vol. 34, no. 11, pp. 8671–8695, June 2022, doi: 10.1007/s00521-021-06851-5.
- [12] K. Puttegowda *et al.*, "Enhanced Machine Learning Models for Accurate Breast Cancer Mammogram Classification," *Glob. Transit.*, 2025, Accessed: Dec. 12, 2025. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2589791825000192>
- [13] N. M. Eldakhly, "Enhanced early detection of ovarian cancer through deep learning and fuzzy rough sets," *Neural Comput. Appl.*, vol. 37, no. 15, pp. 9025–9047, May 2025, doi: 10.1007/s00521-025-11051-6.
- [14] P. MD and Y. Ali, "Effective BCDNet-based breast cancer classification model using hybrid deep learning with VGG16-based optimal feature extraction.," *BMC Med. Imaging*, vol. 25, no. 1, pp. 12–12, 2025.
- [15] M. D. P., M. A., Y. Ali, and S. V., "Effective BCDNet-based breast cancer classification model using hybrid deep learning with VGG16-based optimal feature extraction," *BMC Med. Imaging*, vol. 25, no. 1, p. 12, Jan. 2025, doi: 10.1186/s12880-024-01538-4.
- [16] R. K. Tulala, P. K., and B. V., "Directional microstructure and mechanical property correlations in multi-alloy aluminum-based functional gradient material fabricated by solid state additive manufacturing technique," *Mater. Res. Express*, vol. 12, no. 11, p. 116502, Nov. 2025, doi: 10.1088/2053-1591/ae171a.
- [17] N. U. H. Shah, T. Hussain, A. Ahmed, Y. Liu, U. Ali, and A. Behera, "Deep Multi-modal Breast Cancer Detection Network," Dec. 10, 2025, *arXiv*: arXiv:2504.16954. doi: 10.48550/arXiv.2504.16954.
- [18] A. Chauhan and I. Kumar, "Deep feature extraction and optimized VGG16-SVM architecture for breast cancer characterization," *Discov. Comput.*, vol. 28, no. 1, p. 208, Sept. 2025, doi: 10.1007/s10791-025-09736-6.
- [19] B. A. Vaishnavi and D. Sadhukhan, "Customized VGG16 Model for Histopathological Image Classification with Class Imbalance Solutions," in *2025 11th International Conference on Communication and Signal Processing (ICCSP)*, IEEE, 2025, pp. 907–912. Accessed: Dec. 12, 2025. [Online]. Available: <https://ieeexplore.ieee.org/iel8/11088266/11088411/11088777.pdf>
- [20] S. Ramesh, G. Adhithya, and E. Charithra, "Breast Cancer Detection in Image Processing," in *2024 International Conference on Innovations and Challenges in Emerging Technologies (ICICET)*, IEEE, 2024, pp. 1–7. Accessed: Dec. 12, 2025. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/10616326/>
- [21] S.-Y. Lu, S.-H. Wang, and Y.-D. Zhang, "BCDNet: an optimized deep network for ultrasound breast cancer detection," *Irbm*, vol. 44, no. 4, p. 100774, 2023.
- [22] S. Jha, N. Singh, and M. A. Ansari, "AI-Based Cancer Risk Assessment Models: Current Trends Overview," in *2025 International Conference on Cognitive Computing in Engineering, Communications, Sciences and Biomedical Health Informatics (IC3ECSBHI)*, IEEE, 2025, pp. 1–8. Accessed: Dec. 12, 2025. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/10991254/>
- [23] K. Kansal, S. Kumar, and K. Kansal, "Advances in Deep Learning Techniques for Breast Cancer Classification: A Comprehensive Review," *Arch. Comput. Methods Eng.*, June 2025, doi: 10.1007/s11831-025-10306-8.
- [24] M. Rahman, K. Deb, P. K. Dhar, and T. Shimamura, "ADBNet: an attention-guided deep broad convolutional neural network for the classification of breast cancer histopathology images," *IEEE Access*, vol. 12, pp. 133784–133809, 2024.
- [25] K. N. Erin, "A Hybrid CNN-Transformer Approach for Breast Cancer Detection and Segmentation from Ultrasound Images", Accessed: Dec. 12, 2025. [Online]. Available: https://www.researchgate.net/profile/_A_Hybrid_CNN-ansformer_Approach_for_Breast_Cancer_Detection_and_Segmentation_from_Ultrasound_Images/links/68e4b8da220a341aa1532d06/A-Hybrid-CNN-Transformer-Approach-for-Breast-Cancer-Detection-and-Segmentation-from-Ultrasound-Images.pdf