Multi-Modal Attention-Based Deep Learning Architecture for Breast Cancer Stratification with Uncertainty Estimation

Saran Rohit D V,
Department of Computational Intelligence,
SRM Institute of Science and Technology,
Chennai, India
sr4876@srmist.edu.in

A. Alice Nithya
Department of Computational Intelligence,
SRM Institute of Science and Technology,
Chennai, India
alicenia@srmist.edu.in

Abstract— **The universal number of women diagnosed with breast cancer is estimated to be around three lakhs in 2026, of which 16% of women are less than 50 years at the time of diagnosis. The rate of survival is generally influenced by the time stages of the disease, age of the patient, genetic influence and diagnostic accuracy. Though imaging is one of the parameters, the confirmation of diagnosis is generally done by histopathological examination (Biopsy) and management based on Immunohistochemistry (IHC) only. Multi-modality system is preferred over Single-modality systems, as the later has the disadvantage of modality-specific noise coupled with low sensitivity. In this paper, a multimodal deep learning framework integrating digital mammography with Ultrasound images is developed by applying dual stream convolutional neural networks complemented with cross model attention fusion and Mote Carlo (MC) dropout. The dataset consists of both mammogram and ultrasound images, and they have been processed by two algorithmic models which are Modified Relation and Margin-Based Deep Learning Network (MReMarNet) and the Multi modal attention model. The proposed method aims to improve accuracy while maintaining reliability of two different datasets, Mini-DDSM and BUSI. Upon training, the results of the Multi modal attention model have a clear advantage over the MReMarNet model, thereby showing better results with rates of 99.00%, 98.00% and 96.00%, 94.00% for the BUSI and Mini-DDSM datasets respectively.**

Keywords—Breast cancer, Mammography, Ultrasound, BUSI, Mini-DDSM, data sets, Convolutional Block Attention Module

I. INTRODUCTION

Breast cancer is becoming a life-threatening non-communicable disease affecting women universally. The stage of breast cancer at the time of diagnosis at which, plays a significant role in improving the survival rate and reducing the mortality rate. Mammogram and Ultrasound imaging techniques are part of the investigations available in the early diagnosis of breast cancer. But manual diagnosis of the images can vary due to image quality, which may lead to errors and misdiagnosis [1]. In [2], Ahmed focuses on mammographic diagnosis using advanced pretrained MLP-Mixer models which are useful for Mini-DDSM based classification benchmarking.

Huang, J. et al. [1] in their research have used a deep learning method based on segmentation of ultrasound images for breast cancer and Strong BUSI-related in contrast to our multimodal application. They have reported attention mechanisms highly relevant for BUSI diagnosis pipelines. Similarly, Shaukat et al.,[3] have used a complementary deep with traditional fusion, with ensemble classifiers for breast cancer detection with digital mammogram and ultrasound images.. Recently, Convolutional Neural Network models, which are related to Deep Learning, have demonstrated considerable ability in the detection of tumors and medical image analysis [4]. Hou J. et al [5] have used deep learning with powerful feature extraction and representation capabilities for modelling complex non-linear relationships. In another study by Sreelakshmi et al [6], have introduced a deep learning model in which digital imaging complements effectively in assessing clinical risk factors, achieving high accuracy(98%) on the DDSM data set . Dependence on feature extraction can be minimized manually by integrating CNN models in contrast to single modal data that has the disadvantage of limiting the ability to obtain significant information.

The fast-paced deep learning technique in artificial intelligence has been a major influence in complementing the development of additional mechanisms reported by Ahmed et. al., in [7]. Furthermore, uncertainty estimation techniques, such as the Monte Carlo Dropout, have been introduced to enhance prediction reliability. Nasser et al. [8] have used models based on deep learning in the diagnosis of breast cancer. In this work, they have introduced a multimodal attention-based deep learning framework for breast cancer detection, which involves dual-stream CNN architectures that have been provided.

Lee et. al., [9] in their study on similar topic reported that Digital Mammography is effective for detecting microcalcifications and structural distortions, whereas in patients with dense breast ultrasound is complementary in distinguishing benign from malignant diseases, though Biopsy is the final decision maker. Matharasi et al., [10] insists that despite rapid strides achieved with computer in the design, and processing, single-modality systems have the disadvantage of overlooking these complementary diagnostic aids.

Recent developments in attention mechanisms and uncertainty-aware learning have opened new possibilities for robust medical AI systems. In [11], Guo et al., introduces a multi-modal architecture, that extracted modality-specific features using dual CNN backbones that enhanced spatial and channel-level representation via CBAM. This also used cross-modal relationships applying a dynamic multi-modal attention fusion mechanism, Monte Carlo Dropout estimated and predicted uncertainty. [12]. The advantage of the Monte Carlo dropout was to generate random predictions with interpretations of them as samples from a distribution probability. The dropout layers were kept active during testing, by creating a Bayesian approximation without retraining. It facilitates these models to provide confidence intervals for predictions by performing multiple stochastic forward passes for a single input to generate an output distribution. The novelty lied in the integration of cross-modal attention and uncertainty estimation within a unified framework trained on both BUSI and Mini-DDSM datasets. Yan et. al., [13] in their retrospective analysis of similar study on 663 patients had concluded that this method approach complimented radiologists in breast lesion classification accurately and significantly contributed to clinician in arriving a decision, and improving consistency in diagnosis. Lekamlage et al., [14] stated that characteristic features based on CAD systems, Classical classifiers contrary to Deep CNNs such as ResNet and DenseNet have demonstrated better performance in mammography and ultrasound classification independently. Classification of breast cancer using uncertainty-sensitive cross-modal fusion architecture has been widely applied in the interpretation of breast imaging with histopathology and clinical data. The author had applied deep learning methods in determining the patients' age from the Mammogram images in identifying missing data in their study. From this model, they were able to extract age attributes which had been found useful in training and testing AI models in breast cancer detection.

Farooq et al., [15] has introduced a new model in their research, incorporating clinical materials available on cognitive domains weights with uncertainty attention mechanisms that combines confidence estimates in a robust manner. Similarly, Atrey et al., [16] demonstrated positive impact of clubbing deep learning and traditional machine learning using mammogram and ultrasound images in cancer breast diagnosis. In their study on breast lesion classification using imaging and deep convolutional neural network, Alaa Al Zoubi et al., [17] concluded that CNN models are on par with radiologist-opinion on ultrasound images in the interpretation.

Sahu and Maher [18] in his study have proposed that, modified Relation and Margin Network (MReMarNet) was presented for efficient breast cancer detection and outperformed other networks with the accuracy and coupled benefits of intra-class compactness provided by RU and inter-class separability provided by the FC branch make the system more efficient. Sannasi et al. [19] have utilized ResNet-18. They employed deep features extraction using a ResNet-18 architecture coupled with Improved Crow-Search Optimized Extreme Learning Machine (ICS-ELM) based algorithm. The findings have demonstrated that the computer-based design system is found to be very effective and for the detection in addition to the classification of breast cancer.

II. Proposed Methodology

This is an observational study to compare between the MReMarNet and the Multi-Modal Attention method which is used for the early diagnosis of breast cancer. The following methodologies are used for processing the datasets which are illustrated, as follows: Dual-Stream CNN Backbone, Convolutional Block Attention Module, Multi-Modal Attention Fusion and Monte Carlo Dropout for Uncertainty Estimation

A. Dual-Stream CNN Backbone

This algorithm consists of two parallel CNN branches which process mammogram and ultrasound images separately. Each branch learns about modality-specific features accordingly, since using a single stream CNN may confuse texture patterns and may lead to weaker representations. As a result, this Dual-Stream CNN Backbone architecture is based on ResNet18 which is implemented for a multi-modal feature extraction.

The feature maps for the multimodal inputs are computed using the convoluted equations (1) and (2) as follows,

$$F_m = C_N\,(N_m\,(X_m)) \tag{1}$$

$$F_u = C_N\,(N_u\,(X_u)) \tag{2}$$

Where X_m and X_u denote the input for the mammogram and ultrasound images respectively, and $N_m\,(\cdot)$ and $N_u\,(\cdot)$ represent the feature extraction networks and $C_N\,(\cdot)$ denotes the convolutional operation applied to the features. The outputs F_m and F_u correspond to the mammogram and ultrasound modality.

B. Convolutional Block Attention Module

An attention mechanism that helps a CNN focus on the target and important locations, known as Channel attention and Spatial attention. It does not treat all regions equally and suppresses irrelevant ones while optimising important features.

The Convolutional Block Attention Module is obtained by utilizing sequential application of channel and spatial attention mechanisms sequentially given in equations (3) and (4)

$$F' = M_c\,(F) \otimes F \tag{3}$$

$$F'' = M_s\,(F') \otimes F' \tag{4}$$

In the above equations, F denotes the input feature map, $M_c\,(\cdot)$ denotes the channel attention map, $M_s\,(\cdot)$ denotes the spatial attention map. The operator $\otimes$ denotes the element-wise multiplication. The intermediate feature map F' is obtained by applying channel attention whereas the final refined feature map F'' is obtained after spatial attention.

.

C. Multi-Modal Attention Fusion

A cross-modal attention mechanism has been implemented to dynamically integrate mammographic, and ultrasound features as shown in Fig. 1. Additionally, the proposed fusion module models the inter-modality dependencies allowing adaptive feature weighing based on diagnostic cues. With this mechanism, the Mammogram stream extracts structural features that include mass shapes, margins and calcifications whereas the Ultrasound stream extracts the texture and density features including internal echo patterns. The multimodal features are fused with a weighted combination as shown in equation (5),

$$F_{fused} = \alpha \cdot F_m + \beta \cdot F_u \tag{5}$$

Where F_m and F_u denote the Mammogram feature representation and the Ultrasound feature representation, α and β are the attention weights for the mammogram and ultrasound maps respectively. The fused feature map integrates complementary information from both sources for improved representation.

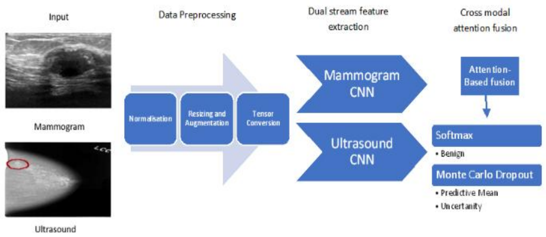

Fig. 1. Multi-Modal Attention Model based architecture

D. *Monte Carlo Dropout for Uncertainty Estimation*

In normal dropout methods, neurons are normally turned off during training. This prevents overfitting. On the other hand, Monte Carlo dropout is enabled during inference to estimate the uncertainty. By performing multiple random probabilistic passes with dropout enabled, the mean and variance are computed resulting in classification probability and uncertainty estimation. This is a strong addition to Medical AI. In the view of Ultrasound and Mammogram, the input is run multiple times, and different neurons are dropped each time. The predictions are taken as samples from a distribution. With the variance being low, the model predictions are stable, and the model predictions will be unsure with a high variance.
The predictive mean in equation (6) is anaysed as follows,

$$\mu = \frac{1}{T}\sum_{t=1}^{T}\hat{y}^{(t)} \qquad (6)$$

Where μ denotes the mean prediction, T is the Total number of forward passes and $\hat{y}^{(t)}$ indicates represents the prediction from the t[th] run
As for the variance, the uncertainty in equation (7) is calculated as follows,

$$\sigma^2 = \frac{1}{T}\sum_{t=1}^{T}\left(\hat{y}^{(t)} - \mu\right)^2 \qquad (7)$$

Where σ^2 indicates the variance, T is the Total number of forward passes, and $\hat{y}^{(t)} - \mu$ indicates the difference of each prediction from the mean.
A higher variance indicates greater uncertainty in the model's predictions.

The proposed multimodal framework was trained using mammogram and ultrasound images corresponding to the non-malignant and cancer classes. Both the modalities were processed simultaneously during each forward pass to enable joint feature learning and cross-modal interaction within the dual-stream architecture. Cross entropy loss based strategic model optimization, an intentional function for binary divisions, the model parameters were optimized using the Adam (Adaptive Moment Estimation) optimizer. As Adam combines the advantages of both momentum and adaptive learning rate methods, it was chosen owing to faster convergence, stability in multimodal architecture, and robustness to sparse gradients.

The training was conducted using gradient descent of mini-batch where a batch-size is defined in the implementation, as the mini-batch training offers an improved gradient stability, efficient GPU utilization, and a faster convergence compared to full batch training.

The training was performed using the GPUs from the Google Cloud platform. Leveraging the parallel processing capabilities has reduced the training time drastically and has enabled the efficient handling of dual-stream feature extraction, attention based fusion operations and Monte Carlo dropout sampling. For this implementation, the mammogram and the ultrasound images are passed through two separate CNN pathways, the features are extracted through the attention based fusion module. Furthermore, the fused representation was fed into the final classification layer and cross entropy loss was measured against the ground-truth level. Finally, the backpropagation has updated all the learnable parameters jointly.

Therefore, this simultaneous operation emphasizes joint feature learning, cross-modal dependency modelling and end-to-end training stability. To prevent overfitting, the dropout layers are incorporated within the fully connected layers. During training, the neurons are deactivated in a predetermined manner and promote robust feature learning and prevent co-adaptation of neurons. Sensitivity is emphasized due to its clinical importance in cancer detection.

IV. Result and Discussion

A. *Dataset*

The dataset consists of the ultrasound images of women ranging between 25 and 75 years of age taken from the BUSI (Breast Ultra-sound Imaging) dataset, which were collected around 2018. The dataset consists of around 780 images which are classified into three parts: Benign, Cancer and Normal, where the average image size is of 500 x 500 pixels. The BUSI dataset has been primarily used to train models to detect cancer lesions in breast ultrasound images.

On the other hand, the Mini-DDSM (Digital Database for Screening Mammograms) is a smaller, lightweight version of the larger DDSM dataset. This dataset consists of the mammographic images of women between 30 to 90 years of age. Owing to the large size in the normal DDSM dataset, the Mini version is used for training the algorithms. The Mini-DDSM dataset is also split into different parts: Benign, Cancer and Normal. The images are compressed with generated lossless JPEG encoding. A thorough evaluation applying this, considering various performance indicators, covering accuracy, sensitivity, and specificity. Through these parameter evaluations, the model's capability to assess and stratify mammography images and clinical diagnosis breast cancer is confirmed. The filtered subset of DDSM, known as

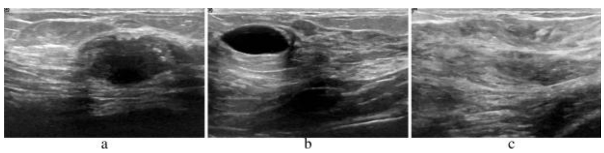

Fig. 2. Ultrasound images of each class of (a) Malignant, (b) Benign and (c) Normal from the BUSI dataset.

the Mini-DDSM Dataset with digital mammography images classified into Benign, Malignant or NAD (No abnormal diagnosis) using BI-RADS. (Figs. 2 & 3).

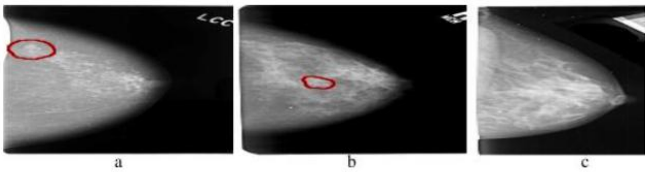

Fig. 2. Mammography images of each class of (a) Malignant, (b) Benign and (c) Normal from the mini-DDSM dataset.

B. System Configuration

To ensure the accurate execution of all results, a Google Colab notebook was used with a T4 cloud GPU to process the datasets. This system is designed to carry out experiments with the greatest accuracy and precision. The training set includes the dataset images.

C. Comparitive analysis with the MReMarNet method

The following comparisons have been made between the MReMarNet [19] and the Multi-Modal Attention method.

The attention method focuses on tumour regions rather than the background tissues which leads to improved detection of microcalcifications, masses and irregular boundaries. Furthermore, it learns complimentary information. In addition, the Monte Carlo method improves robustness, generalization and uncertainty estimation which comes in useful for AI in medical diagnosis.

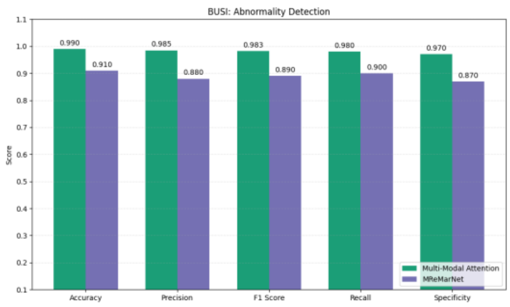

Fig. 4 Comparison with MReMarNet and Multi-Modal Attention Model on BUSI Dataset for detecting Abnormality

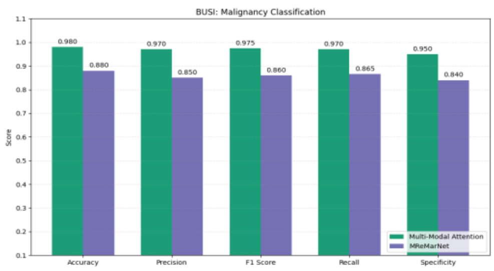

Fig. 5 Comparison with MReMarNet and Multi-Modal Attention Model on BUSI Dataset for Malignancy Classification

In the end, the Multi-modal attention method is the better option as it incorporates attention mechanisms, multimodal feature fusion and Monte Carlo dropout, which improves feature learning and reduce classification errors. As a result, it achieves a higher score on accuracy, precision on the BUSI and the Mini-DDSM datasets.

Since carcinoma of the breast is one of the leading causes of cancer mortality among women around the world, Mammography is widely used for screening; however, dense breast tissue can obscure lesions [20]. Ultrasound imaging complements mammography by providing better visualization of soft tissue structures. Radiologists routinely integrate information from both modalities for improved diagnostic confidence [21]. Recent deep learning systems have demonstrated strong performance in image-based classification, yet most are unimodal and lack uncertainty estimation [22].

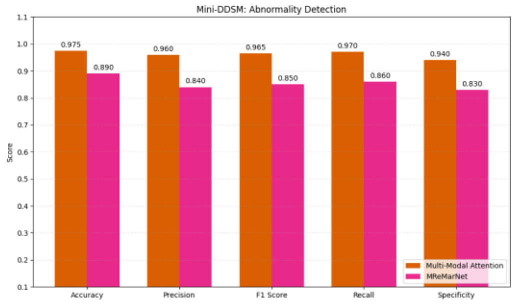

Fig. 6 Comparison with MReMarNet and Multi-Modal Attention Model on Mini-DDSM Dataset for detecting Abnormality

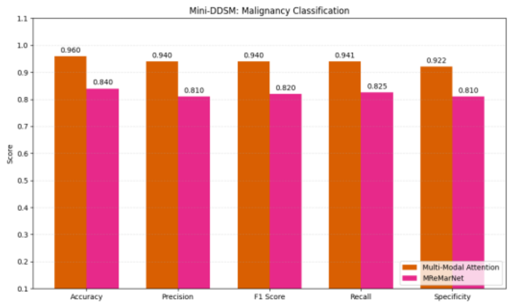

Fig. 7 Comparison with MReMarNet and Multi-Modal Attention Model on Mini-DDSM Dataset for Malignancy Classification

TABLE I. PERFORMANCE ANALYSIS COMPARISON BETWEEN MREMARNET AND MULTI-MODAL ATTENTION FRAMEWORK METHODS TO IDENTIFY ABNORMALITY USING THE BUSI DATASETS.

Metric	Multi-Modal (%)	MReMarNet (%)
Accuracy	99.00	91.00
Precision	98.50	88.00
F1-Score	0.9825	0.8900
Recall	98.00	90.00
Specificity	97.00	87.00

TABLE II. PERFORMANCE ANALYSIS COMPARISON BETWEEN MREMARNET AND MULTI-MODAL ATTENTION FRAMEWORK METHODS TO CLASSIFY MALIGNANCY USING THE BUSI DATASETS.

Metric	**Multi-Modal (%)**	**MReMarNet (%)**
Accuracy	98.00	88.00
Precision	97.00	85.00
F1-Score	0.9751	0.8600
Recall	97.02	86.50
Specificity	95.04	84.00

TABLE III. PERFORMANCE ANALYSIS COMPARISON BETWEEN MREMARNET AND MULTI-MODAL ATTENTION FRAMEWORK METHODS TO IDENTIFY ABNORMALITY USING THE MINI- DDSM DATASETS.

Metrics	**Multi-Modal (%)**	**MreMarNet (%)**
Accuracy	97.50	89.00
Precision	96.00	84.00
F1-Score	0.9650	0.8500
Recall	97.00	86.00
Specificity	94.00	83.00

TABLE IV. PERFORMANCE ANALYSIS COMPARISON BETWEEN MREMARNET AND MULTI-MODAL ATTENTION FRAMEWORK METHODS TO CLASSIFY MALIGNANCY USING THE MINI- DDSM DATASETS.

Metrics	**Multi-Modal (%)**	**MReMarNet (%)**
Accuracy	96.00	84.00
Precision	94.00	81.00
F1-Score	0.9404	0.8200
Recall	94.08	82.50
Specificity	92.16	81.00

Clinical deployment demands, Multimodal feature integration, Attention-guided lesion localization, Reliable confidence estimation. This work introduces a multimodal attention-driven framework that integrates mammography and ultrasound imaging while quantifying model uncertainty. Mammography and ultrasound provide complementary diagnostic signals. Attention mechanisms improve representation learning. Uncertainty estimation enhances model reliability. The integration of CBAM and cross-modal attention improved feature discrimination. MC Dropout enabled uncertainty-aware predictions. The Monte Carlo Dropout has successfully identified high-uncertainty cases, particularly in borderline benign-malignant lesions. The Attention heatmaps demonstrated improved lesion localization, enhancing interpretability. Key Observations include Improved classification stability. Reduced false negatives, better generalization across modalities., a meaningful uncertainty differentiation between confident and ambiguous cases. The multimodal attention framework has achieved a higher accuracy than unimodal baselines which has improved sensitivity for malignant lesions and reduced false negatives.

In comparison with the MReMarNet (Modified Relation and Margin Network) which employs static feature fusion and deterministic prediction, the combined methods of the CBAM, Cross modal attention and MC Dropout have a slightly slower inference speed but the interpretability is better with attention maps and uncertainty. Overall, it is dynamic, adaptive and has led to clinically safe decision making. This enhances inter-modality dependency modelling and improves clinical reliability through probabilistic inference.

To further enhance discriminative feature learning within each modality stream, an attention refinement mechanism based on the Convolutional Block Attention Module (CBAM) is incorporated in our the proposed dual-stream architecture. CBAM enables the network to emphasize diagnostically relevant regions while suppressing redundant background information. This procedure has the best potential of deep learning-based approaches in USG of breast based on computer-aided diagnostic systems, resulting in a reliable, speedy, and decision complementary tool for early diagnosis. Standard convolutional layers treat all spatial and channel information equally. However, lesion-related patterns occupy limited and specific regions of the image. Therefore, attention refinement is necessary to guide the network toward clinically meaningful features.

V. Conclusion

In this work, it is illustrated that the combined features of CBAM, Cross modal attention and MC Dropout show a better and a marginal improvement over the Modified Relation and Margin Network in the detection of tumours present in the mammogram and the ultrasound datasets where the biggest advantage is uncertainty modelling. It is hoped that this will help radiologists make helpful decisions by using these methods.

VI. References

[1] J. Huang et al., "EMGANet: Edge-Aware Multi-Scale Group-Mix Attention Network for Breast Cancer Ultrasound Image Segmentation," in IEEE Journal of Biomedical and Health Informatics, vol. 29, no. 8, pp. 5631-5641, Aug. 2025, doi: 10.1109/JBHI.2025.3546345.

[2] H.O. A. Ahmed and A. K. Nandi, "High Performance Breast Cancer Diagnosis From Mammograms Using Mixture of Experts With EfficientNet Features (MoEffNet)," in IEEE Access, vol. 12, pp. 133703-133725, 2024, doi: 10.1109/ACCESS.2024.3461360.

[3] Shaukat, A., Farhan, S., Javed, R. et al. HyFusion-X: hybrid deep and traditional feature fusion with ensemble classifiers for breast cancer detection using mammogram and ultrasound images. Sci Rep 15, 38314 (2025). https://doi.org/10.1038/s41598-025-22262-1

[4] Sarvamangala, D.R., Kulkarni, R.V. Convolutional neural networks in medical image understanding: a survey. Evol. Intel. 15, 1–22 (2022). https://doi.org/10.1007/s12065-020-00540-3

[5] Hou, J., Zhang, R., Xie, Y. et al. Multimodal deep learning for cancer prognosis prediction with clinical information prompts integration. npj Digit. Med. 9, 76 (2026).

[6] Sreelekshmi, V., Pavithran, K. & Nair, J.J. An integrated model for early breast cancer prediction using microcalcifications and patient risk factors. Discov Artif Intell 6, 30 (2026)

[7] Ahmed, S.F., Alam, M.S.B., Hassan, M. et al. Deep learning modelling techniques: current progress, applications, advantages, and challenges. Artif Intell Rev 56, 13521–13617 (2023). https://doi.org/10.1007/s10462-023-10466-8

[8] Nasser M, Yusof UK. Deep Learning Based Methods for Breast Cancer Diagnosis: A Systematic Review and Future Direction. Diagnostics (Basel). 2023 Jan 3;13(1):161. doi: 10.3390/diagnostics13010161. PMID: 36611453; PMCID: PMC9818155.

[9] Sreelekshmi, V., Pavithran, K. & Nair, J.J. An integrated model for early breast cancer prediction using microcalcifications and patient risk factors. Discov Artif Intell 6, 30 (2026).

[10] T. Matharasi, et al P. Ganguly, I. "Enhancing Breast Cancer Detection with AI: Ensuring Image Security and Diagnostic Precision Through Deep Learning Techniques," 2025 13th International Symposium on Digital Forensics and Security (ISDFS), Boston, MA, USA, 2025, pp. 1-5, doi: 10.1109/ISDFS65363.2025.11011981.

[11] Guo X, Yang F, Ji L (2026) DEF-Net: A dual-modal feature enhancement and fusion network for infrared and visible object detection. PLoS One 21(4): e0345815.

[12] Jhonatan Contreras, Thomas Bocklitz, Enhancing decision confidence in AI using Monte Carlo dropout for Raman spectra classification, Analytica Chimica Acta, Volume 1332, 2024, 343346,ISSN 0003-2670.

[13] Yan Y, Xu Y, Fang G, He X, Qian Y, Zhu W. Multimodal deep learning for breast tumor classification: Integrating mammography and ultrasound for enhanced diagnostic accuracy. J Appl Clin Med Phys. 2026;27:e70464. M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[14] C.D. Lekamlage, F. Afzal, E. Westerberg and A. Cheddad, "Mini-DDSM: Mammography-based Automatic Age Estimation," in the 3rd International Conference on Digital Medicine and Image Processing (DMIP 2020), ACM, Kyoto, Japan, November 06-09, 2020.

[15] Farooq, A., Das, C., Mishra, D. (2026). Uncertainty-Aware Cross-Modal Attention for Breast Cancer Classification in Ultrasound Imaging. In: Zhang, T., et al. Artificial Intelligence and Imaging for Diagnostic and Treatment Challenges in Breast Care. Deep-Breath 2025. Lecture Notes in Computer Science, vol 16142. Springer, Cham.

[16] Atrey K. Singh B. K.Bodhey N. K. (2024). Multimodal breast cancer classification with mammogram and ultrasound fusion using ResNet18 and SVM. J. Med. Imaging.11:021405. doi: 10.1117/1.JMI.11.2.021405

[17] Al Zoubi, A., Lu, F., Zhu, Y. et al. Classification of breast lesions in ultrasound images using deep convolutional neural networks: transfer learning versus automatic architecture design. Med Biol Eng Comput 62, 135–149 (2024).

[18] A. Sahu and S. Meher, "A Modified Relation and Margin-Based Deep Learning Network for Automatic Breast Cancer Detection," in IEEE Transactions on Computational Biology and Bioinformatics, vol. 22, no. 4, pp. 1682-1689, July-Aug. 2025, doi: 10.1109/TCBBIO.2025.356546Sahu and S. Meher, "A Modified Relation and Margin-Based Deep Learning Network for Automatic Breast Cancer Detection," in IEEE Transactions on Computational Biology and Bioinformatics, vol. 22, no. 4, pp. 1682-1689, July-Aug. 2025, doi: 10.1109/TCBBIO.2025.3565465.

[19] S. R. Sannasi Chakravarthy and Harikumar Rajaguru Chakravarthy et al Automatic DA, Automatic Detection and Classification of Mammograms Using Improved Extreme Learning Machine with Deep Learning, IRBM (2021)

[20] Subha S., Amudha Bhomini P. A Comprehensive Review of Image-Based Breast Cancer Detection Techniques: Challenges and Perspectives. Oper. Res. Forum 6, 131 (2025).

[21] Sahar Almahfouz Nasser, Ashutosh Sharma, Anmol Saraf, "Towards improving breast cancer detection through multi-modal image generation,Ultrasonics,Volume 153,2025,107655,ISSN 0041-624X.

[22] Gao S, Liu J, Li L, Yang D, Miao Y, Zhang X, Han Q, Shi Y, Wu J, Zhang K. Application of deep learning technology in breast cancer: a systematic review of segmentation, detection, and classification approaches. Biomed Eng Online. 2026 Jan 4;25(1):19. doi: 10.1186/s12938-025-01502-5. PMID: 41486277; PMCID: PMC12866484.