

An Intelligent Ultrasound Image–Based Deep Learning Framework for Early Breast Cancer Diagnosis Using Transposed Self-Attention and Optimized Spiking Neural Networks

Chalamala Harika Divya

chalamalahardhikdivya@gmail.com

Aditya University

M Vamsikrishna

Aditya University

Research Article

Keywords:

Posted Date: October 28th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9921814/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

An Intelligent Ultrasound Image–Based Deep Learning Framework for Early Breast Cancer Diagnosis Using Transposed Self-Attention and Optimized Spiking Neural Networks

Mrs.Chalamala Harika Divya^{1*}, Dr. M. Vamsikrishna²

^{1}Assistant professor, Department of AIML, Aditya University, Surampalem, Kakinada
Andhrapradesh, India*

*²Professor & HoD, Department of Computer Applications, School of Engineering,
Aditya University, Surampalem, Kakinada,AP, INDIA*

^{1}Corresponding author email: chalamalahardhikdivya@gmail.com*

Abstract

Breast cancer has been ranked as one of the highest causes of death in women all over the world, where early and accurate diagnosis with the help of ultrasound image is the main factor that can enhance the possibility of survival. Nevertheless, the current computer-aided diagnosis methods are associated with most of the shortcomings, including sensitivity to speckle noise, inability to preserve edges, insufficient feature encoding, and inability to generalize on different datasets. To overcome them, the current study proposed a new Entropy Weighted Guided Filtering Convolutional Transposed Self-Attention Optimized Spiking Network (EWGF-CTSA-OSN) to achieve effective breast cancer classification. This framework combines the use of entropy based weighted guided filtering to effectively suppress noise and maintain edges before a Convolutional Transposed Self Attention mechanism is used to capture both local and global contextual features. An adaptive masked attention spiking recurrent network that is biologically inspired, further improves classification with a model temporal dependency and critical tumor area focus. Also, Dung Beetle Optimization algorithm is utilized to polish the model parameters to enhance convergence and the overall performance. Experimental performance on BUSI and Ultrasound data show better performance with an accuracy of 99.20% and 98.90, precision of 99.19% and 99.10, recall of 99.24% and 99.20, F1-score of 99.21% and 99.00, and AUC values of 0.9933 and 0.9947, respectively.

Keywords: *Breast Cancer Classification, Ultrasound Imaging, Deep Learning, Attention Mechanism, Spiking Neural Networks, Metaheuristic Optimization*

1. Introduction

Breast cancer has continued to be one of the most serious health problems faced by the world and one of the major causes of cancer related fatalities in women across the world [1]. The detection of the illness at an early stage is significant in enhancing the survival, slowing the disease, and the quality of life of the patients [2]. The clinical research has shown that the diagnosis of the disease at an early stage greatly reduces the risk of mortality especially where the disease has been detected before it has acquired metastasis [3]. Nonetheless, breast cancer is quite aggressive and can spread quickly to essential organs, including the bone, liver, lungs, and brain, which is why the early identification of a disease is the key to successful treatment and prevention of complications [4, 5].

Ultrasound imaging has been found to be one of the most popular technique among other diagnostic modalities mainly because it is non-invasive, cost-effective, real-time, and appropriate even in dense breast tissues [6]. In spite of these benefits, ultrasound images are usually characterized by a number of shortcomings including speckle noise, low contrast, and complicated anatomy of the tissues, which complicate manual analysis, and may be subject to subjectivity [7]. As a result, it has become necessary to have smart automated systems that will be able to support clinicians in correctly interpreting medical images and enhance diagnostic accuracy [8].

The recent innovations in Deep Learning (DL) and Machine Learning (ML) have changed the sphere of medical image analysis in a rather impressive manner [9]. Specifically, Clinical Decision Support Systems (CDSS) built with deep learning have proven to be remarkably useful in improving the detection of breast cancer due to the ability of such CDSS to extract meaningful and high-level features within medical images [10]. Such systems do not only enhance the accuracy of diagnostic but also decrease the workload on the healthcare professionals as these systems present the accurate and data-driven information [11]. The performance of the methods based on the DL approach is better compared to the more traditional methods, such as ensemble classifiers or voting methods, as these methods are capable of detecting complex patterns and nuances in tumor characteristics [12].

In addition to that, the efficiency in the detection process may be increased by employing alternative methods for screening breast cancer, such as mammography, MRI, thermography, microwave imaging, and MRE [13, 14]. Although every modality has its advantages,

ultrasound imaging is very practical and easy to access, particularly in environments that are resource limited [15].

It is inspired by such obstacles and opportunities and suggests an advanced DL-based model to diagnose breast cancer by using ultrasound images. The system proposed combines the effective preprocessing, attention-based feature extraction, and refined classification processes to improve the detecting capabilities. The framework intends to reduce false detection, enhance the accuracy of the framework as well as the computational efficiency through the addition of adaptable learning techniques and optimization methods. Finally, the study will lead to the creation of a credible, scalable, and real-time clinical decision support system to detect breast cancer in its early stages, which can be used in the future to achieve improved clinical results and patient care. This work has novelty and contributions as explained below,

- A new hybrid framework (EWGF-CTSA-OSN) is introduced on the basis of using entropy-weighted guided filtering, transposed self-attention, spiking neural networks and bio-inspired optimization in one architecture to classify breast ultrasound images.
- To improve the quality of post-processing results, an improved preprocessing method is presented, which is an entropy-weighted guided image filtering technique to help reduce the speckle noise and to maintain non-essential critical edges and lesion boundaries.
- A Convolutional Transposed Self-Attention (CTSA) mechanism is created to aggregate all global contextual dependencies along channel dimensions with lower complexity in computation and higher representation of features.
- A biologically inspired Adaptive Masked Attention Spiking Recurrent Network is postulated to simulate the dynamics of the spatio-temporal dynamics and enhance classification by following the behaviour of neuronal firing.
- The Dung Beetle Optimization algorithm is integrated in order to tune the important model parameters dynamically in order to enhance convergence, stability and generalisation of the complete system.

The paper has included the following structure: Section 1 is the introduction section, Section 2 is the review of the related works, Section 3 outlines the proposed methodology, Section 4 is the discussion of the experimental results, and finally, Section 5 is the conclusion of the study and future research directions.

2. Literature survey

In 2022, Qi, X., et al. [16] presented a Mobile-GAN-based Automated Breast cancer Diagnosis System of ultrasound images. The method processes the images of the ultrasound reports captured by a mobile phone in three steps: noise reduction and image reconstruction by using GAN-based synthesis, classification into malignant and non-malignant through the use of a CNN, and identifying anomalies by using a GAN to reduce false negatives. The system can perform at the level of the experts and in 18,225 ultrasound images and 2,416 reports, and it could be used to screen in the cloud and on mobile gadgets. One of the weaknesses is the reliance on the quality of image captures and internet connectivity of quality to cloud process.

In 2022, Cho, S.W., et al. [17] presented a Multistage Breast Tumor Segmentation Framework based on BTEC-Net and RFS-UNet in ultrasound imaging. In the first step, the approach uses a breast tumor ensemble classification network (BTEC-Net) in order to establish the presence of a tumor in an image. Abnormal images are then subjected to a residual feature selection UNet (RFS-UNet) to obtain precise tumor regions. The method allows complete automated diagnosis and efficient localization of tumors and less computational burdens. One drawback is that there could be the spread of errors in case there is misclassification at the first stage of classification.

In 2023, Boulenger, A., et al. [18] proposed a CNN model based on VGG to detect triple-negative breast cancer (TNBC) in ultrasound images during the preoperative stage automatically. This approach involves ultrasound imaging and labeling based on IHC, stratified cross-validation with k folds, t-SNE and saliency maps to facilitate interpretation. The model was trained on 831 images of 145 patients, where it had high performance with an AUC of 0.86, 85 percent accuracy, and 86 percent sensitivity and specificity, with a clear differentiation of subtypes. One of the limitations is a relatively small and single-center dataset that may limit the generalizability.

In 2023, Raza, A., et al. [19] presented a DeepBraestCancerNet (DBCNet) system to detect and classify breast cancer. It was 24-layer architecture with six convolutional layers, nine inception modules, and one fully connected layer, based on clipped rectified linear unit (Rectified Linear Unit) activation functions, batch normalization and cross-channel normalization. The findings revealed that the DBCNet classification accuracy of 99.35 and 99.63 on a publicly available dataset was even better than state-of-the-art models. The drawback however was the high training time and complexity of computing, caused by the deep architecture.

In 2024, Madhu, G., et al. [20] designed a system called UCapsNet (U-Net Capsule Network) to segment and classify breast tumors with a high segmentation accuracy to enhance the accuracy of early diagnosis. The framework was two-stage and an improved U-Net was used to perform tumor segmentation with increased filter counts and skip connections, and then to classify them with a Capsule Network (CapsNet) utilizing dynamic routing to maintain spatial hierarchies. The findings revealed UCapsNet obtained a precision of 98.12, recall of 99.52, and accuracy of 99.22 on the Breast Ultrasound Image (BUSI) dataset as compared to such models as VGG-19, DenseNet, MobileNet, ResNet-50, and Xception. The drawback, however, was a risk of increased complexity in computations and increased time of training.

In 2025, Patro, B.D.K [21] described a method called Deep Learning Framework (DLF) that utilizes Convolutional Neural Networks (CNNs) in the detection of breast cancer (BC) through a feature selection technique and fusion technique. It was done using multi-level artificial neural networks and learning and classifying BC images to be further enhanced using convolutional recursive networks to increase detection and scalability in the early stages. The findings indicated that DLF attained the highest accuracy of 94.93, the highest precision of 93.66, the highest recall of 89.21, and the highest F1-score of 98.86 that exceeds the existing practices. The drawback however, was average recall rate and reliance on big training data to work best in spite of better automation.

Table 1: Comparative analysis of literature survey

Reference	Method	Result	Advantage	Disadvantage
Qi, X., et al. (2022) [16]	Mobile-GAN-based Automated Breast Cancer Diagnosis System	Achieved expert-level diagnostic performance on 18,225 images and 2,416 reports	Enables mobile-based screening with noise reduction, classification, and anomaly detection	Dependent on image capture quality and stable internet connectivity
Cho, S.W., et al. (2022) [17]	Multistage Framework using BTEC-Net and RFS-UNet	Effective tumor segmentation and	Fully automated pipeline with efficient tumor	Error propagation risk from initial

		localization with reduced computational load	detection and segmentation	classification stage
Boulenger, A., et al. (2023) [18]	VGG-based CNN for TNBC identification	AUC: 0.86, Accuracy: 85%, Sensitivity: 86%, Specificity: 86%	Provides subtype classification with interpretability using t-SNE and saliency maps	Limited generalizability due to small, single-center dataset
Raza, A., et al. (2023) [19]	DeepBraestCancerNet (DBCNet)	Accuracy: 99.35% and 99.63%	High classification accuracy outperforming state-of-the-art models	High computational complexity and longer training time
Madhu, G., et al. (2024) [20]	UCapsNet (U-Net + Capsule Network)	Precision: 98.12%, Recall: 99.52%, Accuracy: 99.22%	Combines strong segmentation and classification for improved diagnosis	High computational cost and increased training time
Patro, B.D.K. (2025) [21]	Deep Learning Framework (DLF) with CNN	Accuracy: 94.93%, Precision: 93.66%, Recall: 89.21%, F1-score: 98.86%	Automated feature learning and improved scalability	Moderate recall and requires large training dataset

Table 1 is the comparison of recent methods of breast cancer detection in terms of performance and features. Models such as DBCNet and UCapsNet were the most accurate, and Mobile-GAN made it possible to diagnose one using the mobile. The majority of the techniques enhanced automation and accuracy but it has a high computational cost, data dependency, and generalization problems thus indicating the trade-off between accuracy, efficiency, and feasibility of actual application.

2.1 Problem statement

Despite the specific progress in breast cancer detection based on the usage of ultrasound images and deep learning, the existing practices do possess several limitations, which are crucial and restrict their use in clinical environments. The issue of speckle noise, low contrast, and complex tissue structure in the ultrasound images is the challenge that many traditional and modern computer-assisted diagnostic systems have as it can easily destroy important lesion boundaries and be misinterpreted. The classical convolutional neural network and ensemble models usually fail to capture the long-range spatial interactions and small tumor structures, and thus, cause the growth in false positives and false negatives. Further, most of the existing methods lack advanced attention protocol and learn frameworks that are based on biology and thus, they are unable to focus on the critical regions and extrapolate effectively especially in the detection of cancer in its early stages. The other critical weakness is bad hyperparameter optimization which will cause slow convergence, high cost of computation and low real time application. Moreover, a lot of systems do not work with a wide range of datasets and imaging conditions in the most robust manner. These deficiencies indicate the necessity of having a more intelligent, noise insensitive, attention based, and computationally efficient framework that can enhance diagnostic accuracy, minimize errors, and enable timely clinical decision making during breast cancer detection.

3. Methodology

The current study proposed a new EWGF-CTSA-OSN to classify ultrasound images of breast cancer with high precision. First, BUSI and other ultrasound datasets have been clinically validated images that are used to guarantee the diversity and strength of data. The preprocessing step is an improved weighted guided image filtering algorithm with entropy-based edge-aware weighting to be able to reduce speckle noise and maintain significant lesion boundaries and structural information. The Convolutional Transposed Self-Attention mechanism is then used with an encoder-decoder architecture that operates to extract precise hierarchical features by

extracting not only local spatial features, but also global contextual relationships which are further accomplished with help of gated and coordinate attention mechanisms. The resulting discriminative features are then processed by a biologically inspired adaptive masked attention spiking recurrent network which simulates the temporal behavior and highlights essential tumor areas by attention masking and approximates neuronal spiking behavior to make a better decision. Lastly, the DBOA is utilized to optimally adjust the main model parameters including the learning rate, synaptic weights, threshold values and attention mask thereby improving classification accuracy, convergence rate and generalization. The combined system provides an effective and dependable diagnostic support tool through the classification of images of breast ultrasound robustly into normal, benign, and malignant groups. The schematic architecture of the developed model is given in Figure 1.

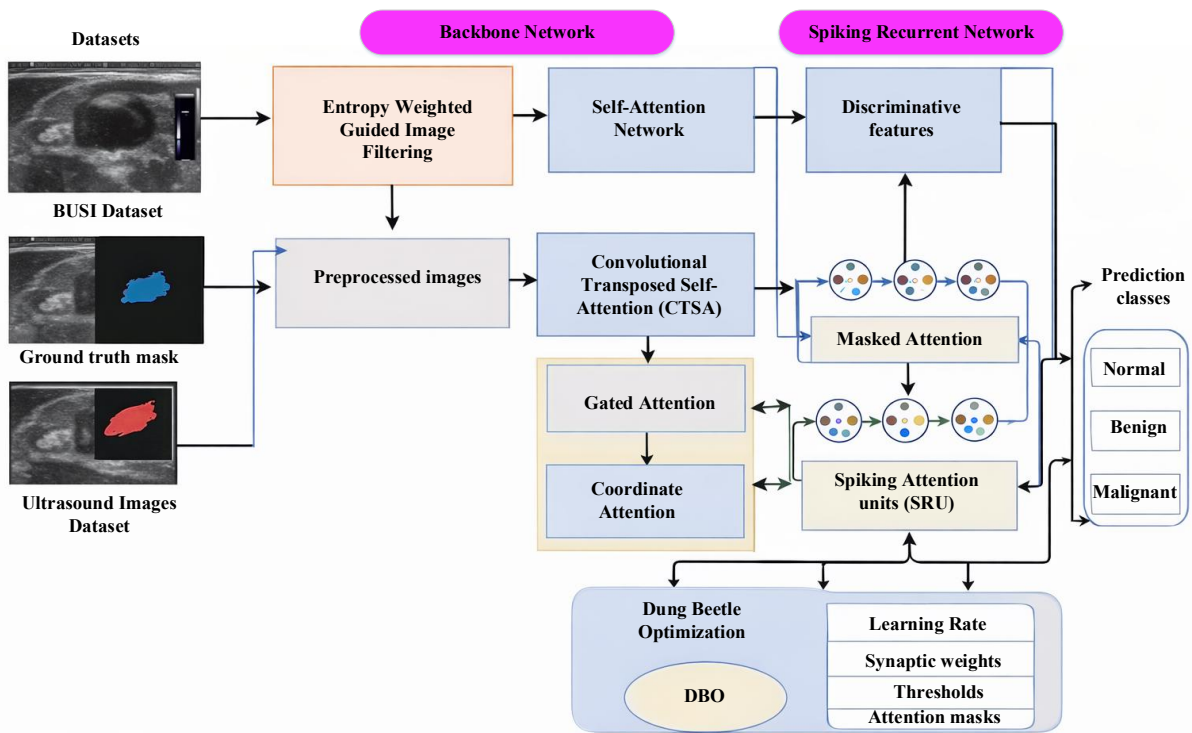


Figure 1: The schematic architecture of the developed

3.1 Data collection

To attain dependability of data and diversity, breast ultrasound images are collected in two widespread benchmark datasets i.e. BUSI (Breast Ultrasound Images Dataset) [22] and Ultrasound Images Dataset [23] in the study. These sets of clinically proven images of three classes (normal, benign, and malignant) can be used in the supervised learning of the detection of cancer using these sets. BUSI dataset is high quality of annotated images with definite

ground truth labeling, and the auxiliary ultrasound dataset can be used to add variability on imaging conditions and tissue properties. Through the collaboration of these datasets, the suggested model will become stronger and it will be able to generalize its functionality since it will be subjected to numerous tumor patterns and actual clinical scenarios.

3.2. Image Preprocessing using Weighted Guided Image filtering (WGIF)

The framework proposes an enhanced WGIF algorithm [24] to handle the issue of automatic speckle noise, low contrast, and blurred edges of the lesions that come along with the use of ultrasound imaging on image preprocessing. The main goal of this step is to produce a noise free, preservation of structure image, which maintains valuable diagnostic information and eliminates irrelevant variations. Whereas in the traditional filters, the proposed approach adds an edge-aware weighting mechanism that is accompanied by entropy-based optimization that enhances the noise suppression and the preservation of edges.

3.2.1 Edge-Aware Weighting Mechanism

The filtering process starts with the estimation of a local edge-aware weighting function that assesses the importance of each pixel according to the variance of the pixel. The traditional method of averaging-based weighting is replaced by an entropy evaluation method (EEM), which assigns weights to pixels according to the entropy values. The normalization measure can be expressed as below according to Equation (1).

$$\tau(i) = \frac{\frac{\sigma_{l,1}^2(k) + \gamma}{\sigma_{l,1}^2(i) + \gamma} - \min}{\max - \min} \quad (1)$$

where $\sigma_{l,1}^2(k)$ indicates the local variance at the target pixel k , $\sigma_{l,1}^2(i)$ refers to the local variance at neighboring pixel i , and γ is a small positive constant added to prevent division instability. The entropy for each pixel is subsequently calculated as follows in Equation (2).

$$E_i = -\frac{1}{\ln(N)} \sum_{i=1}^N \frac{\tau(i)}{\sum \tau(i)} \ln \left(\frac{\tau(i)}{\sum \tau(i)} \right) \quad (2)$$

where N is the number of pixels within the local window. Based on the entropy values, the adaptive weights are determined as follows in Equation (3).

$$W_i = \frac{1 - E_i}{\sum_{i=1}^N (1 - E_i)} \quad (3)$$

These weights emphasize edge pixels (high variance) and suppress smooth regions, enabling better structural preservation. The refined edge-aware weighting is given by Equation (4).

$$\hat{\psi}_I(k) = \sum_{i=1}^N W_i \frac{\sigma_{I,1}^2(k) + \gamma}{\sigma_{I,1}^2(i) + \gamma} \quad (4)$$

where $\hat{\psi}_I(k)$ represents the improved weighting factor for pixel k .

3.2.2 Edge-Preserving Filtering Model

To ensure an optimal balance between smoothing and edge preservation, a first-order edge-aware constraint is incorporated into the filtering cost function is given in Equation (5).

$$E(a_k, b_k) = \sum_{i \in w_k} \left[(a_k I_i + b_k - p_i)^2 + \frac{\epsilon}{\hat{\psi}_I(k)} (a_k - T_k)^2 \right] \quad (5)$$

Here I_i denotes the guidance image pixel, p_i is the input image pixel, a_k and b_k indicates the linear coefficients, w_k denotes the local window centered at pixel k , and ϵ is a regularization parameter. The term T_k is defined based on the gradient magnitude M_k as follows in Equation (6).

$$T_k = \frac{e^{M_k} - e^{-M_k}}{e^{M_k} + e^{-M_k}} \quad (6)$$

where $M_k = \sqrt{G(k_x)^2 + G(k_y)^2}$, with $G(k_x)$ and $G(k_y)$ representing gradient components along horizontal and vertical directions. This function ensures that $T_k \rightarrow 1$ for edge regions and $T_k \rightarrow 0$ for smooth regions. The optimal coefficients are computed as follows in Equations (7) and (8).

$$a_k = \frac{\frac{1}{|w|} \sum_{i \in w_k} I_i p_i - \mu_k \bar{p}_k + \frac{\epsilon}{\hat{\psi}_I(k)} T_k}{\sigma_k^2 + \frac{\epsilon}{\hat{\psi}_I(k)}} \quad (7)$$

$$b_k = \bar{p}_k - a_k \mu_k \quad (8)$$

where μ_k and σ_k^2 are the mean and variance of the guidance image within window w_k and $\bar{p}_k$ is the mean of the input image. Finally, the filtered output is obtained as follows in Equation (9).

$$q_i = \frac{1}{|w|} \sum_{k \in w_i} (a_k I_i + b_k) \quad (9)$$

or using entropy-based weighting is given in Equation (10).

$$q_i = \sum_{k \in w_i} W_k (a_k I_i + b_k) \quad (10)$$

where q_i is the output pixel value and W_k represents adaptive weights.

This enhanced form of the WGIF technique has been able to successfully remove the speckle noise in the images, at the same time maintaining the critical boundaries and details of the structures. The application of entropy-based weights and gradient constraints has ensured that the edges remain sharp, thus enhancing the preprocessing stage, which now provides noise-free, enhanced contrast ultrasound images.

3.3 Feature Extraction using Convolutional Transposed Self-Attention (CTSA)

In the proposed framework, feature extraction is performed using an advanced CTSA [25] mechanism embedded within a Transformer-based encoder–decoder architecture. Initially the preprocessed ultrasound image $I_N \in R^{h \times w \times 1}$ is passed through a 3×3 convolution layer to generate shallow feature maps $F_S \in R^{h \times w \times c}$, which represent low-level spatial and intensity information. These superficial features are then passed through a multi-level encoder-decoder network to learn hierarchical representations. The encoder uses a hierarchical representation by reducing the spatial resolution and increasing the channel depth. The decoder uses skip connections to reconstruct high-resolution features. This process produces deep feature maps $F_D \in R^{h \times w \times c}$ enriched with contextual and spatial information. Figure 2 presents the architecture of CTSA.

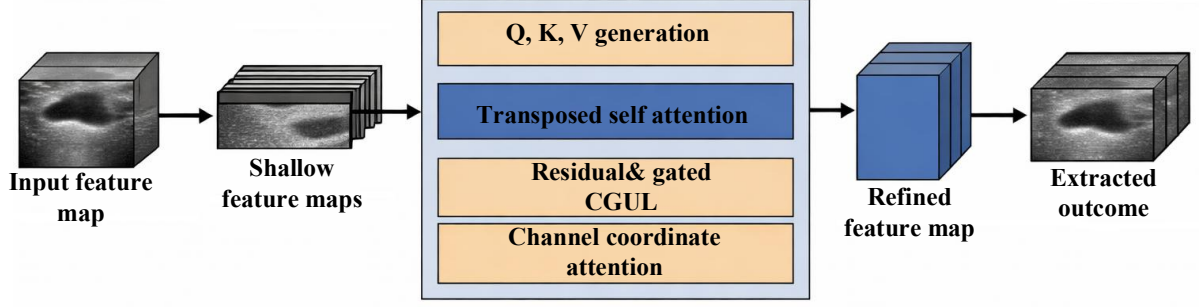


Figure 2: Architecture of CTSA

The core component of this stage is the CTSA block, which efficiently models global dependencies by computing attention along the channel dimension instead of the spatial dimension. Given an intermediate feature map $I \in R^{h \times w \times c}$ query, key, and value representations are generated using point-wise and depth-wise convolutions is given in Equation (11).

$$\begin{aligned} Q &= DConv_{3 \times 3}(Conv_{1 \times 1}(I)), \\ K &= DConv_{3 \times 3}(Conv_{1 \times 1}(I)), \\ V &= DConv_{3 \times 3}(Conv_{1 \times 1}(I)) \end{aligned} \quad (11)$$

These tensors are reshaped into $\hat{Q} \in R^{hw \times c}$, $\hat{K} \in R^{c \times hw}$, and $\hat{V} \in R^{hw \times c}$ to compute transposed self-attention as follows in Equation (12).

$$Attention(Q, K, V) = V \cdot \text{soft max} \left(\frac{\hat{K}^T \hat{Q}}{\alpha} \right) \quad (12)$$

where α is a learnable scaling parameter. This reduces the computational complexity considerably. At the same time, global contextual relationships in feature channels are preserved. The output of the attention is further enhanced by applying an additional depth-wise convolution and then added to the input through a residual connection as shown in Equation (13).

$$\hat{I} = Conv_{1 \times 1}(Attention(Q, K, V) + DConv_{3 \times 3}(V)) + I \quad (13)$$

This residual learning is helpful for stable training and maintaining the original information.

In order to further enhance the quality of feature representation, a Convolutional Gated Linear Unit (CGLU) is utilized to incorporate a gating mechanism to highlight more important features while suppressing irrelevant ones. The transformation is presented in Equation (14).

$$\hat{I} = W_3 (\sigma(DConv_{3 \times 3}(W_1 I)) \otimes W_2 I) + I \quad (14)$$

where W_1, W_2, W_3 are learnable weights, σ is the GELU activation function, and $\otimes$ denotes element-wise multiplication. This mechanism enhances spatial awareness and feature diversity. Additionally, a Channel Coordinate Attention Block (CCAB) is used to refine features by combining channel attention and spatial (coordinate) attention. The refined feature map is expressed as follows in Equation (15).

$$\hat{I}'' = M_h \otimes M_v \otimes (M_c \otimes \hat{I}) \quad (15)$$

where M_c, M_h , and M_v represent channel, horizontal, and vertical attention maps, respectively. This step ensures that the model focuses on critical areas of the tumor while still considering the position.

Overall, the process of feature extraction using the CTSA method has been effective in capturing long-range dependencies, identifying areas of the tumor, as well as global and local context information. This has resulted in a high-level discriminative feature map that is rich in information for classification purposes.

3.4 Classification using Adaptive Masked Attention Spiking Recurrent Network (AMASR-Net)

In the proposed framework, classification is done by employing a novel AMASR-Net model that uses attention mechanisms and biologically inspired spiking neural dynamics. This model is expected to effectively capture both spatial and temporal dependencies in the features of the ultrasound images to ensure improved classification accuracy. The high-level feature maps learned in the previous stage are considered as input embeddings. They are first processed by employing an adaptive attention mechanism to highlight the important features related to the tumor. The features are then passed to the spiking recurrent module to perform the final classification. Figure 3 shows the architecture of the AMASR-Net model.

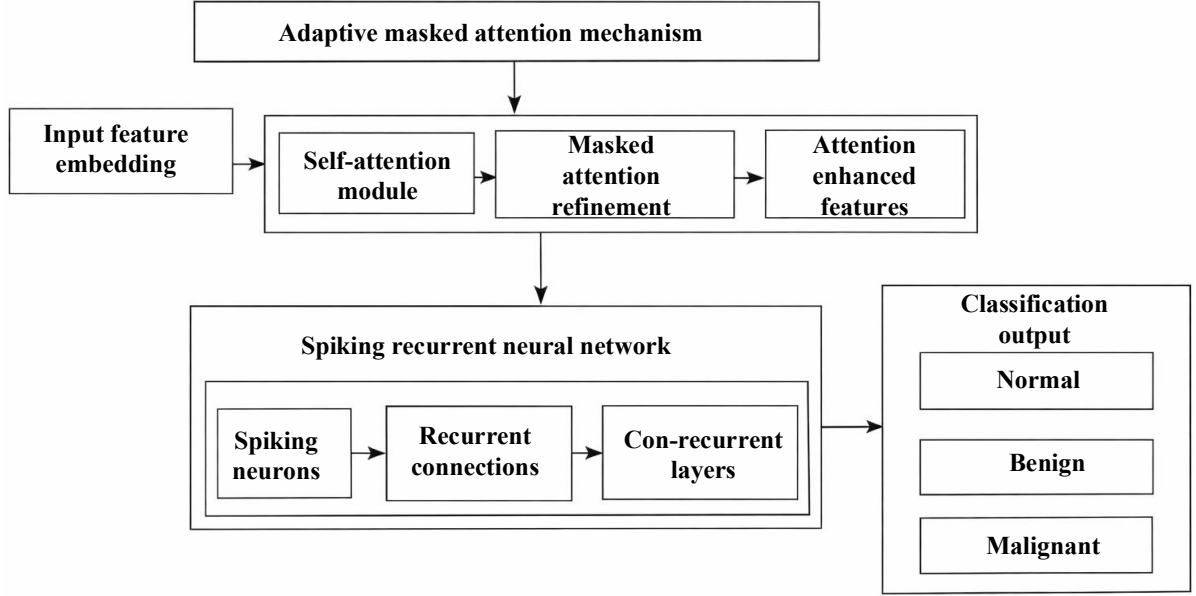


Figure 3: Architecture of AMASR-Net

3.4.1 Adaptive Masked Attention Mechanism

The classification process begins with a self-attention operation that computes contextual relationships within the feature representations [26]. Given an input feature embedding $E \in R^{n \times d}$, where n denotes the number of feature elements and d represents the embedding dimension, the attention is computed as follows in Equation (16).

$$att = \text{softmax} \left(\frac{Q \cdot E^T}{\sqrt{d}} \right) \quad (16)$$

where $Q \in R^{n \times d}$ is the query matrix obtained through a linear transformation of the global feature embedding, and E^T is the transpose of the embedding matrix. The scaling factor $\sqrt{d}$ stabilizes gradients during training. To focus on the most relevant local tumor features, a masked attention strategy is applied. The refined attention scores are computed as follows in Equations (17) and (18).

$$scores = \frac{Q' \cdot (E_l')^T}{\sqrt{d}} \quad (17)$$

$$\alpha_{ij} = \frac{\exp(scores_{ij})}{\sum_{z=1}^n \exp(scores_{iz})} \quad (18)$$

where E'_i represents the processed local feature embeddings, Q' is the corresponding query representation, and α_{ij} denotes the normalized attention weight assigned to feature j with respect to feature i . The final attention-enhanced feature representation is obtained as follows in Equation (19).

$$att_out = \alpha \times E \quad (19)$$

This mechanism dynamically assigns higher weights to important tumor regions while suppressing irrelevant background information, enabling more precise classification.

3.4.2 Spiking Recurrent Neural Processing

Then, the attention-refined features are fed into a spiking recurrent neural network, which mimics the dynamics of a biological neuron. This model can learn to capture temporal dependencies since it has recurrent connections, which allow past states to affect current predictions. The dynamics of a neuron's membrane potential are given by the following equation, which is Equation (20).

$$\tau \frac{dV(t)}{dt} = -(V(t) - V_{rest}) + X(t) \quad (20)$$

where $V(t)$ indicates the membrane potential at time t , V_{rest} is the resting potential, $X(t)$ is the input stimulus, and τ is the time constant controlling decay. In discrete form, the neuron dynamics are expressed as follows in Equation (21-23).

$$H(t) = f(V(t-1), X(t)) \quad (21)$$

$$S(t) = g(H(t) - V_{th}) \quad (22)$$

$$V(t) = H(t)(1 - S(t)) + V_{rest}S(t) \quad (23)$$

where $H(t)$ is the intermediate membrane potential, $S(t)$ is the spike output (binary signal), V_{th} is the firing threshold, $f(\cdot)$ denotes the state update function, and $g(\cdot)$ is a gating function that determines spike generation. A spike is released when the neuron reaches the threshold of the membrane potential, and the neuron resets nearly replicating the neural processes in biology.

To further improve learning of temporal features, convolutional recurrent are used, that is capable of jointly learning spatial and temporal feature. This is capable of assisting the network

to memorize historical feature data and exhibiting more strength in identifying minute tumor changes.

The attention and spiking recurrent processing enable the AMASR-Net to be able to learn discriminative features of the ultrasound images. The resultant output layer results in the transformation of the learnt spatio-temporal features into 3 classes namely normal, benign and malignant. This is the first classification outcome, and it is based upon the capability of the model to discriminate between various breast tissue conditions in a computationally efficient and biologically significant way.

3.5 Parameter Optimization using Dung Beetle Optimization (DBO)

The maximization of the parameters in the proposed framework is done by the DBO algorithm [28], a nature-guided metaheuristic technique that employs the clever foraging and survival behaviors of the dung beetles. The algorithm is applied to control the most significant parameters of the classification model (learning rate, synaptic weights, neuron threshold values and attention masks) in a way that allows optimality. The intelligent foraging strategy of the DBO algorithm helps to efficiently search the space to improve the performance of the model.

3.5.1 Initialization and Population Modeling

In DBO, each dung beetle represents a candidate solution $X_i(m)$, where i denotes the i^{th} beetle and m represents the current iteration. Each position vector encodes a set of model parameters to be optimized. The population is divided into four groups: rolling, breeding, foraging, and stealing beetles, each contributing differently to the exploration and exploitation of the search space.

3.5.2 Rolling Behavior (Exploration Phase)

Rolling dung beetles simulate global exploration by updating their positions based on previous movements and environmental changes is given in Equations (24) and (25).

$$X_i(m+1) = X_i(m) + w \cdot k \cdot X_i(m-1) + b \cdot \Delta X \quad (24)$$

$$\Delta X = |X_i(m) - X_{worst}| \quad (25)$$

where $X_i(m)$ is the current position, X_{worst} is the worst solution in the population, $w \in \{-1, 1\}$ is a direction factor, $k \in (0, 0.2]$ is a deflection coefficient, and $b \in (0, 1)$ is a constant. When obstacles occur, beetles change direction using Equation (26).

$$X_i(m+1) = X_i(m) + \tan(\theta) |X_i(m) - X_i(m-1)| \quad (26)$$

where $\theta \in [0, \pi]$ controls directional change.

3.5.3 Breeding Behavior (Local Exploitation)

Breeding beetles refine solutions by searching within a dynamically shrinking region around the best solution X^* is given in Equation (27) and (28)

$$L_b^* = \max(X^*(1-R), L_b), \quad U_b^* = \min(X^*(1-R), U_b). \quad (27)$$

$$B_i(t+1) = X^* + b_1(B_i(t) - L_b^*) + b_2(B_i(t) - U_b^*) \quad (28)$$

where L_b and U_b are lower and upper bounds, R is a control parameter related to iteration progress, and b_1, b_2 are random coefficients. This phase enhances fine-tuning of parameters.

3.5.4 Foraging Behavior (Guided Search)

Foraging beetles search near the globally best solution X_b to intensify exploitation is given in Equation (29) and (30).

$$K_l = \max(X_b(1-Q), l_b), \quad K_u = \min(X_b(1-Q), u_b) \quad (29)$$

$$X_i(m+1) = X_i(m) + \beta_1(X_i(m) - K_l) + \beta_2(X_i(m) - K_u) \quad (30)$$

where β_1 is a normally distributed random variable and $\beta_2 \in (0, 1)$. This step improves convergence toward optimal parameter values.

3.5.5 Stealing Behavior (Diversity Enhancement)

Stealing beetles introduce randomness to avoid local optima by moving toward the best solution is given in Equation (31).

$$X_i(m+1) = X_b + \eta \cdot \gamma \cdot (|X_i(m) - X_b^i| + |X_i(m) - X_b|) \quad (31)$$

where X_b is the global best solution, γ is a random vector, and η is a scaling constant. This mechanism increases population diversity and prevents premature convergence.

Each iteration has the evaluation of the fitness of a beetle (solution) in terms of classification performance measures, including accuracy and loss. It is a process of successive updating of positions to converge or to attain a maximum number of iterations. It is in this way that DBO is able to find the best combination of model parameters.

Consequently, this optimization step can significantly enhance the accuracy of the classification, the speed at which the system reaches the convergence and make it more generalized. The output of this step is a collection of optimized model parameters that are employed to make robust and reliable predictions of breast cancer in future predictions.

4. Results and Discussion

The Results and Discussion section assesses the performance of the proposed EWGF-CTSA-OSN model with regard to the effectiveness of the model in classifying breast cancer using ultrasound images. The performance of the model is checked based on accuracy, precision, recall, F1-score, and efficiency. Extensive testing is done on benchmark images to ensure the ability of the proposed framework to withstand, be able to generalize, and be able to handle noise. Also, a comparative study of the existing state-of-the-art techniques shows that the proposed technique performs better, particularly in low-contrast image processing and speckle noise images. Table 2 represents the implementation parameters.

Table 2: Implementation Parameters

Parameter	Value / Description
Dataset	BUSI + Ultrasound Dataset
Image Size	224×224
Preprocessing Method	Weighted Guided Image Filtering (WGIF)
Filter Window Size	5×5
Regularization Parameter (ϵ)	0.01
Feature Extraction	CTSA (Transformer Encoder–Decoder)
Convolution Kernel Size	3×3
Activation Function	GELU
Attention Mechanism	Transposed Self-Attention

Classifier	AMASR-Net (Spiking Recurrent Network)
Neuron Model	Leaky Integrate-and-Fire (LIF)
Learning Rate	0.001 (optimized using DBO)
Batch Size	32
Number of Epochs	100
Optimization Algorithm	Dung Beetle Optimization (DBO)
Loss Function	Cross-Entropy Loss
Hardware	GPU (NVIDIA/Equivalent)
Evaluation Metrics	Accuracy, Precision, Recall, F1-score

4.1 Dataset Description

The quality, diversity, and reliability of the dataset in terms of training and evaluating are key factors in ensuring the performance of any computer-aided diagnosis system. This study uses two publicly available datasets of breast ultrasound images to ensure a robust learning and better generalization. These datasets are clinically validated image datasets that have ground truth image annotations, allowing breast tumors to be supervised classified with accuracy into various categories. The Multiple datasets are used to measure the differences in imaging conditions, the shape of the lesion and tissue properties, which will increase the functionality of the suggested model.

4.1.1 Breast Ultrasound Images Dataset (BUSI)

Breast ultrasound images Dataset (BUSI) is a popular benchmark dataset used to analyze breast cancer. It was gathered in 2018 and consists of ultrasound images of women aged between 25 and 75 years. The data set consists of 780 images of 600 female patients, with an average of 500×500 pixels of the image in the PNG format. A ground truth mask is present for every image and is annotated by medical specialists and assist in recognizing lesions. The data is divided into three categories, namely, normal, benign, and malignant, which is appropriate to multi-class classification exercises. Since BUSI includes annotations of high quality, and classes representation is balanced, the dataset may be used as a stable source to create and test diagnostic models on the deep learning ground.

4.1.2 Ultrasound Images Dataset

The second dataset will be an image of breast ultrasound, which is gathered using various clinical cases to accommodate a variety of imaging circumstances and patient groups. This data comprises histologically verified benign and malign cases and radiologist-defined segmentation masks which mark the areas of lesions. Ultrasound imaging has been an important tool in the early detection of breast cancer because it is effective in visualization of the soft tissue and differentiation in the types of lesion. The presence of the expert-annotated masks contributes to the creation of the correct and reproducible computer-aided diagnosis. This data improves the strength of the proposed model by including the data in more centers and different clinical conditions, which minimizes the effects of dataset bias and in the end improves their applicability in real life.

4.2 Performance Evaluation

The performance evaluation section evaluates the effectiveness and reliability of the proposed model in the correct classification of breast ultrasound images. Classification performance is measured using various standard measures, such as accuracy and precision, recall and F1-score. This test is necessary to make sure that the model is not only highly accurate but also consistent, robust and generalized over the various data samples.

4.2.1 BUSI Dataset Performance Evaluation

The BUSI dataset performance evaluation is carried out in order to evaluate the effectiveness of the proposed model on a well-structured and clinically annotated benchmark data. Considering that BUSI has well-defined normal, benign, and malignant classes, which are characterized by ground truth masks, a robust testing ground exists to assess classification ability and discrimination quality. This testing provides an example of the ability of the model to represent specific tumor features and regular ultrasound imaging conditions.

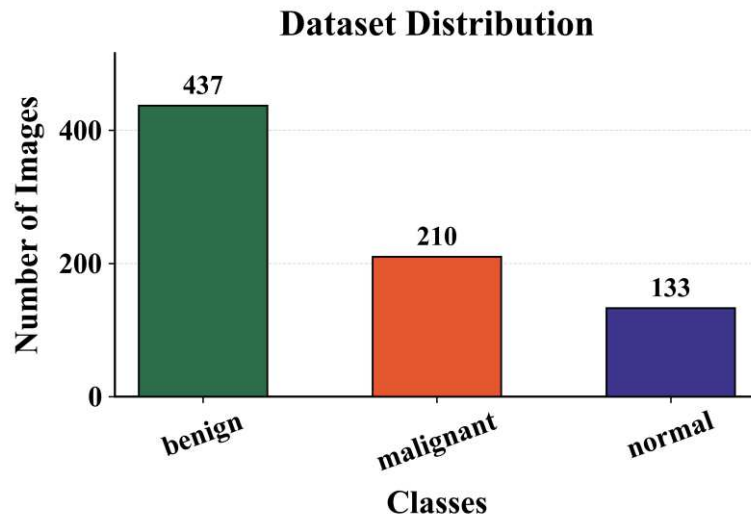
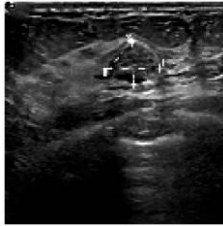
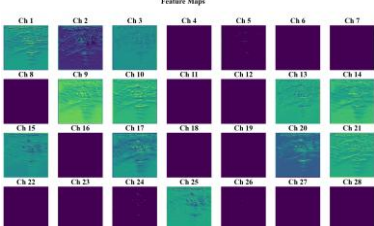
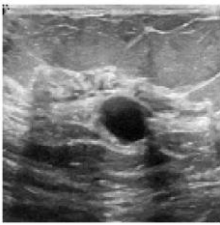
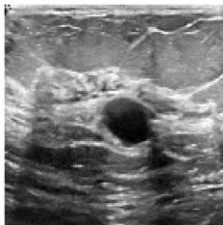
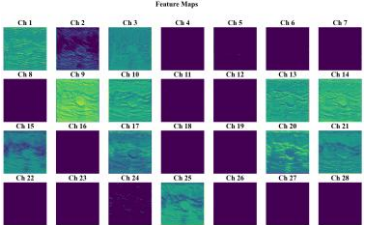


Figure 4: Distribution of Breast Ultrasound Image Classes

Figure 4 depicts the representation of the distribution of images for each class in the dataset, where there are three classes: benign, malignant, and normal. The largest number of images, which is 437, is for the dominant class, which is benign, followed by the number of images for the malignant class, which is 210. The least number of images, which is 133, is for the normal class. Therefore, there is a need to train the model well to avoid bias and ensure accurate classification for all classes.

Table 3: Step-wise Visual Results for Diagnostic Classification

Original	Processed	Feature Extraction	Classification	Class
				Benign
				Benign

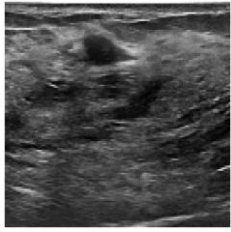
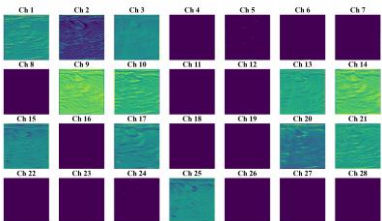
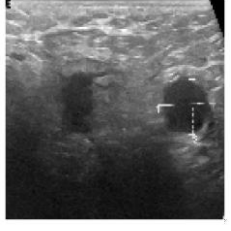
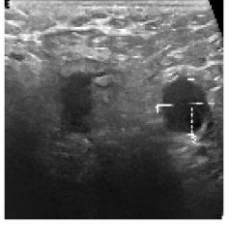
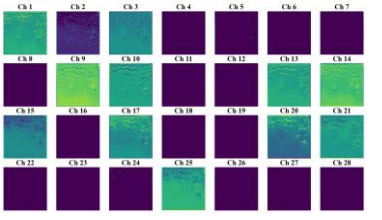
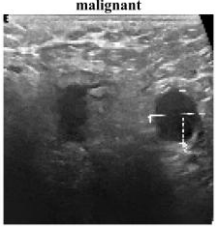
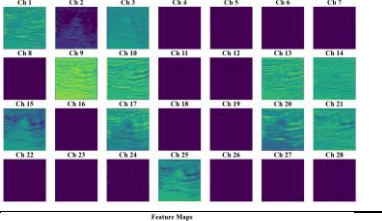
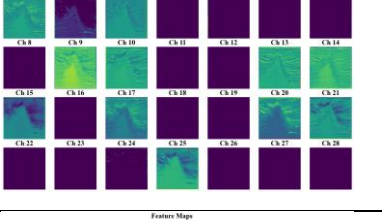
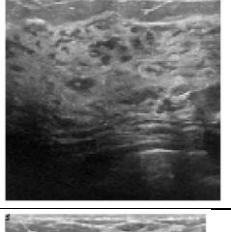
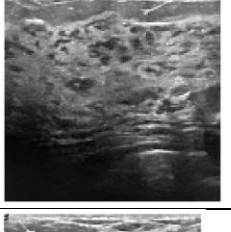
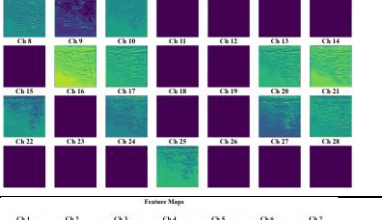
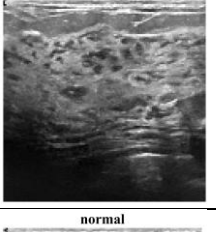
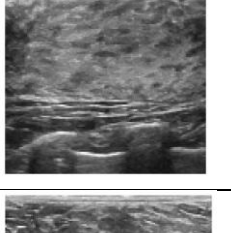
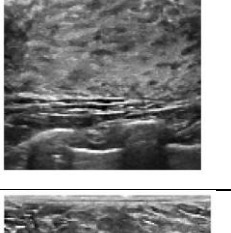
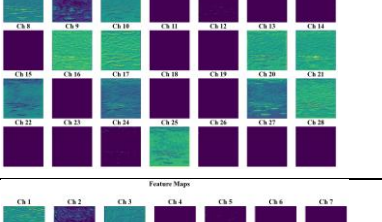
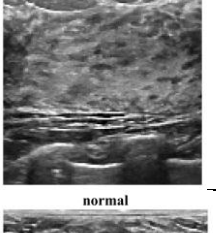
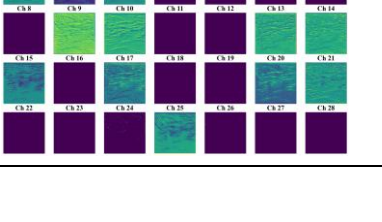
		 <p>Feature Maps</p> <p>Ch 1 Ch 2 Ch 3 Ch 4 Ch 5 Ch 6 Ch 7</p> <p>Ch 8 Ch 9 Ch 10 Ch 11 Ch 12 Ch 13 Ch 14</p> <p>Ch 15 Ch 16 Ch 17 Ch 18 Ch 19 Ch 20 Ch 21</p> <p>Ch 22 Ch 23 Ch 24 Ch 25 Ch 26 Ch 27 Ch 28</p>		Benign
		 <p>Feature Maps</p> <p>Ch 1 Ch 2 Ch 3 Ch 4 Ch 5 Ch 6 Ch 7</p> <p>Ch 8 Ch 9 Ch 10 Ch 11 Ch 12 Ch 13 Ch 14</p> <p>Ch 15 Ch 16 Ch 17 Ch 18 Ch 19 Ch 20 Ch 21</p> <p>Ch 22 Ch 23 Ch 24 Ch 25 Ch 26 Ch 27 Ch 28</p>		Malignant
		 <p>Feature Maps</p> <p>Ch 1 Ch 2 Ch 3 Ch 4 Ch 5 Ch 6 Ch 7</p> <p>Ch 8 Ch 9 Ch 10 Ch 11 Ch 12 Ch 13 Ch 14</p> <p>Ch 15 Ch 16 Ch 17 Ch 18 Ch 19 Ch 20 Ch 21</p> <p>Ch 22 Ch 23 Ch 24 Ch 25 Ch 26 Ch 27 Ch 28</p>		Malignant
		 <p>Feature Maps</p> <p>Ch 1 Ch 2 Ch 3 Ch 4 Ch 5 Ch 6 Ch 7</p> <p>Ch 8 Ch 9 Ch 10 Ch 11 Ch 12 Ch 13 Ch 14</p> <p>Ch 15 Ch 16 Ch 17 Ch 18 Ch 19 Ch 20 Ch 21</p> <p>Ch 22 Ch 23 Ch 24 Ch 25 Ch 26 Ch 27 Ch 28</p>		Malignant
		 <p>Feature Maps</p> <p>Ch 1 Ch 2 Ch 3 Ch 4 Ch 5 Ch 6 Ch 7</p> <p>Ch 8 Ch 9 Ch 10 Ch 11 Ch 12 Ch 13 Ch 14</p> <p>Ch 15 Ch 16 Ch 17 Ch 18 Ch 19 Ch 20 Ch 21</p> <p>Ch 22 Ch 23 Ch 24 Ch 25 Ch 26 Ch 27 Ch 28</p>		normal
		 <p>Feature Maps</p> <p>Ch 1 Ch 2 Ch 3 Ch 4 Ch 5 Ch 6 Ch 7</p> <p>Ch 8 Ch 9 Ch 10 Ch 11 Ch 12 Ch 13 Ch 14</p> <p>Ch 15 Ch 16 Ch 17 Ch 18 Ch 19 Ch 20 Ch 21</p> <p>Ch 22 Ch 23 Ch 24 Ch 25 Ch 26 Ch 27 Ch 28</p>		Normal
		 <p>Feature Maps</p> <p>Ch 1 Ch 2 Ch 3 Ch 4 Ch 5 Ch 6 Ch 7</p> <p>Ch 8 Ch 9 Ch 10 Ch 11 Ch 12 Ch 13 Ch 14</p> <p>Ch 15 Ch 16 Ch 17 Ch 18 Ch 19 Ch 20 Ch 21</p> <p>Ch 22 Ch 23 Ch 24 Ch 25 Ch 26 Ch 27 Ch 28</p>		Normal

Table 3 shows a visual comparative analysis of the experimental workflow for Benign (rows 1 to 3) and Malignant (rows 4 to 6) ultrasound samples. The workflow is described in terms of

raw input, enhancement, feature extraction with attention maps, and finally, the classification, which shows the effectiveness of the model in recognizing unique patterns to make accurate classifications.

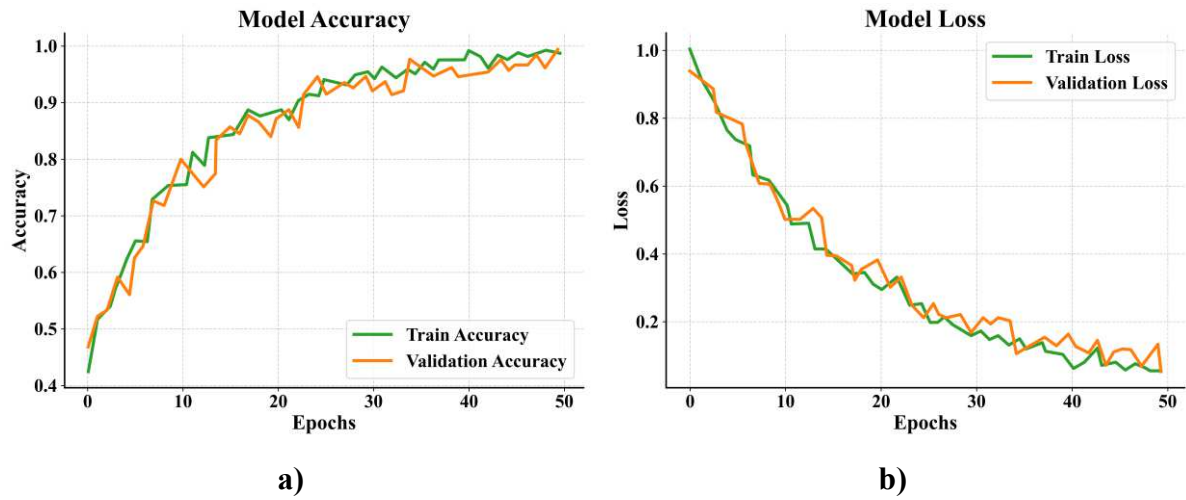


Figure 5: Training and Validation Performance of the Proposed Model

The figure 5 illustrates (a) model accuracy and (b) model loss with respect to training epochs. In (a), both training and validation accuracy are increasing steadily with values very close to 98-99%. In (b), both training and validation loss are decreasing with respect to epochs, which represents that the model is converging properly. The similarity in training and validation curves represents that there is no overfitting in this model.

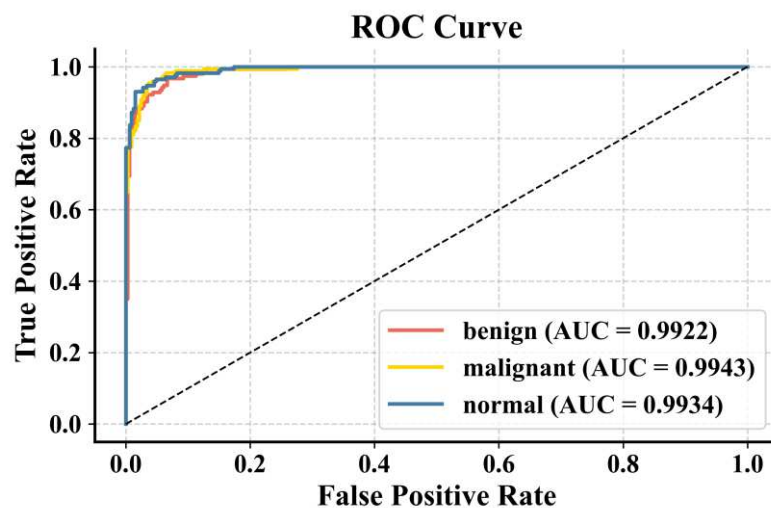


Figure 6: ROC Curve Analysis for Multi-Class Breast Cancer Classification

The above figure, 6, shows the ROC curves for the benign, malignant, and normal classes. All the curves are located near the top left, which is a good position, indicating good classification

results. The AUC values are high, indicating good discriminative power. The values are 0.9922, 0.9943, and 0.9934 for benign, malignant, and normal, respectively. The model has high true positive rates with low false positive values, indicating that the model is reliable and effective in distinguishing various types of breast tissues.

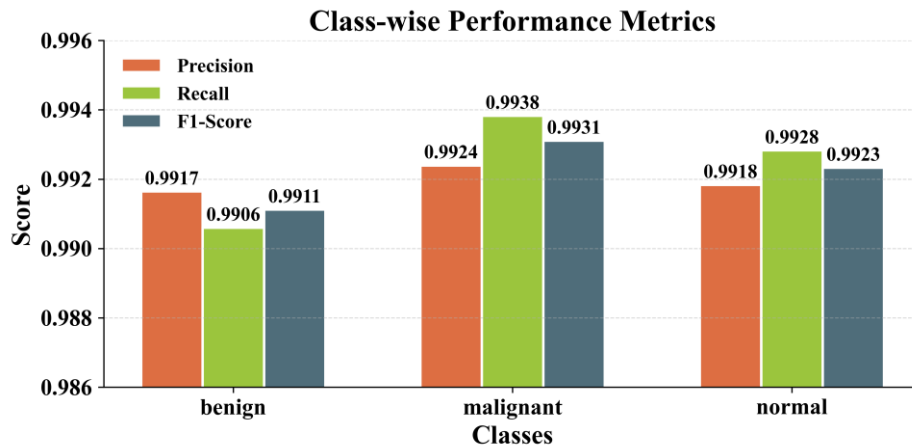


Figure 7: Class-wise Performance Evaluation of the Proposed Model

Figure 7 shows the precision, recall, and F1-score values for three classes, which are benign, malignant, and normal. The malignant class has the highest precision, recall, and F1-score values, which are 0.9924, 0.9938, and 0.9931, respectively. The normal class has high precision, recall, and F1-score values, while the benign class has relatively low values in terms of recall (0.9906). The model has highly consistent values in terms of accuracy.

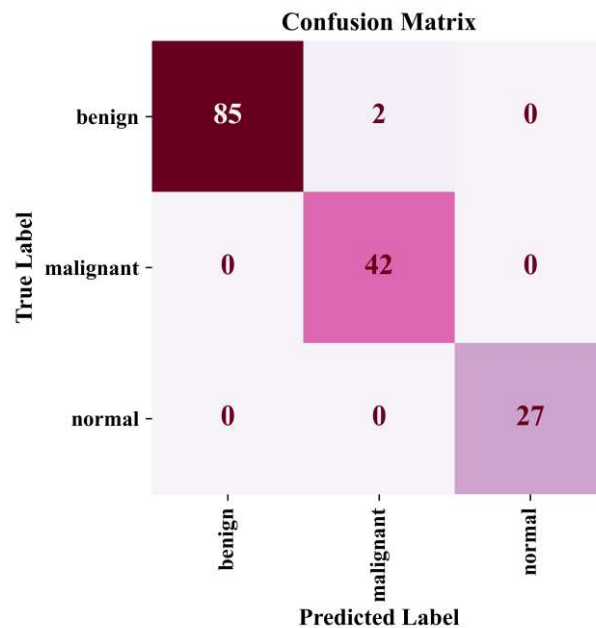


Figure 8: Confusion Matrix of the Proposed Classification Model

Figure 8 shows the classification results for benign, malignant, and normal classes. The model correctly classified 85 benign, 42 malignant, and 27 normal samples. Only 2 benign samples are incorrectly classified as malignant, while no other misclassifications are observed. The results show that the model has excellent accuracy with near perfect predictions for malignant and normal classes, while benign class is almost correctly classified.

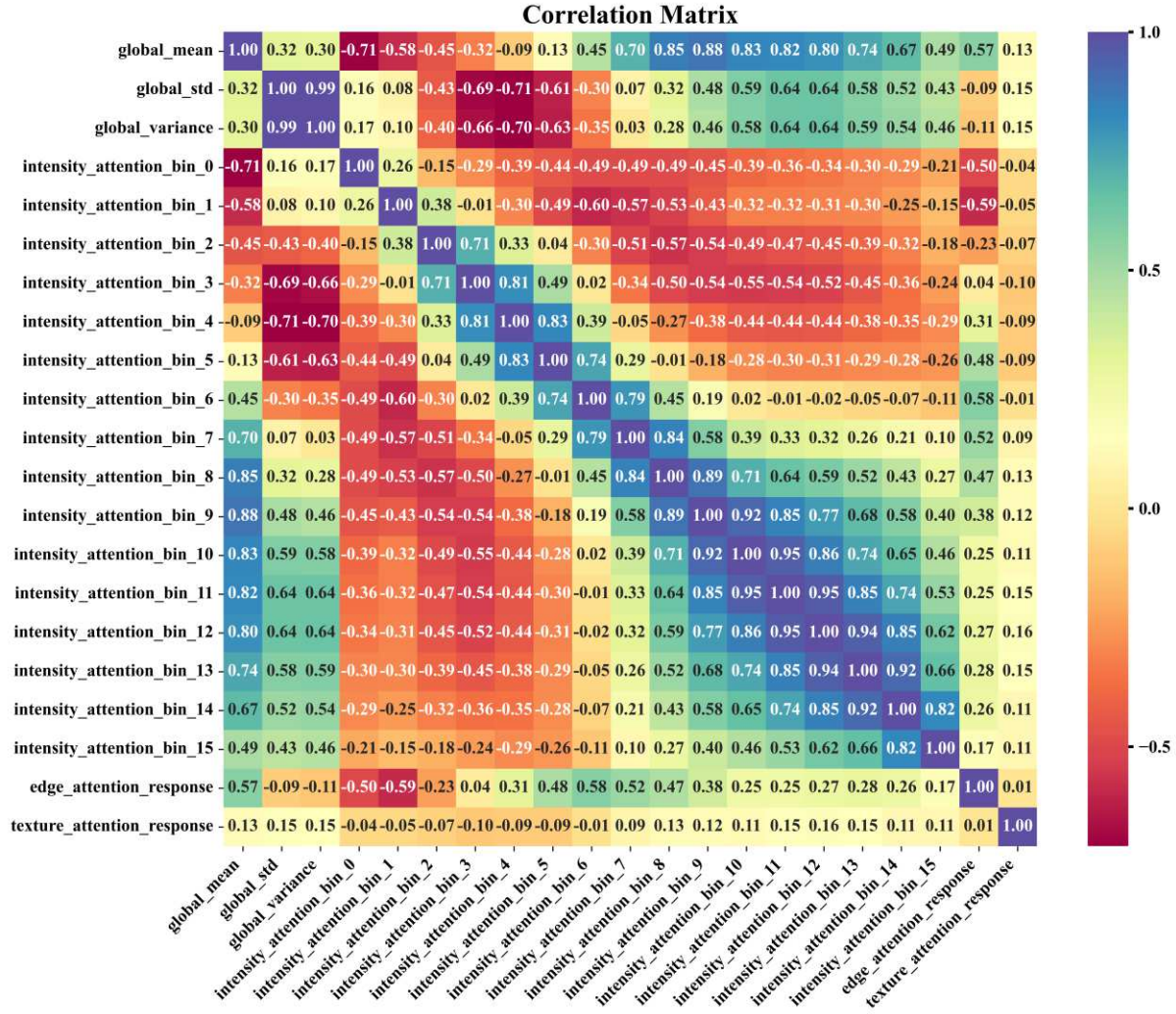


Figure 9: Correlation Matrix of Extracted Ultrasound Image Features

The figure 9 represents a correlation heatmap showing correlations with global statistical features, intensity attention bins, and attention response features. High positive correlations are found among intensity attention bins (especially 8 to 15). Global variance and global standard deviation show a very high correlation value (≈ 0.99), which is due to statistical dependency. Negative correlations are found between global mean and low intensity attention bins

(especially bin 0 with -0.71). Edge attention response features show moderate correlations with attention bins, while texture attention response features show low correlations.

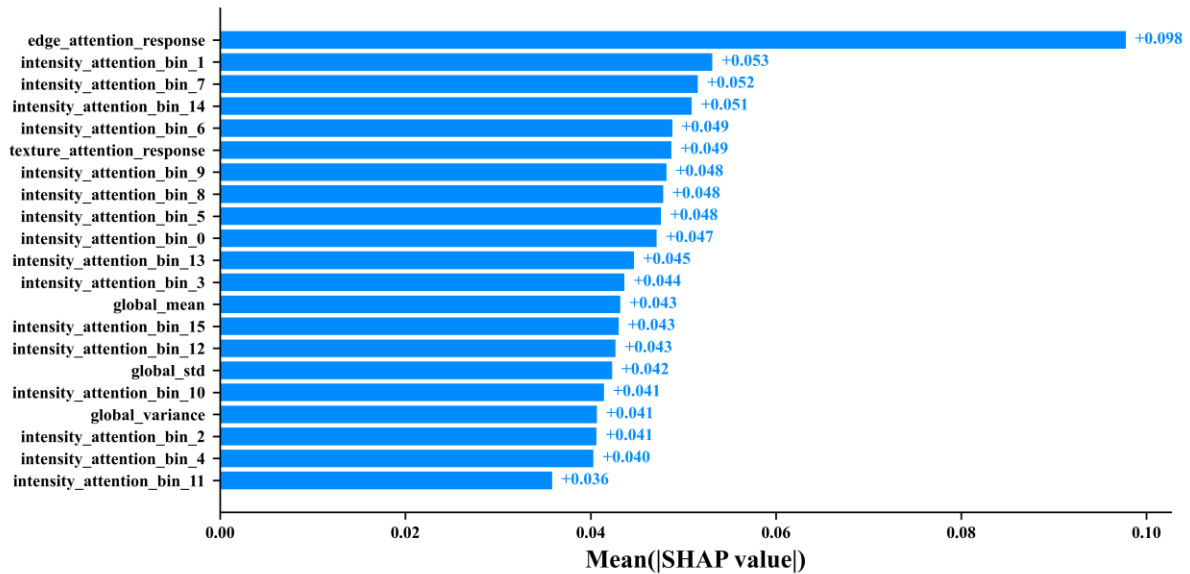


Figure 10: SHAP-Based Feature Importance Analysis

Figure 10 shows the mean SHAP values of the extracted features, which indicate their contribution to the classification model for breast cancer. The edge attention response feature has the highest value (0.098), which shows the highest influence on the model. The intensity attention bin 1 feature (0.053), intensity attention bin 7 feature (0.052), and intensity attention bin 14 feature (0.051) have high values, showing their influence on the model. Global statistical features such as global mean (0.043), global std (0.042), and global variance (0.041) have moderate values, showing their moderate influence on the model. Features such as intensity_attention_bin_11 (0.036) have lower ranks, showing their lesser influence on classification outcomes.

4.2.2 Ultrasound Images Dataset Performance Evaluation

The performance evaluation of the proposed model on the Ultrasound Images dataset is intended to check the generalization capability of the proposed model in different and real-world situations. In this dataset, images are included from different sources with different image qualities and lesions. This provides us with the opportunity to check the flexibility of the proposed model.

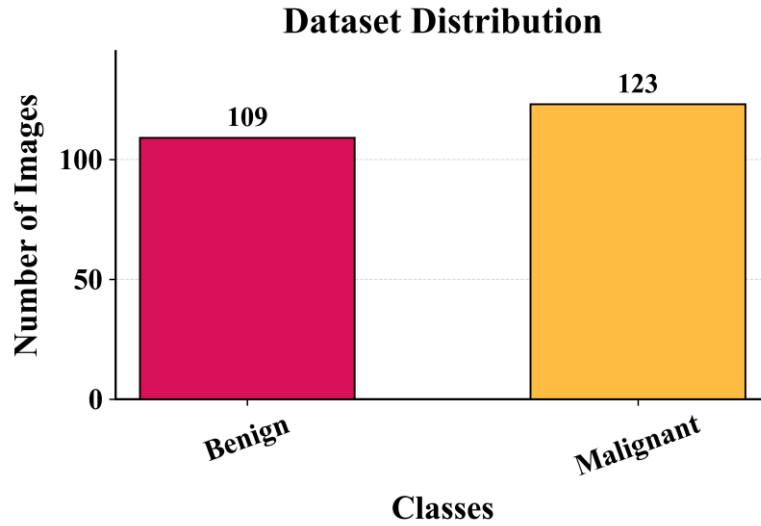


Figure 11: Distribution of Benign and Malignant Ultrasound Images

The figure 11 shows the class-wise distribution of the ultrasound images that are employed in the dataset. Out of the total images in the dataset, there are 109 images that are classified as benign, and there are 123 images that are classified as malignant. Thus, there is a slightly higher number of malignant images when compared to benign images. The distribution of the images is quite balanced in the dataset, which is always beneficial for training classification models. With a well-balanced dataset, bias can be avoided in the classification model for both benign and malignant tumor classes.

Table 4: End-to-End Image Processing and Classification Pipeline

Original	Processed	Feature Extraction	Classification	Class
		<p>Feature Maps</p>	<p>Benign</p>	Benign
		<p>Feature Maps</p>	<p>Benign</p>	Benign

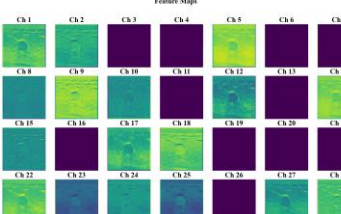
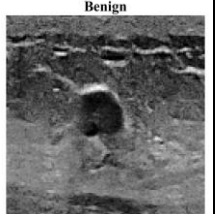
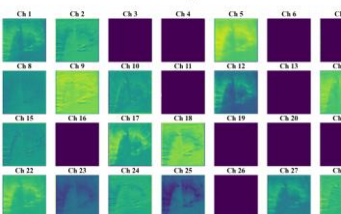
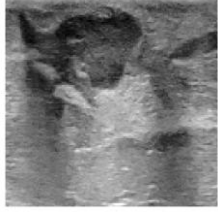
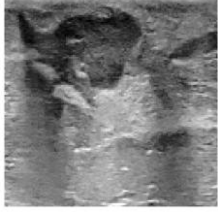
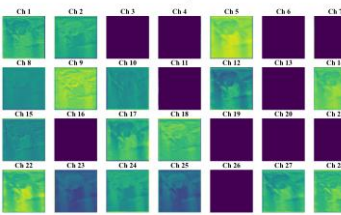
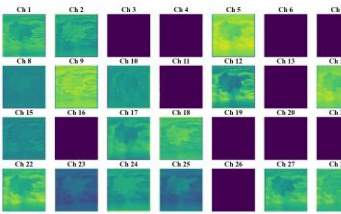
				Benign
				Malignant
				Malignant
				Malignant

Table 4 visualizes the experimental workflow for Benign and Malignant case studies. The workflow clearly shows the progression from Original Ultrasound images, through the processing and feature extraction stages, using attention-based heatmaps. The classification stage verifies the accurate labeling of the images, thus showing the model's potential in picking up critical features.

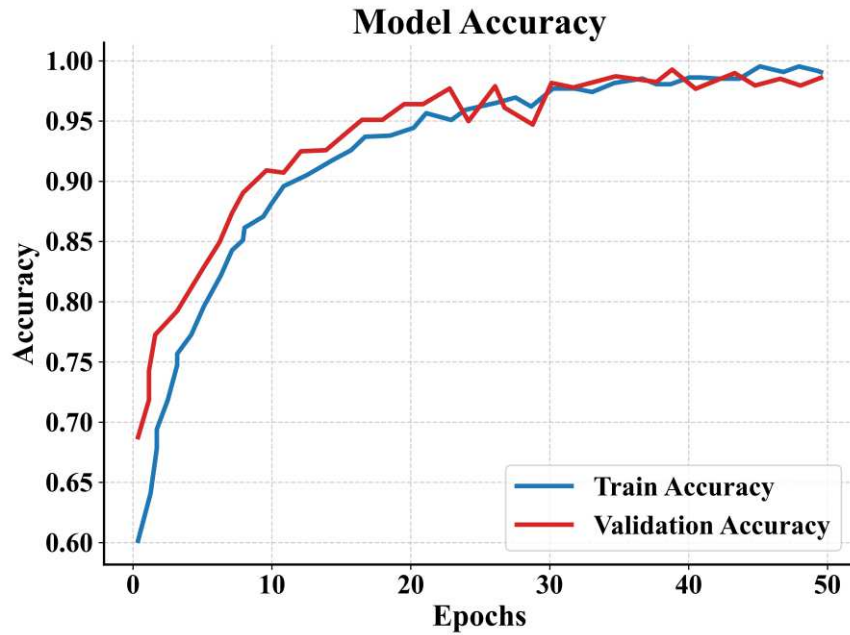


Figure 12: Training and Validation Accuracy over Epochs

The figure 12 shows the training and validation accuracy over 50 epochs. At first, both training and validation accuracy are improving at a rapid pace. The training accuracy increases from 0.60, and validation accuracy increases from 0.69. After some time, i.e., at around 25 epochs, both training and validation accuracy are steadily improving. Some fluctuations are noticed in validation accuracy, but they are close to training accuracy. Finally, at the last epochs, both training and validation accuracy are close to 0.98 or 0.99. This shows that the model is learning well.

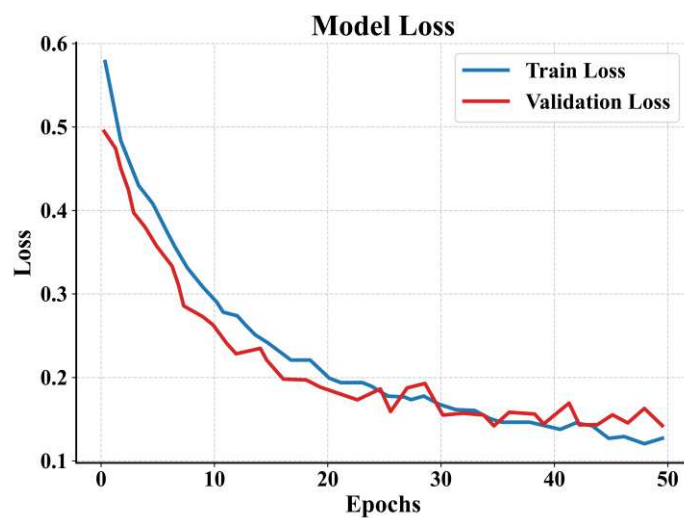


Figure 13: Training and Validation Loss over Epochs

The figure 13 shows the variation of training loss and validation loss over 50 epochs. The training loss starts at a higher point at approximately 0.58, while the validation loss starts at approximately 0.49. After that, all loss values drop significantly over the initial epochs. As training progresses, all loss values gradually drop. The loss values become stable after epoch 25. Some variations are noticed in the validation loss curve, while the values remain close to training loss. At the end of all epochs, all loss values drop to approximately 0.12 to 0.15, indicating effective learning, good convergence, and minimal overfitting.

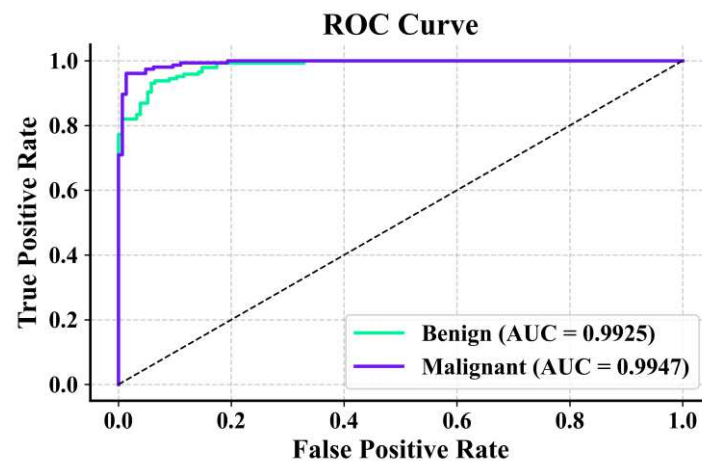


Figure 14: Receiver Operating Characteristic (ROC) Curve for Classification Performance

Figure 14 below depicts the ROC curve for the evaluation of the classification model's performance for two classes: Benign and Malignant. The ROC curve is a plot of the True Positive Rate (Sensitivity) versus the False Positive Rate ($1 - \text{Specificity}$). Both ROC curves indicate that the classification model has exceptional prediction accuracy since they are close to the top left of the plot. This is further supported by the Area Under the Curve (AUC) values for each class. For the Malignant class, the AUC value is higher at 0.9947 compared to the Benign class's value of 0.9925.

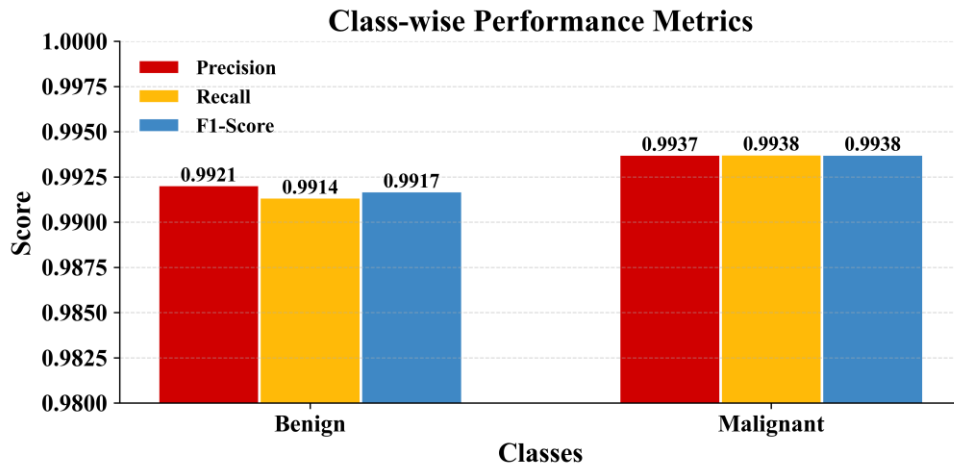


Figure 15: Class-wise Performance Metrics Comparison

Figure 15 shows the classification performance in terms of the Benign and Malignant classes using the three primary statistical measures, namely Precision, Recall, and F1-Score. It shows that the model has high reliability since all the measures are significantly higher than 0.9900. Though the Malignant class has slightly higher reliability than the Benign class, with an F1-Score of 0.9938 compared to 0.9917 in the Benign class, the slight variation in Precision and Recall in both classes indicates a perfectly balanced model. This is seen in the high value of Recall and the low probability of false alarm.

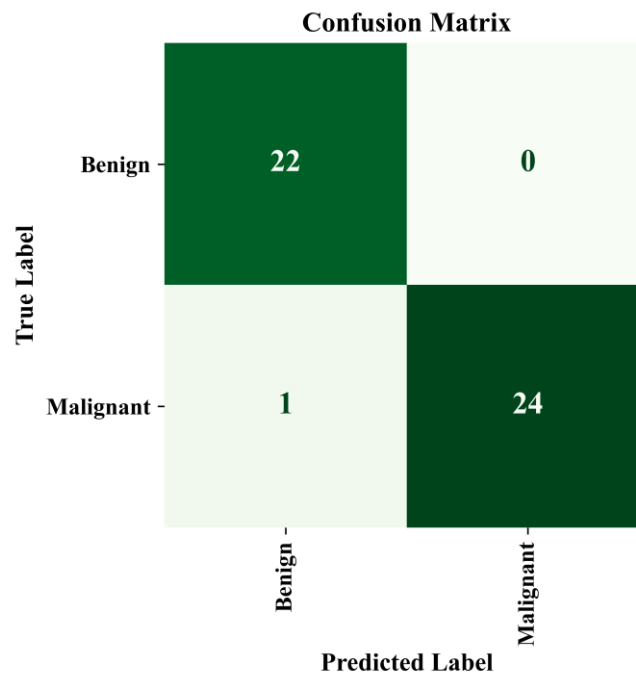


Figure 16: Confusion Matrix for Binary Classification

Figure 16 is a detailed illustration of the accuracy of the model's classification, as shown in the actual labels and the outcomes of the model's predictions. The diagonal line represents the high accuracy of the model's predictions, where there are 22 Benign and 24 Malignant instances classified as True Positives and True Negatives, respectively. However, it is worth noting that the model was able to achieve zero false positives for the Benign class. There is only one false classification in the entire matrix, where a Malignant instance is classified as Benign, representing a minor error in the False Negative classification.

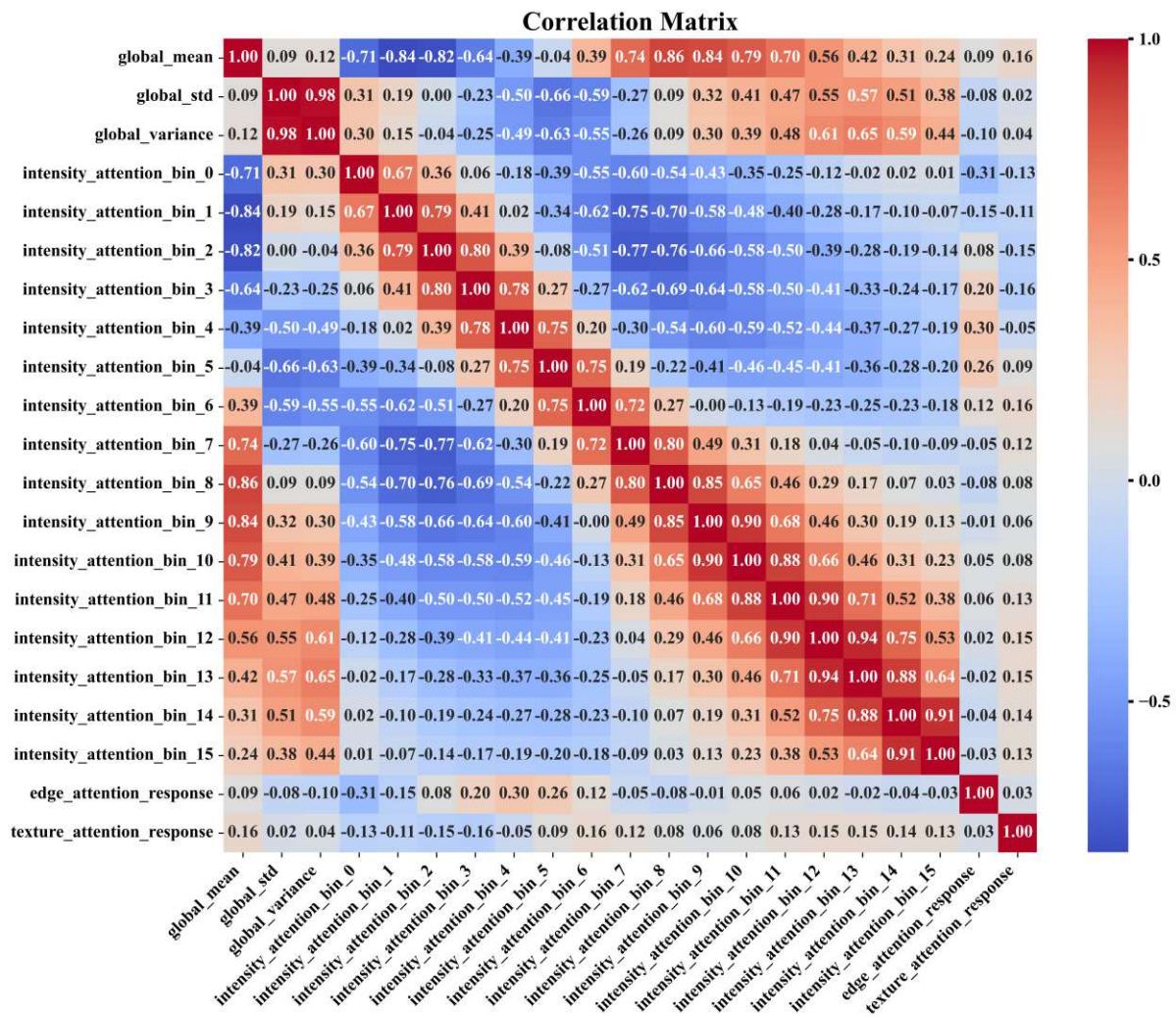


Figure 17: Feature Correlation Matrix Heatmap

Figure 17 illustrates a color map which represents the Pearson correlation coefficients between various features. The features included are global statistical features as well as features based on localized intensity attention bins. Strong positive correlations are seen between consecutive intensity attention bins, i.e., 12-15. On the other hand, global mean features are found to have strong negative correlations with mid-range attention bin features. This figure is important for

determining multi collinearity as well as understanding feature relationships that influence the diagnostic classification model.

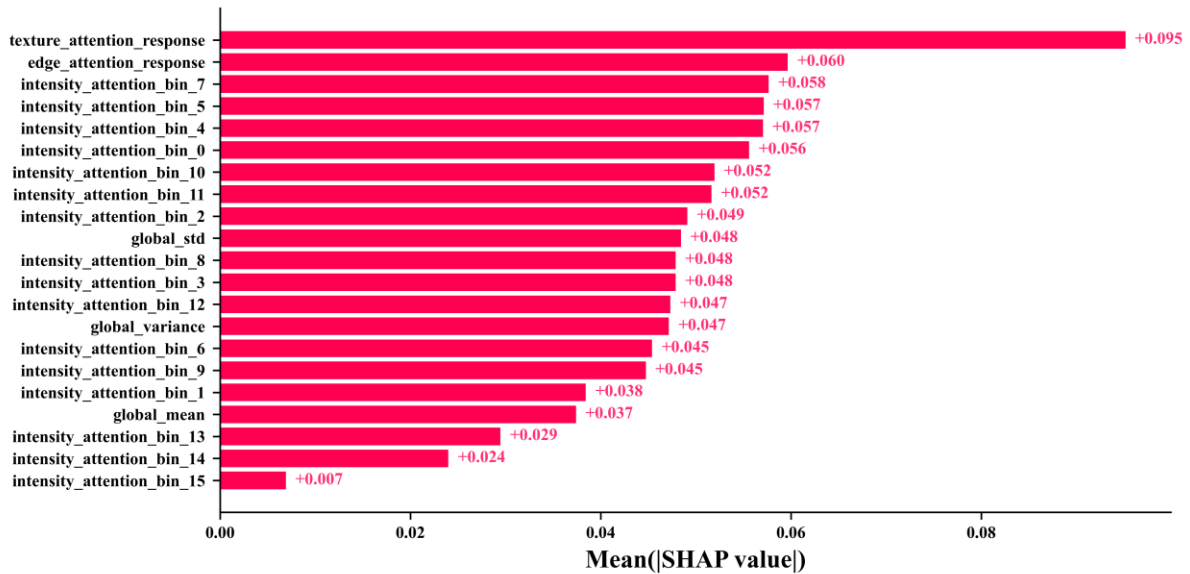


Figure 18: Global Feature Importance via Mean SHAP Values

Figure 18 illustrates the global significance of each feature in the diagnostic model, as determined by the average SHAP (SHapley Additive exPlanations) value. The features are listed in order of their average contribution to the prediction magnitude. It is observed that texture_attention_response has the highest significance, at +0.095, followed by edge_attention_response. The high ranking of the attention metrics indicates that localized structural and textural features are the primary contributors in classifying between the classes, as opposed to global statistical properties like the mean or variance.

Table 5: Ablation Study Results

Model Variant	Accuracy (%)	Precision	Recall	F1-Score
Without WGIF	96.80	0.967	0.965	0.966
Without CTSA	97.40	0.973	0.972	0.972
Without Spiking Network	97.90	0.978	0.977	0.977
Without DBO Optimization	98.30	0.983	0.982	0.982
Proposed EWGF- CTSA-OSN	99.20	0.9919	0.9924	0.9921

The ablation study of the performance of each component of the proposed model is shown in Table 5. The WGIF component has the most impact when removed from the model, as it achieves the lowest classification accuracy of 96.80%. The absence of the CTSA component and the spiking neural network also causes a decline in performance. The absence of the optimization of the DBO component causes a decline in performance to 98.30%. The proposed model achieves the highest performance of 99.20%.

Table 6: Diagnostic Performance

Metric	BUSI Dataset	Ultrasound Dataset
Accuracy	99.20	98.90
Precision	99.19	99.10
Recall	99.24	99.20
F1-Score	99.21	99.00
AUC (Average)	0.9933	0.9936

The overall performance of the suggested model is shown in Table 6 for both datasets. The model's accuracy is found to be exceptionally high for the BUSI dataset at 99.20%, while it is equally impressive for the Ultrasound dataset at 98.90%. The precision, recall, and F1 scores for both datasets are above 99%, indicating that the classification is perfectly balanced. The AUC values of 0.9933 and 0.9936 for the BUSI and Ultrasound datasets, respectively, clearly indicate that the model is capable of perfectly discriminating between different tissue classes.

Table 7: Performance Comparison with Existing Methods on BUSI and Ultrasound Datasets

Method	Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC
CNN-Based CAD Model		92.30	91.80	90.50	91.10	0.945
ResNet-50		94.70	94.10	93.60	93.80	0.962
DenseNet-121		95.80	95.20	94.90	95.00	0.971

Attention U-Net	BUSI	96.90	96.50	96.10	96.30	0.982
Hybrid CNN + Transformer		97.80	97.40	97.20	97.30	0.988
Proposed EWGF-CTSA-OSN		99.12	99.17	99.06	99.11	0.9933
CNN-Based CAD Model	Ultrasound Dataset	91.50	90.90	90.20	90.50	0.938
ResNet-50		93.80	93.10	92.70	92.90	0.955
DenseNet-121		95.10	94.60	94.20	94.40	0.968
Attention-Based CNN		96.70	96.20	95.90	96.00	0.981
Transformer-Based Model		97.60	97.10	96.80	96.90	0.989
Proposed EWGF-CTSA-OSN		98.95	99.05	98.90	98.97	0.9947

Comparison of proposed models with existing models on given datasets is shown in Table 7. For the BUSI dataset, accuracy of traditional CNN models is 92.30%, while accuracy of advanced models such as DenseNet-121 and Attention U-Net is 95.80% and 96.90%, respectively. The accuracy of the proposed model is higher, at 99.12%, with an AUC value of 0.9933. Similarly, in the case of the Ultrasound dataset, the accuracy of the traditional CNN is 91.50%, while the accuracy of the Transformer-based models is 97.60%. The accuracy of the proposed model is higher, at 98.95%, with an AUC value of 0.9947.

4.2.3 Statistical Analysis

In order to validate the performance of the proposed model, statistical analysis is carried out using statistical tests such as mean performance, standard deviation, and significance tests in comparison to other techniques. The performance of the proposed model is found to be higher

in terms of mean performance and less variable in terms of standard deviation. Furthermore, a paired t-test is carried out to validate the significance of the performance of the proposed method.

Table 8: Statistical Comparison of Proposed Model with Existing Methods

Model	Dataset	Mean Accuracy (%)	Std. Deviation	p-value
CNN-Based CAD Model	BUSI	92.30	1.85	0.004
ResNet-50	BUSI	94.70	1.42	0.003
DenseNet-121	BUSI	95.80	1.20	0.002
Attention U-Net	BUSI	96.90	1.05	0.001
Hybrid CNN + Transformer	BUSI	97.80	0.92	0.001
Proposed EWGF-CTSA-OSN	BUSI	99.20	0.45	—
CNN-Based CAD Model	Ultrasound Dataset	91.50	2.10	0.005
ResNet-50	Ultrasound Dataset	93.80	1.60	0.004
DenseNet-121	Ultrasound Dataset	95.10	1.30	0.003
Attention-Based CNN	Ultrasound Dataset	96.70	1.10	0.002
Transformer-Based Model	Ultrasound Dataset	97.60	0.95	0.001
Proposed EWGF-CTSA-OSN	Ultrasound Dataset	98.90	0.52	—

A statistical comparison of various models in two datasets is given in the table 8. The suggested model has the best mean accuracy (99.20% for BUSI and 98.90% for Ultrasound) and the

smallest standard deviation (0.45 and 0.52) which means the stable performance. Reduced p-values (less than 0.05) of other models support the fact that the introduced method is much more effective than the current methods that work with higher reliability and consistency.

4.3 Discussion

The suggested EWGF-CTSA-OSN emphasizes its efficiency in overcoming major biases related to the breast ultrasound image classification such as speckle noise, low contrast, and features ambiguity. Entropy-based weighted guided filtering is an important addition to image quality as it conserves the outline of vital lesions and removes noise. The CTSA module provides a better representation of features, both in terms of local spatial and global contextual dependence, whereas the spiking recurrent network provides biologically plausible temporal learning, which further increases the robustness of classification. Moreover, Dung Beetle Optimization algorithm is efficient to optimize the model parameters resulting in increased convergence and better generalization. The proposed framework outperforms the traditional form of deep learning models in terms of accuracy and overfitting as well as it can work with heterogeneous sets of data. Besides the improvements in accuracy, the model is reasonably computationally efficient with optimized attention operations, and efficient parameter tuning, thus it is applicable in the real world in clinical practice. Table 9 shows the computed effectiveness of the proposed model.

Table 9: Computational Efficiency Analysis of the Proposed Model

Component	Time Complexity	Execution Time (ms)	Memory Usage (MB)
WGIF Preprocessing	$O(N)$	12	45
CTSA Feature Extraction	$O(N \times C)$	28	120
Spiking Recurrent Network	$O(T \times N)$	35	150
DBO Optimization	$O(P \times I \times D)$	22	80
Overall Model	Moderate	97 ms	395 MB

The computational analysis reveals that the proposed model has balanced the trade-off between the performance and efficiency with the total execution time of 97 ms and average memory

consumption, which makes it feasible to be implemented into clinical decision-support systems.

5. Conclusion

The presented Entropy Weighted Guided Filtering Convolutional Transposed Self-Attention Optimized Spiking Network (EWGF-CTSA-OSN) proves to be an extremely useful and powerful structure of breast cancer classification with the help of ultrasound images. The model with advanced preprocessing, attention-based feature extraction, spiking neural dynamic, and Dung Beetle Optimization presents highest diagnostic performance accuracy of 99.20 and 98.90 with high precision, recall and AUC values greater than 0.99 in the BUSI and Ultrasound data respectively. The framework guarantees high convergence rate, good generalization and sound detection of normal, benign and malignant cases, which makes it appropriate to the real-life clinical decision support. Nonetheless, the suggested model is associated with some limitations such as the high complexity of the calculations because of the integration of several modules and the reliance on the high-quality annotated datasets. Also, the performance can change in the case of very low quality or invisible imaging conditions. Future research involves trying to minimize computational resource usage, lightweight architecture, increase multi-center data, and experiment with real-time deployments with explainable AI integration to continue to make clinical usage and scalability more practical.

Compliance with Ethical Standards

Conflict of interest

The authors declare that they have no conflict of interest.

Human and Animal Rights

This article does not contain any studies with human or animal subjects performed by any of the authors.

Informed Consent

Informed consent does not apply as this was a retrospective review with no identifying patient information.

Funding: Not applicable

Conflicts of interest Statement: Not applicable

Consent to participate: Not applicable

Consent for publication: Not applicable

Availability of data and material:

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Code availability: Not applicable

References

1. Dessie, Z. G., & Zewotir, T. (2025). Global determinants of breast cancer mortality: a comprehensive meta-analysis of clinical, demographic, and lifestyle risk factors. *BMC Public Health*, 25(1), 2640.
2. Religioni, U., Barrios-Rodríguez, R., Requena, P., Borowska, M., & Ostrowski, J. (2025). Enhancing therapy adherence: impact on clinical outcomes, healthcare costs, and patient quality of life. *Medicina*, 61(1), 153.
3. Morgan, E., O'Neill, C., Shah, R., Langselius, O., Su, Y., Frick, C., ... & Arnold, M. (2024). Metastatic recurrence in women diagnosed with non-metastatic breast cancer: a systematic review and meta-analysis. *Breast Cancer Research*, 26(1), 171.
4. Dhiman, R., Bazad, N., Mukherjee, R., Himanshu, Gunjan, Leal, E., ... & Pandey, R. P. (2024). Enhanced drug delivery with nanocarriers: a comprehensive review of recent advances in breast cancer detection and treatment. *Discover Nano*, 19(1), 143.
5. Tagde, P., Najda, A., Nagpal, K., Kulkarni, G. T., Shah, M., Ullah, O., ... & Rahman, M. H. (2022). Nanomedicine-based delivery strategies for breast cancer treatment and management. *International Journal of Molecular Sciences*, 23(5), 2856.
6. Bhatti, E., & Kaur, P. (2025). Advances in sensor technologies for breast cancer detection: a comprehensive review of imaging and non-imaging approaches. *Discover Artificial Intelligence*, 5(1), 168.
7. Golemati, S., & Cokkinos, D. D. (2022). Recent advances in vascular ultrasound imaging technology and their clinical implications. *Ultrasonics*, 119, 106599.
8. Oyeniyi, J., & Oluwaseyi, P. (2024). Emerging trends in AI-powered medical imaging: enhancing diagnostic accuracy and treatment decisions. *International Journal of Enhanced Research In Science Technology & Engineering*, 13(4), 81-94.
9. Rana, M., & Bhushan, M. (2023). Machine learning and deep learning approach for medical image analysis: diagnosis to detection. *Multimedia tools and applications*, 82(17), 26731-26769.
10. Sandhu, J. K., Sharma, C., Kaur, A., Pandey, S. K., Sinha, A., & Shreyas, J. (2025). Development of a clinical decision support system for breast cancer detection using ensemble deep learning. *Scientific Reports*, 15(1), 26098.
11. Jiang, S., Wang, T., & Zhang, K. H. (2023). Data-driven decision-making for precision diagnosis of digestive diseases. *BioMedical Engineering Online*, 22(1), 87.

12. Nakach, F. Z., Idri, A., & Goceri, E. (2024). A comprehensive investigation of multimodal deep learning fusion strategies for breast cancer classification. *Artificial Intelligence Review*, 57(12), 327.
13. Wang, L. (2023). Microwave imaging and sensing techniques for breast cancer detection. *Micromachines*, 14(7), 1462.
14. Andreadis, T., Gasteratos, A., Seimenis, I., & Koulouriotis, D. (2025). Application of CAD Systems in Breast Cancer Diagnosis Using Machine Learning Techniques: An Overview of Systematic Reviews. *Bioengineering*, 12(11), 1160.
15. Abhyankar, G., & Raman, R. (2024, April). Decision Tree Analysis for Point-of-Care Ultrasound Imaging: Precision in Constrained Healthcare Settings. In *2024 International Conference on Inventive Computation Technologies (ICICT)* (pp. 782-787). IEEE.
16. Qi, X., Yi, F., Zhang, L., Chen, Y., Pi, Y., Chen, Y., Guo, J., Wang, J., Guo, Q., Li, J. and Chen, Y., 2022. Computer-aided diagnosis of breast cancer in ultrasonography images by deep learning. *Neurocomputing*, 472, pp.152-165.
17. Cho, S.W., Baek, N.R. and Park, K.R., 2022. Deep Learning-based Multi-stage segmentation method using ultrasound images for breast cancer diagnosis. *Journal of King Saud University-Computer and Information Sciences*, 34(10), pp.10273-10292.
18. Boulenger, A., Luo, Y., Zhang, C., Zhao, C., Gao, Y., Xiao, M., Zhu, Q. and Tang, J., 2023. Deep learning-based system for automatic prediction of triple-negative breast cancer from ultrasound images. *Medical & biological engineering & computing*, 61(2), pp.567-578.
19. Raza, A., Ullah, N., Khan, J. A., Assam, M., Guzzo, A., & Aljuaid, H. (2023). DeepBreastCancerNet: A novel deep learning model for breast cancer detection using ultrasound images. *Applied Sciences*, 13(4), 2082.
20. Madhu, G., Meher Bonasi, A., Kautish, S., Almazyad, A. S., Mohamed, A. W., Werner, F., ... & Shokouhifar, M. (2024). UCapsNet: A Two-Stage deep learning model using U-Net and capsule network for breast cancer segmentation and classification in ultrasound imaging. *Cancers*, 16(22), 3777.
21. Patro, B. D. K. (2025). Improved early detection accuracy for breast cancer using a deep learning framework in medical imaging. *Computers in Biology and Medicine*, 187, 109751.
22. <https://www.kaggle.com/datasets/sabahesaraki/breast-ultrasound-images-dataset>
23. <https://qamebi.com/breast-ultrasound-images-database/>

24. Jia, H., Yin, Q., & Lu, M. (2023). Weighted guided image filtering with entropy evaluation weighting. *Computers & Graphics*, 117, 114-123.
25. Shen, Z., Qin, F., Ge, R., Wang, C., Zhang, K., & Huang, J. (2025). IDTransformer: Infrared image denoising method based on convolutional transposed self-attention. *Alexandria Engineering Journal*, 110, 310-321.
26. Lin, T., & Joe, I. (2023). An adaptive masked attention mechanism to act on the local text in a global context for aspect-based sentiment analysis. *IEEE access*, 11, 43055-43066.
27. Xu, Q., Fang, X., Li, Y., Shen, J., Ma, D., Xu, Y., & Pan, G. (2024, October). Rsn: Recurrent spiking neural networks for dynamic spatial-temporal information processing. In *Proceedings of the 32nd ACM international conference on multimedia* (pp. 10602-10610).
28. Fang, R., Zhou, T., Yu, B., Li, Z., Ma, L., & Zhang, Y. (2025). Dung beetle optimization algorithm based on improved multi-strategy fusion. *Electronics*, 14(1), 197.