IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154124

Availability-Aware Multimodal Deep Learning for Breast Cancer Diagnosis with Missing Modalities

Indu P. K[1], Dr. G Beni[2], Dr. D Rene Dev[3]

Dept. of Computer Science and Engineering, Noorul Islam Centre for Higher Education, TamilNadu, India[1]

Dept. of Computer Science and Engineering, Noorul Islam Centre for Higher Education, TamilNadu, India[2]

Dept. of Electronics and Communication Engineering. T John Institute of Technology , Bengaluru, India[3]

Abstract: The diagnosis of breast cancer frequently gains from multimodal imaging; however, comprehensive multimodal data are commonly inaccessible in standard clinical practice due to workflow challenges, budget restrictions, and uneven access to imaging resources. Many current multimodal deep learning models have restricted clinical use since they presume that all imaging modalities are accessible during inference. To overcome this constraint, we introduce a multimodal deep learning framework that is aware of missing modalities, incorporating modality availability modeling alongside reliability-informed decision support for breast cancer detection. The system utilizes modality-specific ResNet-101 encoders for both mammography and ultrasound, along with a fusion module that is aware of availability and that dynamically modifies the contribution of each modality based on its presence. A parallel reliability estimation head forecasts diagnostic confidence, allowing uncertainty-informed clinical recommendations instead of imposing binary choices in unclear situations. A training approach consisting of two stages with random modality dropout is employed to enhance resilience in cases where one or more imaging modalities are absent. Experimental findings indicate that the suggested framework exhibits consistent performance in the presence of missing modalities, while also attaining robust diagnostic discrimination, achieving an AUC of as high as 0.98. In contrast to unimodal models, the multimodal framework generated more accurately calibrated predictions, exhibiting a significantly reduced Expected Calibration Error. Reliability-stratified analysis showed that predictions with high confidence were significantly more accurate than those with low confidence, reinforcing the clinical importance of the suggested reliability score. The proposed framework enhances practical, uncertainty-aware multimodal decision support in real clinical environments by explicitly modeling modality availability and diagnostic confidence.

Keywords: Breast cancer; Multimodal imaging; Missing-modality learning; Availability-aware fusion; Diagnostic reliability; Clinical decision support; Imaging informatics; Deep learning.

I. INTRODUCTION

Breast cancer continues to be among the most prevalent cancers globally, with patient prognoses significantly affected by prompt and precise diagnosis [29,32]. Various imaging techniques, such as mammography and ultrasound, are frequently employed in clinical assessments as they offer additional diagnostic insights [6,24,31]. Ultrasound offers detailed characterization of lesions, especially in dense breast tissue, while mammography reveals more extensive structural patterns in the breast.

Despite the clinical advantages of multimodal imaging, comprehensive multimodal data for each patient are seldom accessible in standard clinical practice. The availability of imaging may be restricted due to financial costs, scheduling issues, patient-related factors, and the availability of institutional resources [18,26]. As a result, numerous deep learning–based multimodal diagnostic systems still face challenges in being deployed in clinical environments since they need simultaneous access to every imaging modality during inference [2,17].

This discrepancy between computational assumptions and clinical realities poses a significant obstacle to the adoption of imaging informatics [18]. Moreover, many current systems emphasize diagnostic accuracy mainly but offer restricted information regarding prediction reliability and decision confidence, both of which are crucial for safe clinical decision support [7,12].

To overcome these limitations, we introduce a multimodal framework for breast cancer detection that is aware of missing modalities and explicitly takes into account modality availability while integrating reliability-informed decision support [23]. The suggested system synchronizes algorithmic actions with radiological practice by postponing low-confidence cases for further imaging or follow-up instead of imposing uncertain predictions [18,32].

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026
DOI: 10.17148/IJARCCE.2026.154124

Like multimodal imaging, multimodal deep learning gains advantages from merging complementary information sources. In the suggested framework, modality-specific feature extractors obtain unique diagnostic traits, while the fusion module consolidates these representations into a cohesive feature space for the ultimate classification.

II. RELATED WORKS

Deep learning has significantly improved the analysis of medical images, especially in identifying and categorizing breast cancer through ultrasound and mammography [1,4,31]. Unimodal methods based on convolutional neural networks (CNNs) have shown excellent results; nonetheless, they are still restricted by diagnostic limitations specific to each modality [5,29].

Multimodal learning approaches aim to address these constraints by combining complementary imaging data from various sources. Earlier research has investigated early fusion, late fusion, attention-based fusion, and gating mechanisms to integrate data from mammography, ultrasound, magnetic resonance imaging (MRI), and computed tomography (CT) [14,21,30]. While these methods frequently enhance diagnostic precision, they generally presume that all modalities are present during inference.

Recent studies in medical artificial intelligence have increasingly concentrated on estimating uncertainty and calibrating confidence [7,18]. Prediction uncertainty is assessed through methods like evidential learning and Monte Carlo dropout; however, these approaches typically lack a direct connection to clinical decision-making processes [12,13]. Conversely, the current research emphasizes robustness against missing modalities, reliability assessment, and the integration of decision support to enhance the real-world usability of multimodal artificial intelligence in imaging informatics.

III. MATERIALS AND METHODS

A. Datasets

Datasets of publicly accessible breast imaging were utilized to assess the suggested framework. Ultrasound images were sourced from the BUSI dataset, while mammography images were sourced from the CBIS-DDSM dataset [1,31]. The two datasets include both benign and malignant cases accompanied by image-level diagnostic labels.

Due to the unavailability of patient-matched multimodal data, the two datasets were considered separate modality sources to replicate incomplete imaging situations frequently encountered in clinical settings [21,28]. This design allowed assessment of modality-conditional diagnostic performance in realistic missing-modality scenarios instead of depending on completely paired multimodal inputs.Every dataset was separately split into training (70%), validation (15%), and testing (15%) subsets to guarantee impartial model assessment [5,18].

B. Model Architecture

The proposed framework is composed of four principal components*:*

1. Modality-specific feature encoders
Separate convolutional neural networks are used to extract modality-specific feature representations from mammography and ultrasound images.

2. Availability-aware fusion module
An explicit modality availability vector is introduced to indicate whether each imaging modality is present or absent. This information is used to dynamically weight modality contributions during feature fusion, thereby preventing the model from relying on unavailable inputs.

3. Diagnostic prediction head
The diagnostic prediction head generates a probabilistic malignancy score for each case.

4. Reliability estimation head
The reliability estimation head produces a continuous reliability score that reflects the confidence of the diagnostic prediction

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154124

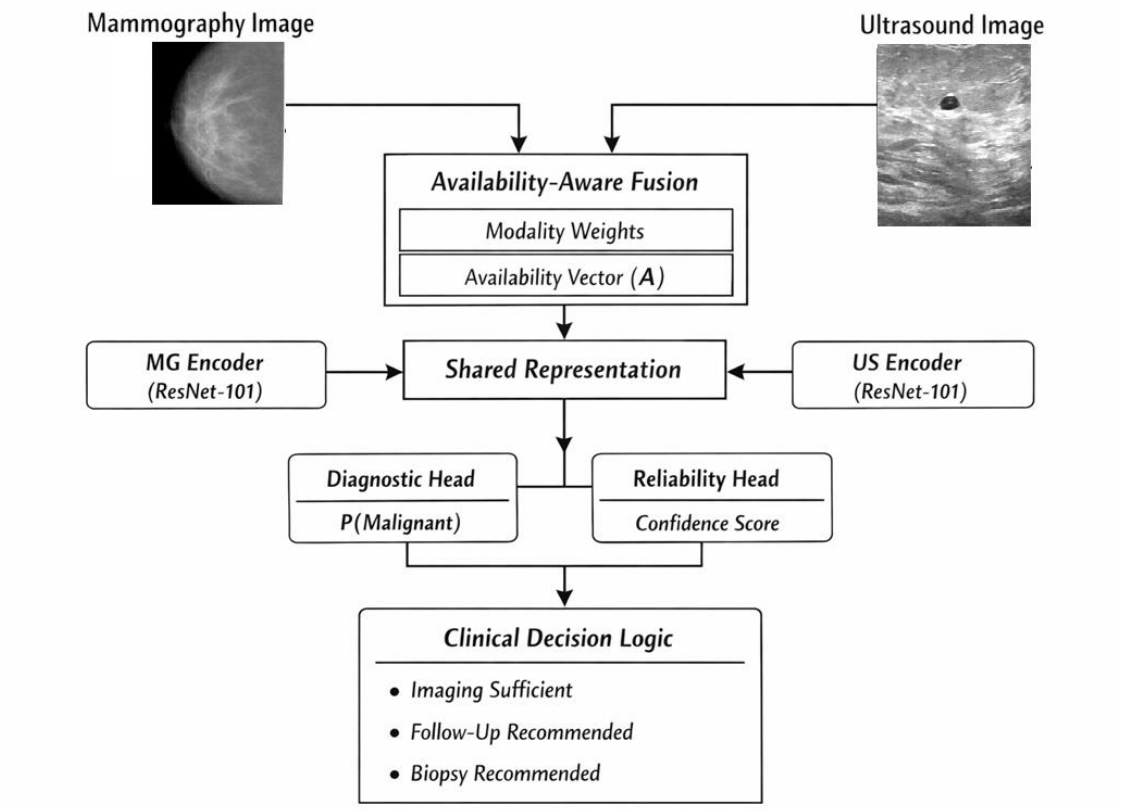

Figure. 1 Architecture of the proposed missing-modality multimodal framework for breast cancer diagnosis

Figure 1 illustrates the overall architecture of the proposed missing-modality-aware multimodal framework for breast cancer decision support. The framework consists of four interconnected components that jointly enable reliable diagnosis and clinically meaningful recommendations when one or more imaging modalities are unavailable.

1) Availability-Aware Fusion Module

The Availability-Aware Fusion module is the key element that allows the suggested architecture to function reliably under conditions of partial multimodal imaging. This module integrates feature representations from the mammography and ultrasound encoders, while explicitly modelling modality presence via an availability vector, as shown in Figure 1

2) Modality Weights

The fusion module acquires weights specific to each modality that dictate the influence of every modality on the overall fused representation, relying on the availability vector and the derived feature representations. By flexibly modifying these weights for each instance, the model highlights useful modalities and diminishes the impact of uncertain or absent inputs [15,28].

When both modalities are present, the fusion module acquires balanced weights that combine complementary diagnostic data from mammography and ultrasound [14,20,30]. When a modality is missing, the model automatically lessens or removes the effect of its associated zero-filled feature representation. This design prevents the fragile behaviour often seen in traditional fusion systems that rely on total multimodal input [21,25–27].

Through the simultaneous modelling of modality availability and feature significance, the fusion module creates a strong shared representation that maintains consistency across different clinical situations. This common representation is subsequently utilized as the input for both the diagnostic prediction head and the reliability estimation head, ensuring that downstream predictions account for both image content and data completeness [7,12,18].

From the viewpoint of imaging informatics, the availability-aware fusion module mirrors actual diagnostic reasoning, where clinicians modify their interpretation and confidence based on the imaging studies present during the assessment [18,32]. By embedding this logic directly into the model structure, the suggested framework enables safer and more flexible integration within standard clinical processes [29].

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026
DOI: 10.17148/IJARCCE.2026.154124

Let

- $\mathbf{f}_M \in \mathbb{R}^d$be the feature vector extracted by the mammography ResNet-101 encoder,
- $\mathbf{f}_{US} \in \mathbb{R}^d$be the feature vector extracted by the ultrasound ResNet-101 encoder.

An explicit modality availability **vector** is defined as:

$$\mathbf{A} = [a_{MG}, a_{US}], a_i \in \{0,1\}$$

where $a_i = 1$indicates the presence of modality i, and $a_i = 0$indicates its absence.

Based on feature content and modality availability, the fusion module learns modality weights[15,25]:

- $[w_{MG}, w_{US}] = \sigma\big(g([\mathbf{f}_M, \mathbf{f}_{US}, \mathbf{A}])\big)$
  where: $g(\cdot)$denotes a lightweight fully connected network, $\sigma(\cdot)$is a sigmoid activation ensuring $w_i \in [0,1]$.

The **fused representation** is computed as:

$$\mathbf{f}_{fusion} = w_{MG} \cdot \mathbf{f}_M + w_{US} \cdot \mathbf{f}_{US}$$

When a modality is unavailable ($a_i = 0$), the learned weighting method effectively suppresses its influence. Regardless of modality availability, this approach guarantees that the fused representation stays stable and informative.

3) Availability Vector (A)

The existence or lack of each imaging method at inference time is represented by a binary availability vector A. In a two-modality context, the availability vector is characterized as

$$A = [a_{MG}, a_{US}]$$

where every element receives a value of 1 if the respective modality is present and 0 if it is absent. This clear encoding allows the model to differentiate between placeholder inputs for absent data and truly informative feature representations [21,28].

4) MG(*Mammography)* Encoder

The mammography encoder is intended to record overall morphological and textural characteristics of breast tissue. Mammography images encompass extensive contextual details, such as tissue density distribution, mass shape, and patterns of architectural distortion.

With its extensive receptive fields and hierarchical residual features, the ResNet-101-based mammography encoder acquires feature representations that characterize: worldwide breast balance and architectural arrangement, characteristics of the lesion's shape and margins, variations in density and spiculated designs, and structural abnormalities linked to cancer. These traits are crucial for risk assessment and large-scale screening since subtle global imaging patterns frequently sway diagnostic choices [6,24,31]. The obtained mammographic features are then sent to the availability-aware fusion module for multimodal integration.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154124

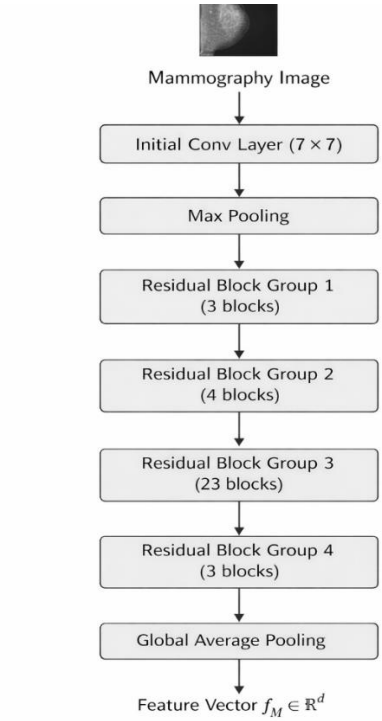

Figure 2: ResNet101-based mammography feature encoder

Figure 2 depicts the architecture of the mammography encoder, which utilizes the ResNet-101 backbone. The encoder analyses every mammography image via a series of convolutional and residual learning phases to identify high-level features pertinent to breast cancer identification.

The initial mammography image undergoes a 7×7 convolutional layer, then goes through batch normalization and nonlinear activation. This phase detects basic elements like edges, gradients, and rough structural designs. A max-pooling layer subsequently lowers spatial resolution while maintaining diagnostically important information.

The obtained feature maps are then handled by four sets of residual bottleneck blocks configured in the conventional 3–4–23–3 layout of ResNet-101. Identity skip connections in each residual block promote stable gradient flow and allow the network to acquire deep hierarchical representations without declining performance.

The initial residual blocks identify intermediate texture patterns, whereas the deeper layers gradually represent structural complexity and advanced morphological features linked to benign and malignant tissues. The deeper residual layers improve abstract feature representations prior to spatial aggregation.

Following the residual phases, global average pooling compresses the spatial feature maps into a feature vector of fixed length $f_M \in \mathbb{R}^d$ .The fully connected classification layer of ResNet-101 is eliminated, enabling the network to operate exclusively as a feature extractor within the multimodal system. Fine-tuning the encoder helps the pretrained backbone adjust specifically to the characteristics of mammographic imaging

5) US *(*Ultrasound) Encoder

Conversely, the ultrasound encoder emphasizes characteristics at the lesion level. Due to the inherent speckle pattern in ultrasound images, this modality highlights intricate anatomical details instead of broader structural context.

The ResNet-101 backbone allows the ultrasound encoder to acquire distinctive representations associated with:

• uneven lesion edges,
• internal echo characteristics,
• acoustic features that follow, and
• variations in local texture.

By analyzing ultrasound images separately, the model maintains modality-specific diagnostic indicators that are especially useful for evaluating lesion malignancy in dense breast tissue.

Employing distinct ResNet-101 encoders for mammography and ultrasound prevents the early feature entanglement and modality dominance that may arise with early fusion approaches [1,4]. Every encoder acquires a modality-adaptive representation that captures the diagnostic features and physical imaging attributes of its associated modality [5,22].

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154124

As illustrated in Figure 1, the availability-aware fusion module then combines these complementary representations to facilitate adaptable multimodal reasoning across various data-availability situations. This structure ensures that clinically significant ultrasound features stay informative even when mammography is not accessible

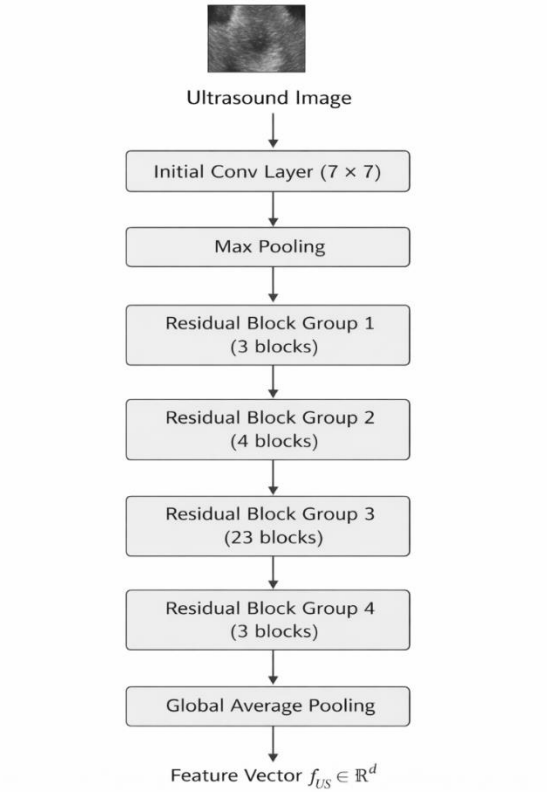

Figure 3: ResNet101-based ultrasound feature encoder

Figure 3 depicts the structure of the ultrasound encoder that utilizes the ResNet-101 backbone. The encoder utilizes a deep residual network to derive hierarchical feature representations that encapsulate lesion-level diagnostic traits from ultrasound images.

The initial ultrasound image undergoes a convolutional layer of 7×7, succeeded by batch normalization, ReLU activation, and max pooling. This phase lowers spatial resolution while capturing basic texture characteristics.

The generated feature maps are subsequently examined through four consecutive sets of bottleneck residual blocks organized in the typical 3–4–23–3 ResNet-101 format. Identity skip connections in each residual block promote stable gradient flow and allow the network to capture both low-level and high-level distinguishing features. This deep residual architecture is especially beneficial in ultrasound imaging, where minor boundary inconsistencies and intricate texture patterns can hold significant diagnostic value.

Following the last group of residual blocks, global average pooling reduces the spatial feature maps into a fixed-length feature vector f_US∈R^d. The initial classification layer of ResNet-101 is taken out, allowing the network to operate strictly as a feature extractor in the multimodal framework.

6) Shared Representation

Following availability-aware fusion, the weighted modality-specific feature vectors are integrated to create a common latent representation that acts as the core information bottleneck of the suggested system [14,15,28]. This collective representation stays strong against incomplete modality inputs while retaining supplementary diagnostic details from all accessible imaging sources [30].

The common depiction is officially described as:

$$z = f_{\text{fusion}} \in \mathbb{R}^d$$

where f_{fusion} represents the result of the availability-aware fusion module. The vector z incorporates multimodal information into a cohesive feature space that remains consistent across various modality-availability patterns.

DOI: 10.17148/IJARCCE.2026.154124

The common representation aligns diverse imaging attributes in a unified semantic space, enabling effective multimodal reasoning. When both mammography and ultrasound data are present, z captures additional global and lesion-specific characteristics [21,25,30]. When a single modality is present, the representation adjusts by highlighting informative aspects while reducing the influence of absent inputs. This design minimizes the instability frequently seen in conventional fusion techniques, in which absent modalities can lead to distorted or inaccurate feature representations [26,27]. The common representation is then utilized by both the diagnostic and reliability branches to generate clinically meaningful predictions.

7) Diagnostic Prediction Head

The diagnostic prediction head translates the common representation into a probabilistic cancer score, referred to as P("malignant"). This head comprises a lightweight fully connected layer and a sigmoid activation function that generates an output within the range [0, 1].
The likelihood of malignancy is calculated as:

$$P(\text{malignant}) = \sigma(W_d z + b_d)$$

where: z is the shared representation, $W_d$and $b_d$are learnable parameters, and $\sigma(\cdot)$denotes the sigmoid activation function.

In contrast to strict binary classification, this probabilistic approach enables a risk-based interpretation, aligning better with clinical workflows where diagnostic choices are frequently influenced by probability thresholds instead of set categorical outcomes.

8) Reliability Estimation Head

At the same time, the reliability estimation head produces a continuous confidence score indicating the trustworthiness of the diagnostic prediction. This branch functions on the identical shared representation and includes a fully connected layer succeeded by a sigmoid activation function to generate the reliability score [7,12,18].
The reliability score is characterized as:

$$r = \sigma(W_r z + b_r)$$

where: $\boldsymbol{z}$ is the shared representation, $\boldsymbol{W_r}$and $\boldsymbol{b_r}$are learnable parameters, and $\boldsymbol{\sigma}(\cdot)$denotes the sigmoid activation function.

In contrast to traditional uncertainty metrics calculated post-prediction, the reliability score is developed simultaneously with the diagnostic task and is fine-tuned to align with the accuracy of predictions [18]. Reduced reliability values signify uncertainty due to incomplete modality availability or unclear imaging characteristics, while increased reliability values reflect higher confidence in the diagnostic results [21,28,32]. This acquired reliability assessment allows the system to differentiate between assured predictions and instances that need further review.

9) Clinical Decision Logic

The concluding phase of the suggested framework is the Clinical Decision Logic module, where model results are converted into understandable and clinically relevant recommendations [3,32]. Figure 1 illustrates that this module integrates the predicted likelihood of malignancy and the trustworthiness score to inform decision-making in alignment with standard radiological practices.
Rather than generating a single binary diagnosis, the framework categorizes each case into a clinically relevant category that represents both diagnostic risk and prediction certainty. This approach permits deferring uncertain situations for further assessment instead of imposing possibly incorrect automated choices.
The decision-making process relies on two outputs from the model:
Probability of Malignancy

The diagnostic prediction head generates P("malignant"), symbolizing the estimated likelihood that the lesion is malignant.
*Reliability Score*The reliability estimation head generates a reliability score r∈[0,1], reflecting the model's confidence in its diagnostic output. By analyzing these two outputs together, the system minimizes the likelihood of overly confident decisions in uncertain situations.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026
DOI: 10.17148/IJARCCE.2026.154124

Categories of Decisions

Each instance is allocated to one of three clinically significant recommendation groups according to established probability and reliability criteria:

Cases that exhibit high reliability and a low likelihood of malignancy are categorized as Imaging Sufficient. In such situations, no urgent further imaging or invasive interventions are advised since the existing evidence indicates a low-risk evaluation.

Cases with low reliability or a moderate likelihood of malignancy are categorized as Follow-Up Suggested. These scenarios exemplify diagnostically ambiguous conditions where further imaging or short-term monitoring may be clinically suitable.

Instances with both a significant likelihood of malignancy and high reliability are categorized as Biopsy Advised. This recommendation indicates robust model evidence of cancer and endorses additional invasive diagnostic assessment per clinical practice guidelines. From an imaging informatics viewpoint, the clinical decision logic converts the suggested framework from a solely predictive model into an applicable decision-support system. Offering organized suggestions rather than standalone probability scores enhances interpretability, aids integration into clinical workflows, and promotes cooperation between humans and AI.

C. Training Strategies

A two-stage training approach is employed to achieve uniform multimodal representation learning and robust performance in less-than-ideal imaging conditions. This method reduces modality dominance and training instability by isolating multimodal fusion optimization from unimodal feature learning.

Step 1: Pretraining the Unimodal Encoder

Every modality-specific encoder is independently trained in the initial phase using solely the imaging data relevant to it. The mammography and ultrasound encoders are set up with ImageNet-pretrained ResNet-101 backbones and adjusted to acquire modality-adaptive representations crucial for breast cancer diagnosis[5,33,19].
Training the encoders separately enables each network to capture diagnostic cues specific to each modality without the influence of other modalities. This phase of mammography emphasizes overall structural patterns, such as architectural distortion and the distribution of tissue density[1,4,11]. Boundary irregularities and internal echogenicity are key characteristics of localized lesions that ultrasound training emphasizes. This phase generates robust and distinctive unimodal feature extractors, providing a solid basis for later multimodal learning[14,30].

Step 2: Multimodal Training with Random Modality Omission

During the second stage, the availability-aware fusion module, the shared latent representation, and the prediction heads are combined with the pretrained unimodal encoders. Multimodal training is subsequently conducted through random modality dropout, where one or more imaging modalities are randomly hidden during training at a specified probability. In every training iteration, the modality availability vector is randomly drawn to replicate real-world situations where either mammography or ultrasound data might not be accessible. If a modality is excluded, the related availability indicator is assigned a value of zero, and its input feature vector is substituted with a placeholder consisting entirely of zeros [15,25,26].

In multimodal training, the diagnostic prediction head and the reliability estimation head are simultaneously optimized. The main classification loss is merged with an auxiliary reliability loss to promote alignment between prediction confidence and the assessed reliability score. Instead of acting as random confidence values, this joint optimization promotes the reliability scores to represent the trustworthiness of predictions.
The suggested training approach offers various benefits:
• Training stability: Distinguishing unimodal from multimodal learning minimizes optimization complexity and safeguards feature learning from the destabilizing effects of early fusion.
• Modality robustness: Randomly dropping modalities exposes the model to various patterns of modality availability, enhancing generalization to actual clinical environments.
• Decreased modality bias: Even if one modality provides more information or is more often accessible during training, the framework minimizes preference for any particular modality.
• The overall loss is defined as:

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026
DOI: 10.17148/IJARCCE.2026.154124

$$L = L_{cls} + \lambda L_{rel}$$

where: L_{cls}= classification loss , L_{rel}= reliability loss , λ= weighting factor

This joint optimization enables the model to produce both accurate predictions and clinically meaningful confidence estimates.

D. Framing Clinical Decisions

Model outputs are translated into clinically meaningful recommendations based on the predicted malignancy probability and the reliability score:

1. High malignancy probability + high reliability: biopsy recommended.
2. Low malignancy probability + high reliability: current imaging considered sufficient.
3. Low reliability (regardless of probability): additional imaging or short-term follow-up recommended.

By deferring ambiguous cases instead of forcing binary decisions, this strategy prioritizes patient safety.

E Evaluation Metrics

To thoroughly assess diagnostic effectiveness, calibration accuracy, and resilience in incomplete imaging scenarios, the suggested framework was evaluated with various complementary metrics. These metrics were chosen to reflect both clinical dependability and predictive precision.

.1) Area Beneath the Receiver Operating Characteristic Curve (AUC)

The area beneath the receiver operating characteristic curve (AUC) served as the main measure for assessing diagnostic separation. AUC evaluates the model's capability to differentiate between benign and malignant instances at every possible classification threshold [7,12]. AUC is especially effective for medical diagnostic tasks where disease prevalence can change because, in contrast to accuracy, it is less affected by class imbalance and is independent of a specific decision threshold [8]. Greater AUC values signify improved discriminative capability.

2) Expected Calibration Error (ECE)

Expected Calibration Error (ECE) was employed to assess the calibration of the predicted probabilities of malignancy. Calibration indicates the degree to which predicted probabilities align with the frequencies of observed outcomes [7,18].

ECE is determined by segmenting predictions into confidence bins and evaluating the weighted average disparity between predicted confidence and actual accuracy for each bin. Reduced ECE values signify more accurately calibrated predictions, which are crucial for dependable clinical decision-making. Assessing ECE aids in establishing if the model's confidence assessments are reliable and relevant in a clinical context [12].

3) Accuracy Stratified by Reliability

To evaluate the efficacy of the suggested reliability estimation approach, predictions were categorized into high-reliability and low-reliability groups based on a preset reliability threshold. The accuracy of classification was subsequently computed independently for every group.

This analysis evaluates the precision of predictions recognized by the model as dependable against those categorized as uncertain. A significant performance disparity between the two groups suggests that the reliability score effectively represents diagnostic confidence instead of merely serving as an inconsequential supplementary output.

IV RESULTS

A Experimental Setup

All experiments utilized publicly accessible breast imaging datasets, such as CBIS-DDSM for mammography and BUSI for ultrasound [1,31]. The image-level data was split into training (70%), validation (15%), and testing (15%) groups. For a fair comparison, all models were trained for the same number of epochs with the same optimization settings. The assessment of performance was conducted by utilizing the area under the receiver operating characteristic curve (AUC), along with metrics based on calibration and reliability [7,18].

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154124

This part provides a thorough assessment of the suggested missing-modality-aware multimodal framework, focusing on diagnostic precision, resilience to partial data, calibration effectiveness, performance based on reliability, and results related to clinical decisions [18,29,32]. This assessment aimed to evaluate both the predictive accuracy and the clinical dependability of the suggested framework in realistic scenarios of missing modality.

B Diagnostic Performance

The suggested framework showed excellent diagnostic capability, reaching an area under the receiver operating characteristic curve (AUC) of 0.98 when both mammography and ultrasound were utilized. This performance surpassed that of traditional multimodal fusion models that rely on full modality availability and unimodal models developed using solely mammography or ultrasound [14,15,20,30].

These findings show that explicitly modeling modality availability and incorporating supplementary imaging data enhances the differentiation between benign and malignant instances. Interestingly, the advancement beyond basic concatenation and attention-based fusion benchmarks indicates that the performance enhancement is not merely a consequence of greater model capacity, but rather stems from the explicit management of modality presence and absence [18,29,32].

From a clinical informatics viewpoint, these results indicate that the suggested framework can efficiently employ multimodal imaging while staying flexible to practical clinical workflows. These results endorse the clinical viability of reliability-conscious multimodal learning in practical missing-modality scenarios.

Table 1. Diagnostic performance comparison

Model	**Modality Assumption**	**AUC**
Mammography-only ResNet-101	MG only	0.82
Ultrasound-only ResNet-101	US only	0.85
Simple concatenation fusion	MG + US (complete)	0.88
Attention-based fusion	MG + US (complete)	0.90
Proposed (Availability-aware, Ours)	MG ± US	**0.98**

Table 1 shows that the proposed framework achieved the highest diagnostic performance among all evaluated methods. This improvement suggests that explicitly incorporating modality availability and diagnostic reliability enhances multimodal breast cancer classification under realistic clinical conditions

C Robustness to Missing Modalities

Experiments involving missing-modality ablation were performed to assess the framework's resilience in situations with incomplete imaging. The model preserved strong diagnostic capability even when one imaging technique was missing, with just a slight decrease in AUC relative to the complete multimodal scenario [21,28,30].

This slight reduction suggests that the availability-aware fusion approach can adjust to missing data without generating unpredictable forecasts.

As anticipated, model performance dropped to chance level when neither imaging modality was present. This behavior demonstrates that the availability-handling mechanism is working properly and that predictions are influenced by significant imaging data instead of bias or artifacts specific to the dataset.

These findings indicate that the framework is appropriate for use in various clinical environments, where access to imaging can differ among patients and facilities. This resilience is crucial for implementing multimodal AI systems in everyday clinical practice, where comprehensive imaging data may not always be accessible.

Table 2. Performance under missing modalities

Modality Availability	**AUC**
Mammography only	0.8443
Ultrasound only	0.8995
Mammography + Ultrasound	**0.9742**
No modality	0.5036

The availability-aware architecture is correct since the model smoothly deteriorates when one modality is unavailable and collapses to chance-level performance when both modalities are absent.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154124

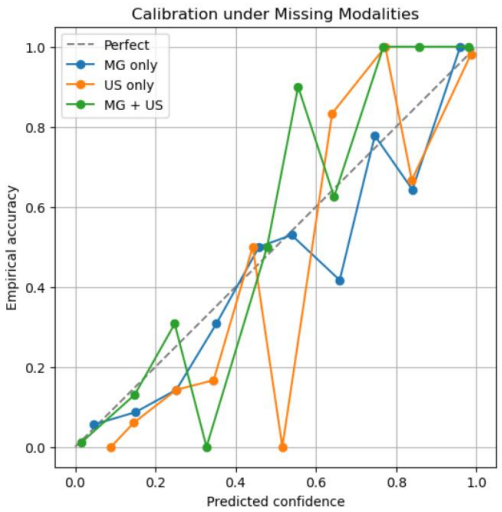

Figure 4: Calibration curves for mammography-only, ultrasound-only, and multimodal inputs.

The calibration behaviour of the suggested architecture is shown in Figure 4 under various modality availability circumstances, such as integrated multimodal input (MG + US), ultrasound-only (US only), and mammography-only (MG only). Perfect calibration, when anticipated confidence precisely equals empirical accuracy, is shown by the dashed diagonal line. By dividing forecasts into confidence intervals and calculating the observed accuracy inside each bin, each curve displays the correlation between anticipated confidence and empirical accuracy.

D Analysis of Calibration

Calibration analysis indicated that multimodal predictions generated by the proposed framework were more accurately calibrated than unimodal predictions, in addition to enhancements in diagnostic discrimination. In contrast to unimodal situations, the Expected Calibration Error (ECE) was reduced by almost 50% when both mammography and ultrasound were utilized [7,18].

In risk-based clinical decision-making, enhanced calibration suggests that estimated malignancy probabilities better represent actual outcome rates [3,18]. Poorly calibrated models might display excessive confidence or insufficient confidence, resulting in unsuitable clinical choices [7].

The enhancements in calibration indicate that incorporating additional imaging data and explicitly modelling the availability of modalities improves both the reliability of probabilities and the accuracy of diagnoses.

Table 3: Calibration under different modality settings

Modality Availability	**ECE ↓**
Mammography only	0.0946
Ultrasound only	0.0739
Mammography + Ultrasound	**0.0496**

Table 3 shows that Combining complementary imaging data increases both accuracy and probability dependability, as evidenced by the significantly better calibration of multimodal predictions compared to unimodal predictions.

E. Reliability-Based Accuracy

To assess the clinical importance of the anticipated reliability scores, test cases were categorized into high-reliability and low-reliability groups. The accuracy for high-reliability predictions was 94%, while low-reliability predictions had an accuracy of 71%.

This significant performance gap suggests that the reliability estimation head reflects diagnostic confidence instead of simply replicating prediction probability. Predictions with high reliability were significantly more precise, while those with low reliability were linked to increased uncertainty and a greater error rate.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026
DOI: 10.17148/IJARCCE.2026.154124

These results reinforce the framework's reliability-driven decision-making process and confirm the appropriateness of utilizing the reliability score to steer subsequent clinical choices. This finding shows that reliability estimation can help enhance safer uncertainty-aware clinical decision making.

Table 4: Reliability-based accuracy stratification

Reliability Group	**Accuracy**
High reliability (≥ 0.7)	**0.94**
Low reliability (< 0.7)	0.71

Table 4 shows that the reliability score gives clinically meaningful confidence estimation, as evidenced by the significantly higher accuracy of predictions marked as reliable compared to low-reliability predictions.

F. Clinical Recommendation Results

The suggested decision logic transformed model outputs into recommendations that are clinically interpretable. According to the estimated likelihood of malignancy and trustworthiness score, each case was classified into one of three actionable groups: imaging adequate, follow-up advised, or biopsy suggested.

Cases assessed as low risk and highly reliable were categorized as having adequate imaging, suggesting the possibility of minimizing unnecessary follow-up imaging and related expenses. Instances with elevated predicted risk and high confidence were marked for biopsy due to compelling indications of cancer. To ensure patient safety, cases deemed low in reliability were postponed for further follow-up instead of being given a conclusive classification.
This behaviour aligns with actual radiological practice and shows that the suggested framework operates as a clinical decision-support system instead of merely a predictive model. These findings illustrate the capability of reliability-aware multimodal AI to enhance safety and provide more clinically actionable evaluations for breast cancer.

model outputs: 1. {'probability': 0.63,
'reliability': 0.41,
'decision': 'Additional imaging recommended'}
2. Indeterminate – short-term follow-up advised: 58
Low suspicion – current imaging sufficient: 103
High suspicion of malignancy – biopsy recommended: 74
3. {probability': 0.2869662046432495, 'reliability': 0.9725335836410522, 'decision': 'Indeterminate – short-term follow-up advised'}
4. {'probability': 0.008555523119866848, 'reliability': 0.9846733808517456, 'decision': 'Low suspicion – current imaging sufficient'}
5. {'probability': 0.07388029992580414, 'reliability': 0.9766407608985901, 'decision': 'Low suspicion – current imaging sufficient'}
6. {'probability': 0.15421432256698608, 'reliability': 0.9728609323501587, 'decision': 'Indeterminate – short-term follow-up advised'}
7. {'probability': 0.9976533055305481, 'reliability': 0.9968603849411011, 'decision': 'High suspicion of malignancy – biopsy recommended'

G. Clinical Consequences

The suggested framework facilitates secure clinical implementation by preserving strong diagnostic effectiveness when one modality is not accessible and by generating precise probability assessments. It additionally postpones ambiguous cases for further imaging instead of imposing possibly inaccurate predictions.

This behaviour supports the incorporation of reliability-guided multimodal AI systems into clinical workflows and closely mirrors actual radiological decision-making [9].

V DISCUSSION

This study's results suggest that the clinical effectiveness of multimodal imaging AI systems can be enhanced by clearly modelling the availability of modalities and the reliability of diagnostics. In contrast to traditional fusion systems that rely on full multimodal input, the suggested framework maintains its robustness in situations with incomplete imaging and produces interpretable confidence estimates that correspond with actual radiological practice.
A significant discovery is that availability-aware fusion allows for smooth performance decline when a particular imaging modality is missing, instead of a complete breakdown. This conduct is crucial for implementation in standard clinical

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026
DOI: 10.17148/IJARCCE.2026.154124

practice, as workflow limitations, patient demographics, and institutional resources might restrict access to mammography or ultrasound. By explicitly incorporating the presence of modality, the framework tailors its reasoning to the accessible data, reflecting how radiologists modify their diagnostic confidence based on the studies at hand.

From an imaging informatics viewpoint, merging reliability estimation with reliability-informed decision-making signifies a crucial advance for enhancing safety in clinical decision support. The noted disparity in performance between high- and low-reliability predictions indicates that the acquired reliability scores genuinely represent diagnostic assurance. This feature enhances human–AI collaboration by permitting the system to postpone uncertain situations for more imaging or follow-up rather than producing potentially untrustworthy automated outcomes.
The framework holds clinical significance as it offers decision support that considers risks and uncertainties instead of concentrating only on predictive precision. This behaviour is crucial for clinical acceptance since radiologists often handle uncertainty and place patient safety above complete automation.

A. Limitations

Even though the outcomes are encouraging, various limitations of this research need to be acknowledged when analyzing the results.
Initially, this research did not utilize patient-matched datasets of mammography and ultrasound across different imaging modalities. While the present design facilitated the evaluation of modality-availability management, subsequent research utilizing fully paired multimodal datasets would permit a more thorough assessment of patient-level fusion and diagnostic reliability.
Secondly, the assessment was conducted at the image level instead of the patient or examination level. In clinical practice, diagnostic choices are usually grounded in various imaging perspectives, ongoing assessments, and relevant patient details. Consequently, evaluation at the image level might not capture the complete clinical advantage of multimodal integration in standard diagnostic procedures.
Third, the thresholds applied in the reliability-guided clinical decision framework were chosen based on empirical evidence. Even though these thresholds distinguished dependable predictions from those that are not, additional refinement and validation remain necessary. Determining ideal threshold values in various clinical environments and patient groups would necessitate prospective clinical research and validation across multiple institutions.
Ultimately, this research depended on datasets that are publicly accessible. While public datasets enhance reproducibility and transparency, it is essential to conduct external validation with institution-specific data and various imaging devices to evaluate generalizability across healthcare systems.

B, Implications of Future Research
The constraints noted in this research also emphasize various avenues for upcoming studies. Enhancing diagnostic performance and personalization could be achieved by expanding the framework to include patient-level modelling, utilizing longitudinal imaging data, and merging clinical metadata like patient age, breast density, and BI-RADS evaluations. Additionally, future clinician-in-the-loop research is essential to assess the impact of reliability-guided decision support on imaging usage, diagnostic efficiency, and patient outcomes in actual clinical environments.

IV. CONCLUSION

This research introduces a multimodal deep learning system for breast cancer decision support that explicitly accounts for imaging accessibility and diagnostic accuracy when certain modalities are absent. The proposed system attains excellent diagnostic performance by integrating modality-specific encoders, availability-aware fusion, and reliability-guided clinical decision logic, while also being resilient to incomplete imaging scenarios.

In addition to enhancing diagnostic distinction, the system offers understandable reliability scores and accurately calibrated probability estimates that facilitate the secure deferral of ambiguous cases. This behaviour, which accounts for uncertainty, enhances human–AI collaboration and closely aligns with actual radiological workflows. These results indicate that reliability-conscious multimodal AI could provide a practical basis for safer and more flexible breast cancer decision support in everyday clinical settings

REFERENCES

[1] Byra, Michal. "*Breast mass classification with transfer learning based on scaling of deep representations*." Biomedical Signal Processing and Control 69 (2021): 102828.
[2] Wang, Lu, et al. "*Trends in the application of deep learning networks in medical image analysis: Evolution between 2012 and 2020*." European journal of radiology 146 (2022): 110069.
[3] Yala, Adam, et al. "*Optimizing risk-based breast cancer screening policies with reinforcement learning.*" Nature

DOI: 10.17148/IJARCCE.2026.154124

medicine 28.1 (2022): 136-143.
[4] Samee, Nagwan Abdel, et al. "*Deep learning cascaded feature selection framework for breast cancer classification: Hybrid CNN with univariate-based approach*." Mathematics 10.19 (2022): 3631.
[5] Li, Mengfang, et al. "*Medical image analysis using deep learning algorithms*." Frontiers in public health 11 (2023): 1273253.
[6] McKinney, Scott Mayer, et al. "*International evaluation of an AI system for breast cancer screening*." *Nature* 577.7788 (2020): 89-94.
[7] Verbakel, Jan Y., et al. "*ROC curves for clinical prediction models part 1. ROC plots showed no added value above the AUC when evaluating the performance of clinical prediction models*." Journal of Clinical Epidemiology 126 (2020): 207-216.
[8] Samawi, Hani M., et al. "*Kullback-Leibler divergence for medical diagnostics accuracy and cut-point selection criterion: How it is related to the Youden index*." J. Appl. Bioinform. Comput. Biol 9 (2020): 2.
[9] Yang, Zhenyu, et al. "*AlphaFold2 and its applications in the fields of biology and medicine*." Signal Transduction and Targeted Therapy 8.1 (2023): 115.
[10] Liu, Wanli, et al. "*CVM-Cervix: A hybrid cervical Pap-smear image classification framework using CNN, visual transformer and multilayer perceptron*." Pattern Recognition 130 (2022): 108829.
[11] Ali, Nairveen, et al. "Automatic label-free detection of breast cancer using nonlinear multimodal imaging and the convolutional neural network ResNet50." *Translational Biophotonics* 1.1-2 (2019): e201900003.
[12] Zhan, Yuejuan, et al. "*Diagnostic accuracy of the artificial intelligence methods in medical imaging for pulmonary tuberculosis: a systematic review and meta-analysis*." Journal of Clinical Medicine 12.1 (2022): 303.
[13] Zhang, Yue, et al. "*Multi-modal tumor segmentation with deformable aggregation and uncertain region inpainting*." IEEE Transactions on Medical Imaging 42.10 (2023): 3091-3103.
[14] Chen, Junwei, et al. "*A deep learning-based multimodal medical imaging model for breast cancer screening*." Scientific Reports 15.1 (2025): 14696.
[15] Kayikci, Safak, and Taghi M. Khoshgoftaar. "*Breast cancer prediction using gated attentive multimodal deep learning*." Journal of Big Data 10.1 (2023): 62.
[16] Nakach, Fatima-Zahrae, Ali Idri, and Evgin Goceri. "*A comprehensive investigation of multimodal deep learning fusion strategies for breast cancer classification*." Artificial Intelligence Review 57.12 (2024): 327.
[17] Kayikci, Safak, and Taghi M. Khoshgoftaar. "Breast cancer prediction using gated attentive multimodal deep learning." *Journal of Big Data* 10.1 (2023): 62.
[18] Park, Seong Ho, et al. "*Methods for clinical evaluation of artificial intelligence algorithms for medical diagnosis*." Radiology 306.1 (2023): 20-31.
[19] Guo, Dinghao, et al. "A multimodal breast cancer diagnosis method based on knowledge-augmented deep learning." Biomedical Signal Processing and Control 90 (2024): 105843.
[20] Al-Tam, Riyadh M., et al. "*Multimodal breast cancer hybrid explainable computer-aided diagnosis using medical mammograms and ultrasound Images*." Biocybernetics and Biomedical Engineering
[21] Jiang, Meiping, et al. "*Multimodal imaging of target detection algorithm under artificial intelligence in the diagnosis of early breast cancer*." Journal of Healthcare Engineering 2022.1 (2022): 9322937.
[22] Maruf, Nazmul Ahasan, Abdullah Basuhail, and Muhammad Umair Ramzan. "*Enhanced Breast Cancer Diagnosis Using Multimodal Feature Fusion with Radiomics and Transfer Learning*." Diagnostics 15.17 (2025): 2170.
[23] Wu, Renjie, et al. "Deep multimodal learning with missing modality: A survey." *arXiv preprint arXiv:2409.07825* (2024).
.[24] Yala, Adam, et al. "*A deep learning mammography-based model for improved breast cancer risk prediction*." Radiology 292.1 (2019): 60-66.
[25] Wen, Jinyu, et al. "*MsgFusion: Medical semantic guided two-branch network for multimodal brain image fusion*." IEEE Transactions on Multimedia 26 (2023): 944-957.
[26] Liu, Hui, et al. "*DDIFN: A dual-discriminator multi-modal medical image fusion network*." ACM Transactions on Multimedia Computing, Communications and Applications 19.4 (2023): 1-17.
[27] Zhou, Yuting, et al. "*Multimodal medical image fusion network based on target information enhancement*." IEEE Access 12 (2024): 70851-70869.
[28] Xu J., Yang G., Zhang Y., et al. "*Multimodal feature learning with adaptive weighting*." *Pattern Recognition*, 2022.
[29] Rahman, Md Atiqur, et al. "*Advancements in breast cancer detection: a review of global trends, risk factors, imaging modalities, machine learning, and deep learning approaches*." BioMed Informatics 5.3 (2025): 46.
[30] Zhang Q., Li K., Wang H., et al. "*Multimodal breast cancer classification using deep learning*." Biomedical Signal Processing and Control, 2023.
[31] Qureshi, Shahzad Ahmad, et al. "*Breast Cancer Detection using Mammography: Image Processing to Deep Learning.*" IEEE Access (2024).
[32] Ahn, Jong Seok, et al. "*Artificial intelligence in breast cancer diagnosis and personalized medicine*." Journal of

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026
DOI: 10.17148/IJARCCE.2026.154124

breast cancer 26.5 (2023): 405.
[33] Zhang, Qi. "*A novel ResNet101 model based on dense dilated convolution for image classification.*" SN Applied Sciences 4.1 (2022): 9.