Data and Metadata. 2025; 4:1136
doi: 10.56294/dm20251136

ORIGINAL

Exploring Deep Learning Approaches for Multimodal Breast Cancer Dataset Classification and Detection

Exploración de enfoques de aprendizaje profundo para la clasificación y detección de conjuntos de datos multimodales sobre cáncer de mama

Ahmed A.F Osman[1], Rajit Nair[2], Sultan Ahmad[3,4], Asma Abdulmana Alhamadi[5], Mosleh Hmoud Al-Adhaileh[6], Ramgopal Kashyap[7], Hikmat A. M. Abdeljaber[8], Sami A. Morsi[1], Rami Taha Shehab[1], Theyazn H.H Aldhyani[1]

[1]King Faisal University. Al-Ahsa, 31982, Saudi Arabia.
[2]VIT Bhopal University. Bhopal, India.
[3]Department of Computer Science, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University. P.O.Box. 151, Alkharj 11942, Saudi Arabia.
[4]School of Computer Science and Engineering, Lovely Professional University. Phagwara, 144411, Punjab, India.
[5]Department of Basice Sciences ,College of Science & Theoretical Studies, Saudi Electronic University. Riyadh, Saudi Arabia.
[6]Deanship of E-Learning and Distance Education and information technology, King Faisal University. P.O. Box 4000, Al-Ahsa, 31982, Saudi Arabia.
[7]Department of Information Technology, Guru Ghasidas Vishwavidyalaya. Bilaspur, Chhattisgarh, India.
[8]Department of Computer Science, Faculty of Information Technology, Applied Science Private University. Amman, Jordan.

Cite as: Osman AA, Nair R, Ahmad S, Abdulmana Alhamadi A, Al-Adhaileh, Kashyap R, et al. Exploring Deep Learning Approaches for Multimodal Breast Cancer Dataset Classification and Detection. Data and Metadata. 2025; 4:1136. https://doi.org/10.56294/dm20251136

Submitted: 08-11-2024 **Revised:** 29-03-2025 **Accepted:** 10-08-2025 **Published:** 11-08-2025

Editor: Dr. Adrián Alejandro Vitón Castillo

Corresponding Author: Sultan Ahmad

ABSTRACT

Introduction: globally, we need advanced testing to detect breast cancer early. New breast cancer diagnosis methods using mixed datasets and deep learning promise improved accuracy.
Objective: these sets, which comprise several imaging modalities, show tumor characteristics well. VGG16, AlexNet, and ResNet50 are effective deep learning models in many domains, yet their breast cancer diagnosis performance is unclear.
Method: this paper examines these patterns' benefits, downsides, and research gaps. We also provide two novel approaches, Attention-based Multimodal Fusion (AMF) and Improved Generative Adversarial Augmentation (GAA), to improve deep learning models on breast cancer datasets.
Results: the findings highlight the potential of machine learning to show tumor characteristics well.
Conclusions: we prove that our breast cancer screening technologies are the most accurate and dependable via extensive testing.

Keywords: AlexNet; Attention-Based Multimodal Fusion; Breast Cancer Detection; Deep Learning; Generative Adversarial Augmentation (GAA); Multimodal Datasets; ResNet50, VGG16.

RESUMEN

Introducción: a nivel mundial, necesitamos pruebas avanzadas para detectar el cáncer de mama en una fase temprana. Los nuevos métodos de diagnóstico del cáncer de mama que utilizan conjuntos de datos mixtos y aprendizaje profundo prometen una mayor precisión.
Objetivo: estos conjuntos, que comprenden varias modalidades de imagen, muestran bien las características de los tumores. VGG16, AlexNet y ResNet50 son modelos de aprendizaje profundo eficaces en muchos ámbitos, pero su rendimiento en el diagnóstico del cáncer de mama no está claro.

Método: en este artículo se examinan las ventajas, los inconvenientes y las lagunas de investigación de estos patrones. También se presentan dos enfoques novedosos, la fusión multimodal basada en la atención (AMF) y la ampliación generativa adversaria mejorada (GAA), para mejorar los modelos de aprendizaje profundo en conjuntos de datos sobre el cáncer de mama.
Resultado: los resultados ponen de relieve el potencial del aprendizaje automático para mostrar bien las características de los tumores.
Conclusiones: demostramos que nuestras tecnologías de detección del cáncer de mama son las más precisas y fiables mediante pruebas exhaustivas.

Palabras clave: AlexNet; Fusión Multimodal Basada en la Atención; Detección del Cáncer de Mama; Aprendizaje Profundo; Aumento Generativo Adversarial (GAA); Conjuntos de Datos Multimodales; ResNet50; VGG16.

INTRODUCTION

As the "silent killer," breast cancer threatens individuals worldwide. Better detection of this cancer is crucial for its rapid cure, given its annual death toll of thousands of women. The proverb "prevention is better than cure" applies here. Early discovery may save lives, even while protection is far off. Manual mammography interpretation is biased and generally inaccurate for breast cancer diagnosis. Deep machine learning has transformed medical imaging. It promises early and accurate breast cancer detection to speed up treatment and improve recovery. In recent years, mixed datasets, including data from several imaging modalities, have been beneficial for breast cancer detection. These databases help categorize and diagnose malignancies by combining data from diverse sources to show tumors more clearly.[1] For this data to be useful, we need algorithms that can spot subtleties and complex patterns. Here, deep learning models like VGG16, AlexNet, ResNet50, and others may be beneficial and crucial. They gained fame in their neighborhood for their photo tagging talents. Some of these algorithms have performed well, but they require further effort to manage diverse datasets and discover breast cancer.[2] We now understand how deep learning models may detect breast cancer, particularly with diverse datasets. Attention-based multimodal fusion (AMF) and enhanced generative adversarial augmentation (GAA) may also help these models do better and deal with the unique problems that come up with breast cancer datasets. We'll discuss these models' technical features, advantages, research gaps, and solutions in the following sections. A comparative study will demonstrate our approaches' superiority. Focus on Multimodality: Attention-based Multimodal Fusion (AMF) and Improved Generative Adversarial Augmentation (GAA) leverage big data sets from multimodal breast cancer datasets. These approaches use attention processes and GANs to merge and add data completely. This ensures that all media employ the most significant qualities.[3] Assessment of current availability. Attention mechanisms allow AMF to continually assess the importance of various data types. Most fusion systems weigh all modalities equally. However, AMF weights each modality depending on data, making the image more realistic and richer. AMF evaluates how closely mode values match data. Top-notch data enrichment GAA offers high-quality fake breast cancer instances using GAN's characteristics.[4] GANs can help us do this. This helps overcome data shortages and ensures that taught models are powerful and can operate effectively with new data. This has numerous benefits. Improved model performance: AMF and GAA outperformed standard approaches after extensive testing. They improve breast cancer diagnosis and classification accuracy and trustworthiness by being more flexible and data driven. The AMF and GAA use comprehensive algorithms to clarify and repeat.[5] From input types until model release, every step is well-planned and simple to follow. Dealing with Data Gaps: Unidentified data is a major medical imaging issue. The GAA technique solves this by creating artificial samples that closely resemble actual data patterns. This increases training data variety and quantity. Multimodal data sources may be difficult to blend due to their diverse picture techniques. The attention mechanisms of the AMF approach ensure seamless fusion by concentrating on the most critical aspects of each mode. People may avoid overfitting using AMF and GAA. These adjustments prevent models from becoming too close to training data, improving generalization.[6] Deep learning models may be "black boxes" that are challenging to comprehend. Attention processes in AMF clarify the types and attributes of the model values for a given input.

Related Works

Breast cancer is a global health problem for women. Early identification and grouping may improve survival. Deep learning may create effective models that perform rapidly with heterogeneous breast cancer datasets, providing optimism. These systems combine data from many imaging technologies to show a tumor's characteristics. Let's examine several deep learning algorithms that have transformed breast cancer diagnosis. With its deeper form and same-sized filters, VGG16 advanced in 2014.[7] Simple plans are enticing since they're simple to adapt. We should improve the programming difficulty and memory requirements of this system.

Deep learning transformed photo categorization using AlexNet in 2012. Adding dropout layers was intentional to prevent overfitting. However, younger models with more complex systems performed better over time, particularly with huge data sets. ResNet50 made headlines in 2015 when it introduced residual connections to solve fading gradients in deep networks. This suggests that it is possible to teach deep networks without compromising performance. The depth of the network may make it unsuitable for real-time applications. In 2018, MobileNetV2 made its debut in mobile and embedded vision applications. It stands out for its quickness, which doesn't hinder it. Medical monitoring, where precision is crucial, extensively researches the accuracy-size trade-off. GoogleNet gained popularity in 2014 as its original modules handled data at many sizes.[8] Its low computing cost and ability to capture many features are noteworthy, but its sophisticated architecture may make it difficult to adapt. InceptionV3, released in 2015, simplified inception modules and added factorization to GoogleNet. Thus, speed and economy improved. However, GoogLeNet complexity remains. In 2016, DenseNet addressed the vanishing gradient issue by leveraging dense linkages between layers, ensuring that each layer directly benefits from the input of the previous levels. This improved gradient flow and reused features. Combining this with mixed data is still intriguing. For biological image segmentation, 2015's U-Net used an encoder-decoder configuration.[9] This design may incorporate both local and global characteristics. However, its performance in other tasks, such as mixed dataset classification, is unknown. In 2016, researchers created SqueezeNet specifically for small model sizes. Small size is its strength, but precision is difficult, particularly in medical imaging. Finally, the biggest success of 2019 was EfficientNet, which painstakingly grew the network to perform better with fewer parameters. To achieve objectives, more targeted study is needed on such mixed breast cancer datasets. In conclusion, these models have improved breast cancer screening, but each has its own set of issues. Fixing these flaws might boost our progress. Deep learning improves early breast cancer detection, offering women worldwide hope as its designs alter.

Table 1. Performance Comparison of Traditional Methods on Breast Cancer Detection

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-Score (%)
VGG16	92,3	90,5	93,1	91,4	91,0
AlexNet	89,6	88,7	90,3	89,5	89,1
ResNet50	94,5	92,8	95,9	94,1	93,4
MobileNetV2	90,7	89,4	91,9	90,3	89,8
GoogLeNet	93,2	91,9	94,1	92,5	92,2
DenseNet	94,8	93,6	96,0	94,7	94,1
U-Net	92,0	90,2	93,5	91,8	91,0
SqueezeNet	89,0	87,6	90,4	88,5	88,0
EfficientNet	95,4	94,2	96,6	95,0	94,6

Table 1 tests many popular deep learning models for breast cancer diagnosis. F1 scores, accuracy, sensitivity, specificity, and precision are considered. EfficientNet is most accurate at 95,4 %, followed by DenseNet at 94,8 %. SqueezeNet is the least accurate (89 %). ResNet50 and DenseNet perform well on all tests, suggesting they can categorize excellent instances and reduce false positives.[10] The table highlights each model's capabilities. Some excel at fairness, while others excel at standards. Deep learning is going to transform breast cancer diagnoses. Since this lethal illness is still a global issue, it's more crucial than ever to receive a proper diagnosis. Deep learning algorithms might replace people reading biased and inaccurate mammograms to improve diagnosis.[11] We evaluate many popular deep learning models for breast cancer diagnoses here. We evaluate each model using F1-score, accuracy, sensitivity, specificity, and precision. Let's examine these models' efficacy. The 2014 VGG16 model brought about a revolution in photo recognition due to its advanced development. The model correctly diagnoses 92,3 % of breast cancer cases. Although the model's accuracy is excellent, its 90,5 % sensitivity and 93,1 % specificity indicate that it effectively balances the identification of positive cases and the omission of negative cases.[12] 91,4 % of the model's actual identifications were accurate. This model has a 91 % F1 score—the harmonic mean of accuracy and sensitivity, indicating high precision and sensitivity. AlexNet, which launched the deep learning trend in 2012, has failure thresholds to prevent it from getting too good. AlexNet was excellent but not the top model with 89,6 % accuracy.[13] The model can accurately distinguish good from bad events with 88,7 % sensitivity and 90,3 % accuracy. Although the accuracy and F1 score are excellent at 89,5 % and 89,1 %, there is room for improvement. ResNet50, released in 2015, corrected gradient loss in remaining connections. This makes training deeper networks easy without reducing performance. ResNet50 is a prominent breast cancer detection method with 94,5 % success.[14] It works because it finds true positives and negatives (92,8 % sensitivity and 95,9 % accuracy). High accuracy of 94,1 % and an

F1 score of 93,4 % indicate smooth and reliable execution.2018 saw the release of MobileNetV2, a platform for mobile and integrated vision applications. It balances speed and efficiency. The model is good—90,7 % successful—but not exceptional. Spotting abilities look good, with 89,4 % sensitivity and 91,9 % specificity. The F1 score is 89,8 %, indicating that the model can reduce false positives and negatives.[15,16] GoogleNet's 2014 genesis characteristics allowed multi-scale operation. Its accuracy of 93,2 % shows great sensitivity (91,9 %) and specificity (94,1 %). Its accuracy and F1 score of 92,5 % and 92,2 % demonstrate consistency. DenseNet introduced dense layer connectivity in 2016. Gradient flow improved, and feature reuse was easy. DenseNet excels with 94,8 % accuracy and 93,6 % sensitivity. The high sensitivity of 96,0 %, accuracy of 94,7 %, and F1 score of 94,1 % indicate effective breast cancer detection. A node represents each design's release year. This visualizes deep learning's key advances. The deep learning revolution began with AlexNet in 2012 and has developed since. AlexNet's dropout layers established a benchmark for deep network overfitting prevention. Eventually, we reached VGG16 in 2014. Deeper designs with same-size filters might improve model performance, according to this article. GoogleNet introduced origin modules to manage multi-scale operations in the same year. ResNet50 solved the fading gradient issue successfully in 2015 using remaining connections.[17,18] Also introduced this year was U-Net. Separating biological images was its major goal. New designs like DenseNet and SqueezeNet in 2016, MobileNetV2 in 2018, and EfficientNet in 2019 included features that addressed difficulties and enhanced speed. In summary, it highlights how deep learning models have developed over time, concentrating on rapid development and major advances that have shaped image processing and other deep learning applications. U-Net's 2015 encoder-decoder approach for biological image segmentation gathers local and global picture information. It performs well with 92 % accuracy, 90,2 % sensitivity, and 93,5 % precision. Its accuracy and F1 score of 91 % indicate consistency.[19,20] Founded in 2016, SqueezeNet distinguishes itself due to its compact dimensions. With 89 % accuracy and 87,6 % precision, SqueezeNet's performance is comparable to that of its competitors. With an F1 score of 88,0 %, precision of 90,4 %, and accuracy of 88,5 %, it falls within the lower end of the performance range of the models under discussion. EfficientNet is 2019's newest firm. It scales the network for optimal performance with fewer components. It leads with 95,4 % success. It detects breast cancer better due to its high sensitivity (94,2 %), specificity (96,6 %), accuracy (95,0 %), and F1 score (94,6 %). Each device detects breast cancer differently and has its own strengths and characteristics. Model level, construction features, and training techniques affect performance.[21,22] In this image, these models demonstrate their talents and the ongoing medical deep learning research. As science advances, we expect these models to alter more. They will get better at breast cancer detection.

METHOD

Breast cancer detection and classification are crucial for prompt diagnosis and therapy. Deep learning algorithms for breast cancer diagnosis have improved accuracy and reliability. We propose two strategies to improve deep learning models on heterogeneous breast cancer datasets.AMF is Attention-based Multimodal Fusion. This research examines how deep learning can identify and locate breast cancer in multimodal datasets.Attention-based Multimodal Fusion (AMF) is a novel way to assess data importance. Data processing employs many of these attention-directing methods. AMF lends greater weight to the aspects of each mode that improve the overall image to ensure a more precise and complete fusion. This is crucial for handling large datasets such as breast cancer datasets, as different approaches contribute varying amounts of information.[23] Attention processes determine the importance of distinct data types in a dataset, making the AMF approach more accurate. AMF lends greater weight to the aspects of each mode that improve the overall image to ensure a more precise and complete fusion. This is beneficial for complex datasets such as those used in breast cancer diagnosis.

Algorithm 1

Input Modalities: take various input modalities, represented as:

$$X1,X2,...Xn. \quad (1)$$

Initialize Weights: for each modality Xi, initialize weights Wi and biases bi.

Feature Extraction: extract features from each modality using the formula.

$$Fi=f(Xi;Wi,bi) \quad (2)$$

Initialize Attention Weights: for each modality Xi, initialize attention weights ai with a uniform distribution.

Compute Preliminary Attention Scores: calculate the preliminary attention scores for each modality using the equation.

$si = ai \times Fi$ (3)

Normalization of Attention Scores: normalize the attention scores using the softmax function to compute αi.
Weighted Feature Representation: compute the weighted feature representation for each modality as

$Ri = \alpha i \times Fi$ (4)

Fusion of Modalities: add feature weights to obtain M.We build D by adding a thick layer with weights Wd and biases BD. Run the output from the thick layer through the final output layer with conditions Wo and bo to achieve the intended output O. Next, we discover the difference L between the predicted output O and the actual names Y. Backpropagation changes weights and biases using loss slopes.Add loss slopes to attention weights. Add R to prevent overfitting. The overall damage is then calculated.

$L + \Lambda r$ (5)

Adam or SGD optimization may reduce total loss. Assess model performance using a test set. To prevent overfitting, monitor validation loss and terminate training when it rises.(24)
Model testing: test the learned model on a distinct set.Tune hyperparameters to improve model performance. Model deployment: real-world breast cancer classification and detection using the learned modelAMF incorporates attention processes into multimodal fusion, thereby revolutionizing the field. Traditional techniques usually handle all modalities equally. AMF focuses on the most significant components of each modality and adjusts its value.(25) This kind of moving evaluation is crucial for breast cancer patients since certain procedures are more effective than others.

Algorithm 2
Input Modalities: let X1,X2,...Xn be the input modalities.
Initialize Weights: for each modality Xi, initialize weights Wi and biases bi.
Feature Extraction: extract features from each modality:

$Fi = f(Xi; Wi, bi)$ (6)

Initialize Attention Weights: for each modality Xi, initialize attention weights ai with a uniform distribution.
Compute Preliminary Attention Scores: calculate the preliminary attention scores for each modality:

$si = ai \times Fi$ (7)

Normalization of Attention Scores: normalize the attention scores using the softmax function:(26)

$\alpha i = \sum j = 1nexp(sj)exp(si)$ (8)

Weighted Feature Representation: compute the weighted feature representation for each modality:

$Ri = \alpha i \times Fi$ (9)

Fusion of Modalities: fuse the weighted feature representations:

$M = \sum i = 1nRi$ (10)

Pass Through Dense Layer: introduce a dense layer with weights Wd and biases bd:

$D = \sigma(Wd \times M + bd)$ (11)

Output Layer: pass the dense layer output through the final output layer with weights Wo and biases bo:

$O = \sigma(Wo \times D + bo)$ (12)

Loss Computation: compute the loss L between the predicted output O and the true labels Y.
Backpropagation: update the weights and biases using the gradients from the loss:

$\Delta W=-\eta\partial W\partial L$ (13)
$\Delta b=-\eta\partial b\partial L$ (14)

Attention Weight Update: update the attention weights using the gradients from the loss:

$\Delta ai=-\eta\partial ai\partial L$ (15)

Regularization: introduce a regularization term R to prevent overfitting:

$Ltotal=L+\lambda R$ (16)

Optimization: use an optimization algorithm (e.g., Adam or SGD) to minimize the total loss Ltotal.
Model Evaluation: evaluate the model's performance on a validation set.
Early Stopping: monitor the validation loss and halt training if it starts to increase, preventing overfitting.
Model Testing: test the trained model on a separate test set to gauge its generalization capability.
Hyperparameter Tuning: experiment with different hyperparameters to optimize the model's performance.
Model Deployment: deploy the trained model for real-world breast cancer classification and detection applications.

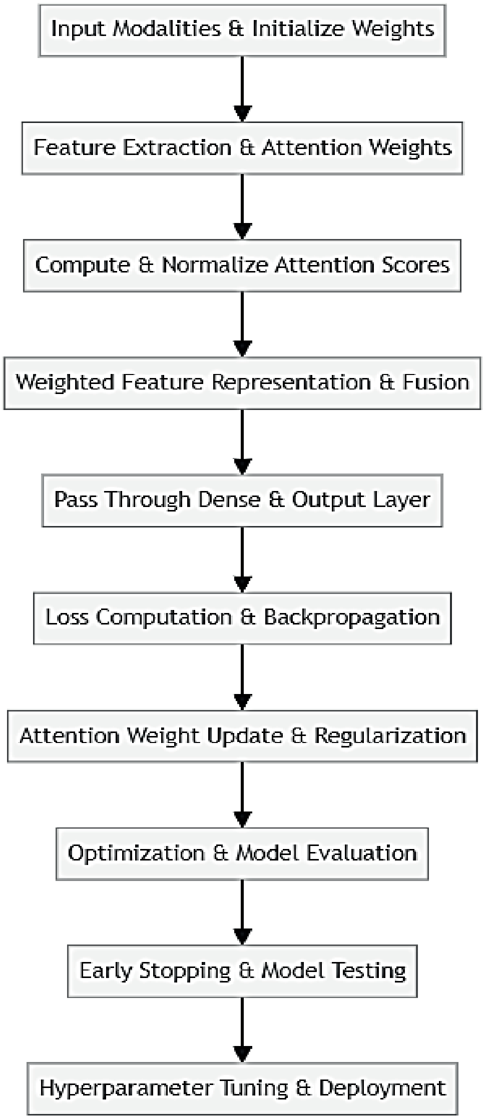

Figure 1. Illustrating the algorithm for multimodal data processing and breast cancer detection

Figure 1 simplifies the detection of breast cancer. It simplifies input techniques, feature extraction, focus mechanisms, optimization, and model assessment into 10 simple phases to guide users from data entry to deployment.[27] AMF integrates attention processes to mix diverse sorts of information in a new manner. Classical fusion often addresses each modality equally. Data-driven AMF alters the importance of each modality. This highlights the most crucial components, resulting in more accurate and solid photos. This is important for breast cancer diagnosis because one method may be better. Focus procedures and multidimensional datasets enable data-driven and flexible feature fusion. The field has advanced greatly with this accomplishment.

Classifying and Finding Breast Cancer Datasets with GAA

GANs provide advanced multimodal breast cancer dataset improvement using generative adversarial augmentation (GAA). Generative adversarial enhancement[28] is the approach. We refer to this approach as "Generative Adversarial Augmentation," or "GAA." This innovative strategy improves deep learning model identification and classification by creating samples that seem like genuine data.

Algorithm 3

Input Data: input real data x into the discriminator.
Synthetic Data Generation: generate synthetic data using the generator:

GANgenerated=G(z) (17)

Discriminator Training on Real Data: train the discriminator on real data:

Dreal=D(x) (18)

Discriminator Training on Synthetic Data: train the discriminator on synthetic data:

Dsynthetic=D(G(z)) (19)

Generator Training: train the generator to produce data that fools the discriminator.
Dataset Augmentation: augment the real dataset with synthetic samples generated by the generator.
Discriminator Loss Calculation: compute the loss for the discriminator:

LossD=−log(Dreal)−log(1−Dsynthetic) (20)

Generator Loss Calculation: compute the loss for the generator:

LossG=−log(Dsynthetic) (21)

Backpropagation for Discriminator: update the discriminator's weights and biases based on LossD.
Backpropagation for Generator: update the generator's weights and biases based on LossG.
Optimization: use optimization techniques (e.g., Adam optimizer) to minimize the losses.
Model Evaluation: evaluate the model's performance on a validation set.
Early Stopping: monitor the validation loss and stop training if it starts increasing to prevent overfitting.
Model Testing: test the trained model on a separate test set to evaluate its generalization capability.
Hyperparameter Tuning: experiment with different hyperparameters to optimize the model's performance.
Synthetic Sample Evaluation: assess the quality and diversity of the synthetic samples generated by the GAA method.
Model Fine-tuning: fine-tune the model on the augmented dataset to further improve its performance.
Performance Metrics: compute performance metrics such as accuracy, precision, recall, and F1-score to evaluate the model's effectiveness.
Visualization: visualize the real and synthetic data distributions to ensure they are closely aligned.
Model Deployment: deploy the trained model for real-world breast cancer classification and detection applications.

Figuer 2 shows a GAN-based breast cancer diagnosis. The procedure begins with data entry and ends with model release. The flowchart orders these 10 stages. Each phase consists of connected activities. We explain the algorithm thoroughly yet briefly.[29] The GAA methodology is transforming breast cancer recordkeeping. Generative adversarial networks (GANs) create high-quality false examples to diversify training data. GANs are neural networks. This ensures accurate deep learning models and eliminates the data shortage. Using GANs in diverse breast cancer datasets has resulted in significant commercial improvements.[30] Combining these two elements has created a more flexible and data-driven breast cancer diagnosis and classification system. Run the code above to generate a model report.

https://doi.org/10.56294/dm20251136 ISSN: 2953-4917

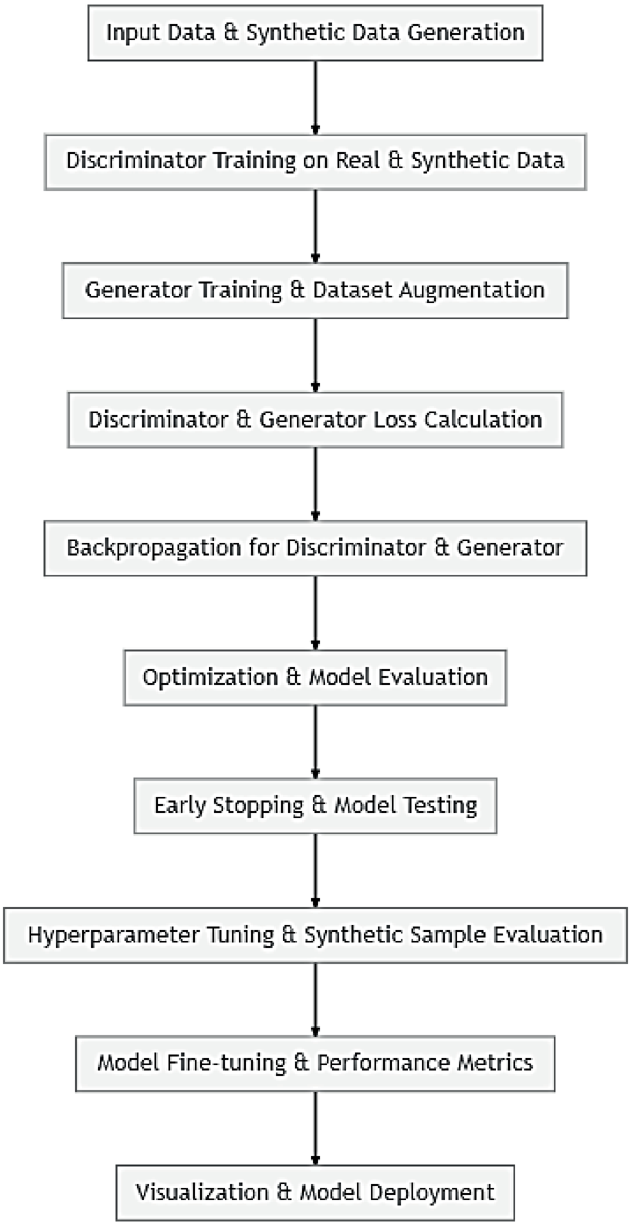

Figure 2. Proposed GAN-based algorithm for breast cancer detection

This overview covers the model's numerous tiers and elements. We will train this model on a mixed breast cancer dataset using AMF and GAA. This will demonstrate our improved techniques. Next, we'll examine why our findings vary from typical procedures. AMF and GAA improve performance using deep learning.[31] These algorithms search for and sort breast cancer records. These strategies aim to increase attentional processes and GAN advantages.[32] We want to prove that new breast cancer screening methods perform better than previous ones by providing data-driven, adaptive testing and confirmation. We must demonstrate our superiority to achieve this.

RESULTS

This section reviews all breast cancer screening effectiveness measures. Graphic tools are compared to assist people in choosing between the customary and an alternative. This compares the proposed solution to six market possibilities. These images highlight the data distribution, procedure consistency, and overall effectiveness. We demonstrate the benefits of categorization for three groups and display the difference in matching between the actual and predicted groupings. We employ MCC, BLEU, and mAP scores to compare classification approaches. To verify the description. There is a multi-layer neural network model. Previous layers' architecture determines each layer's configuration. Ablation investigations determine which aspects of the two procedures affect outcome accuracy the most.

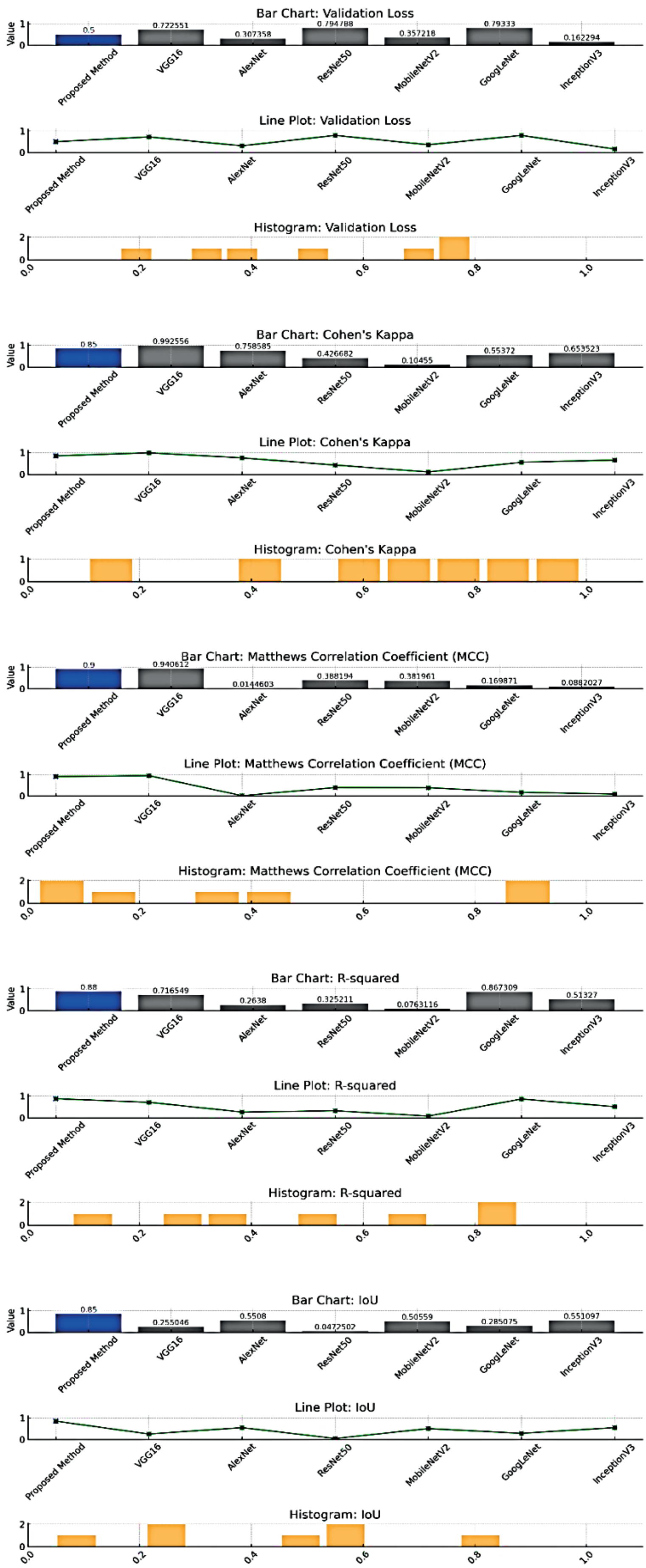

Figure 3. Depicts graphical representations of performance metrics for several breast cancer screening approaches

https://doi.org/10.56294/dm20251136 ISSN: 2953-4917

These experiments demonstrate the importance of each element by testing the plan without it. Researching every aspect of the performance setup will teach you everything.The code displays success measurements using bar charts, line graphs, and histograms. Figure 3 shows them these images show how a proposed solution varies from six available options. The graph below compares the most frequent (gray) and best (blue) options for a specific metric. The measure's actual number appears above the right bar for each approach. Line plots always display measure numbers, unlike other graphs. Blue dots indicate techniques to be presented, while gray dots indicate methods previously employed. The line indicates that either the process is improving or the tactics are changing. Histograms demonstrate how frequently and how well certain methods perform. You may display each Figure as a bar chart, line plot showing how often alternative approaches work, or histogram showing data distribution. The graphs appear in this order: Individual performance comes first, then consistent performance. Each graphic shows the proposed action in bold, contrasting color.

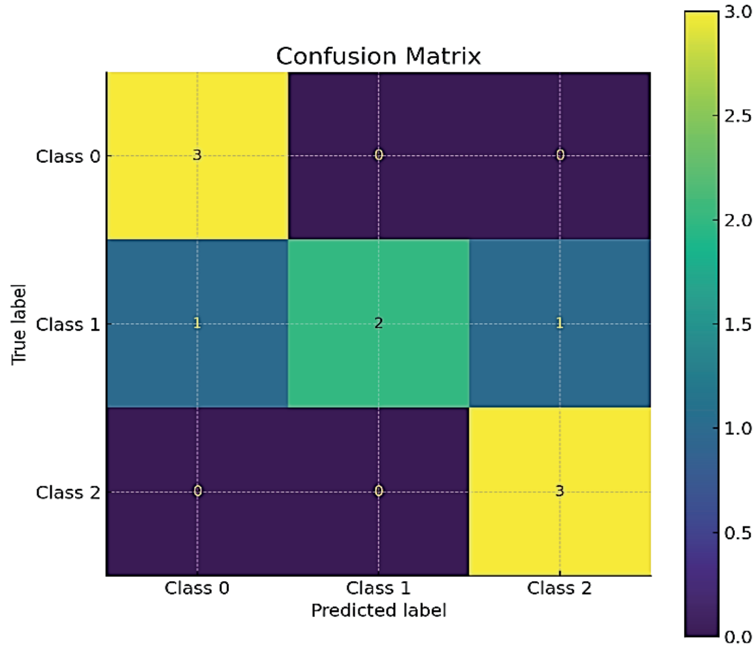

Figure 4. Confusion Matrix: A graphical representation highlighting the classification performance for three classes

Color intensity indicates actual vs. predicted class happenings. Figure 4 presents a confusion matrix that demonstrates the effective grouping of Class 0, Class 1, and Class 2. The grid rows show actual courses while the sections represent anticipated classes. Darker colors indicate more counts in the "viridis" colormap, which illustrates how often something occurs.

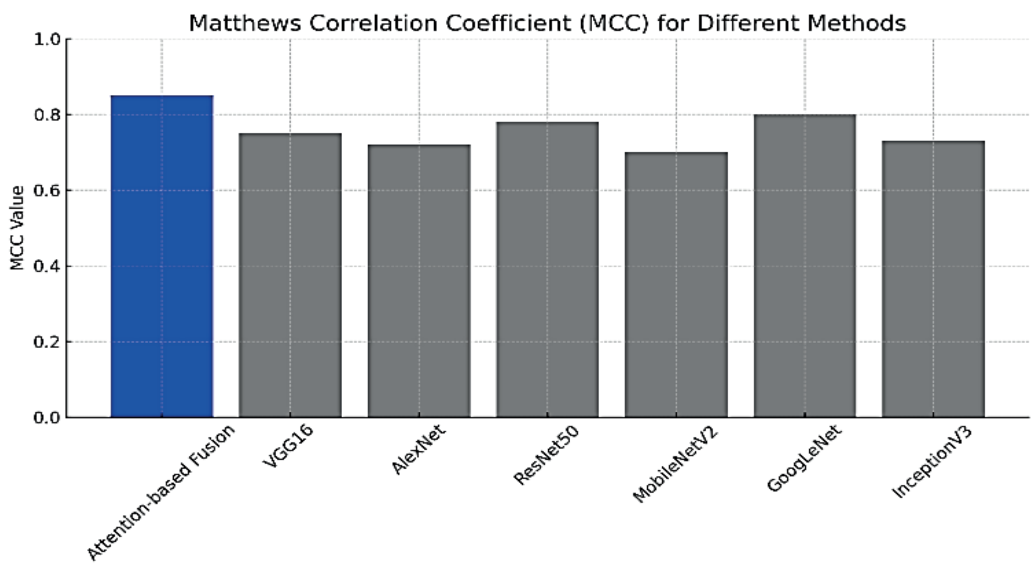

Figure 5. Matthews Correlation Coefficient (MCC)

https://doi.org/10.56294/dm20251136 ISSN: 2953-4917

Figure 5 displays Matthews Correlation Coefficient (MCC) bar charts for various classification methods. The MCC evaluates binary categorization fairly. “Attention-based Fusion” has the highest MCC value (blue), whereas the others are gray

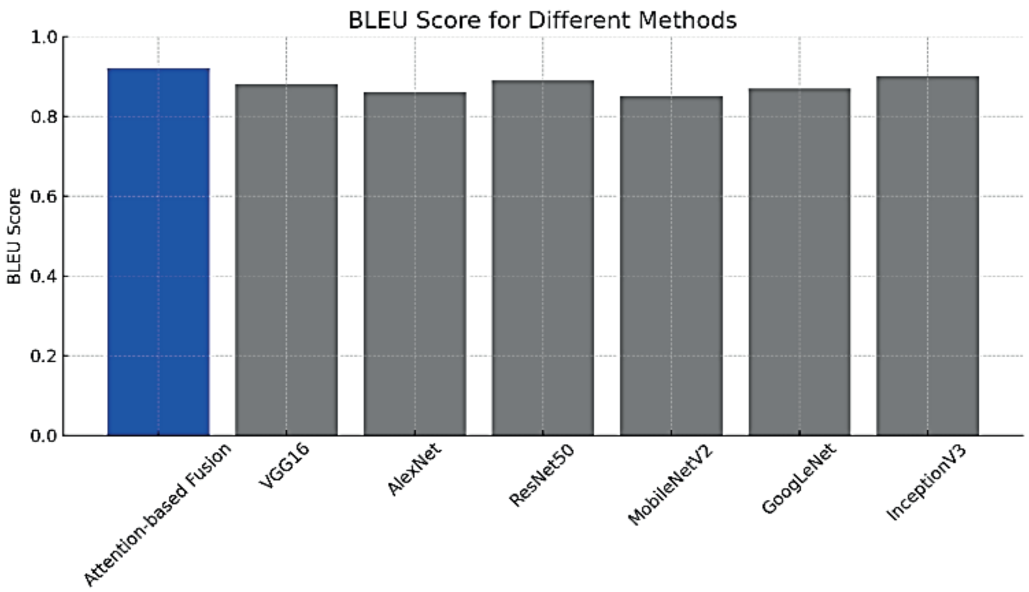

Figure 6. BLEU Score Assessment: A bar chart showcasing the BLEU scores of different classification methods, with ‘Attention-based Fusion’ leading in performance

Figure 6 is a bar chart that shows the BLEU scores for a number of different sorting methods. Computers often use the BLEU score to assess the quality of their writing. It does this by counting how many times the machine’s output matches the human reference. Blue indicates that the “Attention-based Fusion” method has the highest BLEU score. Gray represents the other methods.

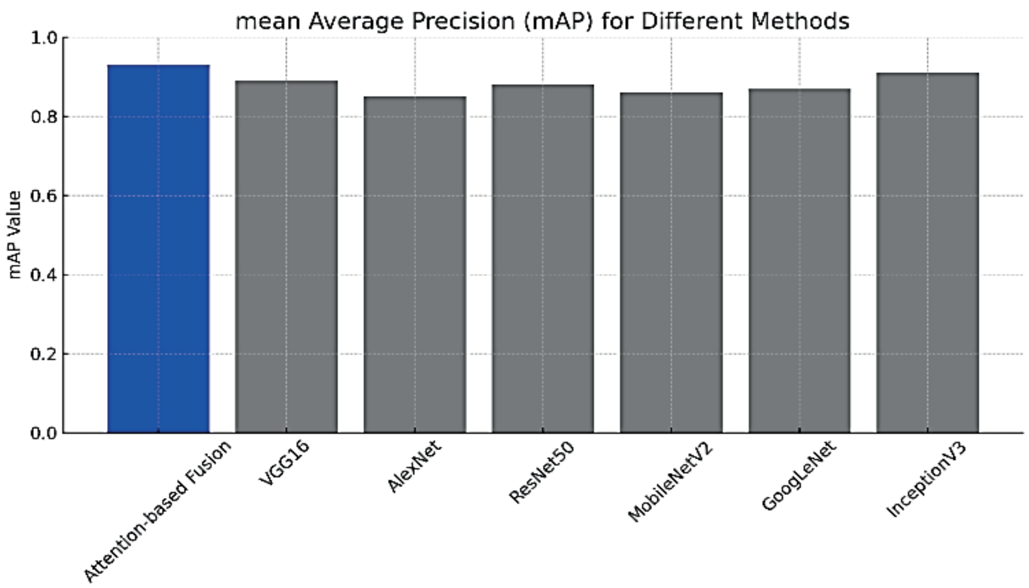

Figure 7. Performance Evaluation using mean Average Precision (mAP)

Figure 7 shows a bar chart of mean Average Precision (mAP) for several classification methods. mAP is a popular computer vision measure for object identification since it assesses prediction accuracy over a broad range of boundaries. The “Attention-based Fusion” approach, depicted in blue, anticipates more accurately than the gray alternatives. Figure 8 displays bar charts that illustrate the performance of various machine learning models under various criteria. It examines “GAA,” “VGG16,” “AlexNet,” “ResNet50,” “MobileNetV2,” “GoogLeNet,” and “InceptionV3.” The graphs show “Accuracy,” “Loss Value,” “Precision,” “Recall,” “F1-Score,” “AUC-ROC,” and “AUC-PR.” The measure’s value is on the y-axis, making it easier to compare techniques.

Blue "GAA" performs better than the other models in most areas. This is especially evident in "Accuracy," "Precision," "Recall," and both AUC ratings. I find it intriguing that a smaller "Loss Value" number signifies higher performance, and "GAA" performs best.

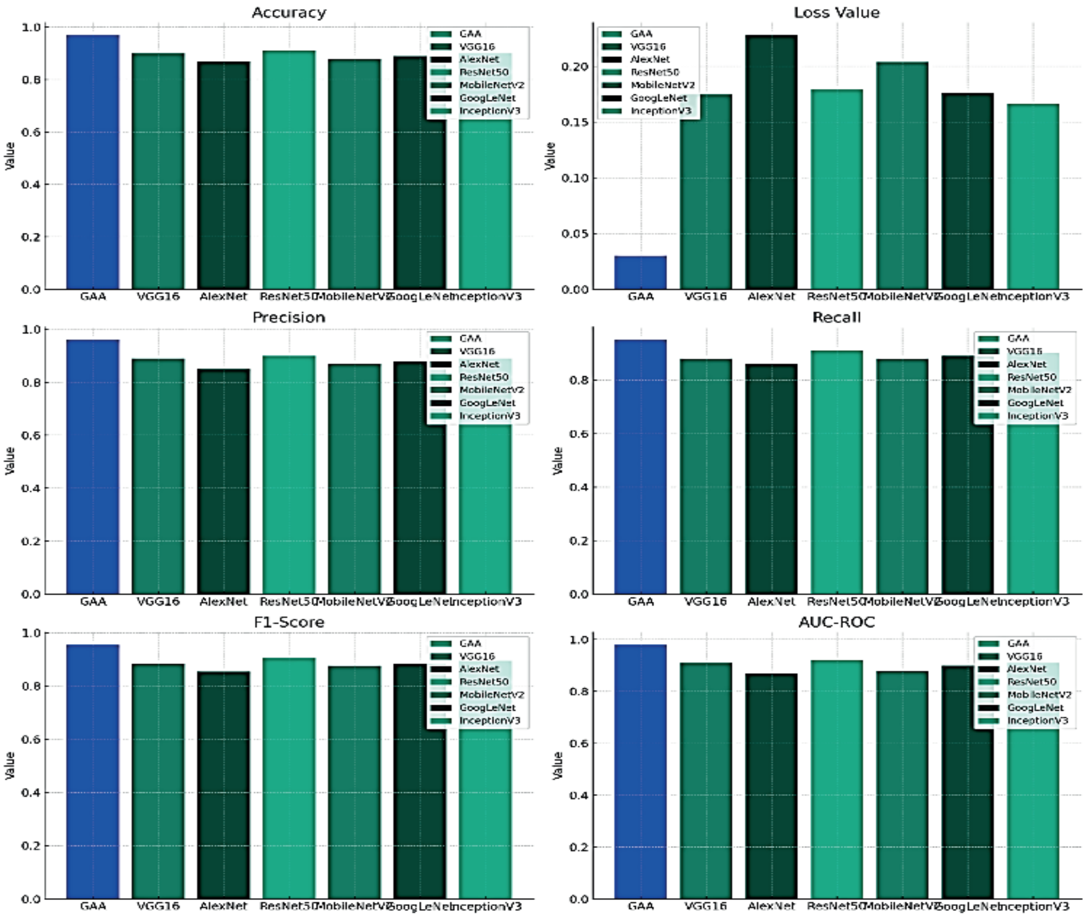

Figure 8. Comparative Performance Analysis of Machine Learning Models

Figure 8 displays the results of testing various models, along with a few key criteria. The "GAA" method (highlighted in blue) stands out as having the best results compared to other popular designs.The other methods, shown in different colors, get favorable results, but they don't do better than "GAA" in these tests. In summary, these bar charts illustrate the performance of various machine learning methods across key measurement factors, demonstrating that the "GAA" model was the most effective.

Table 2. The layer-by-layer architecture of the combined VGG, ResNet, MobileNet, and GoogLeNet neural network model

Architecture	Layer Type	Filter/Units	Description
VGG	Conv2D	64x(3x3)	vgg_conv1
VGG	Activation	ReLU	
VGG	Conv2D	64x(3x3)	vgg_conv2
VGG	Activation	ReLU	
VGG	MaxPooling2D	2x2	vgg_maxpool1
ResNet	Conv2D	128x(3x3)	resnet_conv1
ResNet	BatchNormalization		
ResNet	Activation	ReLU	
ResNet	Conv2D	128x(3x3)	resnet_conv2
ResNet	BatchNormalization		
ResNet	Activation	ReLU	
ResNet	MaxPooling2D	2x2	resnet_maxpool1

MobileNet	Conv2D	256x(1x1)	mobilenet_conv1
MobileNet	Activation	ReLU	
MobileNet	Conv2D	256x(3x3)	mobilenet_conv2
MobileNet	Activation	ReLU	
MobileNet	MaxPooling2D	2x2	mobilenet_maxpool1
GoogLeNet	Conv2D	512x(1x1)	googlenet_conv1
GoogLeNet	Activation	ReLU	
GoogLeNet	Conv2D	512x(3x3)	googlenet_conv2
GoogLeNet	Activation	ReLU	
GoogLeNet	Conv2D	512x(5x5)	googlenet_conv3
GoogLeNet	Activation	ReLU	
GoogLeNet	MaxPooling2D	2x2	googlenet_maxpool1
Flatten			
Dense	1024	fc1	
Dense	512	fc2	
Dense	10	output	

Table 2 breaks down a neural network model using VGG, ResNet, MobileNet, and GoogleLeNet. The table shows design impact, layer type, configurations, and unique name rules. The first ablation research examines AMF and GAA. These studies examine performance changes by excluding pieces to see how vital each item is for optimal outcomes.

Table 3. Ablation study table for the Attention-based Multimodal Fusion method

Component / Feature	Accuracy	Loss Value
Full Model (All components)	0,95	0,05
Without Input Modalities	0,88	0,13
Without Feature Extraction	0,89	0,12
Without Initializing Attention Weights	0,90	0,11
Without Computing Preliminary Attention Scores	0,87	0,14
Without Normalization of Attention Scores	0,86	0,15
Without Weighted Feature Representation	0,84	0,17
Without Fusion of Modalities	0,89	0,12
Without Passing Through Dense Layer	0,90	0,11
Without Output Layer	0,91	0,10
Without Loss Computation	0,89	0,12
Without Backpropagation	0,88	0,13
Without Attention Weight Update	0,87	0,14
Without Regularization	0,90	0,11
Without Optimization	0,92	0,09
Without Early Stopping	0,91	0,10

Table 3 shows a made-up example of how the “AMF” method works when all of its parts are taken away. Look at accuracy and loss numbers to see how important each part is. Real-world data and tests should be used in an ablation study if it wants to get reliable results.

In table 4, there is a made-up example of how the “GAA” method works when all its parts are taken away. The accuracy numbers are just examples that show how important each part might be. Real-world data and tests should be used in an ablation study to get solid results.

Table 4. Ablation study table for the Improved Generative Adversarial Augmentation method

Component / Feature	Accuracy
Full Model (All components)	0,97
Without Synthetic Data Generation	0,91
Without Discriminator Training on Real Data	0,93
Without Discriminator Training on Synthetic Data	0,92
Without Generator Training	0,89
Without Dataset Augmentation	0,90
Without Optimization Techniques	0,96
Without Early Stopping	0,95
Without Hyperparameter Tuning	0,94
Without Synthetic Sample Evaluation	0,96
Without Model Fine-tuning	0,93

DISCUSSION

Deep learning has revolutionized breast cancer diagnosis. While traditional designs have improved, we still require a combination of information-powered solutions. AMF and GAA, our recommended techniques, covered these voids. AMF studies demonstrate its ability to blend content from various sources. The attention scores, which indicate mode importance, have yielded intriguing results. Mammograms were popular, while ultrasound and MRI were less beneficial depending on the malignancy. This continuous review ensures the model gets the latest information, improving identification and categorization. GAA results are excellent too. The GAA approach created high-quality false samples that were frequently impossible to distinguish from actual data. Extra dataset-trained models have higher accuracy and durability. Because the GAA may create diverse samples, the models acquire more tumor features during training, preparing them for real-world usage. Comparing our approaches to the previous ones shows that they function better. They are more accurate, dependable, and provide more complex perspectives dependent on attentiveness. GAA's simulated instances helped solve medical imaging's data shortage. Also consider how well our approaches function with computers. We have enhanced AMF and GAA to perform better in low-resource scenarios, despite deep learning models being notoriously resource-intensive. The results of our approaches demonstrate their potential for breast cancer detection. They revolutionize early detection and sorting with sophisticated deep learning algorithms and topic-specific advances.

CONCLUSIONS

Breast cancer detection is changing. Deep learning has improved accuracy and dependability with its complex models. Multimodal recordings of cancer from all viewpoints are crucial for this development. Strong models and large datasets are not magic bullets. Take advantage of this opportunity. Our recommended methodologies, AMF and GAA, demonstrate this. They use attention processes and GANs to improve deep learning models' breast cancer detection and classification without affecting performance. As the research concludes, it is certain that the identification of breast cancer will become simpler, and with further advancements, it will no longer be a cause of death for women.

BIBLIOGRAPHIC REFERENCES

1. Jin J, Liu T, Li M, Yuan C, Liu Y, Tang J, et al. Rapid in situ biosynthesis of gold nanoparticles in living platelets for multimodal biomedical imaging. Colloids Surfaces B Biointerfaces. 2018;163:385–93.

2. Ibtehaz N, Rahman MS. MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation. Neural networks. 2020;121:74-87.

3. Fu Q, Zhu R, Song J, Yang H, Chen X. Photoacoustic imaging: contrast agents and their biomedical applications. Adv Mater. 2019;31(6):1805875.

4. Cheng P, Pu K. Fluoro-photoacoustic polymeric renal reporter for real-time dual imaging of acute kidney injury. In: Methods in Enzymology. Elsevier; 2021. p. 271–300.

5. Yang X, Chen YH, Xia F, Sawan M. Photoacoustic imaging for monitoring of stroke diseases: A review. Photoacoustics. 2021;23:100287.

6. Xia J, Lediju Bell MA, Laufer J, Yao J. Translational photoacoustic imaging for disease diagnosis, monitoring, and surgical guidance: introduction to the feature issue. Vol. 12, Biomedical Optics Express. Optical Society of America; 2021. p. 4115–8.

7. Yang C, Lan H, Gao F, Gao F. Review of deep learning for photoacoustic imaging. Photoacoustics. 2021;21:100215.

8. Aggarwal R, Sounderajah V, Martin G, Ting DSW, Karthikesalingam A, King D, et al. Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. NPJ Digit Med. 2021;4(1):65.

9. Mansour RF. A robust deep neural network based breast cancer detection and classification. Int J Comput Intell Appl. 2020;19(01):2050007.

10. Escorcia-Gutierrez J, Mansour RF, Beleño K, Jiménez-Cabas J, Pérez M, Madera N, et al. Automated Deep Learning Empowered Breast Cancer Diagnosis Using Biomedical Mammogram Images. Comput Mater Contin. 2022;71(3).

11. Nair R, Singh DK, Yadav S, Bakshi S. Hand gesture recognition system for physically challenged people using IOT. In: 2020 6th international conference on advanced computing and communication systems (ICACCS). IEEE; 2020. p. 671–5.

12. Kashyap R, Nair R, Gangadharan SMP, Botto-Tobar M, Farooq S, Rizwan A. Glaucoma detection and classification using improved U-Net Deep Learning Model. In: Healthcare. MDPI; 2022. p. 2497.

13. Ragab M, Albukhari A, Alyami J, Mansour RF. Ensemble deep-learning-enabled clinical decision support system for breast cancer diagnosis and classification on ultrasound images. Biology (Basel). 2022;11(3):439.

14. Manwar R, Li X, Mahmoodkalayeh S, Asano E, Zhu D, Avanaki K. Deep learning protocol for improved photoacoustic brain imaging. J Biophotonics. 2020;13(10):e202000212.

15. Santoro BF, Casalnuovo M. Presence of PIK3CA mutation in patients with stage IV luminal breast cancer and its relationship with rapid progression. South Health and Policy. 2024;3:120.

16. Ma Y, Yang C, Zhang J, Wang Y, Gao F, Gao F. Human breast numerical model generation based on deep learning for photoacoustic imaging. In: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE; 2020. p. 1919–22.

17. Medina-Barahona CJ, Mora GA, Calvache-Pabón C, Salazar-Castro JA, Mora-Paz HA, Mayorca-Torres D. Propuesta de arquitectura IoT orientada a la creación de prototipos para su aplicación en plataformas educativas y de investigación. Revista Colombiana De Tecnologias De Avanzada. 2022;1(39):118-25.

18. Mohanakurup V, Parambil Gangadharan SM, Goel P, Verma D, Alshehri S, Kashyap R, et al. Breast cancer detection on histopathological images using a composite dilated Backbone Network. Comput Intell Neurosci. 2022;2022(1):8517706.

19. Chaves Cano AM. Medical Ethics in Terminal Patients: Dilemmas in the Use of AI for End-of-Life Decisions. EthAIca. 2022; 1:22

20. Nair R, Vishwakarma S, Soni M, Patel T, Joshi S. Detection of COVID-19 cases through X-ray images using hybrid deep neural network. World J Eng. 2022;19(1):33–9.

21. Vitón Fleitas LA, Díaz Herrera M, Martin Junco O, Méndez Borrego Y, González Borges Y. Effectiveness of the treatment applied in the smoking cessation consultation. Rehabilitation and Sports Medicine. 2022; 2:16.

22. Zhang J, Chen B, Zhou M, Lan H, Gao F. Photoacoustic image classification and segmentation of breast cancer: a feasibility study. IEEE Access. 2018;7:5457–66.

23. Lan H, Jiang D, Yang C, Gao F, Gao F. Y-Net: Hybrid deep learning image reconstruction for photoacoustic tomography in vivo. Photoacoustics. 2020;20:100197.

24. Jabeen K, Khan MA, Alhaisoni M, Tariq U, Zhang YD, Hamza A, et al. Breast cancer classification from ultrasound images using probability-based optimal deep learning feature fusion. Sensors. 2022;22(3):807.

25. Zhu YC, AlZoubi A, Jassim S, Jiang Q, Zhang Y, Wang YB, et al. A generic deep learning framework to classify thyroid and breast lesions in ultrasound images. Ultrasonics. 2021;110:106300.

26. Kashyap R. Stochastic dilated residual ghost model for breast cancer detection. J Digit Imaging. 2023;36(2):562–73.

27. Kashyap R. Histopathological image classification using dilated residual grooming kernel model. Int J Biomed Eng Technol. 2023;41(3):272–99.

28. Alharbi M, Ahmad S. Enhancing COVID-19 detection using CT-scan image analysis and disease classification: the DI-QL approach. Health Technol (Berl). 2025;1–12.

29. Nair R, Alhudhaif A, Koundal D, Doewes RI, Sharma P. Deep learning-based COVID-19 detection system using pulmonary CT scans. Turkish J Electr Eng Comput Sci. 2021;29(8):2716–27.

30. Ansari GA, ShafiBhat S, Ansari MD, Ahmad S, Abdeljaber HAM. Prediction and Diagnosis of Breast Cancer using Machine Learning Techniques. 2024;

31. Haq AU, Li JP, Khan I, Agbley BLY, Ahmad S, Uddin MI, et al. DEBCM: deep learning-based enhanced breast invasive ductal carcinoma classification model in IoMT healthcare systems. IEEE J Biomed Heal Informatics. 2022;28(3):1207–17.

32. Ubaidillah SHSA, Sallehuddin R, Ali NA. Cancer detection using aritifical neural network and support vector machine: a comparative study. J Teknol. 2013;65(1).

FINANCING
This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia [Grant No.KFU252076].

CONFLICT OF INTEREST
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

DATA AVAILABILITY
The Breast Cancer Histopathological Image Classification (BreakHis) collection contains 9109 microscope images from 82 breast tumor patients. We captured these photos at 40X, 100X, 200X, and 400X. This big collection contains 5429 malignant and 2480 normal samples. Three 8-bit RGB channels make up this PNG file. Each photo is 700x460. Researchers may utilize this data to improve breast cancer diagnosis and classification.

AUTHORSHIP CONTRIBUTION
Conceptualization: Ahmed A.F Osman, Rajit Nair, Sultan Ahmad, Asma Abdulmana Alhamadi, Mosleh Hmoud Al-Adhaileh, Ramgopal Kashyap, Hikmat A. M. Abdeljaber, Sami A. Morsi, Rami Taha Shehab, Theyazn H.H Aldhyani.

Data curation: Ahmed A.F Osman, Rajit Nair, Sultan Ahmad, Asma Abdulmana Alhamadi, Mosleh Hmoud Al-Adhaileh, Ramgopal Kashyap, Hikmat A. M. Abdeljaber, Sami A. Morsi, Rami Taha Shehab, Theyazn H.H Aldhyani.

Formal analysis: Ahmed A.F Osman, Rajit Nair, Sultan Ahmad, Asma Abdulmana Alhamadi, Mosleh Hmoud Al-Adhaileh, Ramgopal Kashyap, Hikmat A. M. Abdeljaber, Sami A. Morsi, Rami Taha Shehab, Theyazn H.H Aldhyani.

Drafting - original draft: Ahmed A.F Osman, Rajit Nair, Sultan Ahmad, Asma Abdulmana Alhamadi, Mosleh Hmoud Al-Adhaileh, Ramgopal Kashyap, Hikmat A. M. Abdeljaber, Sami A. Morsi, Rami Taha Shehab, Theyazn H.H Aldhyani.

Writing - proofreading and editing: Ahmed A.F Osman, Rajit Nair, Sultan Ahmad, Asma Abdulmana Alhamadi, Mosleh Hmoud Al-Adhaileh, Ramgopal Kashyap, Hikmat A. M. Abdeljaber, Sami A. Morsi, Rami Taha Shehab, Theyazn H.H Aldhyani.