

Improving Breast Cancer Detection Accuracy using Convolutional Block Attention Modules

Alaa M. Ghazy¹, Hanan M. Amer¹, Hala B. Nafea¹, and Fayez W. Zaki¹

E-mail: alaamohamedghazy@gmail.com, eng_hanan_2007@mans.edu.eg, halabahyeldeen@mans.edu.eg, and fwzaki@mans.edu.eg.

¹Department of Electronics and Communications Engineering, Faculty of Engineering, Mansoura University, Mansoura, Egypt

ABSTRACT:

This study's main objective focuses on enhancing the precision of breast cancer identification. There are two primary stages in this study's framework. First stage: compare the performance of two different convolution neural networks (CNNs) architectures with data splitting size representing 20% of the testing and validation set and 80% of the training data. The convolutional neural networks used in this study are VGG-19 (Visual Geometry Group 19) and ResNet50 (Residual Network 50), both of which employ the Adam optimizer and multilayer perceptron networks to identify the superior architecture. The subsequent phase implements Convolutional Block Attention Modules (CBAM) using optimized structural configurations to refine feature representations and enhance computational effectiveness. Performance assessment encompasses multiple evaluation indicators, namely classification accuracy, precision rates, recall values, F1-measure, and receiver operating characteristic curve area (ROC-AUC). Through the integration of attention-based techniques and the application of pre-trained model knowledge, the framework demonstrates exceptional results using the DDSM database. The results indicated that pretrained model VGG-19 yielded the best accuracy and precision scores with data splitting ratio (80% train, 10% validation, and 10% test) where, the accuracy for testing achieve 97.75%, and precision achieve 100%. After integrating CBAM, the result improved to 99.99% for accuracy and 100% precision outperforming state-of-the-art methods.

Keywords:

Breast Cancer, Mammogram, Convolutional Neural Network, and Convolutional Block Attention Module.

1. INTRODUCTION:

Breast cancer is considered one of the most common malignancies affecting women globally [1]. For better treatment results, early detection is essential. Mammography continues to be the most reliable and accurate screening modality for BC screening programs [2],[3]. It is also considered the most reliable, dependable screening method, and remains the gold standard for community BC screening [4]. Nowadays, MRI and ultrasound are still utilized when combined with mammography, particularly when the density of breast tissue is high, but it can't be replaced. To help radiotherapists, computer-assisted identification and diagnosis (CAD) has been created to improve the accuracy prediction of screening mammography [5]. Deep learning (DL) has been highlighted by many studies as being crucial to the delivery of higher-quality and more secure healthcare [6],[7]. On the other hand, High-dimensional data, such as pictures and videos, may now be analyzed by machines thanks to DL algorithms that were inspired by the structure of human brain [8], [9]. This study aims to enhance BC's classification models' accuracy by integrating (CBAM) with the best CNN architecture. A standard framework is often followed by the entire article: The second section presents relevant literature and previous studies. The third section provides a comprehensive explanation of the suggested methodological approach, encompassing deep learning techniques and classification methods employed in the framework. The fourth section examines the experimental findings, presenting comparative analyses of accuracy metrics, along with descriptions of the utilized dataset, computational requirements, and data partitioning strategies. Additionally, it outlines the evaluation criteria applied to assess model effectiveness. The fifth section offers concluding remarks. The main goal is to improve classification accuracy for mammogram images by leveraging both transfer learning and attention mechanisms. The specific contributions of this research are:

1. Comparative Evaluation: compare and analyze the performance of two different CNN architectures (VGG-19 and ResNet50-V2) on the BC classification task to identify the most suitable backbone model.

2. **Attention Integration:** Incorporate CBAM into the best performance CNN architectures to improve the models' ability to focus on important features, then enhancing overall classification accuracy and robustness of the system.

2. Related Work:

Breast cancer represents a leading cause of women's deaths worldwide. Mammography remains a widely utilized imaging method for early detection and classification of breast cancer. Contemporary deep learning convolutional networks have shown considerable promise in automated mammographic diagnosis. Nevertheless, conventional CNN architectures frequently struggle to prioritize image regions containing critical information, potentially limiting diagnostic accuracy. Attention mechanisms such as CBAM were introduced to improve feature extraction through selective emphasis on important spatial and channel dimensions. Related literature is summarized in Table 1.

The study by Wang and colleagues [10] implemented RcdNet, integrating depthwise separable convolutions and CBAM modules, achieving outstanding results in breast ultrasound analysis: 93.51% classification accuracy, 0.9168 precision value, 0.9495 sensitivity rate, and 0.9290 F1-measure. Generated attention visualizations effectively identified diagnostically relevant lesion regions.

Zeng et al. [11] created a DL model using CBAM ResNet-18 to forecast the expressions of the HER2, ER, and PR receptors from mammograms without the need for manual mass segmentation. The model performed better than the baseline ResNet-18 and VGG-19 variations after being trained using five-fold cross-validation and assessed on an external dataset. It showed excellent potential as an auxiliary diagnostic tool, especially for ER prediction, with AUCs of 0.708, 0.785, and 0.706 for HER2, accuracies of 0.651, 0.845, and 0.678, and F1-scores of 0.528, 0.905, and 0.773 for PR.

Boro et al. [12] created the CBAM-RIUnet model, which improves breast tumor segmentation in ultrasound images by combining residual inception depth-wise separable convolutions with (CBAM). The model outperformed state-of-the-art techniques by successfully suppressing irrelevant features and concentrating on the region of interest, as demonstrated by its Dice score of 89.38%, accuracy of 97.59%, precision of 91.34%, and IoU of 88.71% when evaluated under enhanced breast ultrasound (EBUS) and test-time augmentation (TTA) scenarios.

Alkhalefah et al. [13] suggested a model named MOB-CBAM for BC detection and molecular subtype categorization from mammograms by combining a dual-channel CBAM with the MobileNet-V3 backbone. Using the CMMD database, the model attained 98% classification accuracy for detailed categorization tasks, while achieving 99% across all metrics (accuracy, precision, recall, and F1-measure) with an MCC of 98% for broader classification schemes. The system successfully identified molecular subtypes including Luminal A, Luminal B, HER-2 Positive, and Triple Negative variants. Validation procedures conducted on MIAS and CBIS-DDSM repositories yielded accuracy rates of 97% and 98% correspondingly.

Aggarwal and colleagues [14] introduced an enhanced DeepLabV3+ framework designed for segmenting breast abnormalities in sonographic imagery, incorporating CBAM components within the encoding and decoding phases to improve attention toward relevant characteristics. Using the BUS dataset, the model achieved superior results compared to the baseline DeepLabV3+, reporting precision 0.974, recall 0.933, specificity 0.997, Dice coefficient 0.951, and IoU 0.933, effectively addressing challenges in segmenting small tumors caused by speckle noise, shape variations, and tumor-like regions.

Alashban. [15] proposed a two-stage CAD system combining a modified VGG19 for classifying DBT images and a YOLOv5 model enhanced with CBAM (YOLOv5-CBAM) for lesion detection. The modified VGG19 integrated eight additional layers (four batch normalization and four pooling layers) to improve classification, while CBAM modules were inserted into YOLOv5 after each feature fusion. Using data from 22,032 DBT examinations across 5,060 patients, the system outperformed prior architectures in both classification accuracy and training loss. YOLOv5-CBAM effectively identified and classified lesions as either benign or malignant. The testing accuracy was 93%, training accuracy was 95%, and validating accuracy was 94% when using Batch size 512 with Adam optimizer.

Sengodan. [16] proposed a CBAM-EfficientNetV2 model integrating Convolutional Block Attention Modules with EfficientNetV2 to enhance feature extraction and emphasize clinically significant tissue regions in breast cancer histopathology images. Utilizing transfer learning and CLAHE preprocessing, the BreakHis database, containing images at 40X, 100X, 200X, and 400X magnification levels, served as the evaluation dataset. The model demonstrated exceptional results, attaining 99.01% accuracy and 98.31% F1-measure at 400X magnification, exceeding existing benchmark approaches while preserving processing speed suitable for practical clinical deployment.

Mehmood et al. [17] proposed the CB-Res-RBCMT framework for the diagnosis of breast ultrasonography carcinoma, which blends customized residual CNNs with novel Vision Transformer elements. The RBCMT framework integrates foundational convolutional units, CNN-Transformer hybrid blocks, and sophisticated modules for extracting regional and boundary characteristics to identify differences in contrast and structural patterns. Feature diversity is improved by a channel-boosting

strategy that combines transfer learning-based residual CNN maps with original RBCMT channels. After that, the best pixels are found using a spatial attention block. On the harmonized robust BUSI dataset, the model outperformed existing CNN and ViT methods, attaining performance metrics of 95.57% for F1-measure, 95.63% classification accuracy, 96.42% recall rate, and 94.79% precision value.

Table 1. Show summary of related work.

Author	year	Method	Dataset	Accuracy
Wang et al. [10]	2025	RcdNet (Deep separable conv +CBAM	Breast ultrasound	93.51%
Zeng et al. [11]	2025	CBAM ResNet-18	Mammography (internal + external)	0.651 (HER2), 0.845 (ER), 0.678 (PR)
Boro et al. [12]	2025	CBAM-RIUnet	EBUS + TTA	97.59%
Alkhalefah et al. [13]	2024	dual-channel CBAM)	CMMD, MIAS, CBIS-DDSM	~99%
Aggarwal et al. [14]	2024	DeepLabV3+ +CBAM	BUS dataset	
Alashban [15]	2025	YOLOv5-CBAM + Modified VGG19	22,032 DBT exams	93%
Sengodan [16]	2024	CBAM-EfficientNetV2	BreakHis (multi-mag)	99.01%
Mehmood et al. [17]	2025	CB-Res-RBCMT	Harmonized stringent BUSI	95.63%

3. Methodology:

This study presents a robust framework for BC classification from mammogram images using DL techniques. The proposed method utilizes a publicly available mammography dataset to develop and evaluate the classification models. The framework is designed in two primary phases, as shown in Fig.1.

I. First stage: Comparison between Pre-trained CNNs

In the first stage, the best architecture for BC classification is identified by comparing the performance of two different pre-trained CNNs. The selected models are: VGG-19, and ResNet50. The hyper parameters of Both models are enhanced by using transfer learning to get the best performance of each model and best performance model using a balanced subset of the DDSM dataset with a data splitting ratio representing 20% of the testing and validation set and 80% of the training data. The goal of this stage is leveraging the strengths of each CNN and identify the backbone model that yields the best performance in binary classifying mammographic images.

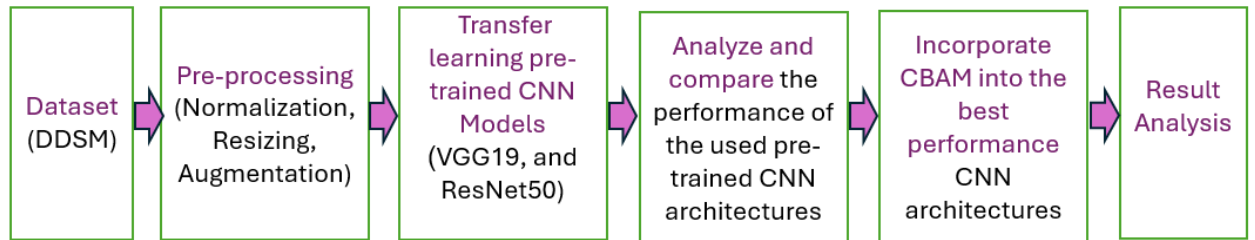


Figure 1. The prime mechanism of the proposed method.

II. Second stage: Integration of CBAM with the best performance CNNs.

After evaluating CNN models individually and getting the best performance model, the second stage focuses on enhancing feature extraction capabilities by integrating the CBAM with the best performance model architecture. CBAM introduces attention mechanisms that enable the networks to focus more effectively on the most informative regions of the mammogram images by applying both channel and spatial attention. This enhancement aims to improve classification accuracy by guiding the CNNs to prioritize relevant features, reducing the risk of overlooking critical image details. A fair comparison is ensured by training and evaluating the modified designs under the same data splitting conditions. Through this two-stage methodology, the study

systematically investigates the impact of attention mechanisms on breast cancer classification performance using DL models.

3.1 The applied CNNs of proposed model.

3.1.1 VGG-19 Architecture:

This deep convolutional network was developed by Oxford University in 2014 for the ImageNet Recognition Challenge [18]. The architecture contains 19 trainable layers comprising 16 convolutional stages utilizing small 3x3 kernels, 2x2 max-pooling operations, and 3 dense layers, enabling extraction of sophisticated hierarchical representations while maintaining computational practicality [19]. Pre-trained using the extensive ImageNet database, the model demonstrated exceptional image classification capabilities.

Although VGG19 is well-known for its depth and simplicity, its high number of parameters make it computationally demanding, which has led to the creation of lighter architectures like MobileNet and EfficientNet [20]. However, VGG19 remains a standard backbone model for many transfer learning tasks, feature extraction processes, and medical image analysis due to its robustness and ease of implementation in frameworks like Keras and PyTorch.

3.1.2 ResNet50 Architecture:

ResNet50, a deep convolutional network introduced by He and colleagues within the Residual Networks series in 2015, comprises 50 trainable layers designed specifically to address vanishing gradient problems [21]. The architecture's key innovation involves implementing skip connections (residual pathways), where a layer's output combines with inputs from layers positioned further in the network. This mechanism enables efficient gradient propagation through numerous layers, reducing training complexity for deeper models while enabling the learning of residual functions. The core residual unit within ResNet50 is expressed mathematically as:

$$y = F(x, W_i) + x \quad (1)$$

Where:

- x represents input,
- W_i are the learnable weights,
- F introduce residual function (series of convolutions, batch normalization, and ReLU),
- y is the output after adding the skip connection.

ResNet50 is structured into convolutional layers, identity blocks, and bottleneck residual blocks to reduce computation while maintaining depth. It uses 1, 3x3, and 1x1 convolutions within these blocks. Because of its excellent accuracy and efficiency, ResNet50 has become known as the most important architecture in detecting objects, image classification, and transfer learning [22]. It is widely available in deep learning libraries such as Keras, PyTorch, and TensorFlow.

3.2 The applied Convolutional Block Attention Modules (CBAM).

Without adding too many network parameters, it is a lightweight, easy-to-use, and efficient feed forward CNN attention strategy that enhances CNN classification performance [23]. The mechanism enables the model to prioritize key image regions through dual attention strategies: channel-based and spatial-based attention. Figure 2 illustrates how spatial attention highlights critical image zones containing relevant features, while the channel attention instructs the network which feature mappings are more significant. CBAM model is easy to add to existing CNN architectures leading to better accuracy in image detection, and classification [24].

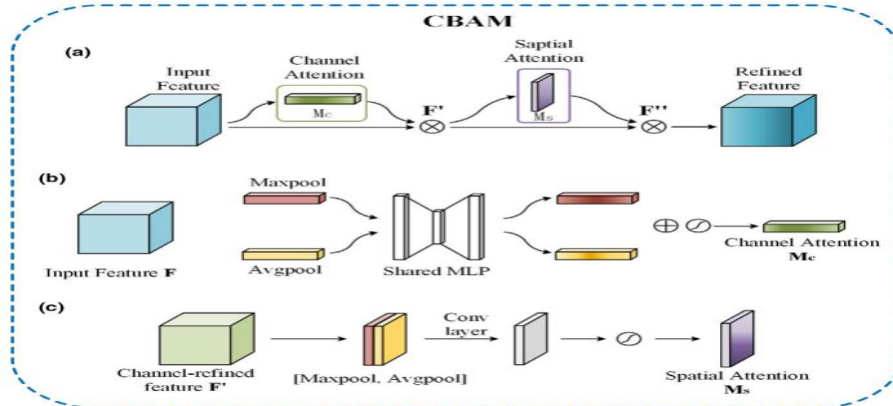


Figure 2. (a) the overall structure of CBAM module includes b and c; (b) channel attention modules; (c) spatial attention module [24].

4. Results Analysis and Discussion:

4.1 Data description

A widely used publicly accessible dataset on the Mendeley website is used to test the suggested approach [25]. On the Mendeley website, every image in this dataset was available in PNG format, with a pixel size of 227×227 . Image augmentation techniques are employed to expand the dataset size, while contrast-limited adaptive histogram equalization (CLAHE) is applied to enhance image contrast. The suggested approach was used in this study just on DDSM datasets. Before augmentation, 2188 mass images were taken from 1319 cases; after augmentation, there were 13128 mass images total, organized into two separate folders.

The benign folder comprised 5970 images and the malignant folder comprised 7158 images. This database isn't balanced. So, in this study, a randomly balanced subset of the DDSM dataset was selected, containing 4000 mass images divided into 2000 benign mass images and 2000 malignant mass images in a balanced manner as illustrated in Table 2. A sample of mammography mass images is shown in Fig.3.

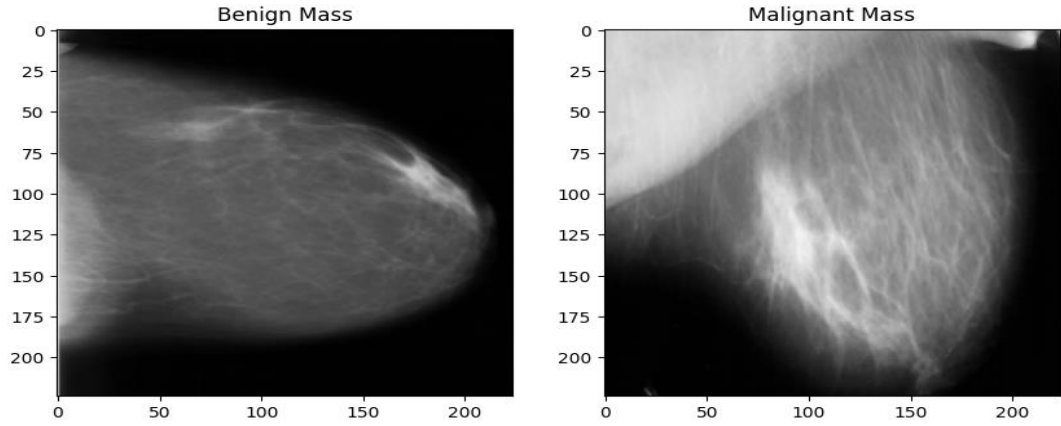


Figure 3. Benign and malignant mass Tumor from DDSM dataset.

Table 2. Number of mammography mass images used.

Benign Images	Malignant Images	Total
2000	2000	4000

4.1.1 Data splitting

The dataset is split into training, validating, and testing sets at random; the training set is 80%, the validation set is 10%, and the testing set is 10%. Table 3 illustrates the number of mammography mass images used.

Table 3. Number of mammography mass images used.

Splitting Ratio	Trained images	Validated images	Tested images	Total
First Scenario (80:10:10)	3200	400	400	4000

4.2 System specifications

All experiments were conducted on Google Collaboratory utilizing an online T4 GPU environment with Python 3, supported by a Core i7 processor and 64 GB RAM configuration. Training employed the Adam optimization algorithm with the following hyperparameters: batch size of eight samples, training duration of twenty epochs, and a learning rate set at 0.0001 to regulate convergence speed.

4.3 Data Preprocessing

The training images are normalized, resized, and enhanced as part of the preprocessing step. To improve model generalizability and enable a smoother training process, normalization involves transforming pixel values to fall within a predefined range between 0 and 1. After that, In order to match the images with the standard input size of all applied pre-trained models, they are reduced to $224 \times 224 \times 3$ pixels. Data augmentation techniques are systematically implemented to enhance dataset variability and strengthen the model's generalization capabilities and overall effectiveness. The augmentation pipeline employs an image generation function configured with the following parameters: 20-degree rotation limit, 10% width and height shifting boundaries, 10% zoom variation, and 10% shear transformation range.

4.4 Performance Evaluation Measures:

Performance assessment metrics were derived from the confusion matrix components: true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN):

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (2)$$

$$F1 - score = 2 \cdot \frac{precision \cdot recall}{precision + recall} \quad (3)$$

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

Additionally, the area under the receiver operating characteristic (ROC) curve and AUC were computed.

4.5 Experimental Results and Discussion

This part presents the classification framework's results utilizing an 80:10:10 data partitioning scheme. The implemented pre-trained architectures (VGG-19 and ResNet50-V2) were refined through transfer learning by incorporating additional layers: a flattening component, a fully connected layer employing ReLU activation, a 0.2 dropout regularization, and softmax classification layers with varying output dimensions appended to the base models. Thus, the models trained and tested by the applied data then improves its classification accuracy performance models by using CBAM module.

4.5.1 Experimental Results for data-splitting [80:10:10]

Applying the proposed classification system using this data-splitting and discussing the experimental results.

I. The experimental result of first stage.

The experimental result of the individual CNN is shown in Table [4,5] which illustrates the classification performance of the two suggested CCNs. In addition, Fig. [4,5] shows the two DL models' accuracy and loss during the training and validating steps over twenty epochs. Finally, Fig. [6,7]. outlines the confusion matrices acquired during experiment, where Classes "0" and "1" denote "benign" and "malignant," respectively.

Table 4. The performance metrics for validation and testing data set on VGG-19 pre-trained models

Pre-trained model VGG-19	Accuracy (%)	Precision (%)	Recall (%)	F1- Score (%)	Roc-AUC (%)
validation	98.75	99.49	98	98.74	99.96
testing	97.75	100	98	97.69	99.96

Table 4 presents the evaluation of the standalone pre-trained VGG-19 architecture before attention module integration. Validation metrics yielded 98.75% accuracy, 99.49% precision, 98% sensitivity, 98.74% F1-measure, and 99.93% AUC value. Test set performance demonstrated 97.75% accuracy, perfect precision at 100%, 98% recall rate, 97.69% F1-score, and 99.96% area under the curve. While these outcomes are robust, the gap between precision and sensitivity indicates the model exhibited high confidence in positive classifications yet failed to detect certain true positive cases.

Pre-trained model ResNet50-V2	Accuracy (%)	Precision (%)	Recall (%)	F1- Score (%)	Roc-AUC (%)
validation	93	98.86	87	92.55	99.59
testing	92.25	100	84.50	91.59	99.90

Table 5. The performance metrics for validation and testing data set on ResNet50-V2 pre-trained models

The pre-trained ResNet50-V2 model was additionally evaluated without attention modules, as presented in Table 5. This model achieved 93% accuracy, 98.86% precision, 87% recall, 92.55% F1-score, and 99.59% ROC-AUC on the validation dataset. For the testing dataset, the model attained 92.25% accuracy, 100% precision, 84.50% recall, 91.59% F1-score, and an exceptional ROC-AUC of 99.90%. These results indicate that the ResNet50-V2 model demonstrated strong performance, though a disparity existed between precision and recall, suggesting that the model emphasized minimizing false positives while occasionally missing certain actual cancer cases.

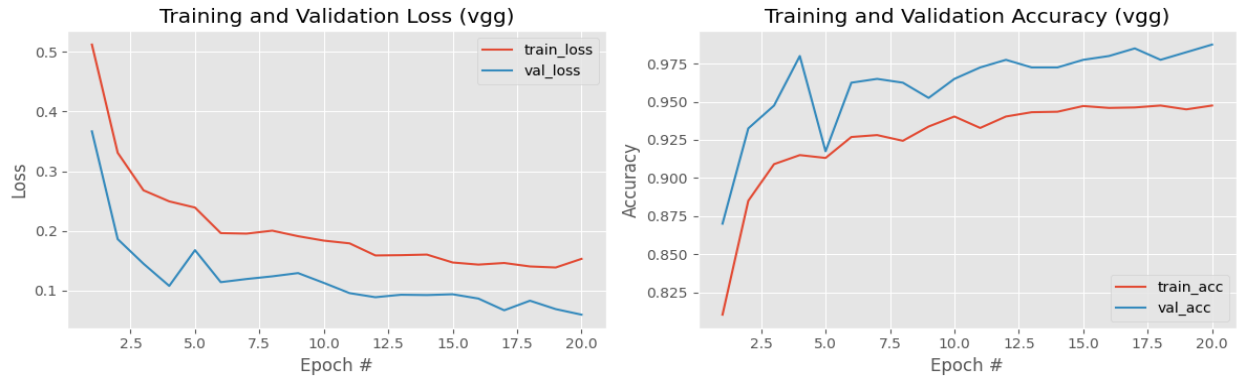


Figure 4. (Train-Validation) accuracy and loss through VGG-19 pretrained model.

Figure 4 illustrates how well the VGG-19 model performed during training and validation throughout 20 epochs. The validation loss falls from around 0.36 to 0.06 and the training loss falls from roughly 0.50 to 0.14 along the loss curves. Good generalization is indicated by a smaller validation loss relative to training loss. According to the accuracy curves, the validation accuracy increases from 88% to around 98%, while the training accuracy increases from roughly 81% to 94%. Throughout the training phase, the validation accuracy remains marginally greater than the training accuracy.



Figure 5. (Train-Validation) accuracy and loss through ResNet50-V2 pretrained model.

Figure 5 displays the ResNet model's performance during 20 epochs. The training loss gradually decreases from about .50 to .14, but the validation loss consistently falls below the training loss, indicating high generalization ability, as seen by the loss curves. While validation accuracy rises from around 90% to peaks near 97%, training accuracy improves from about 81% to over 95% in the accuracy curves. The majority of epochs continue to see validation accuracy marginally surpass training accuracy.

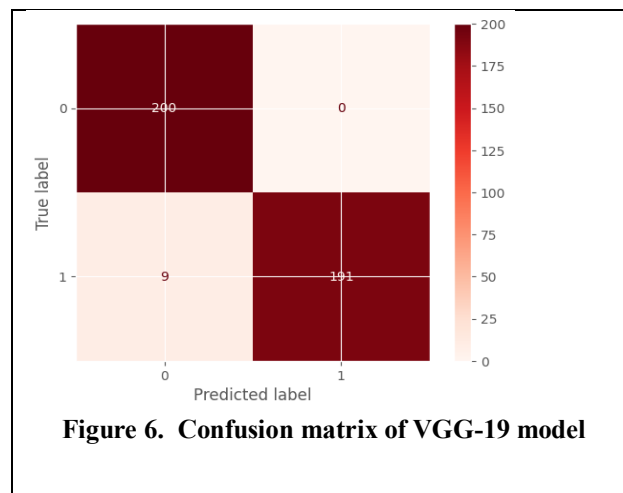


Figure 6. Confusion matrix of VGG-19 model

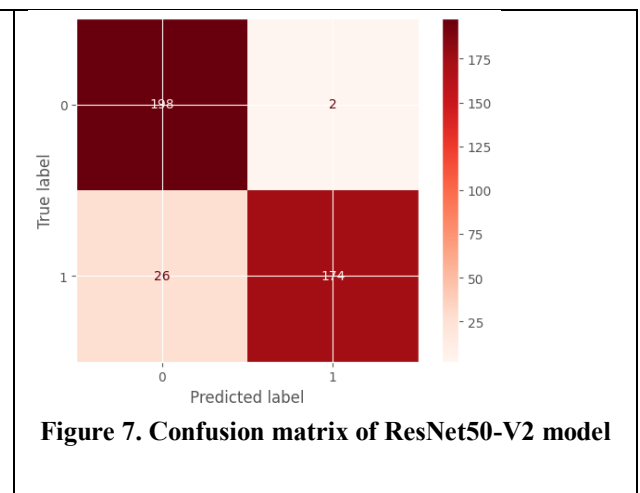


Figure 7. Confusion matrix of ResNet50-V2 model

As illustrated in Figure 6, the confusion matrix for the VGG-19 model demonstrates that all 200 samples from class 0 were accurately classified, with no mistakes. For class 1, the model correctly predicted 191 samples but misclassified 9 samples as class 0. which means the model is very accurate, especially for class 0, and makes only a few errors for class 1.

Figure 7 presents the ResNet50-V2 confusion matrix, revealing 198 accurate predictions for class 0 with 2 misclassifications, while class 1 achieved 174 correct classifications with 26 instances incorrectly labeled as class 0. These results indicate inferior performance relative to the VGG-19 architecture.

II. The experimental result of second stage

After evaluating the performance of each CNN individually, the CBAM module is applied with VGG-19 to enhance extracted feature and increase the performance of the system. The result is shown in Table 6, in addition fig.8.

Table 6. The performance metrics after applying CBAM module for validation and testing dataset on VGG-19 pre-trained models

Pre-trained VGG-19 model after using CBAM block	Accuracy (%)	Precision (%)	Recall (%)	F1- Score (%)	Roc-AUC (%)
validation	99.75	99.75	99.75	99.74	99.99
testing	99.99	100	100	100	100

As demonstrated in Table 6, incorporating the CBAM attention mechanism into the pre-trained VGG-19 architecture enhanced its capability to extract meaningful features from mammographic data. During validation, the integrated model attained 99.75% across accuracy, precision, and recall metrics, with an F1-measure of 99.74% and AUC score of 99.99%. Test set evaluation yielded exceptional results, achieving perfect scores (100%) for all performance indicators: accuracy, precision, recall, F1-score, and ROC-AUC. These outcomes confirm that the VGG-19-CBAM combination successfully identifies critical mammographic characteristics, enabling highly precise breast cancer detection.

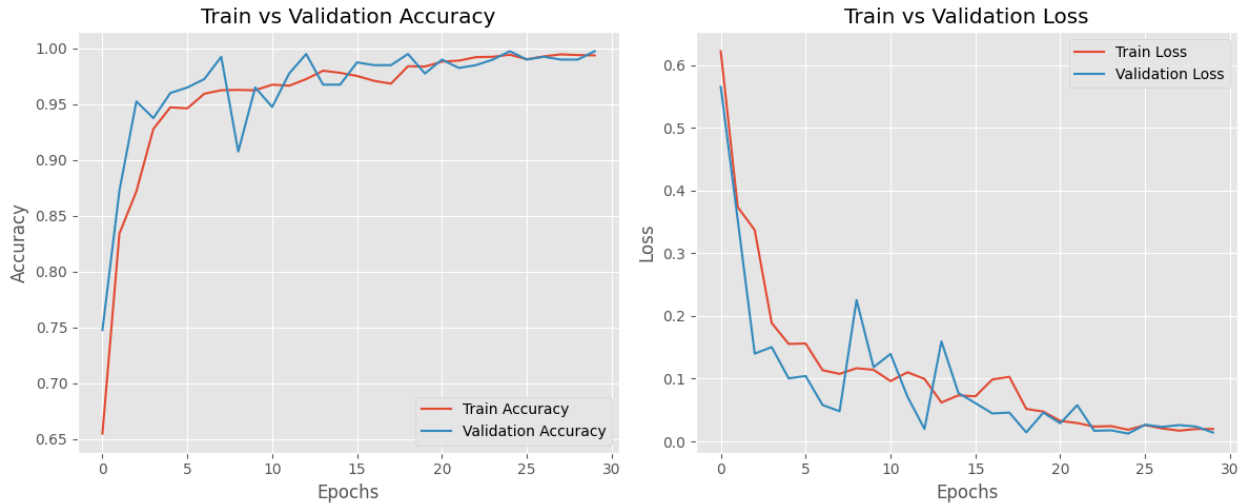


Figure 8. (Train-Validation) accuracy and loss through applying CBAM with VGG-19 pretrained model.

as illustrated in figure 8. In the initial epochs, both the training and validation accuracy curves increase quickly until leveling out around 1.0 and showing minimal difference from one another. The loss curve demonstrates that, with the validation loss changing slightly, both training and validation loss decrease off dramatically in the early epochs and thereafter remain close to zero. These findings indicate the effective generalization capability of the VGG-19 model with CBAM module, which attained minimal loss and elevated accuracy for both training and validation datasets.

CONCLUSION.

This research presents a sophisticated deep learning approach for mammogram-based breast cancer diagnosis employing two established convolutional neural networks (VGG-19 and ResNet50), subsequently enhanced through CBAM integration with the superior performing model. The methodology comprises two distinct stages. The initial phase evaluates both CNN architectures using an 80% training allocation and 20% for validation and testing combined, identifying the optimal network. The subsequent phase incorporates CBAM into the selected architecture to refine feature extraction and elevate classification performance.

Experimental findings revealed that VGG-19 enhanced with CBAM surpassed both individual CNN implementations and comparable studies from existing literature. The integrated VGG-19-CBAM configuration attained 99.99% classification accuracy with perfect precision (100%) on test data, alongside complete AUC coverage (100%). These outcomes emphasize how attention modules strengthen the network's capability to identify crucial mammographic regions, thereby improving diagnostic precision.

Additionally, the findings validate that combining pre-trained models with attention modules successfully addresses difficulties posed by unbalanced classes and intricate mammographic patterns. CBAM integration enhanced the network's responsiveness to essential characteristics while maintaining consistent performance throughout various evaluation metrics including accuracy, precision, sensitivity, F1-measure, and ROC-AUC scores.

Subsequent research will emphasize model validation using expanded heterogeneous databases like INbreast and CBIS-DDSM for assessing transferability, implementing interpretability methods (such as Grad-CAM or SHAP) to enable visual explanations supporting clinical decisions, adapting the framework for multiple category classification encompassing various breast abnormalities and tissue irregularities, and integrating the system into real-time computer-aided diagnosis platforms to support radiological early detection practices.

Declaration

Competing interests

The investigations presented in this publication were not affected by any identified competing financial interests or personal relationships of the authors.

Data Availability

The data are available on the following URL <https://data.mendeley.com/datasets/ywsbh3ndr8/5>.

Funding

This investigation did not obtain any specific funding from public agencies.

Credit authorship contribution statement.

A. M. G.: Methodology, Software, Writing – original draft.

H. M. A.: Conceptualization, Methodology, Supervision, Writing – review & editing.

H. B. N.: Conceptualization, Supervision, Writing – review & editing.

F.W. Z.: Conceptualization, Writing, Supervision – review & editing.

References

- [1] S. Sung et al., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA Cancer J. Clin.*, vol. 71, no. 3, pp. 209–249, 2021, doi: 10.3322/caac.21660.
- [2] A. D. A. Maidment and M. J. Albert, "Digital Mammography," *Radiologic Clinics of North America*, vol. 48, no. 5, pp. 903–923, 2010, doi: 10.1016/j.rcl.2010.06.007.
- [3] M. J. Yaffe, "Mammography and breast cancer screening: Past, present and future," *Radiology*, vol. 273, no. 2S, pp. S28–S52, 2014, doi: 10.1148/radiol.14140370.
- [4] Roslidar, R., Rahman, A., Muharar, R., Syahputra, M. R., Arnia, F., Syukri, M., ... & Munadi, K. A review on recent progress in thermal imaging and deep learning approaches for breast cancer detection. *IEEE access*, 8, 116176-116194. 2020.
- [5] Guo, Y., Shang, X., & Li, Z. Identification of cancer subtypes by integrating multiple types of transcriptomics data with deep learning in breast cancer. *Neurocomputing*, 324, 20-30. 2019.
- [6] Coccia, M. Deep learning technology for improving cancer care in society: New directions in cancer imaging driven by artificial intelligence. *Technology in Society*, 60, 101198. 2020.
- [7] Kumar, Y., Koul, A., Singla, R., & Ijaz, M. F. Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda. *Journal of ambient intelligence and humanized computing*, 14(7), 8459-8486. 2023.
- [8] Han, H., Li, Y., & Zhu, X. Convolutional neural network learning for generic data classification. *Information Sciences*, 477, 448-465. 2019.
- [9] Rasekh, M., Karami, H., Wilson, A. D., & Gancarz, M. Performance analysis of MAU-9 electronic-nose MOS sensor array components and ANN classification methods for discrimination of herb and fruit essential oils. *Chemosensors*, 9(9), 243. 2021.
- [10] X. Wang, Y. Zhang, L. Li, and J. Chen, "RcdNet: Deep separable convolution with CBAM for breast ultrasound lesion classification," *Frontiers in Bioengineering and Biotechnology*, vol. 13, pp. 1–12, Jan. 2025. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fbioe.2025.1526260/full>
- [11] S. Zeng, H. Chen, R. Jing, W. Yang, L. He, T. Zou, P. Liu, B. Liang, D. Shi, W. Wu, Q. Lin, Z. Ma, J. Zha, Y. Zhong, X. Zhang, G. Shao, and P. Gong, "An assessment of breast cancer HER2, ER, and PR expressions based on mammography using deep learning with convolutional neural networks," *Scientific Reports*, vol. 15, Art. no. 4826, pp. 1–10, Feb. 2025. [Online]. Available: <https://doi.org/10.1038/s41598-025-03709-z>.
- [12] L. O. Boro and G. Nandi, "CBAM-RIUnet: Breast Tumor Segmentation With Enhanced Breast Ultrasound and Test-Time

- Augmentation,” *Ultrasound Quarterly*, vol. 47, no. 1, 2025, doi: [10.1177/01617346241276411](https://doi.org/10.1177/01617346241276411).
- [13] F. S. Alkhalefah, M. S. Alshehri, M. Alshamrani, and A. A. Alshehri, “MOB-CBAM: A dual-channel attention-based deep learning generalizable model for breast cancer molecular subtypes prediction using mammograms,” *Computer Methods and Programs in Biomedicine*, vol. 241, p. 107853, May 2024, doi: 10.1016/j.cmpb.2024.107853
- [14] S. Aggarwal, M. Garg, A. Kumar, and R. Kapila, “Breast lesions segmentation from ultrasound images using DeepLabV3+ model with channel and spatial attention mechanism,” *Discover Sustainability*, vol. 5, p. 217, 2024, doi: 10.1007/s43621-024-00424-x.
- [15] Y. Alashban, “Breast cancer detection and classification with digital breast tomosynthesis: a two-stage deep learning approach,” *Diagnostic and Interventional Radiology*, vol. 31, no. 3, pp. 206–214, 2025, doi: 10.4274/dir.2024.242923. (E-pub Dec. 9, 2024).
- [16] N. Sengodan, “CBAM-EfficientNetV2 for Histopathology Image Classification using Transfer Learning and Dual Attention Mechanisms,” *arXiv preprint arXiv:2410.22392*, Oct. 2024. [Online]. Available: <https://arxiv.org/abs/2410.22392>
- [17] A. Mehmood, Y. Hu, and S. H. Khan, “A Novel Channel Boosted Residual CNN-Transformer with Regional-Boundary Learning for Breast Cancer Detection,” *arXiv preprint arXiv:2503.15008*, Mar. 2025. [Online]. Available: <https://arxiv.org/abs/2503.15008>.
- [18] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, “Image segmentation using deep learning: A survey,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 7, pp. 3523–3542, Jul. 2022, doi: 10.1109/TPAMI.2021.3059968.
- [19] G. Litjens et al., “A survey on deep learning in medical image analysis,” *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017, doi: 10.1016/j.media.2017.07.005.
- [20] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, “A survey of the recent architectures of deep convolutional neural networks,” *Artif. Intell. Rev.*, vol. 53, pp. 5455–5516, 2020, doi: 10.1007/s10462-020-09825-6.
- [21] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.
- [22] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, “A survey of the recent architectures of deep convolutional neural networks,” *Artif. Intell. Rev.*, vol. 53, pp. 5455–5516, 2020, doi: 10.1007/s10462-020-09825-6.
- [23] X. Li, Z. Wang, H. Wang, and W. Tang, “CBAM-Enhanced Deep Learning for Breast Cancer Histopathological Image Classification,” *Computers in Biology and Medicine*, vol. 151, p. 106262, 2023, doi: 10.1016/j.combiomed.2022.106262.
- [24] K. Kim, S. Lee, and H. Jang, “A Lightweight Attention Module for Efficient CNNs in Medical Imaging,” *Sensors*, vol. 23, no. 2, p. 612, 2023, doi: 10.3390/s23020612.
- [25] Huang, M. L., & Lin, T. Y. Dataset of breast mammography images with masses. *Data in brief*, 31, 105928. 2020.