ISSN: 2458-8989

Natural and Engineering Sciences

NESciences, 2026, 11 (3): 215-230 doi: 10.28978/nesciences.263018

Attention-Enhanced EfficientNet-B0 with CBAM and Grad-CAM for Explainable Breast Ultrasound Image Classification

Amal Obaid[1]*, Bilal Ozturk[2]

[1,2] Department of Computer Engineering, Istanbul Aydin University, Istanbul, Turkey
E-mails: amalobaid@stu.aydin.edu.tr[1], bilalo@aydin.edu.tr[2]

Abstract

Breast cancer is one of the most common diseases affecting women worldwide, and early diagnosis plays an important role in improving treatment outcomes and reducing mortality. Among the available imaging techniques, breast ultrasound is widely used because it is safe, non-invasive, and particularly effective for examining dense breast tissue. However, interpreting ultrasound images remains difficult due to image noise, low contrast, and the visual similarity between benign and malignant lesions.

This study presents an explainable deep learning framework for breast ultrasound image classification by combining EfficientNet-B0 with the Convolutional Block Attention Module (CBAM). In addition, Gradient-weighted Class Activation Mapping (Grad-CAM) was incorporated to provide visual explanations for the model's predictions and improve their interpretability.

The proposed framework was evaluated using the publicly available BUSI dataset, which contains 780 breast ultrasound images divided into three classes: normal, benign, and malignant. The dataset was split into 70% for training, 15% for validation, and 15% for testing. To evaluate the effectiveness of the proposed method, two baseline models, ResNet50 and EfficientNet-B0, were implemented and compared with the proposed EfficientNet-B0 + CBAM model. ResNet50 was trained for 2 epochs, while EfficientNet-B0 and EfficientNet-B0 + CBAM were trained for 5 epochs using the Adam optimizer with a learning rate of 0.0001 and a batch size of 16.

The experimental results showed that the proposed model achieved a test accuracy of **81.20%**, compared with **79.49%** obtained by the baseline EfficientNet-B0 model. In addition, the recall for the malignant class increased from **0.6452** to **0.6774**, demonstrating an improved ability to identify malignant breast lesions. The generated Grad-CAM visualizations also showed that the model focused on clinically relevant regions of the ultrasound images, providing useful visual explanations for its predictions.

Overall, the results indicate that integrating CBAM with EfficientNet-B0 can improve classification performance while Grad-CAM enhances the transparency of the decision-making process. These findings suggest that the proposed framework has the potential to support radiologists by providing accurate predictions together with interpretable visual explanations.

Keywords
Breast Cancer, Deep Learning, EfficientNet-B0, CBAM, Ultrasound Classification, Grad-CAM

Available online: 01/09/2026

1. Introduction

Breast cancer remains one of the most serious health challenges affecting women worldwide and continues to be a major cause of cancer-related deaths (Liu et al., 2025). Because of this, detecting the disease as early as possible is essential for improving treatment options and increasing survival rates.

**Corresponding Author: Amal Obaid, E-mail: amalobaid@stu.aydin.edu.tr*

Several medical imaging techniques are available for breast cancer diagnosis, including mammography, magnetic resonance imaging (MRI), and ultrasound (Chen et al., 2025).
Among these imaging modalities, breast ultrasound is widely used in clinical practice, especially for women with dense breast tissue. It is considered a safe, non-invasive, and relatively low-cost imaging technique. However, interpreting ultrasound images is still a difficult task because they usually contain speckle noise, low image contrast, and considerable visual similarity between benign and malignant lesions (Liu et al., 2025; Chen et al., 2025).
With the rapid development of artificial intelligence, deep learning has become one of the most successful approaches for medical image analysis. Convolutional neural networks (CNNs) have shown excellent performance in image classification because they can automatically learn useful image features without relying on handcrafted feature extraction methods (Lepcha et al., 2025). More recently, transformer-based architectures have also attracted attention in medical imaging due to their ability to capture global image relationships and long-range dependencies. Vision Transformers (ViTs), in particular, have achieved encouraging results in different medical imaging tasks, including classification, detection, and segmentation (Aburass et al., 2025).
Many recent studies have investigated attention mechanisms and explainable artificial intelligence to improve breast ultrasound image classification. EfficientNet-based models combined with explainability techniques have reported promising performance for breast ultrasound diagnosis (Latha et al., 2024). Other researchers have improved the classification of benign and malignant breast tumors by developing enhanced deep learning models (Liu et al., 2025; Wei et al., 2024). In addition, transformer-based ensemble models have demonstrated the capability of Vision Transformers to perform multi-class breast ultrasound classification effectively (Yıldırım et al., 2025). These studies indicate that attention mechanisms, transformer architectures, and explainable AI are becoming increasingly important for improving both classification accuracy and model interpretability.
Although considerable progress has been achieved, distinguishing benign from malignant breast lesions remains a challenging problem because many ultrasound images share similar visual characteristics. From a clinical perspective, reducing false-negative predictions is particularly important since classifying a malignant lesion as benign may delay diagnosis and treatment, which can negatively affect patient outcomes.
To address this challenge, this study proposes a breast ultrasound image classification framework based on EfficientNet-B0 (Tan & Le, 2019), enhanced with the Convolutional Block Attention Module (CBAM) to improve feature representation (Woo et al., 2018). In addition, Grad-CAM is employed to generate visual explanations that help interpret the model's predictions.
The main contribution of this work is the integration of EfficientNet-B0, CBAM, and Grad-CAM within a single framework for breast ultrasound image classification. While many previous studies have mainly focused on improving overall classification accuracy, the proposed approach gives additional attention to malignant lesion detection and the interpretability of model predictions. By combining an attention mechanism with explainable artificial intelligence, the proposed framework aims to improve classification performance while providing visual evidence that may increase the transparency of the decision-making process and support future clinical applications.

2. Related Work

2.1 CNN-Based Breast Ultrasound Classification

Deep learning has become one of the most widely used approaches for breast ultrasound image classification. In particular, convolutional neural networks (CNNs) have demonstrated strong performance because they can automatically learn meaningful image features without relying on manually designed descriptors (Lepcha et al., 2025). Many studies have applied different CNN architectures to classify breast ultrasound images. EfficientNet-based models have shown promising performance, while other advanced CNN models have also reported improvements in distinguishing benign and malignant breast lesions (Liu et al., 2025; Wei et al., 2024; Latha et al., 2024). Despite these encouraging results, many existing CNN-based studies mainly emphasize classification accuracy, with less attention given to model interpretability and clinical applicability.

2.2 Attention-Based Deep Learning Approaches

To improve feature representation, many researchers have incorporated attention mechanisms into deep learning models. These mechanisms enable the network to focus on the most informative image regions while reducing the influence of less relevant features. Recent studies have shown that attention-based architectures can improve breast cancer classification performance (Mustafa et al., 2025; Ganie et al., 2025). One of the most popular attention mechanisms is the Convolutional Block Attention Module (CBAM), which combines channel attention and spatial attention while introducing only a small computational overhead (Woo et al., 2018). Although attention mechanisms improve feature extraction, many published studies still provide limited explanation of how their models reach the final prediction.

2.3 Transformer-Based Approaches

Transformer-based models have recently emerged as an alternative to conventional CNN architectures in medical image analysis. Vision Transformers (ViTs) are capable of learning global image relationships and long-range dependencies, making them suitable for different medical imaging applications (Aburass et al., 2025). Previous studies have reported successful use of transformer ensemble models for multi-class breast ultrasound image classification (Yıldırım et al., 2025). In addition, hybrid architectures that combine EfficientNet with Vision Transformers have achieved encouraging diagnostic performance (Qahtan Mohammed et al., 2026). Nevertheless, transformer-based approaches generally require larger datasets and higher computational resources than traditional CNN models.

2.4 Explainable AI in Breast Cancer Imaging

As deep learning models become more accurate, understanding how they make predictions has become increasingly important. Explainable artificial intelligence (XAI) techniques help improve the transparency of deep learning systems by identifying the image regions that influence model decisions. Among these techniques, Grad-CAM is one of the most commonly used visualization methods because it highlights the areas that contribute most to the final prediction. Previous research has also discussed both the advantages and the limitations of Grad-CAM in medical imaging applications, emphasizing the need for careful interpretation of attention maps (Suara et al., 2023). Several recent studies have successfully combined explainability methods with breast cancer diagnosis systems (Talaat et al., 2024; Alom et al., 2025; Zou & Miao, 2025). Other researchers have investigated Grad-CAM++, hybrid explainability methods, and different XAI strategies for breast cancer detection (Ameen, 2026; Ahmad et al., 2025). Although these studies demonstrate the value of explainable AI, many of them focus either on improving prediction performance or on providing visual explanations, without achieving a balanced combination of both. In addition, interpretable deep transfer learning approaches have demonstrated that combining classification performance with explainability can improve the usefulness of breast ultrasound classification models (Abbadi et al., 2025).

2.5 Research Gap

Although substantial progress has been achieved in breast ultrasound image classification, there is still a need for lightweight models that combine high classification performance with effective malignant lesion detection and meaningful visual explanations. Existing studies often emphasize either predictive accuracy or explainability, while fewer approaches address both aspects within a computationally efficient framework. Therefore, this study proposes a lightweight framework based on EfficientNet-B0, the Convolutional Block Attention Module (CBAM), and Grad-CAM explainability. The proposed approach aims to improve classification performance, increase malignant lesion detection, and provide visual explanations that may support clinical interpretation of breast ultrasound images. Furthermore, recent review studies have highlighted the growing importance of artificial intelligence and explainable deep learning methods for breast cancer diagnosis across different imaging modalities (Mashekova et al., 2025).

3. Materials and Methods

In this section, the employed dataset and data preparation process, the proposed model and training methodology will be discussed in detail.

3.1 Dataset

Experiments on the Breast Ultrasound Images (BUSI) dataset were carried out (Liu et al., 2025), (Chen et al., 2025). The BUSI dataset contains grayscale ultrasound image data for breast classification

problems and includes three categories: normal, benign, and malignant. The dataset contains 780 images in total; specifically, there are 133 normal, 437 benign, and 210 malignant images.
The dataset used for BUSI shows some signs of class imbalance, where images classified as “benign” are the largest class, then followed by “malignant” and “normal.” The imbalance could affect learning and introduce biases towards classes that have more samples. In order to reduce this problem, a stratified splitting technique was applied to the data, ensuring similar class proportions across the training, validation, and testing sets. No class balancing methods such as oversampling or class weighting were utilized during training.
For the purpose of fair evaluation, a stratified split among the dataset's images was employed with a split ratio of 70% training set, 15% validation set, and 15% test set using a fixed random seed.

3.2 Preprocessing and Data Augmentation

All images were resized to 224 × 224 pixels to match the input requirements of the pretrained convolutional neural network architectures. Since the original BUSI images are grayscale, they were converted into three-channel images to ensure compatibility with ImageNet-pretrained models.
Data normalization was performed using the standard ImageNet mean and standard deviation values (Shobayo & Saatchi, 2025; Halder et al., 2024). Furthermore, Gaussian filtering was applied as a preprocessing technique to reduce image noise. Preliminary experiments were conducted to compare Gaussian filtering, CLAHE, and Gaussian+CLAHE enhancement methods.
To evaluate the effectiveness of the image enhancement techniques, three quantitative image quality metrics were employed: Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Signal-to-Noise Ratio (SNR). These metrics were calculated by comparing the original ultrasound images with their corresponding enhanced images.
The Mean Squared Error (MSE) was calculated as:

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(I_i - K_i)^2$$

where I_i represents the pixel value in the original image, K_i represents the corresponding pixel value in the enhanced image, and N is the total number of pixels.
The Peak Signal-to-Noise Ratio (PSNR) was computed using:

$$PSNE = \log_{10}\left(\frac{255}{\sqrt{MSE}}\right)$$

where 255 represents the maximum possible pixel intensity value for 8-bit images.
The Signal-to-Noise Ratio (SNR) was calculated as:

$$SNR = 10\log_{10}\left(\frac{Singal\ Power}{Noise\ Power}\right)$$

where:
Signal Power = mean(I^2)
Noise Power = mean (($I - K$) ^{2})
Lower MSE values indicate less distortion introduced during preprocessing, whereas higher PSNR and SNR values indicate better image quality preservation and more effective noise reduction.
As shown in Table 3.1, Gaussian filtering achieved better performance compared to other enhancement techniques.

Table 3.1: Image quality evaluation of enhancement methods using MSE, PSNR, and SNR

Method	**MSE**	**PSNR (dB)**	**SNR (dB)**
Gaussian	18.8590	36.1566	27.5124
CLAHE	1005.6584	18.1926	9.5484
Gaussian + CLAHE	913.5431	18.6241	9.9798

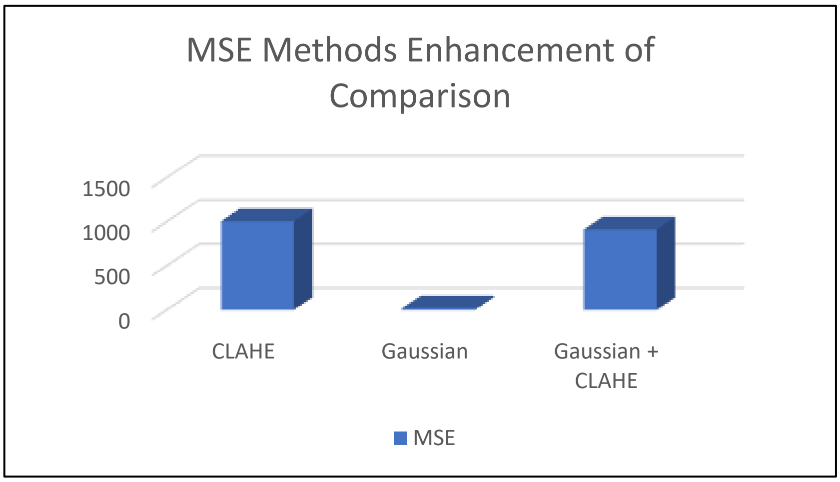

Figure 3.1: Comparison of image enhancement methods based on MSE values.

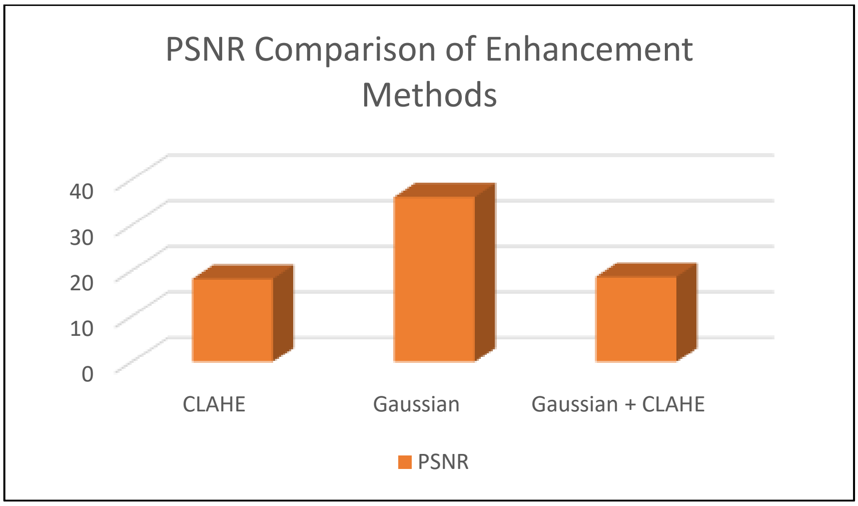

Figure 3.2: Comparison of image enhancement methods based on PSNR values.

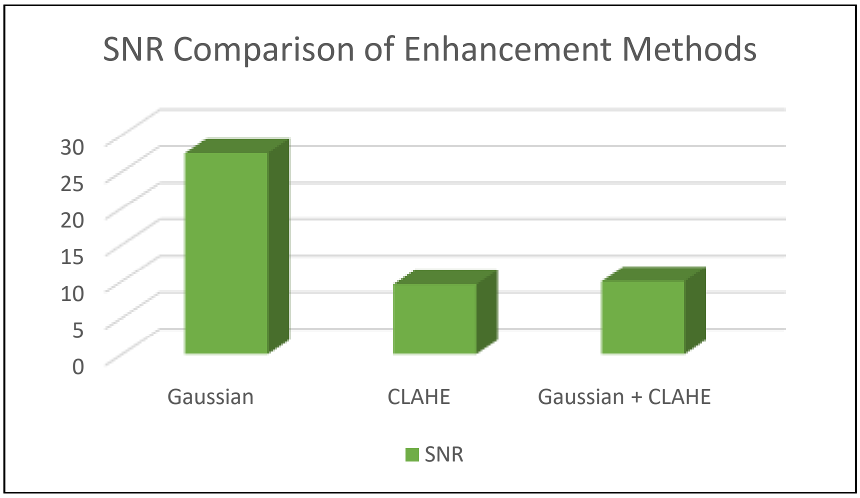

Figure 3.3: Comparison of image enhancement methods based on SNR values.

As shown in Table 3.1 and Figures 3.1–3.3, the Gaussian filtering method achieved the best performance among the evaluated techniques. The lowest MSE (18.8590), indicating minimal distortion from the original image, and the highest PSNR (36.1566 dB) and SNR (27.5124 dB), which reflect superior image quality and noise reduction.

To improve model generalization and reduce overfitting, data augmentation techniques were applied to the training set, including horizontal flipping, small rotations, and brightness/contrast adjustments. Data augmentation was not applied to the validation and test sets in order to ensure an unbiased evaluation of model performance.

3.3 Model Architecture

In the current research work, EfficientNet-B0 has been chosen as the main backbone model due to its well-balanced architecture and efficient performance in processing image classification problems (Tan & Le, 2019). Pretrained weights from ImageNet have been utilized to initialize the backbone model.

For better feature extraction, the Convolutional Block Attention Module (CBAM) was integrated into the EfficientNet-B0 architecture (Woo et al., 2018). CBAM combines channel attention and spatial attention mechanisms to emphasize the most informative features while suppressing less relevant information.

Specifically, the CBAM module was inserted after the final feature extraction stage of the EfficientNet-B0 backbone and before the classification layer. The feature maps generated by EfficientNet-B0 were first refined through channel attention to identify the most important feature channels and then processed through spatial attention to highlight the most relevant image regions. The refined feature maps were subsequently passed to the classification layer for prediction of the three BUSI classes (normal, benign, and malignant).

The attention mechanism introduces only a limited computational overhead because CBAM consists of lightweight channel and spatial attention operations. This design allows the model to improve feature representation while maintaining the computational efficiency that characterizes EfficientNet-B0.

The overall pipeline of the proposed approach is illustrated in Figure 3.4.

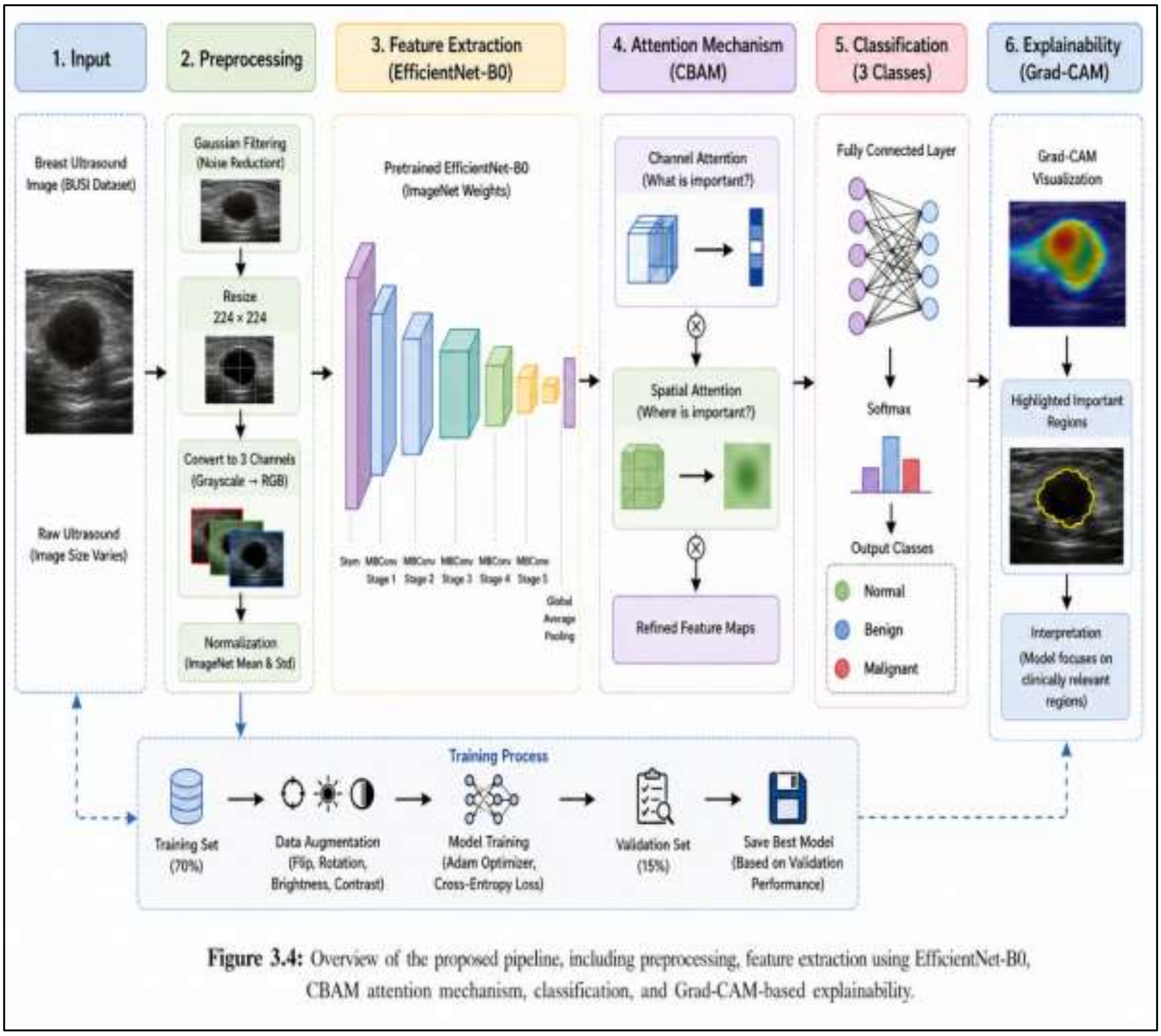

Figure 3.4: Overview of the proposed pipeline, including preprocessing, feature extraction using EfficientNet-B0, CBAM attention mechanism, classification, and Grad-CAM-based explainability.

As illustrated in Figure 3.4, the model processes the input image through a sequence of feature extraction and attention stages before classification.

3.4 Training Procedure

Training was performed using the Adam optimization algorithm with a fixed learning rate of 1×10^{-4}. No learning-rate scheduling strategy was applied during training. A cross-entropy loss function was chosen because of the multi-class nature of the classification task.

ResNet50 (He et al., 2016) was trained for 2 epochs during the baseline evaluation, while the proposed EfficientNet-B0 with CBAM was trained for 5 epochs. A batch size of 16 was used, and the best model was selected based on validation accuracy.

To assess model convergence and training stability, training loss, validation loss, and validation accuracy were monitored throughout the training process. As illustrated in Figures 3.5–3.7, both EfficientNet-B0 and EfficientNet-B0 + CBAM exhibited a consistent decrease in training and validation loss across epochs, indicating stable learning behavior. In addition, validation accuracy showed a progressive improvement during training, reaching its highest values in the final epochs. These observations suggest that both models achieved stable convergence despite the limited number of training epochs. The best-performing model was selected based on validation accuracy and saved as the final checkpoint for subsequent evaluation.

Training was done by using the backpropagation technique and validation was periodically conducted for performance estimation.

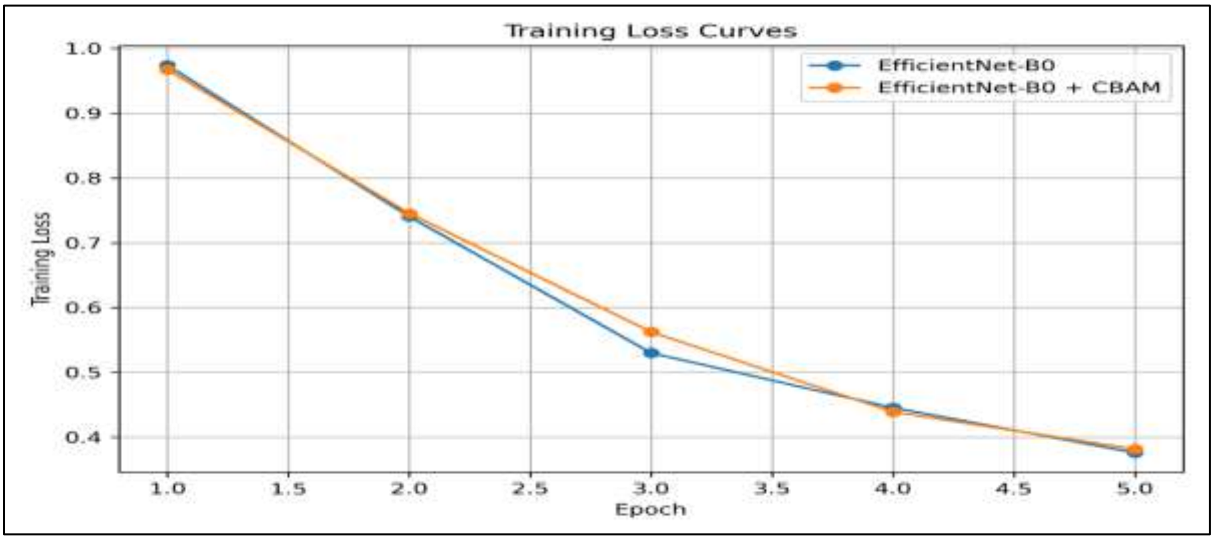

Figure 3.5. Training loss curves for EfficientNet-B0 and EfficientNet-B0 + CBAM during the five training epochs.

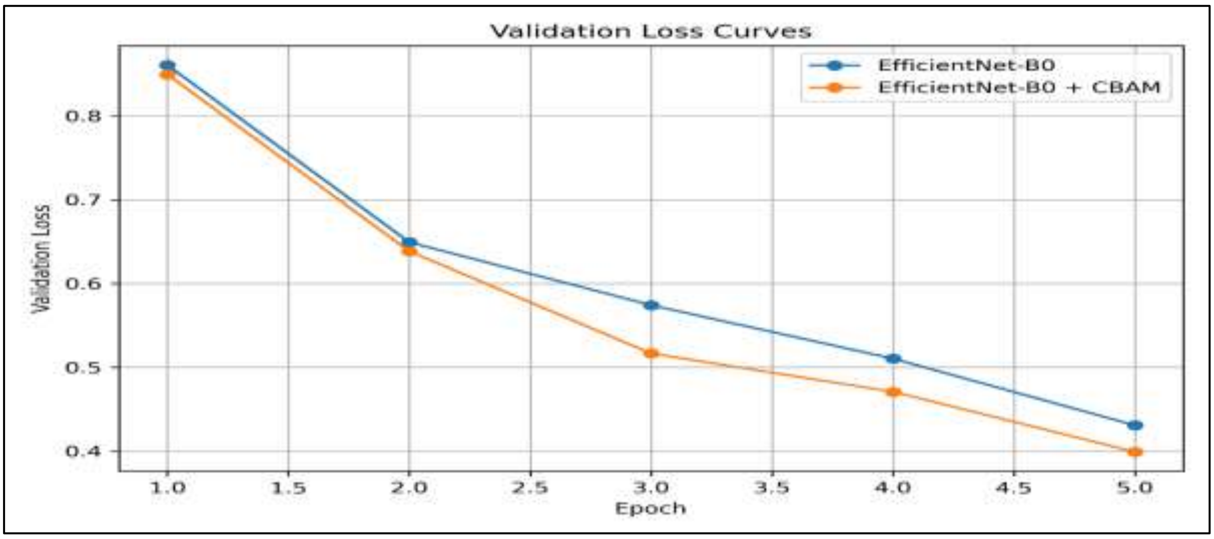

Figure 3.6. Validation loss curves for EfficientNet-B0 and EfficientNet-B0 + CBAM during training.

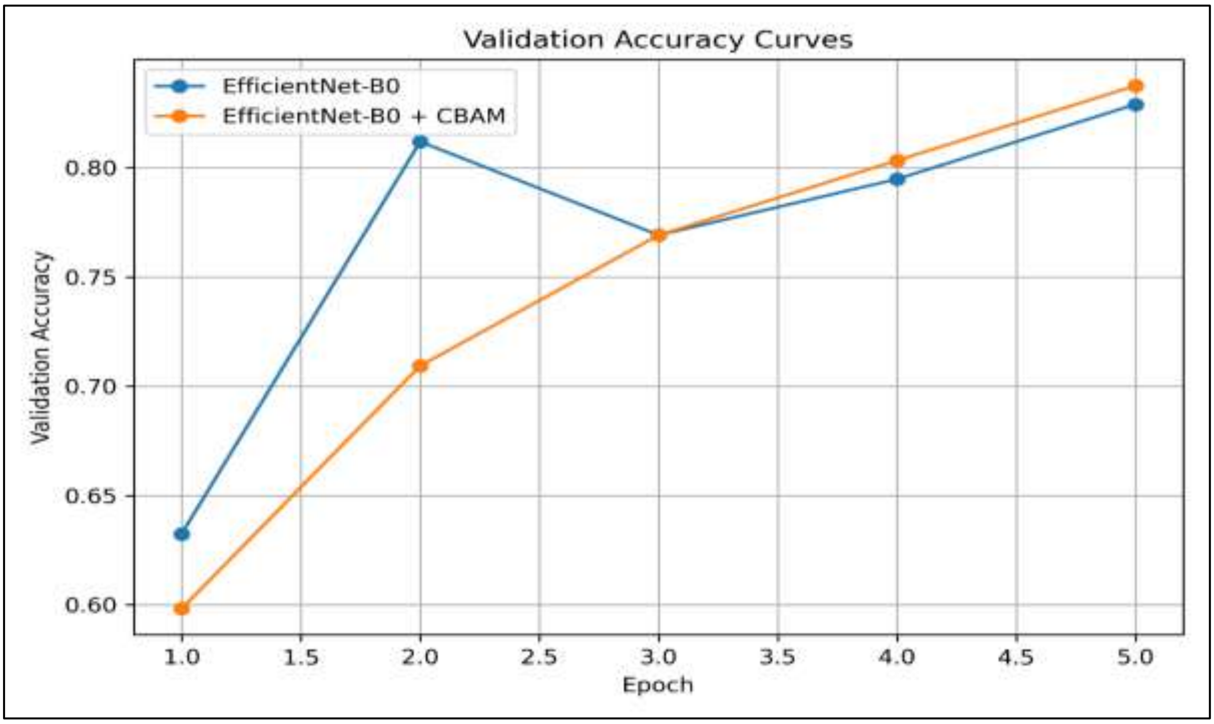

Figure 3.7. Validation accuracy curves for EfficientNet-B0 and EfficientNet-B0 + CBAM during training.

As shown in Figures 3.5–3.7, both EfficientNet-B0 and EfficientNet-B0 + CBAM demonstrated stable convergence throughout the training process. Training and validation losses consistently decreased across epochs, while validation accuracy gradually improved. The EfficientNet-B0 + CBAM model achieved slightly higher validation accuracy and lower validation loss in the final epochs, indicating the positive contribution of the attention mechanism.

3.5 Evaluation Metrics

For the performance evaluation of the proposed model, several metrics were used: Accuracy, Precision, Recall, and F1-Score. As the detection of malignant cases is crucial from the clinical point of view, the focus was made on Recall of malignant cases in particular.
Also, confusion matrix was used to analyse misclassification between classes.

3.6 Explainability Method

To get insights regarding model predictions, Grad-CAM technique was used. With its help, the most relevant parts of input images can be identified as the source of prediction made by the model.

4. Results and Discussion

This section highlights the results generated via the BUSI dataset and discusses the performance of the proposed model vis-à-vis the baseline architectures.

4.1 Baseline Model Performance Comparison

The performance of the baseline models, including ResNet50 and EfficientNet-B0, is summarized in Table 4.1. The comparison with the proposed EfficientNet-B0 + CBAM model is presented separately in Table 4.3.

Table 4.1: Baseline Model Performance on BUSI Dataset

Model	Validation Accuracy (%)	Test Accuracy (%)	Role
ResNet50	68.38	69.23	Comparative baseline
EfficientNet-B0	83.76	79.49	Primary baseline

As shown in Table 4.1, EfficientNet-B0 outperformed ResNet50 in terms of overall accuracy, achieving 79.49% compared to 69.23%. This improvement can be attributed to the more efficient feature scaling strategy employed by EfficientNet.

4.2 Class-wise Performance Analysis

To provide a more detailed evaluation, the classification performance for each class is presented in Table 4.2.

Table 4.2: Classification Performance per Class

Class	Precision	Recall	F1-score
Normal	0.6923	0.9000	0.7826
Benign	0.8462	0.8333	0.8397
Malignant	0.7692	0.6452	0.7018

As shown in Table 4.2, the model achieved strong performance for the benign class, with high precision and recall values. However, the malignant class remains more challenging, as reflected by a relatively lower recall.

These results correspond to the baseline EfficientNet-B0 model. The improvement achieved by the proposed EfficientNet-B0 + CBAM model is presented separately in Table 4.3.

4.3 Confusion Matrix Analysis

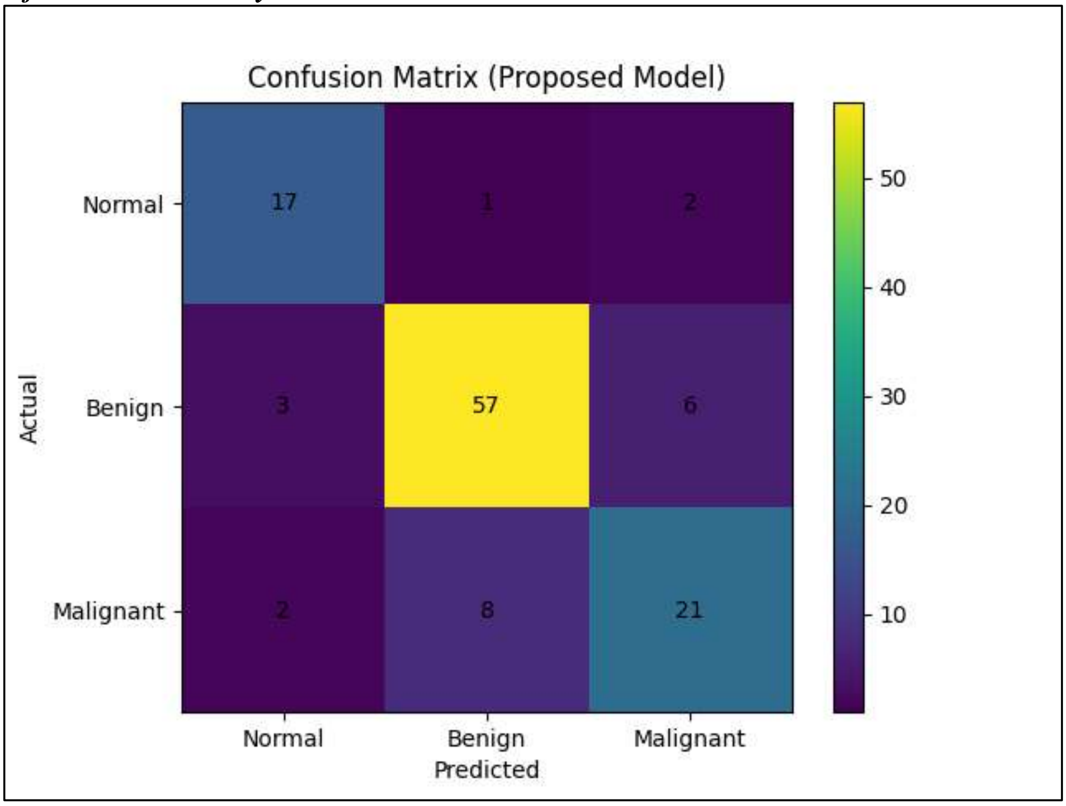

Figure 4.1: Confusion matrix of the proposed EfficientNet-B0 + CBAM model

As shown in Figure 4.1, most normal and benign cases are correctly classified. However, some confusion is observed between benign and malignant classes, which is expected due to their visual similarity in ultrasound images.

The reduction in misclassified malignant cases compared to the baseline model indicates that the proposed approach improves the model's ability to detect cancerous lesions.

4.4 Proposed Model Improvement

Table 4.3 presents a comparison between the baseline EfficientNet-B0 model and the proposed EfficientNet-B0 + CBAM model.

Table 4.3: Effect of CBAM on Model Performance

Model	Test Accuracy (%)	Malignant Recall	Malignant F1
EfficientNet-B0	79.49	0.6452	0.7018
EfficientNet-B0 + CBAM	81.20	0.6774	0.7000

As shown in Table 4.3, the integration of CBAM improved the test accuracy from 79.49% to 81.20% and increased malignant recall from 0.6452 to 0.6774.

4.5Additional Benchmarking Comparison

Table 4.4: Comparison with Additional Baseline Architectures

Model	Test Accuracy (%)	F1-score	Malignant Recall
ResNet50	69.23	0.6341	0.6129
MobileNetV3-Large	68.38	0.6894	0.6452
DenseNet121	84.62	0.8429	0.6774
EfficientNet-B0	79.49	0.7018	0.6452
EfficientNet-B0 + CBAM	81.20	0.7000	0.6774

To further validate the effectiveness of the proposed framework, additional benchmarking experiments were conducted using MobileNetV3-Large and DenseNet121 under the same experimental settings.
As shown in Table 4.4, DenseNet121 achieved the highest overall classification accuracy among the evaluated architectures. The proposed EfficientNet-B0 + CBAM model improved test accuracy and malignant recall compared with the baseline EfficientNet-B0 model, while also providing Grad-CAM-based explainability and attention-enhanced feature learning through CBAM.

4.6Additional Performance Metrics

To provide a more comprehensive evaluation of classification performance, additional metrics including sensitivity, specificity, balanced accuracy, Matthews Correlation Coefficient (MCC), and Cohen's Kappa were computed for the proposed model.

Table 4.5 Additional Performance Metrics of the Proposed Model

Metric	Value
Accuracy	81.20%
Sensitivity (Macro Recall)	0.7970
Specificity (Macro)	0.8930
Balanced Accuracy	0.7970
MCC	0.6783
Cohen's Kappa	0.6780

To provide a more comprehensive evaluation of clinical classification performance, additional metrics were computed for the proposed model. The model achieved a macro sensitivity of 0.7970 and a macro specificity of 0.8930, indicating good capability in identifying both positive and negative cases. Furthermore, the balanced accuracy reached 0.7970, while the MCC and Cohen's Kappa values of 0.6783 and 0.6780, respectively, indicate substantial agreement between predicted and ground-truth labels.

4.7Explainability Analysis

To better understand the model's behavior, Grad-CAM was applied to visualize the regions that contribute to the classification decisions.

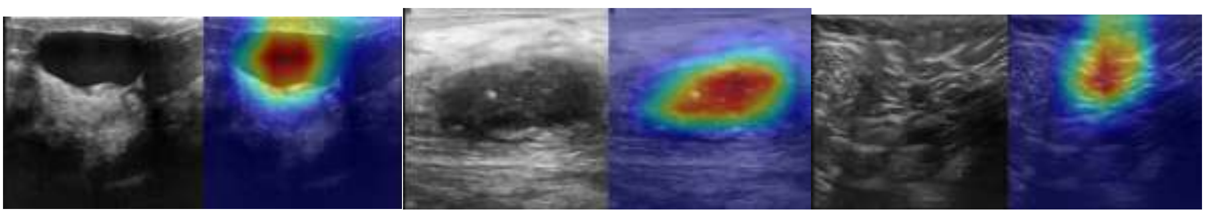

Figure 4.2: Grad-CAM visualization results for selected test samples.

As illustrated in Figure 4.2, correctly classified cases show focused activation over lesion regions, indicating that the model relies on clinically relevant features. In contrast, misclassified cases exhibit attention in less relevant areas, which may explain incorrect predictions.
These observations suggest that the model does not only achieve improved performance but also demonstrates reasonable interpretability.

4.8Discussion
The results overall suggest that the EfficientNet-B0 with CBAM architecture improves classification performance compared with the baseline EfficientNet-B0 model and achieves competitive performance among the evaluated architectures.
The integration of CBAM contributed to improved malignant lesion detection, while Grad-CAM enhanced model interpretability by highlighting diagnostically relevant image regions.
The enhanced malignant recall rate is particularly significant from a clinical perspective because reducing false-negative predictions may help decrease the risk of missed cancer cases during screening. Early identification of malignant lesions is critical for timely treatment planning and improved patient outcomes.
The improvement in malignant recall from 0.6452 to 0.6774 indicates that fewer malignant lesions were incorrectly classified as benign. From a clinical perspective, reducing false-negative cases is particularly important because missed cancer diagnoses may delay diagnosis, treatment initiation, and negatively affect patient prognosis. Even a modest improvement in malignant lesion detection can contribute to earlier clinical intervention and improved patient outcomes. Therefore, the observed increase in malignant recall achieved by the proposed model may have practical significance in breast cancer screening applications.
Although the results are positive, differentiating between benign and malignant lesions proves difficult due to their similarities in appearance. Further research could focus on exploring more sophisticated networks or utilizing bigger data sets.
These findings confirm the potential of incorporating attention mechanisms for improved image classification.

5. Conclusion

This study presented a deep learning approach for the classification of breast ultrasound images using the BUSI dataset. Different CNN architectures were evaluated to investigate their classification performance, and EfficientNet-B0 demonstrated better results than ResNet50 under the adopted experimental settings. To further improve feature extraction, the EfficientNet-B0 architecture was enhanced by incorporating the Convolutional Block Attention Module (CBAM).
The experimental results indicate that the proposed framework performed better than the baseline ResNet50, MobileNetV3-Large, and EfficientNet-B0 models, while achieving competitive performance compared with DenseNet121. In addition, the proposed model improved the detection of malignant lesions and provided visual explanations of its predictions through the Grad-CAM technique.
Besides evaluating the classification performance, this study also focused on model interpretability. The Grad-CAM visualizations showed that, in correctly classified cases, the model mainly concentrated on lesion-related regions, whereas different attention patterns were observed in misclassified images. These observations suggest that the proposed framework provides meaningful visual evidence that can support the interpretation of its predictions.
Although the proposed approach achieved encouraging results, distinguishing between benign and malignant breast lesions remains challenging because of their similar appearance in ultrasound images. Future research may focus on investigating more advanced network architectures, larger datasets, and additional techniques to further improve classification performance.
Overall, the findings of this study demonstrate that combining EfficientNet-B0 with the CBAM attention mechanism and Grad-CAM explainability provides an effective framework for breast ultrasound image classification. The proposed approach improves both classification performance and model interpretability, making it a promising solution for supporting future computer-aided diagnosis systems in clinical practice.

Acknowledgments
The authors would like to express their sincere gratitude to Istanbul Aydin University for providing academic support and a research environment that contributed to this work. The authors would also like to thank the Organization for Women in Science for the Developing World (OWSD) for their support and contribution.

Funding Source
This research received no external funding.

Author Contributions
Conceptualization, methodology, implementation, data analysis, and manuscript preparation were performed by Amal Obaid. Supervision, guidance, and manuscript review were performed by Dr. Bilal ÖZTÜRK.

Data Availability Statement
The dataset used in this study is publicly available. The BUSI dataset can be accessed through publicly available repositories.

Conflicts/Competing Interests
The authors declare no conflict of interest.

Compliance with Ethical Standards
This study used a publicly available dataset and did not involve direct participation of human subjects or animals. Therefore, ethical approval was not required.

References

1. Abbadi, M., Himeur, Y., Atalla, S., & Mansoor, W. (2025). Interpretable deep transfer learning for breast ultrasound cancer detection: A multi-dataset study. arXiv. https://doi.org/10.48550/arXiv.2509.05004
2. Aburass, S., Dorgham, O., Al Shaqsi, J., Abu Rumman, M., & Al-Kadi, O. (2025). Vision transformers in medical imaging: A comprehensive review of advancements and applications across multiple diseases. Journal of Imaging Informatics in Medicine, 38(6), 3928–3971. https://doi.org/10.1007/s10278-025-01481-y
3. Ahmad, I., Amelio, A., Cardone, D., Gill, E. Z., & Scozzari, F. (2025). Explainability in breast cancer detection. CEUR Workshop Proceedings. https://ceur-ws.org/Vol-4121/Ital-IA_2025_paper_115.pdf
4. Alom, M. R., Farid, F. A., Rahaman, M. A., Rahman, A., Debnath, T., Miah, A. S. M., & Mansor, S. (2025). An explainable AI-driven deep neural network for accurate breast cancer detection from histopathological and ultrasound images. Scientific Reports, 15(1), 17531. https://doi.org/10.1038/s41598-025-97718-5
5. Ameen, M. (2026). Explainable Mammogram Analysis with EfficientNetV2 and Grad-CAM++ for Robust Cancer Diagnosis. Diagnostics, 16(1), 105. https://doi.org/10.3390/diagnostics16010105
6. Chen, J., Pan, T., Zhu, Z., et al. (2025). A deep learning-based multimodal medical imaging model for breast cancer screening. Scientific Reports,15, 14696. https://doi.org/10.1038/s41598-025-99535-2
7. Ganie, S. M., Malik, M. B., Aadil, M., Muteeb, G., Farhan, M., & Aatif, M. (2025). Explainable AI based hybrid DRM-Net transfer learning model for breast cancer detection and classification using ultrasound images. Scientific Reports, 15(1), 44170. https://doi.org/10.1038/s41598-025-27934-6
8. Halder, A., Gharami, S., Sadhu, P., et al. (2024). Implementing vision transformer for classifying 2D biomedical images. Scientific Reports, 14, 12567. https://doi.org/10.1038/s41598-024-63094-9
9. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 770–778). https://doi.org/10.1109/CVPR.2016.90
10. Latha, M., Kumar, P. S., Chandrika, R. R., et al. (2024). Revolutionizing breast ultrasound diagnostics with EfficientNet-B7 and explainable AI. BMC Medical Imaging, 24, 230. https://doi.org/10.1186/s12880-024-01404-3

11.Lepcha, D. C., Goyal, B., Dogra, A., Alkhayyat, A., Sahu, P. K., Ali, A., & Kukreja, V. (2025). Deep learning in medical image analysis: A comprehensive review of algorithms, trends, applications, and challenges. CMES - Computer Modeling in Engineering and Sciences, 145(2), 1487–1573. https://doi.org/10.32604/cmes.2025.070964
12.Liu, B., Liu, S., Cao, Z., Zhang, J., Pu, X., & Yu, J. (2025). Accurate classification of benign and malignant breast tumors in ultrasound imaging with an enhanced deep learning model. Frontiers in Bioengineering and Biotechnology, 13, 1526260. https://doi.org/10.3389/fbioe.2025.1526260
13.Liu, J., Pian, L., Chen, J., Zhao, J., Liu, Y., Meng, F., & Zeng, C. (2025). Artificial intelligence in breast ultrasound: A systematic review of research advances. Frontiers in Oncology, 15, 1619364. https://doi.org/10.3389/fonc.2025.1619364
14.Mashekova, A., Zhao, M. Y., Zarikas, V., Mukhmetov, O., Aidossov, N., Ng, E. Y. K., Wei, D., & Shapatova, M. (2025). Review of artificial intelligence techniques for breast cancer detection with different modalities: Mammography, ultrasound, and thermography images. Bioengineering, 12(10), 1110. https://doi.org/10.3390/bioengineering12101110
15.Mustafa, M. A., Erdem, O. A., & Söğüt, E. (2025). Hybrid optimization and explainable deep learning for breast cancer detection. Applied Sciences, 15(15), 8448. https://doi.org/10.3390/app15158448
16.Qahtan Mohammed, Z., Tuama Alhussainy, A., Salman Jasim, I., & Mohamed Moubark, A. (2026). A hybrid model for ultrasound image-based breast cancer diagnosis using EfficientNet-V2 and Vision Transformer. Diagnostics, 16(8), 1176. https://doi.org/10.3390/diagnostics16081176
17.Shobayo, O., & Saatchi, R. (2025). Developments in deep learning artificial neural network techniques for medical image analysis and interpretation. Diagnostics, 15(9), 1072. https://doi.org/10.3390/diagnostics15091072
18.Suara, S., Jha, A., Sinha, P., & Sekh, A. A. (2023). Is Grad-CAM explainable in medical images? arXiv. https://doi.org/10.48550/arXiv.2307.10506
19.Talaat, F. M., Gamel, S. A., El-Balka, R. M., Shehata, M., & ZainEldin, H. (2024). Grad-CAM enabled breast cancer classification with a 3D Inception-ResNet V2: Empowering radiologists with explainable insights. Cancers, 16(21), 3668. https://doi.org/10.3390/cancers16213668
20.Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In K. Chaudhuri & R. Salakhutdinov (Eds.), Proceedings of the 36th International Conference on Machine Learning (ICML) (Vol. 97, pp. 6105–6114). PMLR. https://proceedings.mlr.press/v97/tan19a.html
21.Wei, J., Zhang, H., & Xie, J. (2024). A novel deep learning model for breast tumor ultrasound image classification with lesion region perception. Current Oncology, 31(9), 5057–5079. https://doi.org/10.3390/curroncol31090374
22.Woo, S., Park, J., Lee, J. Y., & Kweon, I. S. (2018). CBAM: Convolutional block attention module. In V. Ferrari, M. Hebert, C. Sminchisescu, & Y. Weiss (Eds.), Computer Vision – ECCV 2018 (Lecture Notes in Computer Science, Vol. 11211, pp. 3–19). Springer. https://doi.org/10.1007/978-3-030-01234-2_1
23.Yıldırım, T. T., Yaman, O., Kılıç, İ., Taşar, B., Timurkaan, E. S., & Aydoğdu, N. (2025). Multi-class classification of breast ultrasound images using vision transformer-based ensemble learning. Diagnostics, 15(17), 2235. https://doi.org/10.3390/diagnostics15172235
24.Zou, Y., & Miao, P. (2025). Explainable AI-enabled hybrid deep learning architecture for breast cancer detection. Frontiers in Immunology, 16, 1658741. https://doi.org/10.3389/fimmu.2025.1658741