Journal of Applied Engineering and Technological Science
Vol 6(1) 2024: 561-578

A COMBINED MOBILENETV2 AND CBAM MODEL TO IMPROVE CLASSIFYING THE BREAST CANCER ULTRASOUND IMAGES

Muhammad Rakha[1], Mahmud Dwi Sulistiyo[2*], Dewi Nasien[3], Muhammad Ridha[4]
School of Computing, Telkom University, Bandung 40257, Indonesia[1,2,4]
Department of Informatic Engineering, Institut Bisnis dan Teknologi Pelita Indonesia, Pekanbaru, 28127, Indonesia[3]
mrakhaa@student.telkomuniversity.ac.id[1], mahmuddwis@telkomuniversity.ac.id[2*], dewinasien@lecturer.pelitaindonesia.ac.id[3], mridha@student.telkomuniversity.ac.id[4]

Received: 08 April 2024, Revised: 11 June 2024, Accepted: 25 July 2024
Corresponding Author

ABSTRACT
Breast cancer is the main cause of death in women throughout the world. Early detection using ultrasound is very necessary to reduce cases of breast cancer. However, the ultrasound analysis process requires a lot of time and medical personnel because classification is difficult due to noise, complex texture, and subjective assessment. Previous studies were successful in ultrasound classification of breast cancer but required large computations and complex models. This research aims to overcome these shortcomings by using a lighter but more accurate model. We integrated the CBAM attention module into the MobileNetV2 model to improve breast cancer detection accuracy, speed up diagnosis, and reduce computational requirements. Gradient Weighted Class Activation Mapping (Grad-CAM) is used to improve classification explanations. Ultrasound images from two databases were combined to train, validate, and test this model. The test results show that MobileNetV2-CBAM achieves a test accuracy of 93%, higher than the complex models VGG-16 (80%), VGG-19 (82%), InceptionV3 (80%), and ResNet-50 (84%). CBAM is proven to improve MobileNetV2 performance with an 11% increase in accuracy. Grad-CAM visualization shows that MobileNetV2-CBAM can better focus on localizing important regions in breast cancer images, providing clearer explanations and assisting medical personnel in diagnosis.
Keywords: *MobileNetV2, CBAM, Image Classification, Breast Cancer, Ultrasound.*

1. Introduction

Breast cancer is prevalent among women and ranks as a leading cause of female mortality worldwide (Nasien et al., 2022). The World Health Organization (WHO) (Organization, 2021) reported that in 2020, approximately 2.3 million women were diagnosed with breast cancer, resulting in 685,000 deaths globally. By the end of 2020, 7.8 million women were living with a breast cancer diagnosis from the past five years. Given the substantial incidence of breast cancer, it is imperative to prioritize early detection through screening to mitigate the prevalence of the disease. Statistics show that around 40% reduction in deaths from breast cancer can be achieved if early detection is carried out (Arleo EK, Hendrick RE, Helvie MA, n.d.). This explains why regular breast cancer screening is essential (A Samah et al., 2020).

One of the examination procedures for detecting breast cancer is using the ultrasound screening method. Ultrasound screening is a common technique for diagnosing breast cancer, as it can differentiate between solid masses and fluid-filled areas of the breast. Also, it can identify irregular shapes and additional cancer blood vessels (Sahiner et al., 2007). That is why ultrasound has recently become a significant concern for dense breasts (Disha et al., 2009). However, medical personnel usually need a lot of time and energy to read or analyze ultrasound results. In addition to the increasing volume of ultrasound in hospitals, the classification process is also tricky due to spot noise, complex textures, and subjective judgment of medical personnel (Uysal & Köse, 2023) (Hossain et al., 2023). Currently, many methods can be used to assist medical personnel in classifying breast cancer from ultrasound images.

Artificial Intelligence plays an essential role in breast cancer classification. Deep learning has provided better results in image processing with the ability to automatically learn complex features from images (Liu et al., 2017). One of the uses that can be applied in image processing methods is a field of computer vision that can help classify types of breast cancer into specific categories. Research (Shehab et al., 2022), an analysis of 200 publications spanning from 2000 to 2022 focused on the utilization of machine learning in medical application. The research explains that machine learning can examine a lot of data, find exciting relationships, and perform pattern recognition, ultimately enhancing the effectiveness and precision of diagnostic systems across various diseases, including cancer detection. Research (Sebastian & Peter, 2022) discusses AI in cancer research; it explains that humans have limitations in analyzing large amounts of data in a short time; the integration of AI into cancer research overcomes the mistakes of medical experts in failing to diagnose and cure cancer, and the results are pretty encouraging. AI powered systems have the potential to assist pathologists in making more precise cancer diagnoses, thereby lowering error rates. As a result, the integration of AI, particularly through machine learning, in automation is significantly enhancing USG based breast screening processes.

In the context of ultrasound classification of breast cancer, several studies have demonstrated significant efficacy. (Xiao et al., 2018) compared transfer learning, CNN3, and conventional machine learning. These findings show that the transfer learning model using inceptionV3 provides superior results compared to other models. This underscores the challenges in developing CNN architectures to achieve optimal performance, which depends on several factors, such as the number of layers in the network, the number of filters in a layer, the convolutional size, the depth of the network, and the large number of available datasets (Liu et al., 2017). Research (Hossain et al., 2023; Uysal & Köse, 2023) also applied transfer learning to ultrasound breast cancer classification. The results show that transfer learning can significantly improve predictions, producing more accurate results using relatively small data sets. Integrating AI, particularly through machine learning and deep learning, has significantly improved the breast cancer diagnostic process. Using these techniques in medical image processing, such as ultrasound, helps overcome human limitations in analyzing large and complex data, thereby providing more accurate and efficient results in breast cancer diagnosis.

However, despite significant progress, the research must improve accuracy and efficiency. The use of more complex and deep models, such as VGG-16, InceptionV3, and ResNet-50, requires large computational resources (Bal-Ghaoui et al., 2023; Hossain et al., 2023; Tasnim & Hasan, 2023), this approach sometimes fails to improve network efficiency in size and speed. On the other hand, (Mukhlif et al., 2022; Wu et al., 2023) suggest further research to use lightweight models such as MobileNet or MobileNetV2, which can work well if set up correctly because transfer learning models are still widely used with computing high on medical images. In addition, although deep learning systems can provide correct diagnosis results, doctors are often unable to understand the basis or reasoning behind the decisions produced by the system, which gives rise to a new challenge called the black box problem (Fujioka et al., 2020; Plass et al., 2023; Starke & Poppe, 2022). This lack of clarity hinders acceptance and trust in this technology among medical professionals. Based on these problems, this research aims to fill the existing gaps by proposing a new model, namely the use of MobileNetV2 combined with the Convolutional Block Attention Module (CBAM) module to improve the accuracy of good diagnosis, speed up the analysis process, and increase the explanatory power of breast cancer classification. Ultrasound uses Gradient Weighted Class Activation Mapping (Grad-CAM). This approach is expected to provide a more efficient and acceptable solution to the medical community.

Therefore, it is very important to use image processing techniques to classify breast cancer on ultrasound. One image processing method currently popular and proven to provide high performance is the MobileNetV2 model (Sandler et al., 2018) as well as an attention module that utilizes CBAM (Woo et al., 2018). This module can identify the most relevant regions and significantly increases accuracy. Using breast cancer ultrasound images from two databases, a strategic approach has been used to assess the overall performance of the general model. Given the limited size of the dataset, even after combining images from both databases, applying transfer learning with established pre-training is considered advantageous. This approach allows the classifier to leverage existing knowledge and avoids the need to train models from scratch. Grad-CAM is also used to generate heat maps to visualize features by

calculating classification gradients and feature maps in the final convolution to achieve a focus on differentiating normal, benign, and malignant cancers in ultrasound breast cancer.
The primary contributions of this study include:

- We propose MobileNetV2-CBAM, which is a lightweight deep learning model, to show that using computationally high deep learning models is not an efficient solution.
- We show the effectiveness of the attention module in extracting essential and critical features to improve model performance in small datasets.
- We explore the benefits of the attention module for studying representations of breast cancer types.
- We use Grad-CAM to generate heat map visualizations, which help understand and explain the reasoning behind model decisions, thereby increasing interpretability and clinician confidence in the automated diagnosis system.

2. Literature Review

Over the past few years, machine learning has been widely used to solve many problems. For example, it has made a significant contribution to the medical world. (Shehab et al., 2022) conducted a study that reviewed the utilization of machine learning in medical application. This study analyzed the latest research consisting of 200 publications from January 2000 to December 2022 to diagnose disease, predict treatment outcomes, and identify disease risk factors. Five main medical applications are discussed in depth that focus on adapting machine learning, one of which is to solve cancer problems. The research results explain that machine learning can enhance the effectiveness and accuracy of diagnostic systems across various diseases by analyzing extensive datasets, uncovering exciting relationships, and recognizing patterns. Therefore, machine learning can help medical personnel diagnose patient illnesses digitally.

(Xiao et al., 2018) their research compared transfer learning, CNN3 and conventional machine learning in breast cancer classification. According to the results, the transfer learning model utilizing InceptionV3 demonstrated the highest performance, achieving an accuracy of 85.13%, outperforming the performance of other models. This explains the challenges in developing a CNN architecture to achieve good results. It depends on how many layers are in the network, how many filters are in one layer, how many convolutional sizes, network depth, learning speed and others (Liu et al., 2017). Typically, choices are made only by manual experimentation, which certainly requires a certain level of knowledge and expertise in deep learning. The trial and error process can disappoint beginners (Liu et al., 2017). Therefore, manually trying all possibilities to find the best model could be more efficient. Comparing various deep learning models in breast cancer classification is important to understand their advantages and disadvantages. When evaluating models such as MobileNetV2, VGG-16, InceptionV3, and ResNet50 for breast cancer classification, it is important to consider factors such as model performance, computational efficiency, and generalization ability.

Research (Mukhlif et al., 2022) reviews transfer learning on medical images. The research results explain that transfer learning achieves better results in overcoming the lack of medical images during training. Many use transfer learning models in medical images, such as InceptionV3, InceptionResNetV2, and VGGnet. Therefore, this research provides a potential future study using the lightweight MobileNet and MobileNetV2 models, which can work well if tuned well. Larger-scale medical picture analysis has been done by some prior research using computer vision and deep learning approaches. These studies include the analysis of human bone fingerprint images (Nugraha et al., 2022; Sitaba et al., 2023; Syam et al., 2023), interpretation of EKG signal images (Hadiyoso et al., 2022), segmentation of organs within the human body from x-ray images (Kurniawan et al., 2023), and identification of polyp diseases through colonoscopy images (Rafi et al., 2023).

Several studies have implemented transfer learning architectural models for classifying ultrasound breast cancer. This study (Uysal & Köse, 2023) took a different approach, where previous investigations were carried out with mammogram images, but this study uses ultrasound images. Different architectural models, ResNet-50, ResNeXt50, and VGG-16, are used in the analysis. The dataset uses 780 breast cancer ultrasound images (BUSI) from kaggle.com (Al-Dhabyani et al., 2020). This research obtained the highest accuracy on ResNeXt50, namely 85.83%, compared to VGG-16 and ResNet-50, which were lower, 81.11% and 85.4%, due to insufficient normal class samples.

In addition, (Hossain et al., 2023) researched ultrasound breast cancer classification using the VGG-16 model. The datasets used are two datasets, namely (Al-Dhabyani et al., 2020) plus (Rodrigues, 2018), totaling 250 images (100 benign and 150 malignant). The class combines only benign and malignant classes, resulting in 897 images. The research uses median filtering to filter noise in the image. It uses GRAD-CAM visualization to compare with the original image to visualize the targeted region in the final convolution layer. The results achieved a satisfactory accuracy of 91%.

(Hossinq et al., 2022) applied the transfer learning architecture to ultrasound breast cancer classification. The six research models used are MobileNetV2, VGG-16, AlexNet, ResNet, InceptionV3, and Xception. Each model is drilled and then evaluated at each different epoch for analysis. In terms of accuracy, MobileNetV2 performs better than other models, with an accuracy rate of 98.82%. This suggests lower computational models can also provide better results. Another study (Rachburee & Punlumjeak, 2022) applied image detection to a lotus species to expand student opportunities at the University of Technology Thanyaburi (RMUTT) museum. The dataset was taken from five species of lotus, totaling 720 images. Comparisons were made on the VGG-16, ResNet152V2, and MobileNetV2 models. The results show that MobileNetV2 has better performance and accuracy, namely 99.5%, compared to VGG-16 98.1% and ResNet152V2 97.6%. Moreover, MobileNetV2 has the lowest parameters and the best performance regarding time and speed. This research will help implement the model on mobile devices, making it easier for students to utilize and learn without limitations.

Research (Woo et al., 2018) on CBAM attention modules to improve CNN learning ability has attracted extensive research. CBAM aims to enhance representation capabilities by incorporating an attention mechanism, which prioritizes crucial features while and emphasizing irrelevant ones. CBAM is a lightweight module that can integrate into CNN architectures. This research was conducted on two models, namely ResNet-50 and MobileNet. A comparative evaluation was carried out with CBAM and without CBAM. The results show that CBAM leads to notable enhancements in model accuracy without significantly increasing the model size, rendering it well suited for deployment on mobile devices. Research (Niu et al., 2021) has included CBAM in mammogram breast cancer case classification. The proposed method is trained and validated using two databases: the Inbreast databases and DDSM, which include labels for both benign and malignant cases. The research uses CBAM-DenseNet and CBAM-Res2Net, the evaluation results of which are compared with regular Res2Net and regular DenseNet. The results show that the proposed method has better classification performance. Not only that, (Ijaz et al., 2023) also applied VGG+CBAM to the histopathological classification of breast cancer. To reduce the high computation and number of parameters, GAP (Global Average Pooling) was added. This approach helps reduce the network depth compared to the original VGG, mitigating the risks of underfitting and overfitting during training. The research results show that VGG+CBAM outperforms other methods carried out in previous research.

Another study (Shahi et al., 2022) combined CBAM with MobileNetV2 in the case of fruit classification. The results demonstrate that this approach surpasses other methods, offering fewer parameters and achieving higher classification accuracy. Another study (Ma et al., 2022) suggests MobileNetV2 by combining CBAM in identifying corn seed varieties and then comparing it with other CNN models, namely ResNet-50, Xception, E-AlexNet, DenseNet121, MobilenetV3 and MobileNetV2. The results show that MobileNetV2+CBAM obtains an accuracy of 98.21% higher than other models. The model can locate seeds accurately, expand the identification area of corn seed varieties, and reveal important information in the image.

A related work found that MobileNetV2 is a more efficient and effective architectural model for mobile applications due to equivalent accuracy without sacrificing performance. By embedding CBAM, the attention module provides better accuracy and emphasizes important area information in the image. CBAM is a lightweight module that can be included in CNN architectural models and has never been used for ultrasound breast cancer cases. So, this research will use MobileNetV2 combined with CBAM to classify ultrasonographic breast cancer Based on the description of the literature study above it shows that MobileNetV2 stands out because of its efficiency in terms of computing resources while maintaining a competitive level of performance (Fenu, 2021). MobileNetV2, known for its lightweight architecture, can provide advantages in limited computing resources, making it a suitable choice for practical applications in resource-constrained environments. On the other hand, VGG-16, InceptionV3, and ResNet50 have been widely used in various image classification tasks, including breast cancer diagnosis(Silva et al., 2021). These models have demonstrated strong generalization skills and high-performance levels in various medical imaging applications. However, compared to MobileNetV2, they may require more computing resources due to their deeper and more complex architecture. Resource efficiency is important, especially in developing countries where healthcare facilities have limited access to high-performance computing resources (Gao et al., 2023).

The MobileNetV2 pre-trained model has a small number of parameters, namely 3.4 million (Sandler et al., 2018), which is compared with the number of parameters of other models such as VGG-16: 138, VGG-19: 143 million (Simonyan & Zisserman, 2015), InceptionV3: 24 million (Szegedy et al., 2015), and ResNet-50: 25.6 million (Zagoruyko & Komodakis, 2017), of course the MobileNetV2 model is the right solution for devices with limited resources. Therefore, the lightweight nature of MobileNetV2 may make it a preferred choice for implementing breast cancer classification systems in such environments. Although VGG-16, InceptionV3, and ResNet50 have demonstrated good generalization capabilities in previous studies (Silva et al., 2021), MobileNetV2's efficiency in computing resources and generalization performance make it an attractive choice for practical applications in real-world healthcare environments. Research has also demonstrated the efficacy of combining CBAM with the MobileNetV2 lightweight model in image classification tasks. It can empower the development of powerful systems that perform well without excessive computing power or memory resources. So, in this research, we will use MobileNetV2, optimized with CBAM.

3. Research Methods

The methodology flow of this research can be seen in Figure 1 below.

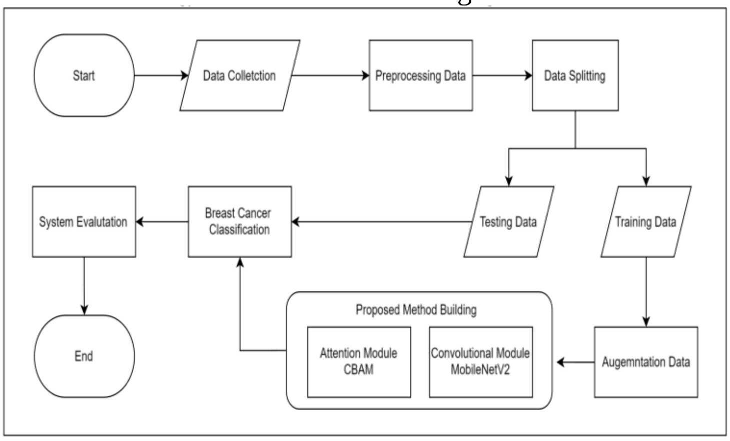

Fig. 1. Flowchart of the research methodology.

The methodology consists of collecting ultrasound breast cancer datasets, data preprocessing, data splitting, data augmentation, proposed method building, classification of breast cancer and system evaluation using a confusion matrix. Each process has its task to achieve the desired goals.

3.1 Data Collection

This research used two publicly available ultrasound breast cancer datasets. These datasets were evaluated for two primary objectives: firstly, to expand the dataset size for training, aiming to reduce overfitting and bias; secondly, to combine three classes (malignant, benign, and normal). The Breast Cancer Ultrasound Image (BUSI) dataset (Al-Dhabyani et al.,

2020) was gathered in 2018 using data from 600 women between the ages of 25 and 75. The images comprised 780 of 437 benign, 210 malignant, and 133 normal.

Each image is in high-quality PNG format and has 500×500 pixels. Mandeley Breast Cancer Ultrasound dataset (Rodrigues, 2018) contains 250 images, including 100 benign and 150 malignant, with a size of 100×100 pixels in BMP format with poor quality and zoom-in. There are more variations in the BUSI dataset than in the Mandeley BUS dataset. These two datasets will be combined as part of the study strategy to evaluate the proposed classification ability for generalization. Table 1 shows the dataset quantities, while Figures 2 and 3 present some samples from BUSI and Mandeley images, respectively.

Table 1 – Quantities of the collected breast cancer ultrasound dataset.

Dataset	Benign	Malignant	Normal	Total
BUSI	437	210	133	780
Mandeley	100	150	-	250
Total	537	360	133	1,030

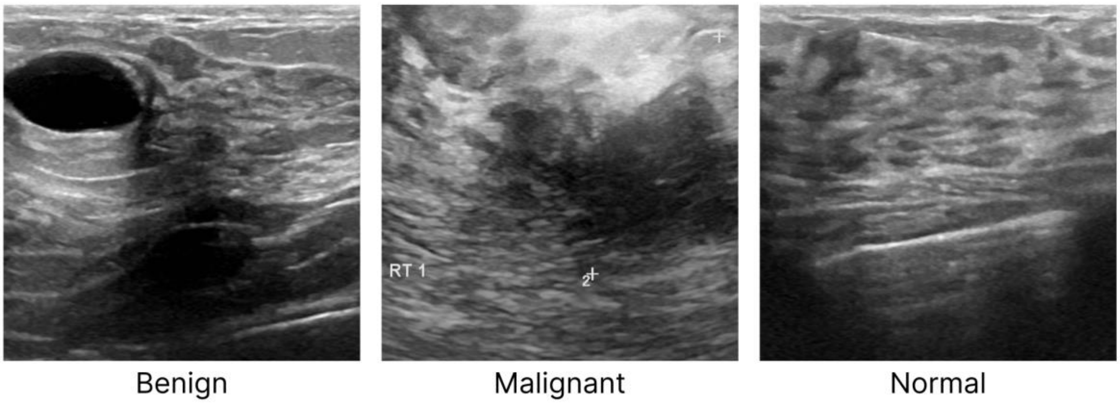

Fig. 2. Ultrasound images from BUSI dataset.

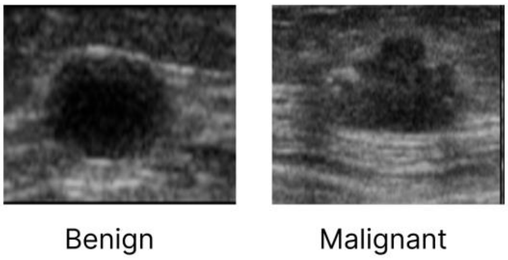

Fig. 3. Ultrasound images from Mandeley dataset.

3.2 Preprocessing Data

In the initial preprocessing phase, we transformed the images from our dataset into JPG format, initially stored as BMP and PNG files, to increase dataset consistency. Additionally, duplicate images were identified within the BUSI dataset, potentially resulting from inherent dataset errors. These duplicate images, comprising one benign and one malignant class image, were excluded from the study to avoid confusion during classifier development. Subsequently, the research encompassed 1,028 breast ultrasound images, comprising 536 benign class images, 359 malignant class images, and 133 normal class images.

Ultrasound breast cancer images inherently contain noise. The median filter, a straightforward yet effective technique for noise reduction in ultrasound images, was applied during preprocessing to enhance image quality. In the median filter operation on ultrasound images, a 5×5 moving kernel's center pixel is replaced by the corresponding kernel's median value. Following the median filter application, the ultrasound images were resized to a consistent size of 224×224×3, aligning with the input image size of the MobileNetV2 model. This step ensures uniform dimensions across all images. Furthermore, the image's pixel intensity range was normalized to 0 – 1, facilitating swift convergence of the model to a solution.

3.3 Data Splitting

The total data after the preprocessing stage was 1,028 ultrasound images. We first took 45 images randomly for testing set: 15 belonged to benign class, 15 were malignant class, and 15 represented normal class. For training and validation sets, the remaining 983 images were utilized, randomly divided into 80% for training and 20% for validation.

3.4 Augmentation Data

Deep learning models typically require a substantial number of images for effective training. Thus, it's imperative to appropriately augment the dataset's variety to enhance the diversity of data utilized. The limited availability of data makes this step very necessary. Increasing data can prevent models from overfitting and improve model generalization. This study considers using two data enhancement strategies: flip and zoom.

Table 2 – Attributes of the data augmentation.

Attribute	Value
Randomized horizontal flip	True
Randomized vertical flip	True
Randomized zoom	0.2

- Random (horizontal and vertical) flips: A technique for randomly rotating images horizontally and vertically, creating additional variation in the training data.
- Random zoom (0.2): Technique to randomly enlarge or reduce the image by up to 20%.

3.5 Building the Proposed Model

This study presents the proposed MobileNetV2-CBAM architectural method for ultrasonographic breast cancer classification. The main blocks of the proposed architectural method are as follows.

3.5.1 MobileNetV2

CNN is the most effective learning algorithm for understanding image material. CNNs consist of input, convolutional, pooling, fully connected, and output layers (Yu et al., 2019). However, A frequent challenge encountered when employing deep CNN models is the lack of training data, as these models demand extensive datasets for optimal performance. Additionally, collecting large data sets takes time. As a result, problems with small data sets have now been addressed using Transfer learning techniques (Pan & Yang, 2010). Transfer learning involves training a CNN model initially on a large dataset and then fine-tuning it to train on a smaller desired data set. Since the pre-trained model already comprehends fundamental features, adopting a transfer learning approach substantially reduces training time compared to training from scratch or without utilizing transfer learning. Many pre-trained algorithms are drilled on large data sets, namely the ImageNet dataset (Russakovsky et al., 2015), such as the model we use in this research, MobileNetV2.

Thus, convolutional operations constitute a significant contribution to computer vision tasks. However, the computation will be costly when the network structure is deeper and larger, like in VGG-16. (Simonyan & Zisserman, 2015). The MobileNetV1 model (Howard et al. 2017) carries the idea of Depthwise Separable Convolution as its main layer, which divides the convolution into two subtasks: Depthwise convolution filters the input data, while pointwise convolution (1×1) merges the filtered values to generate new features. The MobileNetV2 model (Sandler et al., 2018) adds layer expansion, residual connection, and projection layers in addition to deep convolutional, known as block residual resistance. Layer expansion (1×1) expands the number of layers according to an expansion factor of 6, while layer projection reduces the number of layers to a lower channel. Residual connections help the gradient flow through the network. The following is a complete architectural table of the convolution layers in the MobileNetV2 model, which is summarized in Table 3.

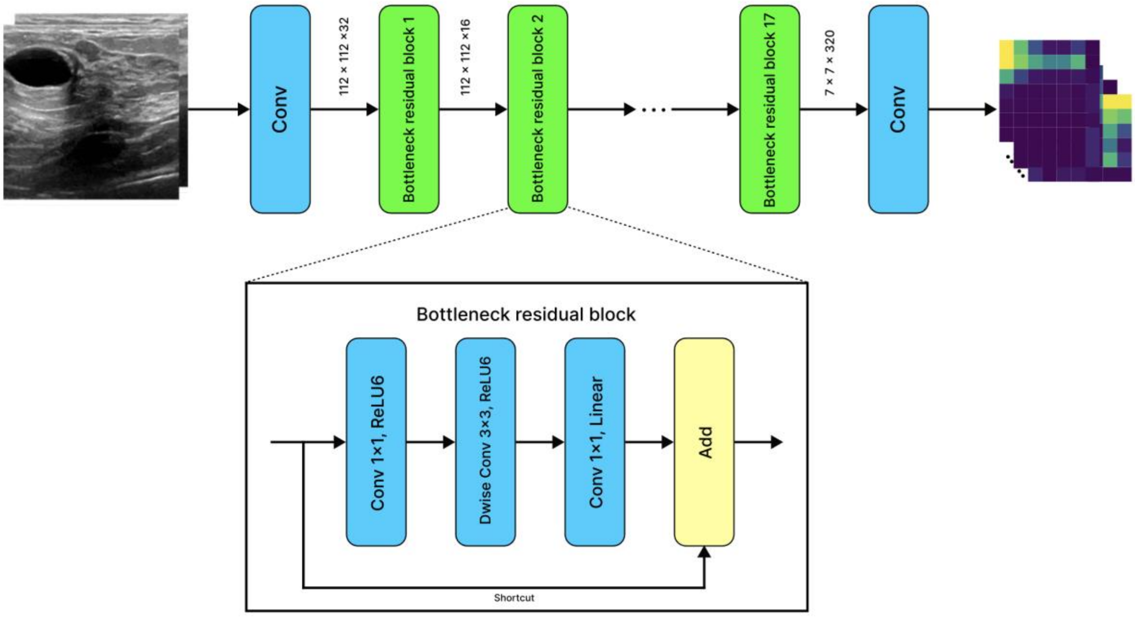

Fig. 4. The architecture of the MobileNetV2 network (Sandler et al., 2018).

Table 3 – The complete architecture of MobileNetV2.

Operator	Input Shape	t	c	n	s
Conv2d	224×224×3	-	32	1	2
Bottleneck	112×112×32	1	16	1	2
Bottleneck	112×112×16	6	24	2	2
Bottleneck	56×56×24	6	32	3	2
Bottleneck	28×28×32	6	64	4	2
Bottleneck	14×14×64	6	96	3	1
Bottleneck	14×14×96	6	160	3	2
Bottleneck	7×7×160	6	320	1	1
Conv2d 1×1	7×7×320	-	1,280	1	1
Avgpool 7×7	7×7×1,280	-	-	1	-
Conv2d 1×1	1×1×1,280	-	K	-	-

Note: The expansion factor, output channels, number of repeats, and step size are indicated by the letters t, c, n, and s.

3.5.2 The Improved Attention Module CBAM

The attention network is inspired by the attention mechanism observed in the human brain when processing images. A person focuses only on essential clues or salient parts to capture the visual structure better rather than looking at all the details of an object. The attention module of this research uses CBAM (Woo et al., 2018), which can help the model improve its learning ability and focus on essential features in the image so that this information can be used to make more accurate predictions. The CBAM module is lightweight and can integrate into a CNN architecture. It comprises both a channel attention module and a spatial attention module. Below is an illustration describing the structure of CBAM.

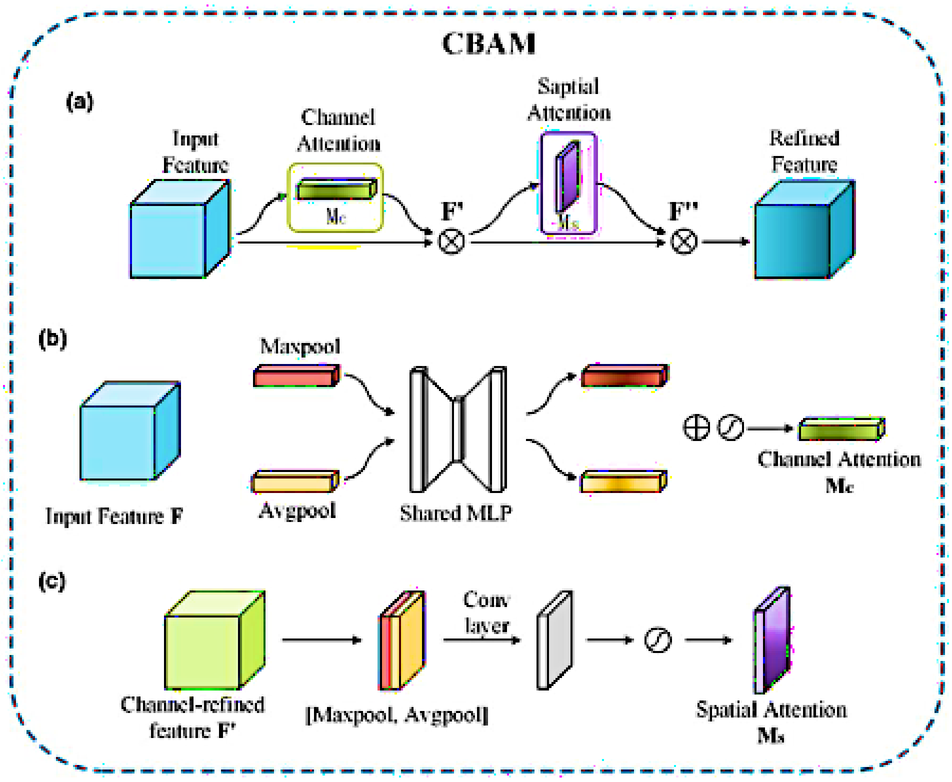

Fig. 5 Structures of the CBAM: (**a**) Overall structure of CBAM module; (**b**) Attention channel module structure; and (**c**) Attention spatial module structure (Niu et al., 2021).

From Figure 5 (a), Initially, the CBAM module takes the feature map F as input, which is then subjected to weighting by the attention channel F'. Subsequently, the resulting feature map F" undergoes further weighting by spatial attention in a sequential. The process formula is as follows:

$$F' = M_c(F) \otimes F$$
$$F'' = M_s(F') \otimes F'$$

Description: $M_c(F)$ represents the output F after being processed through the attention channel
$M_s(F')$ represents the output F' after passing spatial attention.
$\otimes$ indicates the operator for weighted multiplication of feature maps

The symbol $\otimes$ describes the operation of multiplying the appropriate elements, the initial step is to calculate the channel feature map from the input feature map F. Next, and the feature map is multiplied by the input feature map to produce F'. Then, F' is multiplied by the spatial feature map resulting from F' to obtain the final feature map, namely F".

Channel Attention Module

As in Figure 5(b), it focuses on the significant "WHAT" in light of the input image. This module gathers crucial information about various object properties through maximum and average pooling to improve channel attention. The calculation formula is as follows:

$$M_c(F) = \sigma(MLP(\mathit{AvgPool}(F)) + MLP(\mathit{MaxPool}(F)))$$
$$= \sigma(W_1(W_0(F^c avg)) + W_1(W_0(F^c max)))$$

Spatial Attention Module

Figure 5(c) focuses on "WHERE," which represents the informative segment crucial for completing the attention channel. Initially, average pooling and maximum pooling operations are applied across the layers, combined using a 7×7 convolution, resulting in a connected feature map that effectively enhances spatial attention. The calculation formula is as follows:

$$M_s(F') = \sigma(f^{7\times7}([\mathit{AvgPool}(F'); \mathit{MaxPool}(F')]))$$

$$= \sigma(f^{7\times7}([F'^{s}avg; F'^{s}max]))$$

Description: σ is a sigmoid function
W_0 and W_1 indicate MLP
$F^{c}avg$ and $F^{s}avg$ are features of average pooled
$F^{c}max$ and $F^{s}max$ are features of maximum pooled
$f^{7\times7}$ is a convolution operation with 7×7 filters

3.5.3 MobileNetV2-CBAM Architecture

The advantage of MobileNetV2 lies in its architecture, which combines residual connections and linear bottlenecks. This design allows efficient data processing with fewer parameters, making it well-suited for scenarios with limited computing resources. Nevertheless, MobileNetV2 still maintains high execution speed and good accuracy. Incorporating CBAM into the MobileNetV2 network enhances its capability to represent features, enabling improved emphasis on local and global information within the image. This integration resulted in the development of MobileNetV2-CBAM, a model specifically designed to classify different cancer types in ultrasound images accurately. The following MobileNetV2-CBAM network model is depicted in Figure 6.

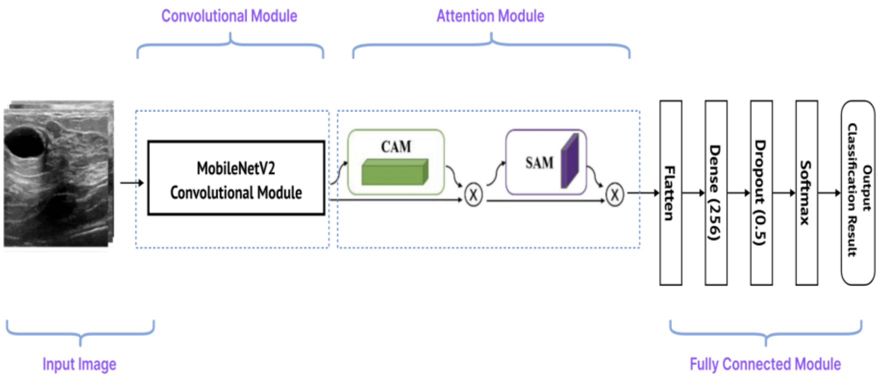

Fig. 7. Proposed architecture of the MobileNetv2-CBAM.

3.5.4 Environment Configuration Model Training

This experiment was implemented using Python with the TensorFlow deep learning libraries on the Google Collaboratory platform. Experiments were run on a computer with AMD Ryzen 5 2500U, including Google Collaboratory RAM and GPU support. Specific configuration parameters are listed in Table 4.

Table 4 – Detailed hyperparameters of the experiment.

Parameter	Value
Image Size	224×224
Epoch	30
Batch Size	32
Learning Rate	0.0001
Optimizer	Adam
Loss	Categorical Cross Entropy

3.5.5 Model Performance Evaluation Metrics

This experiment uses the Confusion Matrix to measure model performance. Confusion Matrix (Tharwat, 2021) evaluates a classification model by tabulating the number of true and false predictions made on the test dataset. It comprises positive and negative values, which are analyzed to derive metrics such as accuracy, precision, recall, and F1 score.

$$\text{accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \times 100\%$$

$$\text{recall} = \frac{TP}{TP + FN} \times 100\%$$

$$\text{precision} = \frac{TP}{TP + FP} \times 100\%$$

$$\text{F1} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \times 100\%$$

Description: TP is predicted to be Positive, and it's True
TN is predicted to be Negative, and it's True
FP is predicted to be Positive, and it's False
FN is predicted to be Negative, and it's False

4. Results and Discussions

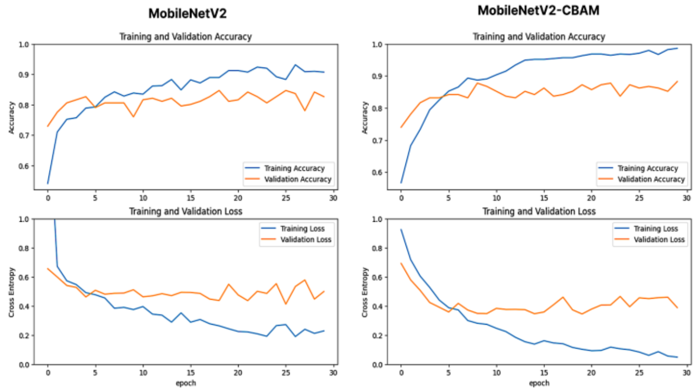

Fig. 8. Accuracy and loss curves during training and validation processes.

As shown in Figure 8, adding the CBAM attention module to MobileNetV2 at epoch 30 significantly enhances the training performance compared with the network without the CBAM attention module. A maximum training accuracy of 98.6% and a validation accuracy of 88.2% are attained by the MobileNetV2-CBAM model. Meanwhile, MobileNetV2 without CBAM obtained 90.7% training accuracy and 82.6% validation accuracy. Therefore, adding the CBAM attention module to the MobileNetV2 network is practical.

The model's performance was evaluated using the testing dataset, with Figure 9 showcasing the Confusion Matrix for ultrasound breast cancer classification comparing MobileNetV2-CBAM and MobileNetV2.

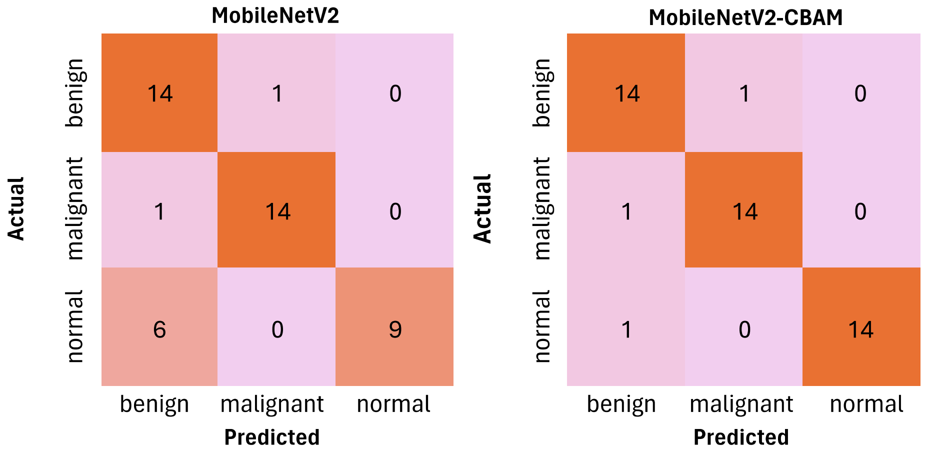

Fig. 9. MobileNetV2-CBAM and MobileNetV2 results shown in confusion matrices.

From the confusion matrix, a testing data set with a total of 45 images, including 15 benign, 15 malignant, and 15 normal, classification using the MobileNetV2-CBAM model succeeded in predicting 14 samples correctly for each class and mispredicting 1 sample for each class. On the other hand, the MobileNetV2 model without CBAM mispredicts the normal class more often, namely six samples. This is because the number of normal samples is smaller, so it tends to recognize better classes with a more significant number of samples, namely the benign class. So, embedding the CBAM attention module on the MobileNetV2 network can help the model improve its learning ability and focus on essential features in the image to make more accurate predictions. The testing data set's quantitative evaluation results are shown in Table 5.

Table 5 – Quantitative evaluation results.

Model	Accuracy	Precision	Recall	F1-Score
MobileNetV2	82%	86%	82%	84%
MobileNetV2-CBAM	93%	93%	93%	93%

The comparison shown in Table 5 indicates that the MobileNetV2-CBAM model produces significant improvements in model performance compared to the MobileNetV2 standard. MobileNetV2-CBAM has higher accuracy, with a value of 93% compared to 82% of MobileNetV2. Additionally, all other evaluation metrics, such as precision, recall, and F1-score, show consistent improvements with higher accuracy. This indicates that MobileNetV2's implementation of the CBAM Attention Module has improved the model's object recognition and classification capabilities.

MobileNetV2-CBAM is compared with high computation models such as VGG-19, VGG-16 (Simonyan & Zisserman, 2015), ResNet-50 (Zagoruyko & Komodakis, 2017) and InceptionV3 (Szegedy et al., 2015) using testing data, to demonstrate performance of ultrasound breast cancer classification. In terms of architecture, VGG-16, VGG-19, ResNet-50 and InceptionV3 are deeper networks than MobileNetV2. VGG-16 and VGG-19 consist of multiple convolutional layers with small 3x3 filters, while ResNet50 introduces skip connections to address the vanishing gradient problem, and InceptionV3 incorporates a start module that balances performance and efficiency by leveraging different kernel sizes in the same layer (Rajinikanth et al., 2020). Meanwhile, MobileNetV2 uses an inverted residual block, which aims to balance performance and efficiency (Khairnar et al., 2024). Figure 10 shows the confusion matrix for the remaining four models.

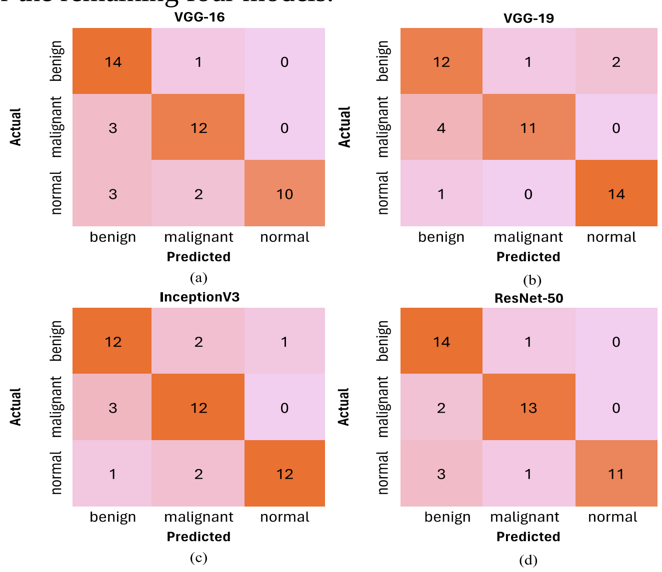

Fig. 10. Comparison of confusion matrices for the other four models:
(a) VGG-16; **(b)** VGG-19; **(c)** InceptionV3; and **(d)** ResNet-50.

As shown in Figure 10, MobileNetV2-CBAM exhibits considerably fewer incorrect identifications in the testing dataset compared to the other models in the experimental comparison. The test results show that the CBAM attention module improves breast cancer classification abilities. Quantitative comparison experiments were carried out to evaluate all methods. The evaluation results are mentioned in Table 6.

Table 6 – Quantitative evaluation results of different models.

Model	Parameters	Accuracy	Precision	Recall	F1-Score
MobileNetV2	3.538.984	82%	86%	82%	84%
VGG-16	138.357.544	80%	83%	79%	81%
VGG-19	143.667.240	82%	83%	82%	82%
InceptionV3	23.851.784	80%	80%	80%	80%
ResNet-50	25.636.712	84%	86%	84%	84%
MobileNetV2-CBAM	3.948.584	93%	93%	93%	93%

From Table 6 shows that MobileNetV2-CBAM obtained the highest overall accuracy, precision, recall, and F1-Score compared to other models. It's essential to recognize that effective model development does not necessarily mean selecting deep or complex models by default. Careful optimization of the model architecture and considering the trade-off between accuracy and computational efficiency is essential in designing the optimal network for a given task. Moreover, the data set used, which consists of a combination of the BUSI and Mendeley Data datasets of 1028 images, is still relatively small and unbalanced between classes (with the number of normal classes being fewer than the benign and malignant classes). If a complex model with many parameters is used, this can lead to overfitting and reduce system performance on the test dataset. In this case, the implementation of CBAM on the MobileNetV2 architecture, which uses depthwise separable convolutions to reduce the number of parameters, proved to be very effective. CBAM helps the model combine information from all stages of feature representation and places greater emphasis on important features generated by MobileNetV2. This can help the model make more accurate predictions. Overall, the combination of MobileNetV2 with CBAM shows that with the right architecture and good optimization, more efficient and lightweight models can outperform more complex models, especially on limited and imbalanced data sets.

We try to see how CBAM helps the network increase its representation power by highlighting regions that are considered important by the network in predicting a class using Grad-CAM. Grad-CAM is a method for visualizing features in class activation heatmaps (Selvaraju et al., 2019). Due to its excellent visual support, Grad-CAM is used to compute classification gradients to determine the localization of target regions in the final convolution. Gradient areas with high colour intensity in the heatmap indicate highly active parts of the feature map and, therefore, influence the classification decision significantly.

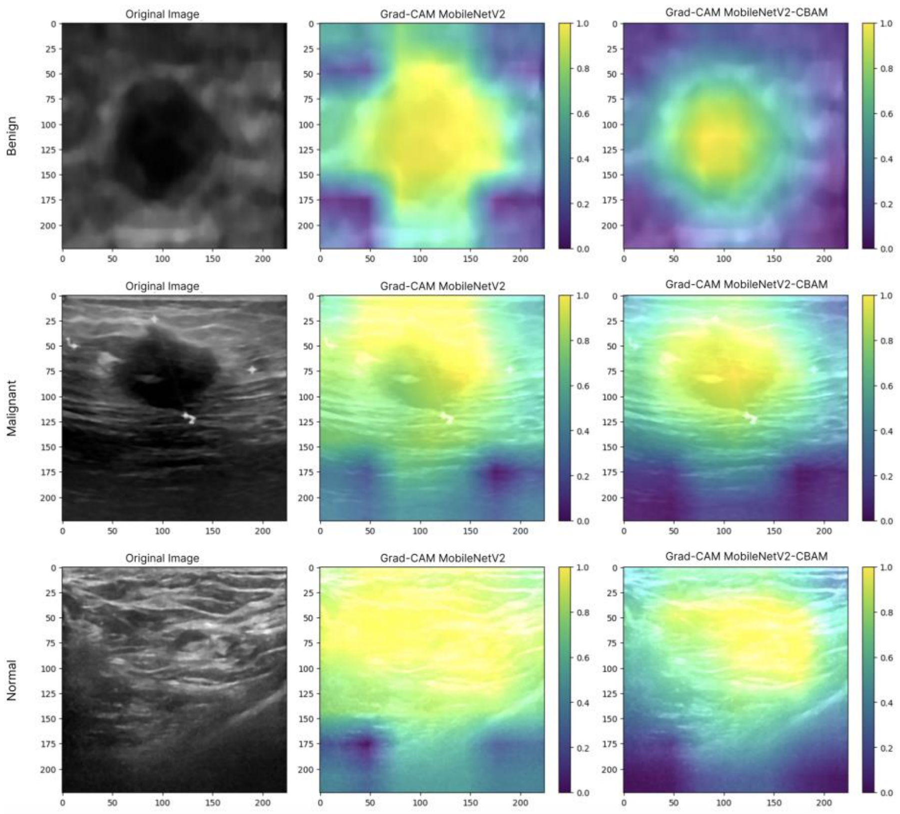

Fig. 11. Comparison of Grad-CAM Visualization MobileNetv2 and MobileNetv2-CBAM

Figure 11 is a comparison of the Grad-CAM MobileNetV2 and MobileNetV2-CBAM visualization results for each class (benign, malignant and normal). For the benign class, the Grad-CAM MobileNetV2 heat map shows strong activation (yellow colour) in the centre of the mass area. This indicates that the model pays attention to the centre of the lesion for classification. Similar to MobileNetV2, but the Grad-CAM MobileNetV2-CBAM heat map is slightly more focused, and the activation area is smoother in the centre of the lesion. So for benign cases, the model tends to focus on the center of the lesion where there are obvious changes in tissue texture and density. Meanwhile, for the malignant class, the Grad-CAM MobileNetV2 heat map shows broader activation, covering the centre and some areas around the lesion. This diffuse activation indicates that the model pays attention not only to the centre but also to the borders and peripheral features of the lesion. Meanwhile, the Grad-CAM MobileNetV2-CBAM heat map shows a broader and more intense activation pattern around the lesion. CBAM adds contextual attention making it more sensitive to features surrounding the lesion that may be relevant for malignant classification. Thus, for malignant cases, the model tends to pay more attention to areas around the lesion that may show characteristics of invasion or speculation that are common in malignant tumors. Meanwhile, the Grad-CAM MobileNetV2 heat map for the normal class shows activation in a large area but is less intense compared to benign or malignant lesions. Areas in yellow reflect areas that may be denser, but no obvious lesions are marked. Similar patterns to MobileNetV2, MobileNetV2-CBAM's Grad-CAM heat map shows slight differences in intensity distribution. Activations were more spread out and indicated that the model was paying attention to the entire normal network without a specific focus on a particular area. For normal cases, the model tends to look at the whole area to ensure there are no suspicious lesions or masses. These heatmaps help in understanding how the model makes decisions and which areas are considered most important for classification. Doctors can use it to verify and understand model decisions in the context of medical ultrasound images.

By comparing the heat maps shown in Figure 11, it can be concluded that after adding the CBAM attention module in MobileNetV2, the model consistently shows stronger and sharper emphasis on important features in the images of each class of ultrasound breast cancer. In contrast, the MobileNetV2 standard tends to highlight features more broadly and less focused, which may reduce the ability to identify features relevant to breast cancer classification correctly.

5. Conclusion

The MobileNetV2-CBAM model, a deep learning-based computer vision technology, is proposed in this study to classify three types of breast cancer on ultrasound (benign, malignant, and normal). From the tests carried out, this model achieved training accuracy of 98%, validation accuracy of 88%, and testing accuracy of 93%, proving its high effectiveness. Compared with the standard model without CBAM, MobileNetV2-CBAM shows superiority in all evaluation indices. With 93% accuracy, this model also outperforms other complex models such as VGG-16 (80%), VGG-19 (82%), InceptionV3 (80%), and ResNet-50 (84%). This model captures high-level information about objects, with the attention module emphasizing salient regions more than the convolutional module. The combination of convolutional and attention modules provides comprehensive information, which can potentially improve the classification of various types of breast cancer on ultrasound. The proposed MobileNetV2-CBAM model can contribute to early detection, more accurate diagnosis, and treatment planning for breast cancer patients. This shows that a well-designed lightweight model can provide significant classification results that are not inferior to complex models. Grad-CAM visualization shows that MobileNetV2-CBAM can provide complete regional information when diagnosing breast cancer and can focus stronger and sharper emphasis on important features in images of each ultrasound breast cancer class. This helps medical personnel obtain clearer and more precise information and improves the quality of clinical decisions.

We suggest that future research enrich the ultrasound breast cancer dataset and apply the MobileNetV2-CBAM model on mobile devices, making it more widely accessible to medical personnel and providing direct benefits in clinical practice. Additionally, future research could explore new attention mechanisms, integrate multi-modal data such as MRI and mammography, or overcome challenges in model interpretability to improve diagnoses' reliability and accuracy further. With this guide, it is hoped that researchers can continue to advance the field of medical imaging and have a greater positive impact on health care.

Acknowledgment

The authors extend their sincere gratitude to the Indonesia’s KEMDIKBUDRISTEK for the generous financial support under the DRTPM Grant numbers 106/E5/PG.02.00.PL/2024, 043/SP2H/RT-MONO/LL4/2024, and 085/LIT07/PPM-LIT/2024.

References

A Samah, A., Nasien, D., Hashim, H., Sahar, J., Majid, H., Yusoff, Y., & Ali Shah, Z. (2020). Application of Deep Learning Method in Facilitating the Detection of Breast Cancer. *IOP Conference Series: Materials Science and Engineering*, *864*, 12079. https://doi.org/10.1088/1757-899X/864/1/012079

Al-Dhabyani, W., Gomaa, M., Khaled, H., & Fahmy, A. (2020). Dataset of breast ultrasound images. *Data in Brief*, *28*, 104863. https://doi.org/https://doi.org/10.1016/j.dib.2019.104863

Arleo EK, Hendrick RE, Helvie MA, S. E. (n.d.). *Comparison of recommendations for screening mammography using CISNET models*. https://doi.org/doi: 10.1002/cncr.30842.

Bal-Ghaoui, M., El Yousfi Alaoui, M. H., Jilbab, A., & Bourouhou, A. (2023). Optimizing Ultrasound Image Classification Through Transfer Learning: Fine-Tuning Strategies and Classifier Impact on Pre-Trained Inner-Layers. In *Informatyka Automatyka Pomiary W Gospodarce I Ochronie Środowiska*. https://doi.org/10.35784/iapgos.4464

Disha, E., Manxhuka-Kerliu, S., Ymeri, H., & Kutllovci, A. (2009). Comparative Accuracy of Mammography and Ultrasound in Women with Breast Symptoms According to Age and Breast Density. *Bosnian Journal of Basic Medical Sciences / Udruženje Basičnih Mediciniskih Znanosti = Association of Basic Medical Sciences*, *9*, 131–136. https://doi.org/10.17305/bjbms.2009.2832

Fenu, G. (2021). Using Multioutput Learning to Diagnose Plant Disease and Stress Severity. In

Complexity. https://doi.org/10.1155/2021/6663442

Fujioka, T., Mori, M., Kubota, K., Oyama, J., Yamaga, E., Yashima, Y., Katsuta, L., Nomura, K., Nara, M., Oda, G., Nakagawa, T., Kitazume, Y., & Tateishi, U. (2020). The utility of deep learning in breast ultrasonic imaging: A review. *Diagnostics*, *10*(12). https://doi.org/10.3390/diagnostics10121055

Gao, W., Wang, D., & Huang, Y. (2023). Designing a Deep Learning-Driven Resource-Efficient Diagnostic System for Metastatic Breast Cancer: Reducing Long Delays of Clinical Diagnosis and Improving Patient Survival in Developing Countries. In *Cancer Informatics*. https://doi.org/10.1177/11769351231214446

Hadiyoso, S., Fahrozi, F., Hariyani, Y. S., & Sulistyo, M. D. (2022). Image Based ECG Signal Classification Using Convolutional Neural Network. *International Journal of Online and Biomedical Engineering (IJOE)*, *18*(04), 64–78. https://doi.org/10.3991/ijoe.v18i04.27923

Hossain, A. B. M. A., Nisha, J. K., & Johora, F. (2023). Breast Cancer Classification from Ultrasound Images using VGG16 Model based Transfer Learning. *International Journal of Image, Graphics and Signal Processing*, *15*(1), 12 – 22. https://doi.org/10.5815/ijigsp.2023.01.02

Hossinq, M. M., Molla, S., Javed Mehedi Shamrat, F. M., Samia, M. M., Saha, R. M., & Khater, A. (2022). Comparative Study On Breast Cancer Classification Using Multiple Convolution Neural Network Architectures. *2022 3rd International Conference on Smart Electronics and Communication (ICOSEC)*, 1081–1088. https://doi.org/10.1109/ICOSEC54921.2022.9951881

Ijaz, A., Raza, B., Kiran, I., Waheed, A., Raza, A., Shah, H., & Aftan, S. (2023). Modality Specific CBAM-VGGNet Model for the Classification of Breast Histopathology Images via Transfer Learning. *IEEE Access*, *11*, 15750–15762. https://doi.org/10.1109/ACCESS.2023.3245023

Khairnar, S., Gite, S., & Thepade, S. D. (2024). *Empirical Performance Analysis of Deep Convolutional Neural Networks Architectures for Face Liveness Detection*. https://doi.org/10.21203/rs.3.rs-3824202/v1

Kurniawan, B., Sulistiyo, M. D., Hadiyoso, S., & Zarifie Bin Hashim, N. M. (2023). Semantic Segmentation of Anatomical Structures in Posterior-Anterior Chest X-Ray Image Using U-Net. *2023 11th International Conference on Information and Communication Technology (ICoICT)*, 142–147. https://doi.org/10.1109/ICoICT58202.2023.10262815

Liu, X., Yang, D., & Gamal, A. El. (2017). Deep neural network architectures for modulation classification. *2017 51st Asilomar Conference on Signals, Systems, and Computers*, 915–919. https://doi.org/10.1109/ACSSC.2017.8335483

Ma, R., Wang, J., Zhao, W., Guo, H., Dai, D., Yun, Y., Li, L., Hao, F., Bai, J., & Ma, D. (2022). Identification of Maize Seed Varieties Using MobileNetV2 with Improved Attention Mechanism CBAM. *Agriculture*, *13*, 11. https://doi.org/10.3390/agriculture13010011

Mukhlif, A. A., Al-Khateeb, B., & Mohammed, M. A. (2022). An extensive review of state-of-the-art transfer learning techniques used in medical imaging: Open issues and challenges. *Journal of Intelligent Systems*, *31*(1), 1085–1111. https://doi.org/doi:10.1515/jisys-2022-0198

Nasien, D., Enjeslina, V., Adiya, M. H., & Baharum, Z. (2022). Breast Cancer Prediction Using Artificial Neural Networks Back Propagation Method. *Journal of Physics: Conference Series*, *2319*(1), 12025. https://doi.org/10.1088/1742-6596/2319/1/012025

Niu, J., Li, H., Zhang, C., & Li, D. (2021). Multi-scale attention-based convolutional neural network for classification of breast masses in mammograms. *Medical Physics*, *48*(7), 3878–3892. https://doi.org/https://doi.org/10.1002/mp.14942

Nugraha, D. B., Rachmawati, E., & Sulistiyo, M. D. (2022). Semantic Segmentation of Whole-Body Bone Scan Image Using Btrfly-Net. *2022 14th International Conference on Information Technology and Electrical Engineering (ICITEE)*, 264–269. https://doi.org/10.1109/ICITEE56407.2022.9954073

Organization, W. H. (2021). *Breast Cancer*. https://www.who.int/news-room/fact-sheets/detail/breast-cancer

Pan, S. J., & Yang, Q. (2010). A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, *22*(10), 1345–1359. https://doi.org/10.1109/TKDE.2009.191

Plass, M., Kargl, M., Kiehl, T. R., Regitnig, P., Geißler, C., Evans, T., Zerbe, N., Carvalho, R., Holzinger, A., & Müller, H. (2023). Explainability and causability in digital pathology. *Journal of Pathology and Clinical Research*, *9*(4), 251–260. https://doi.org/10.1002/cjp2.322

Rachburee, N., & Punlumjeak, W. (2022). Lotus species classification using transfer learning based on VGG16, ResNet152V2, and MobileNetV2. *IAES International Journal of Artificial Intelligence (IJ-AI)*, *11*(4), 1344. https://doi.org/10.11591/ijai.v11.i4.pp1344-1352

Rafi, W., Sulistiyo, M. D., Hadiyoso, S., & Wisesty, U. (2023). Polyp Identification from a Colonoscopy Image Using Semantic Segmentation Approach. *Building of Informatics, Technology and Science (BITS)*, *5*(2), 423−431. https://doi.org/10.47065/bits.v5i2.4083

Rajinikanth, V., Raj, A. N. J., Krishnan, P. T., & Naik, G. R. (2020). A Customized VGG19 Network With Concatenation of Deep and Handcrafted Features for Brain Tumor Detection. In *Applied Sciences*. https://doi.org/10.3390/app10103429

Rodrigues, P. S. (2018). Breast Ultrasound Image. *Mendeley Data*, *1*. https://doi.org/oi: 10.17632/wmy84gzngw.1

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., & Fei-Fei, L. (2015). *ImageNet Large Scale Visual Recognition Challenge*.

Sahiner, B., Chan, H.-P., Roubidoux, M., Hadjiiski, L., Helvie, M., Paramagul, C., Bailey, J., Nees, A., & Blane, C. (2007). Malignant and Benign Breast Masses on 3D US Volumetric Images: Effect of Computer-aided Diagnosis on Radiologist Accuracy 1. *Radiology*, *242*, 716–724. https://doi.org/10.1148/radiol.2423051464

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). *MobileNetV2: Inverted Residuals and Linear Bottlenecks*. 4510–4520. https://doi.org/10.1109/CVPR.2018.00474

Sebastian, A. M., & Peter, D. (2022). Artificial Intelligence in Cancer Research: Trends, Challenges and Future Directions. *Life*, *12*(12). https://doi.org/10.3390/life12121991

Shahi, T. B., Sitaula, C., Neupane, A., & Guo, W. (2022). Fruit classification using attention-based MobileNetV2 for industrial applications. *PLOS ONE*, *17*(2), 1–21. https://doi.org/10.1371/journal.pone.0264586

Shehab, M., Abualigah, L., Shambour, Q., Abu-Hashem, M. A., Shambour, M. K. Y., Alsalibi, A. I., & Gandomi, A. H. (2022). Machine learning in medical applications: A review of state-of-the-art methods. *Computers in Biology and Medicine*, *145*, 105458. https://doi.org/https://doi.org/10.1016/j.compbiomed.2022.105458

Silva, L., Araújo, L. dos S., Souza, V. F., Neto, R. M. B., & Santos, A. (2021). Comparative Analysis of Convolutional Neural Networks Applied in the Detection of Pneumonia Through X-Ray Images of Children. In *Learning and Nonlinear Models*. https://doi.org/10.21528/lnlm-vol18-no2-art1

Simonyan, K., & Zisserman, A. (2015). *Very Deep Convolutional Networks for Large-Scale Image Recognition*.

Sitaba, A. Y., Rachmawati, E., & Sulistiyo, M. D. (2023). Investigating Convolution-Attention Model for Bone Scan Image Segmentation. *2023 10th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*, 344–349. https://doi.org/10.1109/ICITACEE58587.2023.10276939

Starke, G., & Poppe, C. (2022). Karl Jaspers and artificial neural nets: on the relation of explaining and understanding artificial intelligence in medicine. *Ethics and Information Technology*, *24*(3), 26. https://doi.org/10.1007/s10676-022-09650-1

Syam, R. F. K., Rachmawati, E., & Sulistiyo, M. D. (2023). Whole-Body Bone Scan Segmentation Using SegFormer. *2023 10th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*, 419–424. https://doi.org/10.1109/ICITACEE58587.2023.10277219

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2015). *Rethinking the Inception Architecture for Computer Vision*.

Tasnim, J., & Hasan, M. K. (2023). CAM-QUS guided self-tuning modular CNNs with multi-loss functions for fully automated breast lesion classification in ultrasound images. *Physics in Medicine & Biology*, *69*(1), 15018. https://doi.org/10.1088/1361-6560/ad1319

Tharwat, A. (2021). Classification assessment methods. *Applied Computing and Informatics*, *17*(1), 168–192. https://doi.org/10.1016/j.aci.2018.08.003

Uysal, F., & Köse, M. M. (2023). Classification of Breast Cancer Ultrasound Images with Deep Learning-Based Models. *Engineering Proceedings*, *31*(1). https://doi.org/10.3390/ASEC2022-13791

Woo, S., Park, J., Lee, J.-Y., & Kweon, I.-S. (2018). CBAM: Convolutional Block Attention Module. *European Conference on Computer Vision*.

Wu, J.-M., Chen, L., Chen, S.-P., Lee, C.-C., Huang, T.-K., Tseng, N.-C., Tsai, W. S., Lin, Y., Wu, Y.-C., Cheng, B., & Yu, J.-T. (2023). *Classification of Catheters and Tubes on Chest Radiographs Using Light-Weight Deep Convolutional Neural Networks*. https://doi.org/10.21203/rs.3.rs-2463844/v1

Xiao, T., Liu, L., Li, K., Qin, W., Yu, S., & Li, Z.-C. (2018). Comparison of Transferred Deep Neural Networks in Ultrasonic Breast Masses Discrimination. *BioMed Research International*, *2018*, 1–9. https://doi.org/10.1155/2018/4605191

Yu, L., Li, B., & Jiao, B. (2019). Research and Implementation of CNN Based on TensorFlow. *IOP Conference Series: Materials Science and Engineering*, *490*(4), 42022. https://doi.org/10.1088/1757-899X/490/4/042022

Zagoruyko, S., & Komodakis, N. (2017). *Wide Residual Networks*.