Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 7, 2025

AN ENHANCED EFFICIENTNETB0 WITH CBAM ATTENTION FOR ROBUST SKIN CANCER DETECTION

Muhammad Suleman Memon[1], Mumtaz Qabulio[2], Samia Aijaz Siddiqui[3], Sorath Mumtaz Mangi[4], Naeem Junejo[5]

[1]Department of Information Technology, Dadu Campus, University of Sindh msuleman@usindh.edu.pk
[2]Department of Software Engineering, Faculty of Engineering & Technology, University of Sindh, Jamshoro
[3,4]Department of Computer System Engineering, Dawood University of Engineering and Technology
[5]MS IT GC University,
msuleman@usindh.edu.pk[1] mumtaz.qabulio@usindh.edu.pk[2] samiasiddiqui@duet.edu.pk[3], engr.sorath@duet.edu.pk[4] engr.junejonaeem@gmail.com[5]

DOI: https://doi.org/10.5281/zenodo.16731501

Keywords
Skin Cancer Classification, EfficientB0, CBAM, Deep Learning

Article History
Received on 08 May 2025
Accepted on 08 June 2025
Published on 28 July 2025

Corresponding Author: *
Muhammad Suleman Memon

Skin cancer is one of the most common and deadly types of cancer in the world, so it's important to get the right diagnosis quickly. In this study we propose a strong deep learning framework based on EfficientNetB0 and the Convolutional Block Attention Module (CBAM). This will help the model better focus on important areas in dermoscopic images. The model is trained and tested on a dataset of skin cancers containing benign and malignant classes. The size of the dataset for training is 2637 images, and 660 images for the test set. The dataset was obtained from Kaggle. The EfficientNetB0 was selected as the backbone architecture because of the small number of parameters. The batch normalization and dropout layers were used for the model to generalize better and stop it from overfitting. The proposed study CBAM block for the attention mechanism. Grad-CAM visualizations show which parts of an image affect the model's predictions, which makes the decision-making process more open and trustworthy. The model achieved 90% accuracy. The proposed method shows great promise for use in automated skin cancer screening systems and can help dermatologists.

INTRODUCTION

The skin is the vital human organ that serves as the interface between the internal and external environment [1]. Skin cancer is one of the most common types of cancer, and it affects millions of people around the world. According to the World Health Organization (WHO) [2] three out of four cases are caused by skin cancer. Melanoma [3] is a highly lethal form of skin cancer. In the last few years, the cases of skin cancer have risen because of the exposure to UV [4] radiation, which comes from the sun. The treatment of skin cancer starts with the proper diagnosis. The proper and early diagnosis [5] leads to better and timely treatment. There are many skin diseases that can be categorized as malignant melanoma and benign melanoma [6] [7] .The traditional approach of diagnosis with a naked eye leads to inaccurate diagnosis due to several reasons. The skin lesions may look alike. Detecting skin lesions using automated procedures such as image processing, computer vision, and deep learning helps to diagnose the disease timely and accurately. The deep learning approach for the detection of skin lesions is one of the best choices because of its automated feature extraction process. Deep learning architecture handles large sizes of image data. To identify skin cancer using deep learning is an effective choice. Several CNN architectures, such as VGG, ResNet, DenseNet, and more recently EfficientNet, have shown the best performance on a range of medical imaging tasks, including classifying skin lesions.

In this study we proposed a deep learning approach for the classification of skin diseases consisting of benign and malignant. The proposed approach comprised EfficientB0 and the convolutional block attention module (CBAM). The EfficientNetB0 architecture is the variation of EfficientNet. The architecture contains fewer parameters and FLOPs. EfficientNet B0, the first model in the EfficientNet series, was developed to achieve a good balance between model size and performance [8] . This makes it useful for diagnosing and deploying in real time on edge devices. But, like other CNNs, EfficientNet can't always focus on the most important parts of an image. This is very important in medical settings where finding lesions is important. Attention mechanisms [9] have become a powerful way to help neural networks find useful parts of images to get around this problem. One of these mechanisms is the Convolutional Block Attention Module (CBAM). It uses channel and spatial attention in turn to make feature maps better. When used with CNNs, CBAM makes the network better at showing data by showing important features and hiding unimportant ones. When classifying skin lesions, attention modules can help the model focus on patterns and textures that are unique to a disease. This makes the diagnosis more accurate and dependable. We suggested a deep learning framework that uses EfficientNetB0 and CBAM to accurately classify skin cancer. The model is trained and tested on a collection of skin images that display two categories of skin cancer, specifically benign and malignant forms. We made the model more general and less likely to overfit by adding dropout, batch normalization, and data augmentation techniques. The architecture is supposed to find a balance between how complicated the model is and how well it works, making sure that both speed and accuracy are kept. In healthcare, AI models need to be able to be understood as well as be accurate at classifying things. Clinicians need to know why a model makes a prediction to trust it and use it correctly in their work. So, we use several different methods to see and understand how the model works inside. We use Grad-CAM (Gradient-weighted Class Activation Mapping) to make heatmaps that show areas of interest in the input images. These visualizations make it easy to understand what the model is focusing on during inference, which helps to check if it is looking at medically important areas. We look at the learned features of space's global structure in addition to its pixel-level interpretability. We use standard classification metrics like accuracy, precision, recall, and F1-score to measure how well our model works. Our model gets high accuracy and AUC scores for all classes, showing that it can handle the difficulties that come with classifying skin cancer into more than one type. This study contributes to the expanding research on utilizing AI to assist in the detection of skin cancer by integrating a lightweight and efficient backbone (EfficientNetB0) with a robust attention mechanism (CBAM) and a suite of tools for interpreting the outcomes. Our method is better for helping doctors make decisions than traditional black-box models because it shows both performance and reasoning in a way that is easy to see and understand. The proposed model is adaptable for larger The present study [10] employs a hybrid methodology by combining EfficientNetV2L, a cutting-edge convolutional architecture recognized for its compound scaling and effective feature extraction, with LightGBM, a gradient-boosting framework tailored for speed and performance. This ensemble addresses the issues of previous ones by enhancing classification accuracy, reducing overfitting, and effectively handling class variability. The proposed system outperforms previous models in several key metrics achieving a test accuracy of 99.90%, a ROC-AUC of 0.98, and a balanced precision-recall across both malignant and benign classes. This makes it one of the most reliable frameworks for real-time, automated skin cancer detection in medical imaging. The author in [11] employed a deep learning approach for the classification of dermoscopic images. The proposed work was implemented to diagnose two types of disease patterns which include benign and malignant. The suggested work used different deep learning backbone architectures, transfer learning, generative adversarial networks (GANs), and data augmentation. The data augmentation and GANs were used to handle the issue of class imbalance in the dataset. The proposed work also used KNN algorithm. The ANN achieved an accuracy of 86.6% for feature extraction. The study [12] proposed a fusion-based strategy that is the combination of multiple deep learning architectures, which include EfficientNetB0, EfficientNetB2 and ResNet50. The proposed approach combines EfficientNetB0 and EfficientB2 and ResNet50. A fusion mechanism sends the features it has extracted to dense and dropout layers to make them more general and lower their dimensionality. The study used a dataset consisting of 27153 images from Kaggle. The datasets or can be modified for binary classification tasks (e.g., melanoma vs. non-melanoma) according to clinical requirements.

Related Work

dataset was split into training, validation and test. 80% was used for training and 10% for validation and 10% for test. The proposed approach achieved 99.14% accuracy. The study presented by [13] employs a hybrid framework that combines deep learning and quantum learning, integrating a CNN backbone with a Quantum Variational Classifier (QVC). The suggested method captures deep spatial features. The proposed work uses ensemble deep learning along with LSTM, ELM, SSDA, and gray wolf optimization for fine-tuning. The proposed work also uses preprocessing such as CLAHE and WF-based preprocessing, data augmentation, and segmentation methods to improve the performance of the model. The Grad-CAM was used to add explainability, making it useful for real-life clinical settings. The author trained two datasets, one is HAM10000 and ISIC. The model achieved 98.38% and 98.17%, respectively. The author in [14] presented a work using deep learning. This study evaluates pre-trained models (ResNet152V2, EfficientNetV2B0, InceptionResNetV2, and MobileNetV3) alongside a custom CNN (FCDS-CNN) for the classification of skin lesions. The author used a dataset containing 10015 images obtained from Kaggle. The dataset was split 90% for training and 10% for the test set. The dataset contains seven diseases. The proposed approach achieved an accuracy of 96%. The author in [15] proposed ResUNet++ and AlexNet-RF for skin disease segmentation. The author trained models on different datasets which include HAM1000, ISIC2019 and PH2. On HAM1000 the models achieved 92% accuracy, on ISIC2019 the models achieved 94.2% accuracy, and on PH2 the models achieved 93.8% accuracy. The author in [16] employed deep learning and handcrafted

ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 7, 2025

feature extraction to classify eight different diseases of skin. The image preprocessing was also performed to remove hair and other anomalies. For the feature extraction, the InceptionV3 model was used. The dataset ISIC2019 was used for training, containing 25331 images. The proposed model achieved an accuracy of 97.81%. The study [17] created an ASCDC-CSODL method to use deep learning on dermoscopic images to find and classify skin cancer. The BF method was used to get rid of noise in ASCDC-CSODL and the U-Net method to split the data into groups. The suggested method also used MobileNet to get the features. The GRU method was also used to tell the difference between skin cancers.

Methodology

In this study we used EfficientNetB0 along with Convolutional Block Attention module. The skin disease contains different visuals to diagnose the disease correctly. The EfficientNetB0 with fewer parameters is helpful feature extraction without costing too much on computation. On the other hand, CBAM block is helpful to visualize the infected parts that help dermatologists in making right and quick decisions. Figure 1 shows the workflow of the proposed system. First th dataset was obtained from Kaggle. The input images were preprocessed and then EfficientB0 was employed for feature extraction along with CBAM. The model classifies benign and malignant skin diseases.

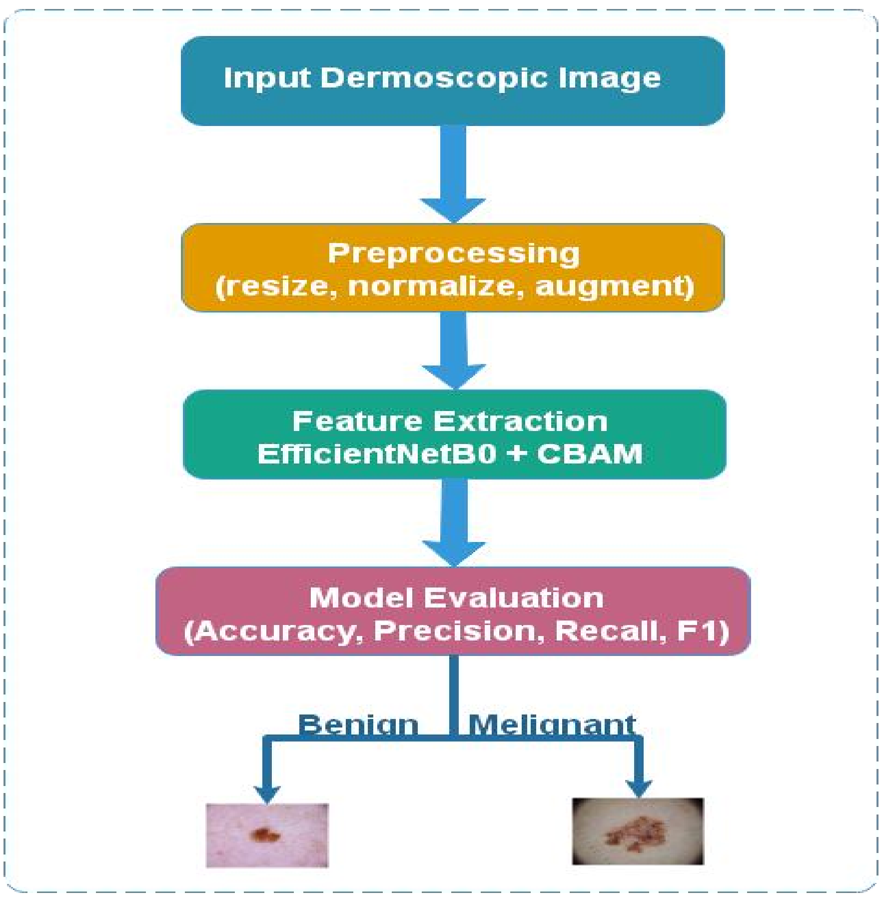

Figure 1. Workflow

Figure 2 shows a methodology for classifying skin lesions using a deep learning method based on the EfficientNetB0 model with the Convolutional Block Attention Module (CBAM) added. The first thing to do is get the data, which means taking pictures of skin lesions with a dermoscope. Most of the time, these images show different kinds of skin problems that might not be bad. The images are preprocessed after they are acquired. This can mean changing their size, normalizing them, reducing noise, and adding new features to make the input data better and more varied. This helps the model understand more about different situations. After being preprocessed, the images are sent to the proposed model, which is a mix of EfficientNetB0 and CBAM. EfficientNetB0 is a new kind of convolutional neural network that uses a compound scaling method to find the right size and accuracy for the model. Architecture learns features even better when you add CBAM to it. CBAM adds attention mechanisms that work on both the channels and the space. These help the network focus on the most important parts of the image and make it better at telling the difference between good and bad features. Then, the model uses the network's output to figure out what kind of lesion it is by putting it into the classification block. The classification can either be benign (not cancerous) or malignant (cancerous). The purpose of this method is to give doctors a very precise, automated way to find skin cancer early. Quickly diagnosing them will help them make better choices and get better results for their patients.

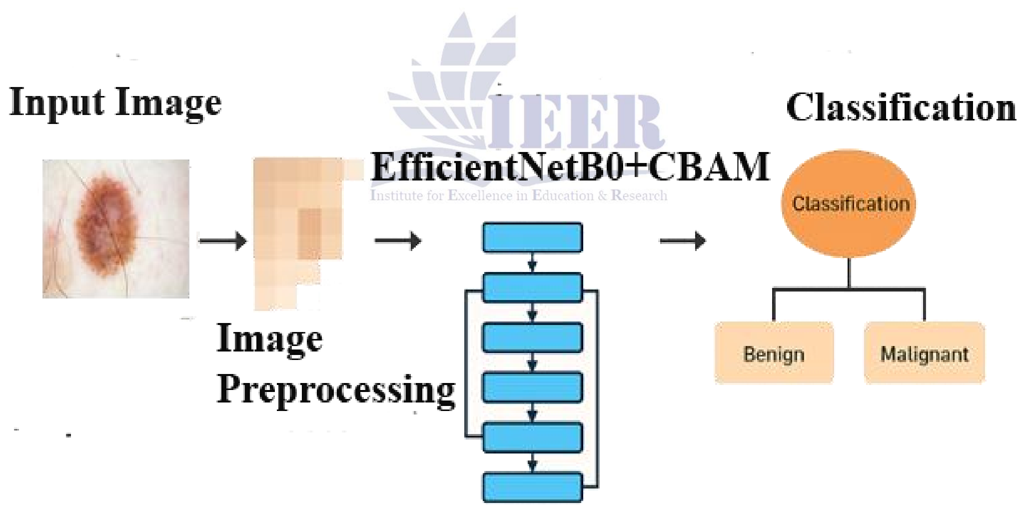

Figure 2. Proposed Methodology

Dataset

We did experiments on a publicly available dermoscopic image dataset, which contains 3397 dermoscopic images of 2 skin disease classes. Figure 3 shows the train and test set. We used image processing such as resizing images to 224×224 pixels so that they would be the right size for the model. We also used normalization and data augmentation.

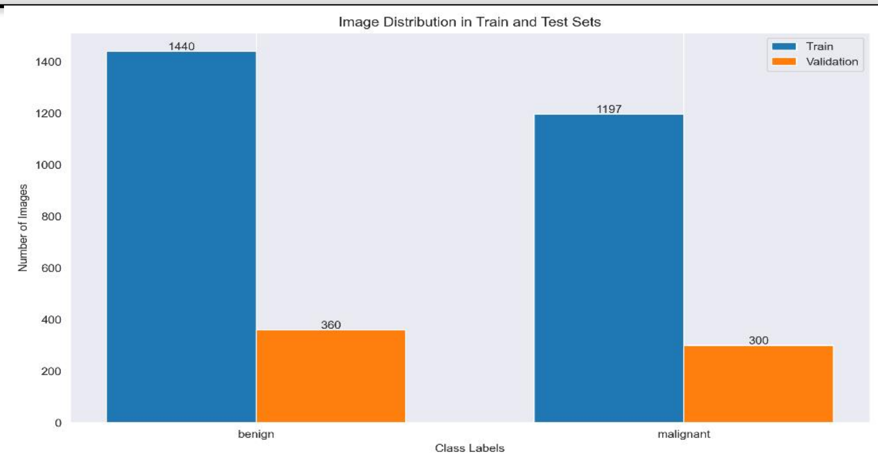

Figure 3. Image Distribution in Train and Test

Figure 4 shows the sample images of skin diseases in the benign and malignant classes.

Figure 4. Sample Skin Disease Images

Figure 5 shows training and validation accuracy. For training the model, early stopping was used. Model was trained on 30 epochs.

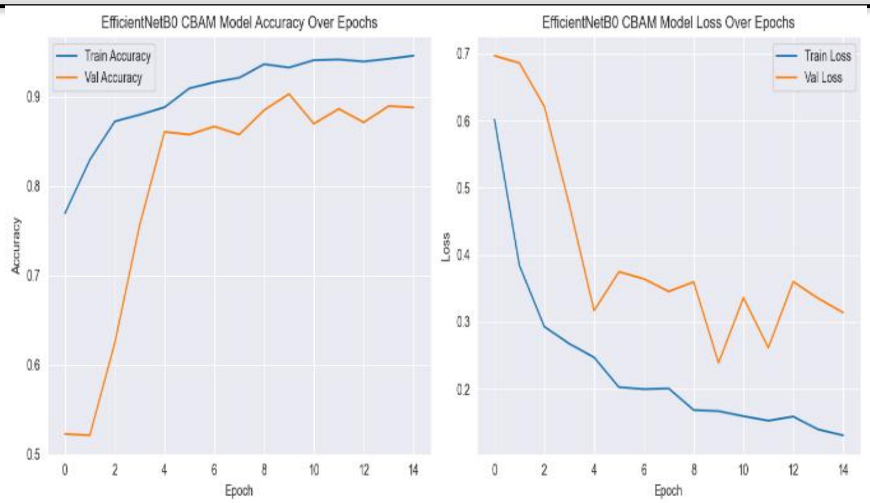

Figure 5. Model Training and Validation Accuracy

Figure 6 shows the number of predicted classes. The model predicted 376 benign and 284 malignant.

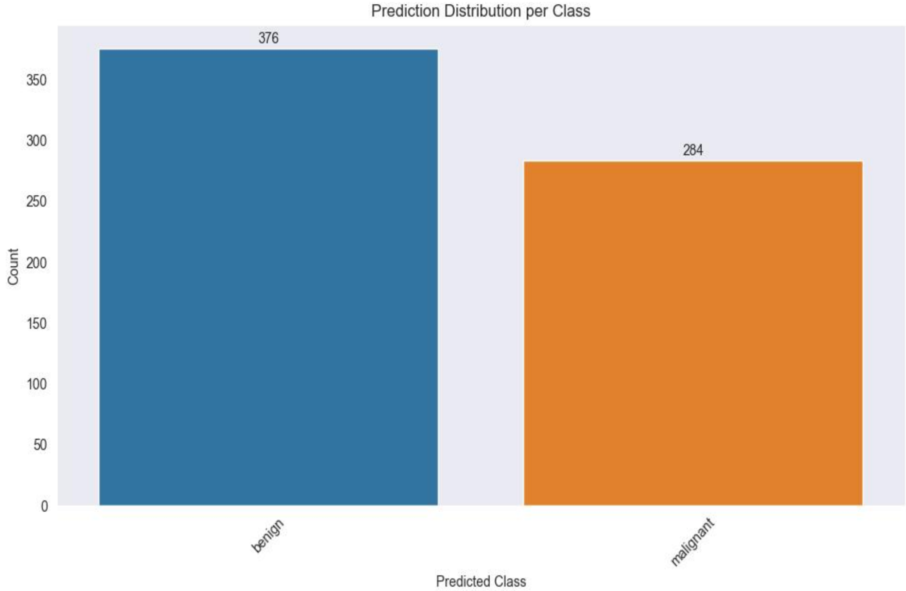

Figure 6. Predicted Class Distribution

Figure 7 visualizes the predicted images along with the confidence score.

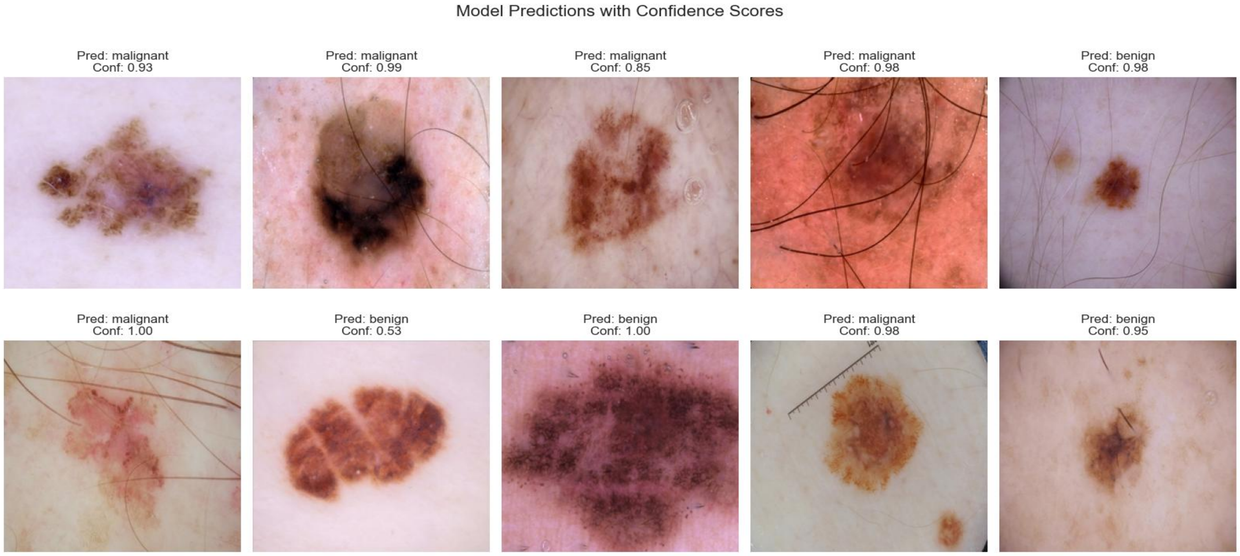

Figure 7. Model Predictions with Confidence Scores

Figure 8 shows the original image along with the Grad-CAM heatmap and Grad-CAM overlay. Grad-CAM (Gradient-weighted Class Activation Mapping) is a well-known explainable AI method that shows which parts of an input image have the most effect on a model's prediction. It makes a heatmap by finding the gradients of the predicted class score with respect to the feature maps of the last convolutional layer. These gradients show which parts of the image had the most effect on the decision by highlighting important spatial areas. The Grad-CAM heatmap is a grayscale or colored intensity map. Warmer colors, like red and yellow, show higher importance, while cooler colors, like blue, show lower importance. The Grad-CAM overlay is when you put this heatmap on top of the original input image to make a visualization that makes sense. Researchers and doctors can use this overlay to see if the model is focusing on the right lesion area or being led astray by features that don't matter. It makes the model clearer, which is especially important in sensitive areas like medical diagnosis.

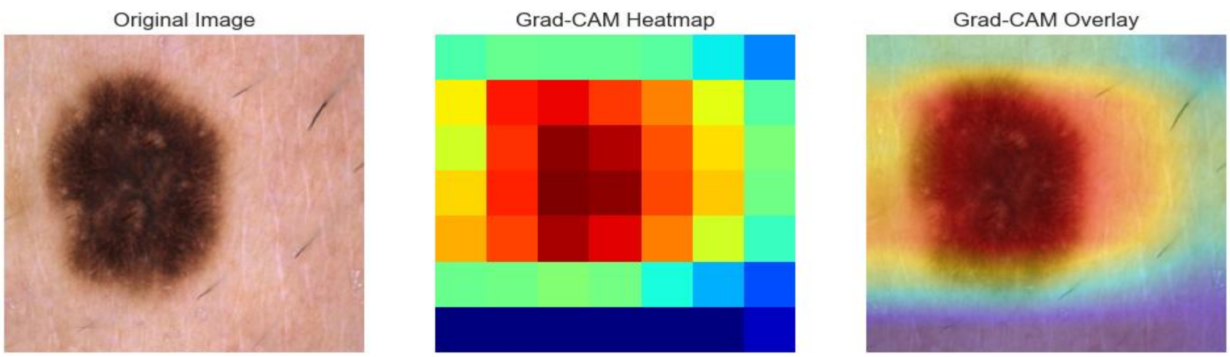

Figure 8. Grad-CAM Heatmap and Grad-CAM Overlay

Table 1: Classification Report

Class	Precision	Recall	F1 Sore	Support
benign	0.89	0.93	0.91	360
malignant	0.92	0.87	0.89	300

Model Accuracy	**90.0%**

Conclusion

The proposed CBAM-enhanced EfficientNetB0 architecture significantly improves the classification of skin cancer, enhancing its efficacy in clinical applications. Because of Grad-CAM heatmaps, the model is both easy to understand and very good at making predictions. In this study, we demonstrated a deep learning framework for the classification of two types of skin cancer. It uses the Convolutional Block Attention Module (CBAM) with the lightweight but powerful EfficientNetB0 architecture. The main goal was to make it easier to understand and sort dermoscopic pictures. We used a dataset of skin cancer with two types of lesions to test the model. The model was able to classify things with 90.0% accuracy. Grad-CAM visualizations demonstrated that the inclusion of CBAM enhanced the model's ability to concentrate on the appropriate lesion regions. Quantitative analysis showed that precision, recall, and F1-score all got better across different classes when compared to the baseline EfficientNetB0 model. In addition to performance, our framework puts a lot of emphasis on transparency and clinical applicability. Grad-CAM made it easier to understand local data by showing which parts of dermoscopic images were most useful for telling lesions apart. The proposed model generally finds a good balance between speed, accuracy, and ease of understanding. This makes it a good choice for use in real-world clinical settings or mobile diagnostic apps. Because its design is modular and scalable, it can be used for more than just dermatology.

1) **References**

2) [1] B. Ozdemir and I. Pacal, "An innovative deep learning framework for skin cancer detection employing ConvNeXtV2 and focal self-attention mechanisms," *Results in Engineering*, vol. 25, Mar. 2025, doi: 10.1016/j.rineng.2024.103692.
3) [2] C. Kavitha, S. Priyanka, M. P. Kumar, and V. Kusuma, "Skin Cancer Detection and Classification using Deep Learning Techniques," in *Procedia Computer Science*, Elsevier B.V., 2024, pp. 2793–2802. doi: 10.1016/j.procs.2024.04.264.
4) [3] D. Meedeniya, S. De Silva, L. Gamage, and U. Isuranga, "Skin cancer identification utilizing deep learning: A survey," Nov. 13, 2024, *John Wiley and Sons Inc.* doi: 10.1049/ipr2.13219.
5) [4] R. Raut, Y. Borole, S. Patil, V. Khan, and D. G. Takale, "Skin Disease Classification Using Machine Learning Algorithms", doi: 10.14704/nq.2022.20.10.NQ55940.
6) [5] A. A. Hussein and A. M. Abdulazeez, "A Review on Utilizing Machine Learning Classification Algorithms for Skin Cancer," *Journal of Applied Science and Technology Trends*, vol. 5, no. 2, pp. 60–71, Jul. 2024, doi: 10.38094/jastt52191.
7) [6] A. H. Roky *et al.*, "Overview of skin cancer types and prevalence rates across continents," Mar. 01, 2025, *Chinese Medical Association.* doi: 10.1016/j.cpt.2024.08.002.
8) [7] N. KAVITHA *et al.*, "A PERSPECTIVE STUDY ON SCALABLE COMPUTATION MODEL FOR SKIN CANCER DETECTION: ADVANCEMENTS AND CHALLENGES," *Scalable Computing*, vol. 26, no. 1, pp. 48–63, 2025, doi: 10.12694/scpe.v26i1.3501.

9) [8] K. Kansal, T. B. Chandra, and A. Singh, "ResNet-50 vs. EfficientNet-B0: Multi-Centric Classification of Various Lung Abnormalities Using Deep Learning 'session id: ICMLDsE.004,'" in *Procedia Computer Science*, Elsevier B.V., 2024, pp. 70–80. doi: 10.1016/j.procs.2024.04.007.
10) [9] H. Liu *et al.*, "A skin disease classification model based on multi scale combined efficient channel attention module," *Sci Rep*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-90418-0.
11) [10] S. M. R. Swapno, S. M. N. Nobel, P. K. Meena, V. P. Meena, J. Bahadur, and A. Appaji, "Accelerated and precise skin cancer detection through an enhanced machine learning pipeline for improved diagnostic accuracy," *Results in Engineering*, vol. 25, Mar. 2025, doi: 10.1016/j.rineng.2025.104168.
12) [11] O. Akinrinade and C. Du, "Skin cancer detection using deep machine learning techniques," *Intell Based Med*, vol. 11, Jan. 2025, doi: 10.1016/j.ibmed.2024.100191.
13) [12] M. Alruwaili and M. Mohamed, "An Integrated Deep Learning Model with EfficientNet and ResNet for Accurate Multi-Class Skin Disease Classification," *Diagnostics*, vol. 15, no. 5, Mar. 2025, doi: 10.3390/diagnostics15050551.
14) [13] J. D. Dorathi Jayaseeli *et al.*, "An intelligent framework for skin cancer detection and classification using fusion of Squeeze-Excitation-DenseNet with Metaheuristic-driven ensemble deep learning models," *Sci Rep*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-92293-1.
15) [14] K. Nawaz *et al.*, "Skin cancer detection using dermoscopic images with convolutional neural network," *Sci Rep*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-91446-6.
16) [15] S. Mustafa, A. Jaffar, M. Rashid, S. Akram, and S. M. Bhatti, "Deep learning-based skin lesion analysis using hybrid ResUNet++ and modified AlexNet-Random Forest for enhanced segmentation and classification," *PLoS One*, vol. 20, no. 1, Jan. 2025, doi: 10.1371/journal.pone.0315120.
17) [16] A. Naeem, T. Anees, M. Khalil, K. Zahra, R. A. Naqvi, and S. W. Lee, "SNC_Net: Skin Cancer Detection by Integrating Handcrafted and Deep Learning-Based Features Using Dermoscopy Images," *Mathematics*, vol. 12, no. 7, Apr. 2024, doi: 10.3390/math12071030.
18) [17] V. A. Rajendran and S. Shanmugam, "Automated Skin Cancer Detection and Classification using Cat Swarm Optimization with a Deep Learning Model," *Engineering, Technology and Applied Science Research*, vol. 14, no. 1, pp. 12734–12739, Feb. 2024, doi: 10.48084/etasr.6681.
19)