VOL. 10. NO. 4 MAY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480 /jitk.v10i4.6541

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

DEVELOPMENT OF SKIN CANCER PIGMENT IMAGE CLASSIFICATION USING A COMBINATION OF MOBILENETV2 AND CBAM

Juni Ismail[1*]; Lili Tanti[2]; Wanayumini[3]

Computer Science, Faculty of Engineering and Computer Science[1,2,]
Universitas Potensi Utama, Indonesia[1,2]
www.potensi-utama.ac.id[1,2]
juniismailll@gmail.com[1*], lilitanti82@gmail.com[2]

Computer Science, Faculty of Engineering and Computer Science[3]
Universitas Asahan, Indonesia[3]
www.una.ac.id[3]
wanayumini@gmail.com[3]

(*) Corresponding Author
(Responsible for the Quality of Paper Content)

Abstract—*Skin cancer is one of the most common types of cancer worldwide, making early detection a crucial factor in improving patient recovery rates. This study compares three classification methods for pigmented skin cancer images using a combination of VGG16 with CBAM, MobileNetV2 with CBAM, and a hybrid VGG16-MobileNetV2 approach with transfer learning. The dataset used in this study is the Skin Cancer ISIC - The International Skin Imaging Collaboration (HAM10000) from Kaggle, which consists of 10,015 images covering seven types of skin cancer. After balancing, the dataset was reduced to 2,400 images with three main classes: Actinic Keratosis (AKIEC), Basal Cell Carcinoma (BCC), and melanoma (MEL), each containing 800 images. This study involves data preprocessing stages such as augmentation, normalization, and image resizing to ensure optimal data quality. The model training process was conducted using the Adam optimizer, a batch size of 16, and an Early Stopping mechanism to prevent overfitting. Evaluation results indicate that the MobileNetV2 with CBAM model achieved the best performance with a validation accuracy of 86%, followed by the VGG16-MobileNetV2 combination at 77%, while VGG16 with CBAM experienced overfitting with an accuracy of 54%. Additionally, the best-performing model demonstrated a precision of 86.53% and a recall of 86.46%, highlighting its superior stability in detecting skin cancer compared to previous single-model approaches. With these results, the developed system can serve as an effective tool for medical professionals in performing early and more accurate skin cancer diagnoses.*

Keywords*: CBAM, CNN, image classification, mobilenet, skin cancer,*

Intisari—*Kanker kulit merupakan salah satu jenis kanker yang paling umum di dunia, sehingga deteksi dini menjadi faktor krusial dalam meningkatkan tingkat kesembuhan pasien. Penelitian ini membandingkan tiga metode klasifikasi untuk gambar kanker kulit berpigmen dengan menggunakan kombinasi VGG16 dengan CBAM, MobileNetV2 dengan CBAM, dan pendekatan hibrida VGG16-MobileNetV2 dengan transfer learning. Dataset yang digunakan dalam penelitian ini adalah Skin Cancer ISIC - The International Skin Imaging Collaboration (HAM10000) dari Kaggle, yang terdiri dari 10.015 gambar mencakup tujuh jenis kanker kulit. Setelah dilakukan penyeimbangan, dataset dikurangi menjadi 2.400 gambar dengan tiga kelas utama: Actinic Keratosis (AKIEC), Basal Cell Carcinoma (BCC), dan melanoma (MEL), masing-masing sebanyak 800 gambar. Penelitian ini mencakup tahap praproses data seperti augmentasi, normalisasi, dan pengubahan ukuran gambar untuk memastikan kualitas data yang optimal. Proses pelatihan model dilakukan menggunakan optimizer Adam, ukuran batch 16, dan mekanisme Early Stopping untuk mencegah overfitting. Hasil evaluasi menunjukkan bahwa model MobileNetV2 dengan CBAM mencapai performa terbaik dengan akurasi validasi*

VOL. 10. NO. 4 MAY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v10i4.6541

sebesar 86%, diikuti oleh kombinasi VGG16-MobileNetV2 sebesar 77%, sementara VGG16 dengan CBAM mengalami overfitting dengan akurasi sebesar 54%. Selain itu, model dengan performa terbaik menunjukkan nilai presisi sebesar 86,53% dan recall sebesar 86,46%, yang menandakan stabilitas superior dalam mendeteksi kanker kulit dibandingkan pendekatan model tunggal sebelumnya. Dengan hasil ini, sistem yang dikembangkan dapat menjadi alat yang efektif bagi para profesional medis dalam melakukan diagnosis kanker kulit secara dini dan lebih akurat.

Kata Kunci: *CBAM, CNN, klasifikasi citra, mobilenet, kanker kulit,*

INTRODUCTION

Skin cancer is one of the most common types of cancer worldwide [1][2][3]. According to the American Cancer Society, more than 5 million cases of skin cancer are diagnosed each year, making it one of the most frequently occurring forms of cancer in humans [4][5][6][7]. Early detection of skin cancer is crucial for increasing the chances of recovery and reducing patient mortality rates. Classification of skin cancer images faces several challenges, including high color and texture variations, imbalance in the number of samples between classes, and difficulty in distinguishing similar categories. Deep learning models require special strategies, such as data augmentation techniques and model architectures that can effectively capture critical features. In recent decades, advancements in computational technology have significantly contributed to skin cancer diagnosis through complex medical image analysis [8][9]. One widely used approach is deep learning, particularly Convolutional Neural Networks (CNNs), which have been proven effective in analyzing and classifying medical images, including pigmented skin cancer images [10][11][12].

One of the most commonly used CNN architectures is VGGNet. VGGNet is well known for its deep structure and strong feature extraction capabilities [13][14][15][16]. This model utilizes small convolutional layers in large numbers to enhance feature extraction effectiveness from medical images [17][18][19]. In addition to VGGNet, MobileNet is also widely used due to its design for higher computational efficiency with a smaller model size. This model is suitable for deployment on resource-constrained devices without significantly compromising accuracy [20][21][22]. Moreover, the Convolutional Block Attention Module (CBAM) is frequently used to improve CNN performance in classification and segmentation tasks [23][24]. CBAM is an attention module consisting of two main components: Channel Attention and Spatial Attention [25][26]. This module works by adjusting feature weights based on their importance in both the channel and spatial domains, thereby enhancing the representation of relevant features [27][28]. By incorporating CBAM into CNN architectures such as *VGGNet* or *MobileNet*, the model can adaptively highlight crucial features in medical images, ultimately improving classification accuracy and feature extraction efficiency.

The related research referenced in this study is the work of Fedryanto Dartiko et al., which developed a skin cancer classification method using a Convolutional Neural Network (CNN) with a hybrid preprocessing approach. The preprocessing techniques used include CLAHE, morphological closing, and median filtering to remove noise caused by fine hairs on the epidermis. The results of the study showed that this method achieved an accuracy of 78.19% with a loss of 0.5324. Although the results were quite good, this study still faced challenges such as limited data availability and overfitting at higher epochs [29]. Another related study, which serves as the primary reference for this research, was conducted by Luqman Hakim et al. [30]. Their study used a CNN model with an eight-layer convolutional architecture and achieved an accuracy of 75%, with the highest precision and recall values in the benign class, at 0.80 and 0.82, respectively, and an F1-score of 0.81. However, their study still had limitations in improving model accuracy and computational efficiency.

The primary issue addressed in this research is the need to enhance accuracy compared to previous studies. Although CNNs have been widely used for skin cancer image classification, challenges remain in improving accuracy without excessively increasing computational costs. To address this issue, this study aims to enhance the classification accuracy of pigmented skin cancer images by comparing the advantages of VGG16 with CBAM, MobileNetV2 with CBAM, and a combination of VGG16 and MobileNetV2. By comparing these three model combinations, the goal is to achieve higher accuracy without compromising computational efficiency. Several previous studies have used CNN in skin cancer classification, such as VGG16 and ResNet, but still face challenges in detecting subtle features in skin lesions. The combination models compared in this study have been applied in various studies, but often have limitations in capturing important features of the image. To overcome these shortcomings, this study proposes a comparison of the combination of VGG16 and CBAM, MobileNetV2 and CBAM, and the combination of VGG16 and MobileNetV2 to improve the accuracy of the model with the best mechanism.

The gap between this study and previous research lies in the approach used to improve classification model accuracy. Previous studies have only utilized a single CNN architecture, whereas this study compares three combinations: VGG16 with CBAM, MobileNetV2 with CBAM, and the combination of VGG16 and MobileNetV2 to enhance classification accuracy without significantly increasing computational complexity. Additionally, this study will evaluate the combined model's performance in improving interpretability for skin cancer diagnosis. The uniqueness of this research lies in the comparative approach of VGG16 with CBAM, MobileNetV2 with CBAM, and the combination of VGG16 with MobileNetV2 in classifying pigmented skin cancer images. This approach has not been widely explored in previous studies and is expected to improve classification accuracy without sacrificing computational efficiency.

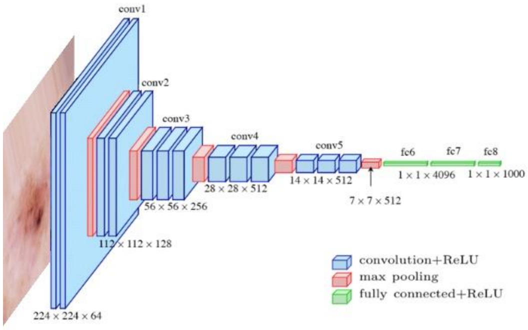
This study contributes to the field of artificial intelligence for medical diagnosis, specifically in improving the accuracy of pigmented skin cancer image classification through a combined deep learning model approach. The results of this study are expected to assist medical professionals in detecting skin cancer more quickly and accurately.

The primary objective of this research is to develop and evaluate a pigmented skin cancer classification method that enhances diagnostic accuracy. Specifically, this study aims to compare the effectiveness of three combination approaches: VGG16 with CBAM, MobileNetV2 with CBAM, and the combination of VGG16 and MobileNetV2. By analyzing these models, the study seeks to determine the most optimal architecture for classifying pigmented skin cancer.

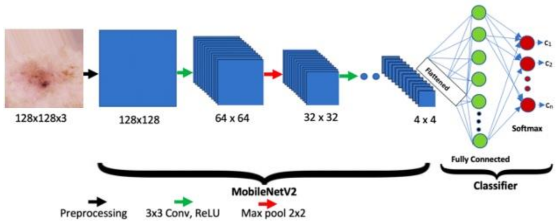
The dataset used is the Skin Cancer ISIC - The International Skin Imaging Collaboration (HAM10000) from Kaggle, consisting of 10,015 images of three primary types of skin cancer: Basal Cell Carcinoma, Melanoma, and Actinic Keratosis.

MATERIALS AND METHODS

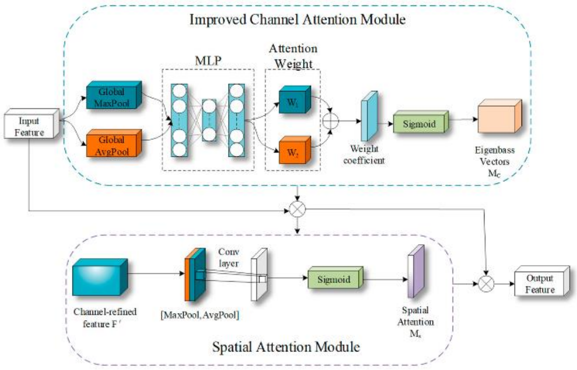
This study was conducted to improve the accuracy of pigmented skin cancer image classification by comparing three models: VGG16 with CBAM, MobileNetV2 with CBAM, and a combination of VGG16 and MobileNetV2. The results of this study are expected to assist medical professionals in detecting skin cancer more quickly and accurately. The architecture of the three models can be seen in Figures 1, 2, and 3.



Source: (Research Results, 2025)
Figure 1. VGG16 architecture in image classification



Source: (Research Results, 2025)
Figure 2. MobileNetV2 architecture in image classification



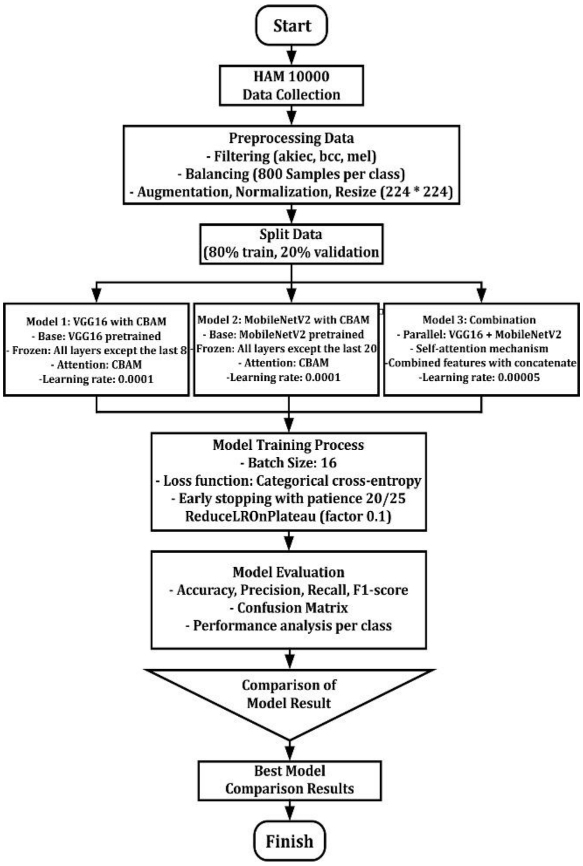
Source: (Research Results, 2025)
Figure 3. Convolutional Block Attention Module Architecture

VGG16, MobileNetV2, and CBAM play an important role in improving the accuracy of pigment skin cancer image classification in this study. VGG16, with 16 small convolutional layers (3x3), is able to capture spatial patterns and textures in depth, although it has a weakness in computational efficiency. MobileNetV2, which is designed for high efficiency, uses *depth wise* separable convolutions and inverted residual blocks to reduce the number of parameters without

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)
VOL. 10. NO. 4 MAY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v10i4.6541

sacrificing performance, making it suitable for implementation on devices with computational limitations. CBAM (Convolutional Block Attention Module) is added as an attention module that enhances feature selectivity with Channel Attention and Spatial Attention, helping the model to focus on more relevant image parts in classification. The combination comparison involving these three architectures is expected to optimally improve the accuracy of skin cancer classification, combining the advantages of deep feature extraction from VGG16, the efficiency of MobileNetV2, and the enhancement of feature representation from CBAM.

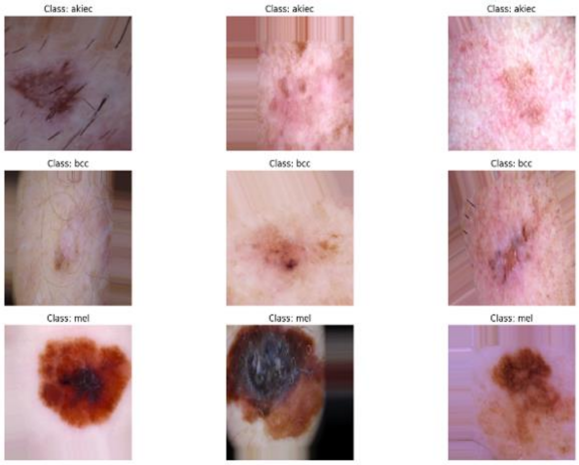
Research Stages



Source: (Research Results, 2025)
Figure 4. Research Stages

1. Data Collection

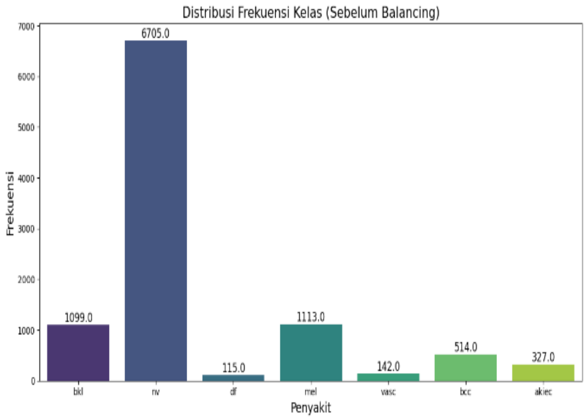
The dataset used in this study is "Skin Cancer ISIC - The International Skin Imaging Collaboration (HAM10000)" from Kaggle. This dataset consists of 10,015 images, but only 2,400 sample images were used in this study. Among the seven available classes in the original dataset, this research focuses on three major types of skin cancer: Actinic Keratosis (AKIEC), Basal Cell Carcinoma (BCC), and melanoma (MEL), in accordance with the study's problem scope. Figure 5 presents examples of each skin cancer class used in this research.



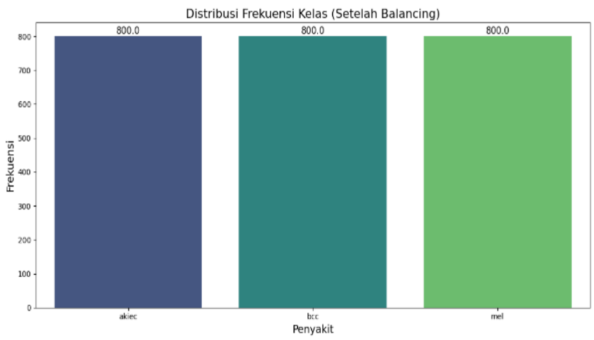
Source: (Kaggle, 2021)
Figure 5. Sample Dataset Of 3 Classes Of Skin Cancer

2. Data Balancing

Before proceeding to the next steps, the dataset was first balanced using a balancing technique to ensure a more proportional distribution. The frequency distribution of classes before and after balancing can be seen in Figures 6 and 7.



Source: (Research Results, 2025)
Figure 6. Data Class Frequency Before Balancing



Source: (Research Results, 2025)
Figure 7. Data Class Frequency After Balancing

Figure 7 illustrates the dataset after the balancing process, which focuses only on three major types of skin cancer: Actinic Keratosis (AKIEC), Basal Cell Carcinoma (BCC), and melanoma (MEL). After balancing, each class contains 800 samples, resulting in a total of 2.400 samples used in this study.

3. Data Preprocessing

Before being used in the deep learning model, the dataset undergoes a series of preprocessing steps to ensure its quality. Data augmentation is applied using various transformations, including rotation (±40°), horizontal and vertical shifts (0.2), scaling (0.2), zooming (0.3), flipping, and brightness adjustment. These techniques enhance data diversity and help reduce the risk of overfitting.
Additionally, normalization is performed by rescaling pixel values from the 0–255 range to 0–1, making it easier for the model to learn. All images are also resized to 224×224 pixels, aligning with the requirements of the CNN architectures used in this study.

4. Dataset Splitting

After the preprocessing stage is completed, the dataset is divided into two parts 80% (1.920 samples) for training and 20% (480 samples) for validation. This division maintains a balanced proportion for each class, ensuring fair representation of all skin cancer types. This approach allows the model to learn effectively from the training data while also enabling a valid evaluation of its generalization ability on unseen data. However, since no separate test set is used, the reported model performance on the validation set might be optimistically biased, as it is also involved in model selection and early stopping during training. Therefore, this data split is an important methodological aspect to consider in this study. This limitation can be addressed in future research by introducing a separate test set that is completely excluded from the training and validation processes. This would provide a more reliable evaluation of the model's performance on entirely new data.

5. Comparison of the Three Combination Methods

To improve accuracy in skin cancer classification, this study compares three combination methods: VGG16 with CBAM, MobileNetV2 with CBAM, and a combination of VGG16 and MobileNetV2. Each method employs a different approach to feature extraction and attention mechanisms in identifying critical areas within the images. Table 1 presents the key comparisons among these three methods.

Table 1. Three Combination Methods

Aspect	VGG16 + CBAM	MobileNetV2 + CBAM	Combination of VGG16 & MobileNetV2
Base Model	VGG16 (pre-trained ImageNet)	MobileNetV2 (pre-trained ImageNet)	VGG16 & MobileNetV2 (run in parallel)
Trainable Layer	Last 8 layers	Last 20 layers	All layers in both models remain trainable
Attention Mechanism	CBAM to increase focus on important features	CBAM to increase focus on important features	Self-attention for adaptive weights on features
Feature Fusion	None (only one model)	None (only one model)	Concatenation of features of both models
Learning Rate	0.0001	0.0001	0.00005
Excess	Stable in feature extraction	Light and fast in computing	Produces a richer feature representation
Lack	Heavier in computing than MobileNetV2	Performance may be lower if the data is complex	More complex and requires greater resources

Source: (Research Results, 2025)

6. Model Training

The training process for the three methods involves a classification stage with multiple Dense Layers, utilizing *ReLU* activation in hidden layers and *softmax* activation in the output layer to generate final predictions. The models are trained with a batch size of 16, balancing efficiency and memory usage. The categorical cross-entropy loss function is used, as it is suitable for multi-class classification tasks. To prevent overfitting, Early Stopping (patience of 20–25 epochs) is applied, stopping training when no significant improvement is detected. Additionally, *ReduceLROnPlateau* is implemented, automatically reducing the learning rate by a factor of 0.1 if validation performance does not improve.

7. Evaluation and Hyperparameter Tuning

After training, the model performance is evaluated using accuracy, precision, recall, F1-score, and a confusion matrix. If the results are not optimal, hyperparameters are fine-tuned, including adjustments to the learning rate, model architecture modifications, and the application of regularization techniques such as Dropout to enhance performance and generalization.

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)
VOL. 10. NO. 4 MAY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v10i4.6541

8. Classification with the Best Model

The best-performing model is selected based on validation accuracy, loss, and training stability. The classification process begins with feature extraction, where input images are processed through the chosen model's architecture. Attention mechanisms such as CBAM or Self-Attention are employed to emphasize critical areas in the images. Extracted features are then passed through Dense Layers with *softmax* activation, producing probabilities for each skin cancer category. The model's evaluation using a confusion matrix, precision, recall, and F1-score demonstrates its capability to distinguish between the three primary skin cancer types—Actinic Keratosis (AKIEC), Basal Cell Carcinoma (BCC), and melanoma (MEL)—with minimal misclassification. However, a limitation of this study is the absence of a separate test set, which could be addressed in future research to further validate the model's performance on entirely unseen data.

RESULTS AND DISCUSSION

The research results represent the stage of obtaining outcomes from the developed model architecture for skin cancer classification. The first step in this process is data preprocessing, where image data is first cleaned and adjusted to be suitable for model training. This phase includes filtering, balancing, data augmentation, normalization, and image resizing to align with the CNN architecture requirements.

The resizing process is performed using the resize function from libraries such as OpenCV or PIL. Once the images have been resized, the next step is dataset splitting. The dataset is divided into two main parts: training data and validation data. The splitting ratio used is 80% for training data and 20% for validation data. The dataset distribution results are presented in Table 2.

Table 2. Dataset Divisions Results

Training Data	Validation Data
1.920	480

Source: (Research Results, 2025)

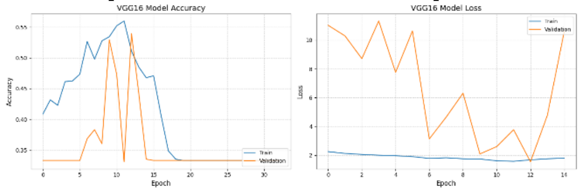
The next stage is data augmentation, where various augmentation parameters are applied to increase the dataset size and enhance model generalization. Several parameters used in this process are presented in Table 3.

Table 3. Augmentation Parameters

Parameter	Value	Description
rotation_range	40	Image rotation in degrees.
width_shift_range	0.2	Horizontal shift up to 20% of the image width.
height_shift_range	0.2	Vertical shift up to 20% of the image height.
shear_range	0.2	Shape distortion in degrees.
zoom_range	0.3	Zoom in or zoom out up to 20%.
horizontal_flip	True	Horizontal flipping of the image.
vertical_flip	True	Vertical flipping of the image.
fill_mode	nearest	Pixel filling method for empty spaces after transformation.
brightness_range	[0.7, 1.3]	Brightness variation of the image.
channel_shift_range	0.15	Color channel shift adjustments.
rescale	1./255	normalization of pixel values in the image.

Source: (Research Results, 2025)

After the data augmentation process, the next stage is model development and training. Each model is trained with 150 epochs and a batch size of 16. The training process stops at different epochs depending on the model's convergence. The training progress is visualized using graphs to illustrate performance trends across epochs.

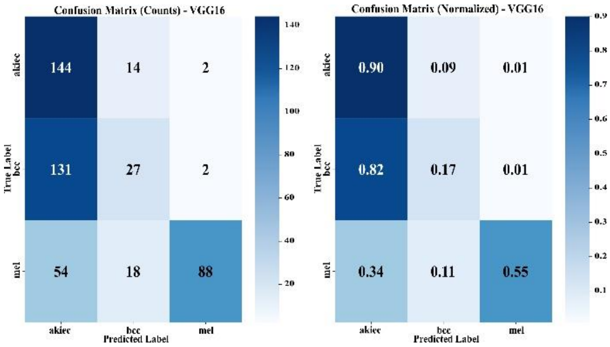


Source : (Research Results, 2025)
Figure 8. VGG16 + CBAM Model Accuracy and Loss Graph

Figure 8 presents the accuracy and loss trends of the VGG16 with CBAM model during training. In the accuracy graph (left), the model exhibits a gradual improvement in training accuracy, but validation accuracy fluctuates sharply after 10 epochs, peaking at 55% before overfitting occurs. In the loss graph (right), training loss remains stable and decreases over time, whereas validation loss fluctuates and rises significantly in the final epochs, indicating the model struggles with generalization. These results suggest overfitting, necessitating mitigation strategies such as architectural adjustments, regularization techniques, or additional data augmentation.

The confusion matrix for VGG16 with CBAM is shown in Figure 9.

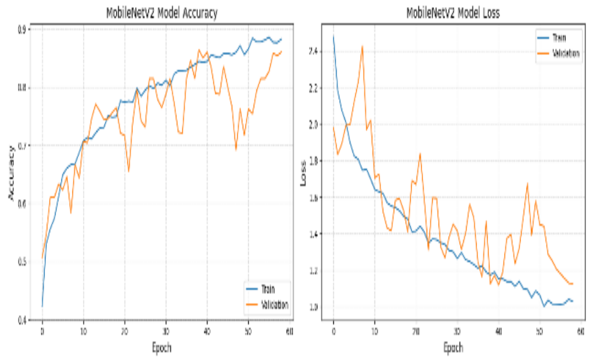
Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri



Source: (Research Results, 2025)
Figure 9. Confusion Matrix VGG16 + CBAM

Figure 9 displays the confusion matrix of the VGG16 with CBAM model for classifying three skin cancer types: Actinic Keratosis (AKIEC), Basal Cell Carcinoma (BCC), and melanoma (MEL). In the absolute matrix (left), the model performs well on AKIEC (144 correct samples) but misclassifies BCC as AKIEC in 131 cases. For melanoma, the model correctly classifies 88 samples, but frequently misclassifies 54 as AKIEC and 18 as BCC. The normalized confusion matrix (right) shows AKIEC achieving the highest accuracy (90%), while BCC struggles with only 17% accuracy, indicating difficulty in distinguishing BCC from AKIEC. melanoma reaches 55% accuracy but is still often confused with other classes. The high misclassification rate between BCC and AKIEC may result from similar visual characteristics between the two types.

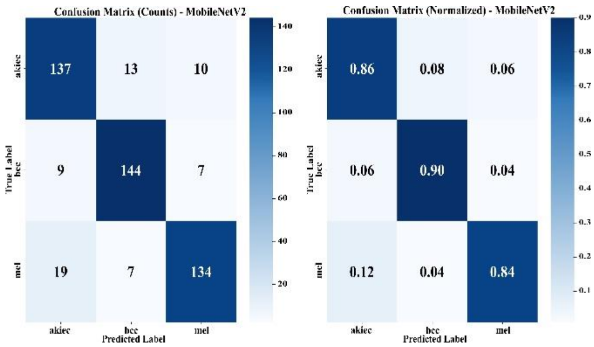
Next, Figure 10 presents the accuracy and loss trends of the MobileNetV2 with CBAM model during training.



Source: (Research Results, 2025)
Figure 10. MobileNetV2 + CBAM Model Accuracy And Loss Graph

Figure 10 presents the accuracy and loss graphs for the MobileNetV2 with CBAM model during training. In the accuracy graph (left), the model exhibits a steady increase in both training and validation accuracy. The validation accuracy reaches 86%, indicating strong model performance in recognizing patterns from the dataset. While there are minor fluctuations in validation accuracy, the overall trend remains positive without significant overfitting. The training process stops at epoch 59, demonstrating that early stopping has been applied to prevent overfitting. In the loss graph (right), both training and validation loss gradually decrease, confirming that the model is effectively minimizing prediction errors. Although validation loss fluctuates slightly more than training loss, it follows a downward trend, showing that the model is still learning without severe performance degradation. By the end of training, validation and training loss converge, reinforcing that the model generalizes well to unseen data.

Next, Figure 11 presents the confusion matrix for the MobileNetV2 with CBAM model to evaluate its classification performance across different skin cancer types.

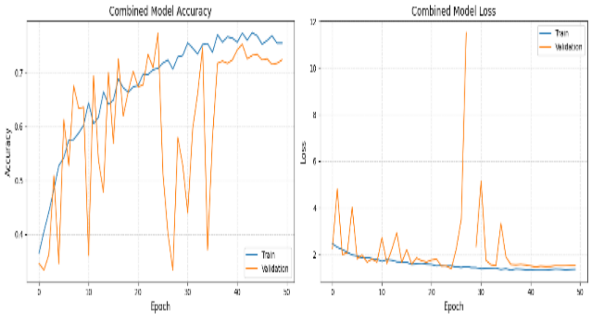


Source: (Research Results, 2025)
Figure 11. Confusion Matrix MobileNetV2 + CBAM

Figure 11 presents the confusion matrix for the MobileNetV2 with CBAM model, illustrating its classification performance. In the absolute confusion matrix (left), the model correctly classifies 137 samples of AKIEC, although 13 are misclassified as BCC and 10 as MEL. The BCC class has 144 correctly predicted samples, with 9 misclassified as AKIEC and 7 as MEL. For the MEL class, 134 samples are correctly classified, but 19 are misclassified as AKIEC and 7 as BCC. In the normalized confusion matrix (right), the model achieves an accuracy of 86% for AKIEC, 90% for BCC, and 84% for MEL. The highest misclassification rate occurs in the MEL class (12% misclassified as AKIEC), indicating possible feature similarities between these classes. Overall, the model demonstrates strong performance, with its highest accuracy in the BCC class (90%).

Next, Figure 12 presents the accuracy and loss graphs for the combined VGG16 and MobileNetV2 model to evaluate its training progress.

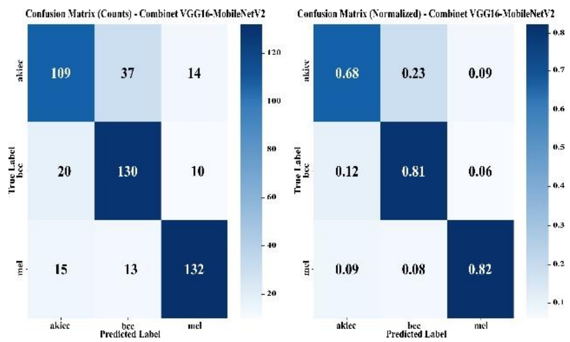
VOL. 10. NO. 4 MAY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v10i4.6541



Source: (Research Results, 2025)
Figure 12. Graph of Accuracy and Loss Model Combination of VGG16 With MobileNetV2

Figure 12 presents the accuracy and loss graphs for the combined VGG16 and MobileNetV2 model during training. In the accuracy graph (left), the model shows a gradual improvement, reaching 77% accuracy on the training data. However, the validation accuracy fluctuates sharply in the early epochs before stabilizing after epoch 30. In the loss graph (right), the training loss steadily decreases, indicating continuous learning. However, the validation loss exhibits extreme fluctuations, especially around epoch 30, before becoming more stable towards the end of training. This pattern suggests the presence of partial overfitting and highlights the model's difficulty in generalizing to unseen data.

Next, Figure 13 presents the confusion matrix for the combined VGG16 and MobileNetV2 model, illustrating its classification performance across the three skin cancer classes.

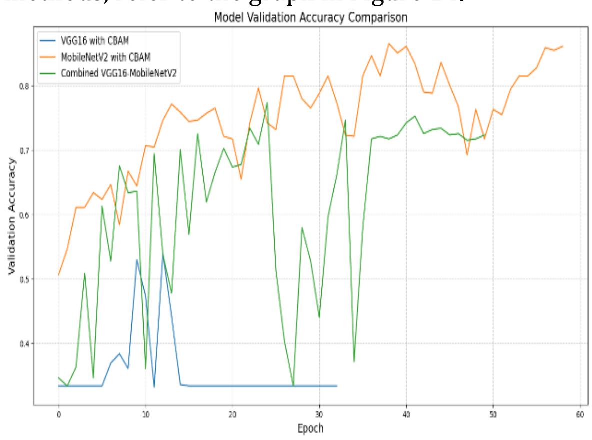


Source : (Research Results, 2025)
Figure 13. Confusion Matrix Combination of VGG16 With MobileNetV2

Figure 13 presents the confusion matrix of the combined VGG16 and MobileNetV2 model for three skin cancer classes AKIEC, BCC, and MEL. In the absolute confusion matrix (left), the model correctly classified 109 AKIEC samples but misclassified 37 samples as BCC and 14 samples as MEL. For the BCC class, the model correctly identified 130 samples, while 20 samples were misclassified as AKIEC and 10 as MEL. Meanwhile, the model successfully classified 132 melanoma samples correctly, though 15 samples were misclassified as AKIEC and 13 as BCC.

In the normalized confusion matrix (right), the AKIEC class achieved a classification accuracy of 68%, but it still exhibited an error rate of 23% misclassified as BCC and 9% as MEL. The BCC class attained an accuracy of 81%, with 12% misclassified as AKIEC and 6% as MEL. Similarly, the melanoma class achieved an accuracy of 82%, but 9% of the samples were misclassified as AKIEC and 8% as BCC. These results indicate that while the combined model performs relatively well, there are still misclassifications due to the similarity of visual features among different skin cancer classes.

For a clearer comparison of the three combination methods, refer to the graph in Figure 14.



Source: (Research Results, 2025)
Figure 14. Comparison Chart of the Three Combination Methods

Figure 14 presents a comparison of the validation accuracy of three combination methods: VGG16 with CBAM, MobileNetV2 with CBAM, and the combination of both (VGG16-MobileNetV2). The graph shows that MobileNetV2 with CBAM (orange line) exhibits a more stable and consistent accuracy improvement trend compared to the other methods, reaching a peak of nearly 87%. The combination of VGG16 and MobileNetV2 (green line) displays greater fluctuations, but overall, it follows an increasing trend and achieves approximately 77% accuracy. Meanwhile, VGG16 with CBAM (blue line) has the lowest performance, with validation accuracy stagnating at around 54% after approximately 10 epochs, indicating limitations in the learning process.

These results suggest that MobileNetV2 with CBAM has better generalization capability than the other methods, while the VGG16-MobileNetV2 combination still shows potential but requires improvements to reduce accuracy fluctuations. Compared with previous studies that only achieved an accuracy of 75%, the proposed method shows an accuracy improvement to 86%. This shows that the integration of CBAM with MobileNetV2 provides significant benefits in improving the model's attention to important features in skin cancer images. A more detailed analysis of model performance can be found in Table 4.

Table 4. Results Of Each Combination

Model	Accuracy	Precision	Recall	F1-score	Training Time
VGG16 + CBAM	53,96 %	61,73 %	53,96 %	51,13 %	23.81 minutes
MobileNetV2 + CBAM	86,46 %	86,53 %	86,46 %	86,46 %	39.79 minutes
VGG16 – MbileNet V2 Combination	77,29 %	77,51 %	77,29 %	77,24 %	51.75 minutes

Source: (Research Results, 2025)

CONCLUSION

Based on the research findings, it can be concluded that the performance of the combination models is influenced by the network architecture used, the number of layers, kernel size, augmentation techniques, and dataset balance. The models encountered difficulties in distinguishing certain classes, particularly between BCC and AKIEC, likely due to similarities in their visual features. Among the three combination methods tested, the MobileNetV2 with CBAM model demonstrated the best performance, achieving the highest validation accuracy of 86%. This model was able to produce more stable classification results compared to the other combinations. Meanwhile, the VGG16 with CBAM model struggled to reach high accuracy, achieving only 54%, with significant overfitting. The VGG16-MobileNetV2 combination achieved a relatively good accuracy of 77% but experienced fluctuations during training. These results indicate that selecting a lighter and more efficient network architecture, such as MobileNetV2, combined with an attention mechanism like CBAM, can improve skin cancer classification performance. For future research, it is recommended to explore alternative pooling techniques such as global pooling, further balance the dataset distribution, and apply fine-tuning to model parameters to enhance classification accuracy and stability.

REFERENCE

[1] K. Urban, S. Mehrmal, P. Uppal, R. L. Giesey, and G. R. Delost, "The global burden of skin cancer: A longitudinal analysis from the Global Burden of Disease Study, 1990–2017," *JAAD Int.*, vol. 2, pp. 98–108, 2021, doi: 10.1016/j.jdin.2020.10.013.

[2] W. Gouda, N. U. Sama, G. Al-Waakid, M. Humayun, and N. Z. Jhanjhi, "Detection of Skin Cancer Based on Skin Lesion Images Using Deep Learning," *healthcare*, 2022, doi: https://doi.org/10.3390/healthcare10071183.

[3] R. R. Saputro, J. Apri, and W. A. Saputra, "Journal of Dinda Klasifikasi Penyakit Kanker Kulit Menggunakan Metode Convolutional Neural Network (Studi Kasus: MELanoma)," *J. DINDA*, vol. 2, no. 1, pp. 52–57, 2022, doi: https://doi.org/10.20895/dinda.v2i1.349.

[4] D. Raval and J. N. Undavia, "A Comprehensive assessment of Convolutional Neural Networks for skin and oral cancer detection using medical images," *Healthc. Anal.*, vol. 3, no. March, p. 100199, 2023, doi: 10.1016/j.health.2023.100199.

[5] M. A. Arshed, S. Mumtaz, M. Ibrahim, S. Ahmed, M. Tahir, and M. Shafi, "Multi-Class Skin Cancer Classification Using Vision Transformer Networks and Convolutional Neural Network-Based Pre-Trained Models," *Inf.*, vol. 14, no. 7, 2023, doi: 10.3390/info14070415.

[6] H. M. Balaha and A. E. S. Hassan, "Skin cancer diagnosis based on deep transfer learning and sparrow search algorithm," *Neural Comput. Appl.*, vol. 35, no. 1, pp. 815–853, 2023, doi: 10.1007/s00521-022-07762-9.

[7] D. Yuliana, Fairuz Dyah Widowati, and Muammar Fawwaz, "Analisis Farmakoekonomi Terhadap Regimen Kemoterapi Pasien Kanker Payudara Di Rumah Sakit Universitas Hasanuddin Makassar," *Media Farm.*, vol. 20, no. 1, pp. 71–83, 2024, doi: 10.32382/mf.v20i1.491.

[8] Z. Li, K. C. Koban, T. L. Schenck, R. E. Giunta, Q. Li, and Y. Sun, "Artificial Intelligence in Dermatology Image Analysis: Current Developments and Future Trends," *J. Clin. Med.*, vol. 11, no. 22, 2022, doi: 10.3390/jcm11226826.

[9] L. Rey-Barroso, S. Peña-Gutiérrez, C. Yáñez, F. J. Burgos-Fernández, M. Vilaseca, and S. Royo, "Optical technologies for the improvement of skin cancer diagnosis: A

VOL. 10. NO. 4 MAY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v10i4.6541

review," *Sensors (Switzerland)*, vol. 21, no. 1, pp. 1–31, 2021, doi: 10.3390/s21010252.

[10] M. S. Ali, M. S. Miah, J. Haque, M. M. Rahman, and M. K. Islam, "An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models," *Mach. Learn. with Appl.*, vol. 5, no. April, p. 100036, 2021, doi: 10.1016/j.mlwa.2021.100036.

[11] D. P. Mawardi, M. Novita, and N. D. Saputro, "Deteksi Awal Klasifikasi Jenis Penyakit Kanker Kulit Dengan Algoritma Convolutional Neural Network (Cnn) Berbasis Mobile Apps," *Adopsi Teknol. dan Sist. Inf.*, vol. 3, no. 2, pp. 1–6, 2024, doi: https://doi.org/10.30872/atasi.v3i2.2305.

[12] K. Vayadande, A. A. Bhosle, R. G. Pawar, D. J. Joshi, P. A. Bailke, and O. Lohade, "Innovative approaches for skin disease identification in machine learning: A comprehensive study," *Oral Oncol. Reports*, vol. 10, no. April, p. 100365, 2024, doi: 10.1016/j.oor.2024.100365.

[13] M. S. Majib, M. D. M. Rahman, T. M. S. Sazzad, N. I. Khan, and S. K. Dey, "VGG-SCNet: A VGG Net-Based Deep Learning Framework for Brain Tumor Detection on MRI Images," *IEEE Access*, vol. 9, pp. 116942–116952, 2021, doi: 10.1109/ACCESS.2021.3105874.

[14] Z. Yang, "Classification of picture art style based on VGGNET," *J. Phys. Conf. Ser.*, vol. 1774, no. 1, 2021, doi: 10.1088/1742-6596/1774/1/012043.

[15] D. Bhatt *et al.*, "Cnn variants for computer vision: History, architecture, application, challenges and future scope," *Electron.*, vol. 10, no. 20, pp. 1–28, 2021, doi: 10.3390/electronics10202470.

[16] S. R. Shah *et al.*, "Comparing Inception V3, VGG 16, VGG 19, CNN, and ResNet 50: A Case Study on Early Detection of a Rice Disease," *Agronomy*, pp. 1–13, 2023, doi: https://doi.org/ 10.3390/agronomy13061633.

[17] S. Iqbal, A. N. Qureshi, J. Li, and T. Mahmood, *On the Analyses of Medical Images Using Traditional Machine Learning Techniques and Convolutional Neural Networks*, vol. 30, no. 5. Springer Netherlands, 2023. doi: 10.1007/s11831-023-09899-9.

[18] L. Gaur, U. Bhatia, N. Z. Jhanjhi, G. Muhammad, and M. Masud, "Medical image-based detection of COVID-19 using Deep Convolution Neural Networks," *Multimed. Syst.*, vol. 29, no. 3, pp. 1729–1738, 2023, doi: 10.1007/s00530-021-00794-6.

[19] F. A. Shidik, K. Musthofa, A. P. Kartiningtyas, and T. Agustin, "Analisis citra medis untuk identifikasi penyakit mata dengan teknologi convolutional neural networks," *Semin. Basional Amikom Surakarta*, no. November, pp. 68–80, 2024.

[20] H. Lokhande and S. R. Ganorkar, "Object detection in video surveillance using MobileNetV2 on resource-constrained low-power edge devices," *Bull. Electr. Eng. Informatics*, vol. 14, no. 1, pp. 357–365, 2025, doi: 10.11591/eei.v14i1.8131.

[21] J. Lee, J. Park, and Y. Lee, "Towards Efficient Cancer Detection on Mobile Devices," *IEEE Access*, vol. PP, p. 1, 2025, doi: 10.1109/ACCESS.2025.3543838.

[22] G. A. Pamungkas and A. G. Persada, "Implementasi Sistem Reverse Vending Machine (RVM) Berbasis IoT Menggunakan MobileNet SSD untuk Deteksi Objek dan Mekanisme Insentif Poin pada Platform Sampahmas," *J. Indones. Manaj. Inform. dan Komun.*, vol. 6, no. 1, pp. 414–426, 2025, doi: https://doi.org/10.35870/jimik.v6i1.1224.

[23] M. Zakariah and A. Alnuaim, "Recognizing human activities with the use of Convolutional Block Attention Module," *Egypt. Informatics J.*, vol. 27, no. August, p. 100536, 2024, doi: 10.1016/j.eij.2024.100536.

[24] N. Sengodan, "CBAM-EfficientNetV2 for Histopathology Image Classification using Transfer Learning and Dual Attention Mechanisms," *arxiv*, no. January, pp. 0–6, 2024, doi: https://doi.org/10.48550/arXiv.2410.22392 Focus to learn more.

[25] M. Yin, Z. Chen, and C. Zhang, "A CNN-Transformer Network Combining CBAM for Change Detection in High-Resolution Remote Sensing Images," *Remote Sens.*, vol. 15, no. 9, pp. 1–26, 2023, doi: 10.3390/rs15092406.

[26] H. Xin and K. Zhang, "Surface Defect Detection With Channel-Spatial Attention Modules and Bi-Directional Feature Pyramid," *IEEE Access*, vol. 11, no. July, pp. 88960–88970, 2023, doi: 10.1109/ACCESS.2023.3303897.

[27] W. Y. Hsu and Y. W. Cheng, "EEG-Channel-Temporal-Spectral-Attention Correlation for Motor Imagery EEG Classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 1659–1669, 2023, doi: 10.1109/TNSRE.2023.3255233.

[28] B. Chen and Z. Dang, "Fast PCB Defect

Detection Method Based on FasterNet Backbone Network and CBAM Attention Mechanism Integrated With Feature Fusion Module in Improved YOLOv7," *IEEE Access*, vol. 11, no. July, pp. 95092–95103, 2023, doi: 10.1109/ACCESS.2023.3311260.

[29] F. Dartiko, R. J. Pradana, R. E. Sari, W. Syahputra, and W. K. Oktoeberza, "Klasifikasi Kanker Kulit Berbasis CNN dengan Metode Hybrid Preprocessing," *Med. Tek. J. Tek. Elektromedik Indones.*, vol. 5, no. 2, pp. 124–132, 2024, doi: 10.18196/mt.v5i2.22675.

[30] Luqman Hakim, Z. Sari, and H. Handhajani, "Klasifikasi Citra Pigmen Kanker Kulit Menggunakan Convolutional Neural Network," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 2, pp. 379–385, 2021, doi: 10.29207/resti.v5i2.3001.