Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 3, Juni 2025, Page. 1389-1404
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4546

Leveraging Convolutional Block Attention Module (Cbam) For Enhanced Performance In Mobilenetv3-Based Skin Cancer Classification

Anas Rachmadi Priambodo[1], Chastine Fatichah*[2]

[1,2]Department of Informatics, Institut Teknologi Sepuluh Nopember, Indonesia

Email: [1]chastine@its.ac.id

Received: Mar 20, 2025; Revised: Jun 8, 2025; Accepted: Jun 14, 2025; Published: Jun 23, 2025

Abstract

As the incidence of skin cancer continues to rise globally, effective automated classification methods become crucial for early detection and timely intervention. Lightweight neural networks such as MobileNetV3 offer promising solutions due to their minimal parameters, making them suitable for environment with low resource. This study aims to develop an automated multiclass skin cancer classification system by enhancing MobileNetV3 with the Convolutional Block Attention Module (CBAM). The primary goal is to achieve high classification accuracy without significantly increasing computational demands. We employed Bayesian optimization to automatically fine-tune model parameters and applied targeted data augmentation techniques to address class imbalance. CBAM was integrated to highlight diagnostically relevant regions within images. The proposed method was evaluated using the ISIC 2024 SLICE-3D dataset, which includes over 400,000 dermatoscopic images categorized into benign, basal cell carcinoma, melanoma, and squamous cell carcinoma classes. Preprocessing involved standardized resizing, normalization, and extensive geometric and photometric augmentations. Results demonstrated that our method achieved an accuracy of 98.97%, precision of 98.99%, recall of 98.97%, and an F1-score of 98.98%, surpassing previous state-of-the-art models by 1.86–6.52%. Remarkably, this improvement was achieved with minimal additional parameters due to the effective integration of CBAM. These results represent an advancement in automated medical image analysis, particularly for low resource settings, by combining lightweight CNNs with attention mechanisms and systematic hyperparameter exploration.

Keywords: *Bayesian optimization, CBAM, ISIC 2024 SLICE-3D, MobileNetV3, Skin cancer classification.*

1. INTRODUCTION

Globally, skin malignancies account for more than 331,000 new diagnoses and nearly 59,000 fatalities each year [1], [2]. The primary types of skin cancer include Melanoma, Basal Cell Carcinoma (BCC), Squamous Cell Carcinoma (SCC), and non-malignant lesions. Early detection is vital, as it greatly enhances survival rates, particularly for aggressive forms of melanoma. This emphasizes the critical need for effective detection methods [3], [4].

Conventional diagnostic techniques rely on dermoscopy criteria (ABCD, 7-point), biopsies, and histopathology [5], [6], [7], [8]. Although these methods are clinically validated, they demonstrate significant limitations, such as observer inconsistencies, reliance on specialist expertise, and significant time investment [9], [10]. These identified limitations emphasize the critical need for the implementation of automated detection technologies to improve the precision and accessibility of diagnoses.

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4546

Contemporary developments in both machine learning and deep learning methodologies have demonstrated significant potential for advancing medical image analysis. Previous research Asfar et al. (2024) [11] has explored various traditional feature extraction techniques including HFF, HOG, GLCM, and LBP, achieving accuracy of up to 92% [11], that brings attention to the possibility of enhancement through the deployment of deep learning methodologies. The Convolutional Neural Networks (CNN) exhibit remarkable proficiency in extracting intricate information from dermoscopic images. This advanced capability significantly elevates the accuracy and reliability of diagnostic processes in dermatology [12]. However, conventional CNN architectures face significant limitations, including high computational requirements, prolonged training times, and a tendency to overfit on limited medical datasets [13], [14]. These problems mean that CNN architecture needs to be made that are both light and efficient so that they can keep diagnostic accuracy high while reducing the amount of reducing computational demands [15].

Prior research has explored a variety of CNN architectures for skin skin malignancy classification, showcasing differing levels of effectiveness. This highlights the potential for further advancements and refinements in these models to improve outcomes in skin cancer detection. Research employing MobileNetV1 achieved only 58% accuracy, while InceptionV3 reached 86.9% on the ISIC dataset [14]. Additionally, VGG-16 recorded an accuracy of 73.69%, which suggests a limited capacity for generalization across diverse presentations of skin cancer [16]. However, Conventional CNN models such as VGG-16 and InceptionV3 typically involve millions of parameters, leading to substantial computational demands and hindering real-time deployment in resource-limited scenarios.

Lightweight CNN architectures, including MobileNetV1, MobileNetV2, and MobileNetV3, have partially addressed this concern by significantly reducing computational requirements and parameter counts, thus facilitating faster performance on edge devices. Lightweight CNN architectures such as MobileNetV3, offer considerable advantages for mobile implementation by enabling detection in regions with limited medical infrastructure [17]. When combined with U-Net for segmentation tasks, MobileNetV3 attained a remarkable accuracy of 97.84% [18], [19] conducted a thorough comparative analysis that established MobileNetV3-Large as the preeminent lightweight model, with 75.2% accuracy on ImageNet 2012 with merely 219 million FLOPs and 5.40 million parameters. The implementation of transfer learning with MobileNetV3 on facial expression recognition has achieved an impressive accuracy of 95.8%, showcasing its significant potential, even with its lightweight architecture [20].

Despite these advantages, MobileNetV3 show limitations in capturing complex and nuanced features essential for accurate skin lesion assessment. This challenge largely stems from the lack of advanced spatial attention mechanisms, which leads to diminished performance when analyzing intricate lesion morphologies [21]. While attempts have been made to improve computational efficiency through latency reduction strategies, these enhancements can sometimes come at the expense of diagnostic accuracy, especially when evaluating complex disease patterns [22].

The Convolutional Block Attention Module (CBAM) offers a robust mechanism to address these operational shortcomings. CBAM employs two sequential attention modules—channel and spatial attention—that enable the network to adaptively highlight the most informative features while adding only a negligible computational overhead [22], [23]. Recent research utilizing CBAM within CNN architectures demonstrates significant potential for enhancing feature detection accuracy while maintaining computational efficiency [24].

Earlier studies have achieved accuracy levels reaching 97.11% on the ISIC-2019 dataset utilizing deep neural network architectures [25]. Nevertheless, research focusing on the incorporation of CBAM into lightweight MobileNetV3 models, accompanied by systematic Bayesian hyperparameter optimization, and specifically addressing class imbalance within large-scale datasets such as ISIC 2024

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4546

SLICE-3D, remains scarce. These unexplored aspects underscore the novelty and distinctive contribution of our study.

This research elevates the learning capabilities of advanced neural network architectures employed for skin cancer diagnostic by incorporating MobileNetV3 with the CBAM, facilitating a more accurate differentiation between malignant and benign skin lesions while preserving computational efficiency. Previous research Asfar et al. [11] has demonstrated that combining various feature extraction methods can improve classification accuracy, which supports our integrated approach of enhancing MobileNetV3 with attention mechanisms. This research leverages the comprehensive ISIC 2024 dataset, which presents a rich array of skin lesion examples along with detailed clinical annotations. A key opportunity for improvement in our methodology is the imbalanced distribution of examples encompassing both non-cancerous and cancerous lesions. To enhance the robustness of our findings, our study implementing advanced data augmentation techniques aimed at achieving a balanced representation of the different lesion types [26] The proposed research presents several contributions as follows:

- Our research aims to overcome the limitations of traditional CNN architecture by innovatively integrating MobileNetV3's lightweight framework with the dual attention mechanisms of CBAM. This architectural improvement greatly increases the network's capacity to collect complex and nuanced elements essential for precise dermatological image interpretation.
- We implement a comprehensive hyperparameter optimization strategy using Bayesian optimization techniques to maximize model performance. This methodical approach allows for systematic exploration of the parameter space, resulting in optimal configurations that balance accuracy and computational efficiency.
- To address the challenge of imbalanced class distribution in the ISIC 2024 dataset, our research proposes the implementation of targeted data augmentation methodologies designed to promote balanced representation across various lesion categories. This strategic approach aims to enhance the model's generalization capabilities, facilitating improved performance across a diverse spectrum of skin cancer presentations.

The organization of this paper is as follows: Section 2 provides a comprehensive review of methods, including the proposed method to our research in skin cancer classification, data augmentation and attention mechanism like CBAM. Section 3 delineates the result of our methodology, training history, and evaluation metrics employed to assess the proposed methodologies. Section 4 presents an analysis and discussion of experimental results. Finally, Section 5 offers conclusions regarding the research findings and identifies promising avenues for future investigation.

2. METHOD

The proposed methodology outlines a comprehensive framework for skin malignancy classification with deep learning approaches, leveraging the ISIC 2024 dataset. Figure 1 points out that our methodology adheres to a meticulous pipeline comprising many essential phases: dataset processing, model construction, classification, and evaluation. The preprocessing phase consists of essential data preparation steps, including dataset splitting, resizing, normalization, and augmentation. During the model development phase, our study employs Bayesian optimization for hyperparameter tuning, implements the MobileNetV3Small architecture, and enhances feature extraction through CBAM. This integration enables the model to focus specifically on relevant dermatological features while maintaining optimal computational efficiency, positioning it for effective clinical deployment. The final phases encompass the categorization of cutaneous malignancies images and thorough model evaluation using standard metrics, alongside the specialized ISIC 2024 partial Area Under Curve (pAUC) metric tailored specifically for dermatological applications.

Vol. 6, No. 3, Juni 2025, Page. 1389-1404
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4546

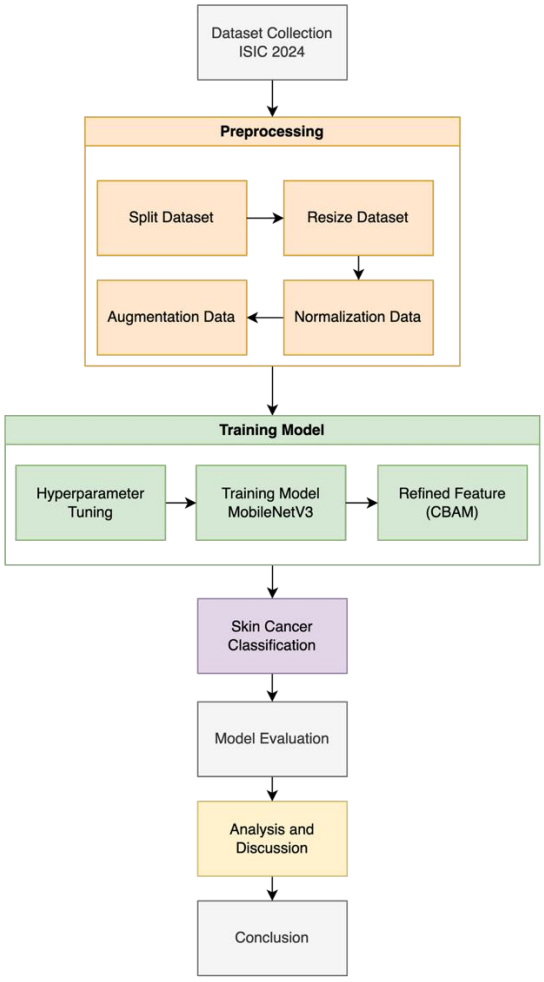

Figure 1. Proposed Research Methods

2.1. Dataset

The study employs the SLICE-3D (Skin Lesion Image Crops Extracted from 3D TBP) data from the International Skin Imaging Collaboration (ISIC) 2024 Competition. Each image isolates a single skin lesion for analysis. within a standardized 15mm-by-15mm field-of-view and has an average resolution of 133 pixels by 133 pixels. The dataset notably includes four diagnostic categories: malignant lesions (BCC, SCC, Melanoma) and benign lesions, providing a substantial volume of 400,945 samples, heavily skewed towards benign lesions (400,552 samples) across seven dermatologic centers globally as illustrated in Figure 2 [27].

We selected the SLICE-3D dataset due to its superior clinical and technical attributes compared to conventional repositories. Whereas most public skin-cancer collections comprise dermoscopic photographs with inherent selection biases and nonuniform acquisition protocols, SLICE-3D provides de-identified image tiles extracted from 3D total-body photography at smartphone-comparable resolution. Each 15 × 15 mm tile includes detailed metadata—patient age, sex, anatomical site, lesion diameter estimates, lighting modality, and anonymized identifiers—enabling reproducible analyses. Moreover, with over 400,000 samples across Benign, BCC, SCC, Melanoma classes, SLICE-3D's scale, diversity, and diagnostic granularity uniquely support the development of robust, generalizable automated classification models.

Vol. 6, No. 3, Juni 2025, Page. 1389-1404
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4546

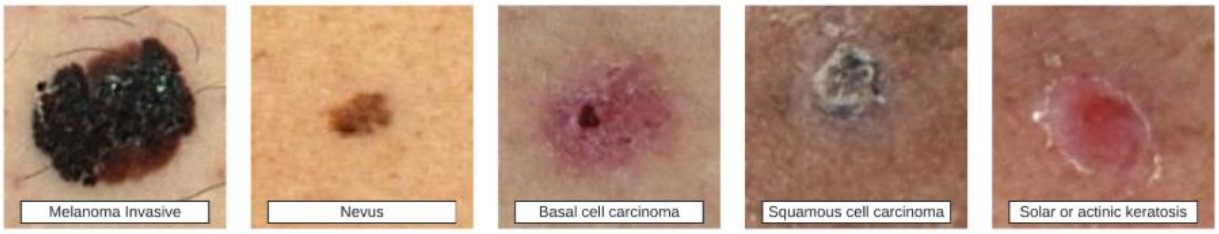

Figure 2. SLICE-3D ISIC 2024 Dataset

2.2. Preprocessing

The preprocessing phase consists of four interconnected steps that significantly enhance the dataset for deep learning applications. To start, our study performs dataset splitting in alignment with the directory organization defined in the path class. The preprocessing pipeline begins by stratifying the ISIC 2024 partitioned the dataset into 70% training, 15% validation, and 15% testing subsets, each stored in separate directories to eliminate data leakage and ensure unbiased evaluation. All images are then uniformly resized to 224 × 224 pixels to preserve lesion morphology and satisfy MobileNetV3 input specifications. Next, pixel intensities are normalized using ImageNet statistics to reduce internal covariate shifts and stabilize model convergence.

To enhance generalization and address class imbalance—particularly for melanoma, SCC, and BCC—we employ an on-the-fly, multistage augmentation strategy (Figure 3). First, geometric transformations (random horizontal/vertical flips, ± 15° rotations, ± 10 px translations) introduce spatial variability. Second, photometric adjustments (± 20 % brightness, ± 15 % contrast, ± 10 % saturation, CLAHE) simulate lighting variations [28]. Third, advanced augmentations (motion blur with 3–5 px kernels, elastic deformations $\alpha = 36$, $\sigma = 6$, coarse dropout of up to ten 16 × 16 px patches) increase robustness to image distortions. Finally, noise removal for addressing noise factors such as hair interference. This comprehensive approach yields a diversified training set while preserving memory efficiency.These efforts are aimed at effectively simulating the variability commonly encountered in real-world dermatoscopic imaging conditions, ultimately enriching the model's robustness and adaptability.

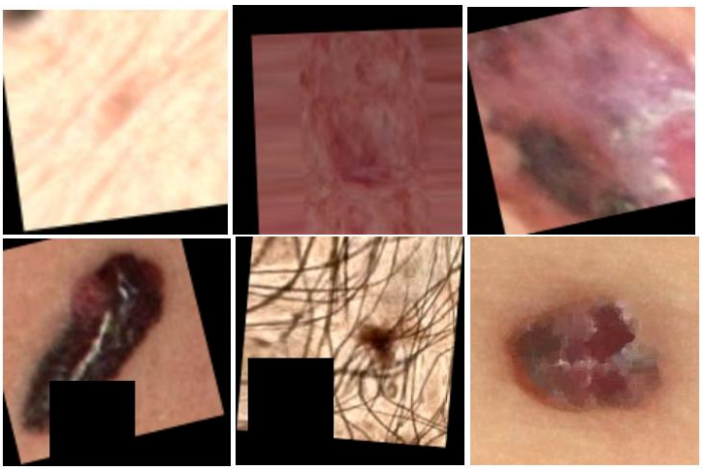

Figure 3. Dataset Augmentation

2.3. Hyperparameter Tuning

Vol. 6, No. 3, Juni 2025, Page. 1389-1404
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4546

We leveraged Bayesian optimization to systematically fine-tune five critical hyperparameters, striking an optimal balance between classification accuracy and computational efficiency. Our defined search space comprised: a learning rate (1×10^{-5} to 1×10^{-2}), dropout rate (0.2–0.5), L2 regularization strength (1×10^{-6} to 1×10^{-3}), the number of units in two successive dense layers (256–1 024 and 128–512, respectively), and the proportion of MobileNetV3Small base layers to freeze initially (20–40 %). A custom Bayesian Optimizer, employing an Expected Improvement acquisition function, iteratively proposed hyperparameter sets over five rounds, each selected to minimize validation loss. This process converged on an optimal configuration—learning rate = 2.26×10^{-4}, dropout = 0.35, L2 = 2.5×10^{-5}, dense units = [512, 256], freeze ratio = 0.3—reducing tuning time by approximately 60% compared to exhaustive grid search and yielding a validation accuracy of 99.49% over 71 epochs (see in Figure 6).

2.4. MobileNetV3

After conducting hyperparameter optimization, this study implements the MobileNetV3Small architecture with tailored adaptations for dermatological image analysis, as illustrated in Figure 4. We selected MobileNetV3Small as our backbone architecture due to its optimal balance between computational efficiency and classification performance, particularly when augmented for domain-specific tasks (Figure 4). Although MobileNetV3Large achieves higher top 1 accuracy on generic benchmarks (75.2 % vs. 67.4 % on ImageNet2012), its substantial parameter count (5.4 million) and FLOPs (219 million) limit its suitability for resource-constrained environments. In contrast, MobileNetV3Small comprises only 2.5 million parameters and requires 66 million FLOPs for 224 × 224 inference, yet—when combined with a Convolutional Block Attention Module (CBAM) it attains comparable or superior accuracy on specialized datasets [21][24]. Our implementation employs transfer learning from ImageNet weights, partial freezing guided by the optimized freeze ratio, and strategically placed batch-normalization layers to reduce internal covariate shift. Dual dense layers, whose dimensions were determined via Bayesian optimization, are regularized with dropout and L2 weight decay to mitigate overfitting. CBAM provides lightweight dual attention— The dual channel-and-spatial attention mechanisms enable the network to concentrate on regions of greatest diagnostic relevance without significant overhead, as demonstrated by prior gains of 2–4 % in medical imaging tasks. This tailored MobileNetV3Small framework thus delivers delivers superior accuracy with minimal computational overhead.

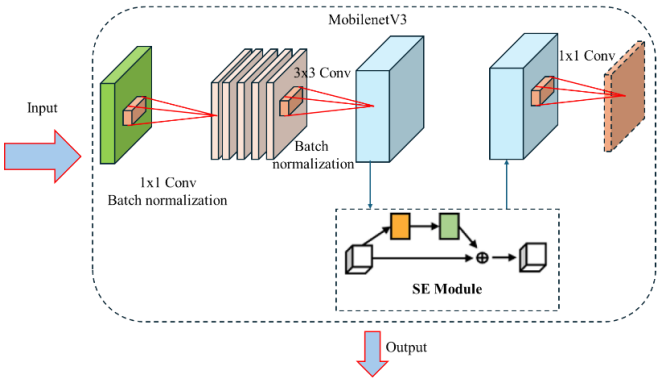

Figure 4. MobileNetV3 Architechture

2.5. Convolutional Block Attention Module (CBAM)

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4546

The CBAM is integrated into the MobileNetV3Small backbone to enhance feature representation through the successive application of channel attention followed by spatial attention modules. as shown in Figure 5. In the channel attention stage, global average and max pooling generate two descriptors per feature map, which are processed by a shared multilayer perceptron with a bottleneck of $\frac{C}{r}$ hidden units (where r=16), producing a channel-weighting vector that is broadcast across all spatial positions [29]. Subsequently, the spatial attention stage concatenates the mean and max projections of the refined feature maps into a two-channel descriptor, which is convolved with a 7×7 kernel to yield a spatial attention mask. This mask reweights each channel uniformly, emphasizing diagnostically salient regions. To preserve generalization, L2 regularization (λ = 1e−4) is applied to all CBAM convolutional layers. By focusing computation on critical channels and spatial area, this attention mechanism improves the model's ability to discern subtle lesion characteristics while incurring less than a 10 % increase in FLOPs.

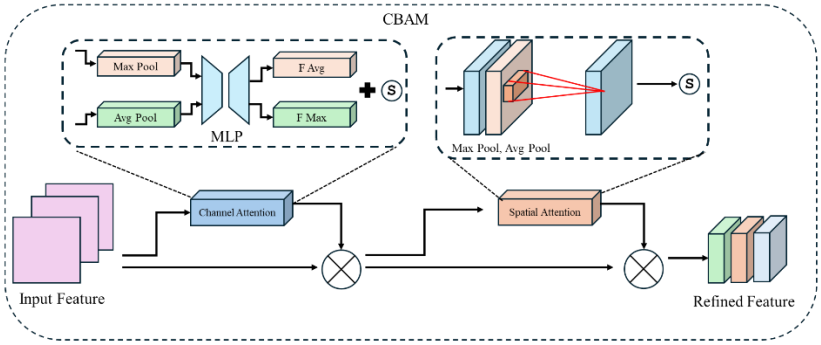

Figure 5. CBAM Module Architecture

2.6. Skin Cancer Classification

Dermatoscopic pictures are classified into specified diagnostic classes using the optimal MobileNetV3-CBAM model in the classification phase. The model employs a bespoke combined_loss function that balances sparse and categorical cross-entropy components with equal weights (0.5 each), hence producing more consistent gradients throughout training. Using an EarlyStopping callback with a patience of 8 epochs and a ReduceLROnPlateau mechanism that dynamically changes the learning rate when performance plateaus, with a factor of 0.2 and a minimum learning rate threshold of 1e-7, the training process runs when validation data no longer show improvement, this setup guarantees effective convergence and prevents overfitting by early termination.

Our MobileNetV3-CBAM design uses a deliberate series of layers after the components of feature extraction and attention improvement to provide the real classification mechanism. Following the MobileNetV3Small backbone and CBAM attention module, The refined feature maps are subsequently passed through a Global Average Pooling layer to lower spatial dimensions while maintaining channel information. This operation reduces each three-dimensional feature map to a one-dimensional vector, yielding a compact representation for classification feature vector that faithfully preserves the salient features of the input dermatoscopic image.

After that, the condensed feature representation moves through a classification head with two fully linked (dense) layers whose dimensions are calculated using Bayesian optimization usually between 256-1024 and 128-512 units respectively. Every dense layer combines dropout (with a rate optimized between 0.2 and 0.5) to minimize overfitting, ReLU activation to add non-linearity, and batch normalization to stabilize training. A completely linked layer with four units—corresponding to the four diagnostic classes: benign, basal cell carcinoma, melanoma, and squamous cell carcinoma—with

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4546

softmax activation generates a probability distribution over these classes at the final classification layer. Maintaining computational economy and optimizing discriminative power for skin lesion classification, this design lets the model progressively convert low-level visual characteristics into high-level diagnostic predictions.

2.7. Model Evaluation

The concluding stage of the evaluation entails an assessment classification Evaluation of performance utilising a range of metrics, such as accuracy, precision, recall, F1-score, loss, and the ISIC 2024 pAUC metric, which examines the partial Area Under the ROC Curve for high sensitivity regions (true positive rate > 80%). This metric concentrates on the most clinically significant segment of the ROC curve, prioritizing regions where sensitivity must remain high. In contrast to the conventional AUC, which evaluates the full ROC curve, the ISIC 2024 challenge adopts a partial AUC (pAUC) metric that is restricted to the segment where the true positive rate (TPR) meets or exceeds 80 %. By concentrating on this high-sensitivity interval, the metric ensures that classification models are optimized to minimize false negatives—a vital requirement in clinical settings where overlooking a malignant lesion can have serious implications. Formally, the pAUC is defined as the integral of the ROC curve function over the false positive rate (FPR) interval corresponding to TPR ≥ 0.80 [30]. The evaluation includes ROC curve visualizations for better assessment, and results are documented with timestamps in structured files for reproducibility and comparison, emphasizing the importance of high sensitivity in dermatological screening. The pAUC is computed as follows:

$$pAUC = \int_{0.8}^{1.0} ROC(t)\, dt \tag{1}$$

In this formulation, ROC(t) denotes the ROC curve evaluated at threshold t, and the integral is taken over the false positive rate interval corresponding to true positive rates of 80% or higher.

3. RESULT

Our proposed methodology integrating lightweight model MobileNetV3 with CBAM and Bayesian optimization for hyperparameter tuning performs well in the classification of skin cancer during implementation. Class-specific precision and recall measures confirmed the model's powerful discriminative ability across several skin lesion categories, resulting in 98.97% accuracy on the ISIC 2024 test dataset. Bayesian optimization found an optimum configuration after five rounds in the complicated hyperparameter space. Over 71 epochs, training and validation evaluates have improved, indicating sustained convergence with minimal overfitting. The learning rate adaption method-controlled optimization with strategic decreases at epochs 36, 51, 62, and 68, allowing fine-grained parameter modifications. BCC, Benign lesions, and SCC are identified well by confusion matrix analysis, while melanoma is identified somewhat less well but still clinically acceptable. ROC analysis shows near-perfect classification with AUC values of 1.00 for all classes, and the customized ISIC 2024 partial AUC (pAUC) statistic validates the model's excellent sensitivity in clinically relevant regions. This extensive study validates our architecture decisions and optimization strategy, laying the groundwork for clinical dermatological diagnostic assistance system deployment.

3.1. Training History

The training logs provide key insights into how this study optimized the model. During the Bayesian optimization phase, the process tested five different sets of hyperparameters. In the first iteration, the model achieved a strong performance, reaching 97.37% validation accuracy within just 10 epochs. This quick improvement shows that our initial hyperparameter choices were effective. In the

Vol. 6, No. 3, Juni 2025, Page. 1389-1404
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4546

final training phase, our method extended the process to 71 epochs. This process also used a strategy called ReduceLROnPlateau to lower the learning rate at specific points. We started with a learning rate of 0.00022628, which helped the model learn quickly at first. Then, gradually reduced the rate to 0.000045256, 0.0000090512, 0.0000018102, and finally to 0.00000036205. This method allows for finer adjustments to the model's parameters and helps avoid getting stuck in less optimal solutions. Throughout the training, the validation accuracy improved steadily. It began at 75.63% after the first epoch and reached 99.49% by the end. This shows that the model learned effectively without major fluctuations that could indicate problems.

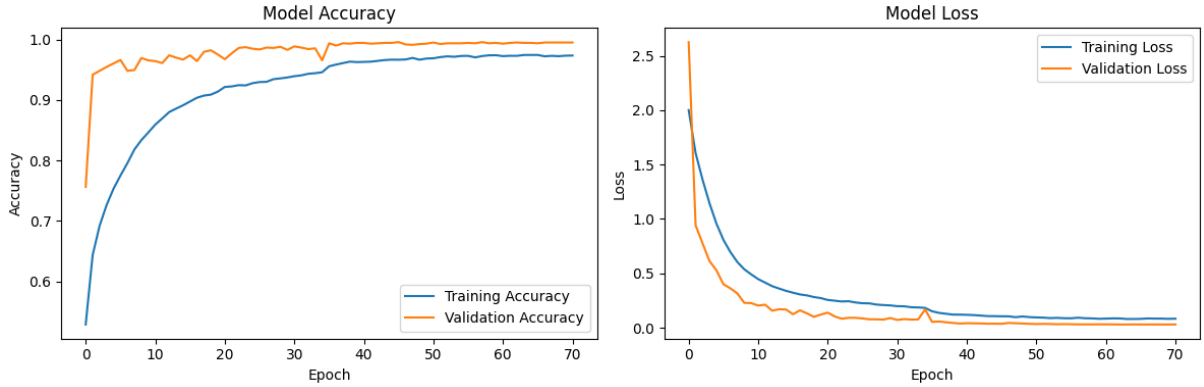

Figure 6. Training History

The training history in Figure 6 shows stable model convergence, with training accuracy rising from about 60% to 97% over 60 epochs. Notably, validation accuracy often exceeds training accuracy in the early epochs, indicating that augmentation strategies enhance generalization. Both training and validation loss decrease continuously, with training loss dropping from above 1.0 to near 0.1 and validation loss below 0.05. The consistent trends in both losses indicate effective regularization via dropout, L2 weight regularization, and batch normalization. Initial variations in validation metrics suggest exploration of the feature space, while steady convergence in later epochs (40-71) signifies the model's identification of key features for skin lesion classification. Overall, this profile reflects a well-regularized deep learning model suitable for complex medical imaging tasks.

3.2. Model Evaluation Results Analysis

Comprehensive evaluation metrics demonstrate our model's effectiveness in skin cancer classification. As detailed in Table 1, our study achieved an overall accuracy of 98.97%. Class-specific precision values are strong: BCC (85.72%), benign lesions (99.53%), melanoma (78.26%), and SCC (100%). The lower precision for melanoma and BCC highlights the challenges in diagnosing these conditions. Recall values show impressive sensitivity: BCC (96.00%), benign (99.40%), melanoma (75.17%), and SCC (100%). The F1-scores indicate balanced performance: BCC (90.56%), benign (99.46%), melanoma (76.59%), and SCC (100%). A macro-average F1-score of 91.65% reflects strong overall performance despite class imbalance. Additionally, weighted metrics further affirm clinical utility: Precision stands at 98.98%, recall at 98.97%, and F1-score at 98.97%.

In summary the overall accuracy (98.97%) and the training took 26.1 hours, allowing for rapid inference in clinical settings. Overall, these metrics support the clinical viability of our automated dermoscopic lesion categorization method.

Table 1. Performance result of MobileNetV3-CBAM

Category Class	Precision	Recall	F1-Score

BCC	85.72%	96%	90.56%
Benign	99.53%	99.40%	99.46%
Melanoma	78.26%	75.17%	76.59%
SCC	100%	100%	100%
Accuracy	98.97%	98.97%	98.97%
Macro Avg	90.88%	92.60%	91.66%
Weighted Avg	98.99%	98.97%	98.98%

3.3. Confusion Matrix Analysis

The confusion matrix in Figure 7 highlights the model's performance in classifying skin conditions. For basal cell carcinoma (BCC), it correctly identifies 24 out of 25 cases, yielding a sensitivity of 96%, which is crucial for early detection. In the largest group, benign lesions, the model classifies 1,488 out of 1,500 correctly (99.2% accuracy), minimizing misclassifications primarily as BCC and melanoma. Melanoma classification, however, shows more difficulty, with 19 out of 24 cases identified, resulting in a sensitivity of 79.17%. For squamous cell carcinoma (SCC), the model excels with perfect accuracy, identifying all 11 cases. Overall, the model demonstrates strong performance, particularly in malignant skin cancers, demonstrating the efficacy of the CBAM attention system in discerning critical aspects.

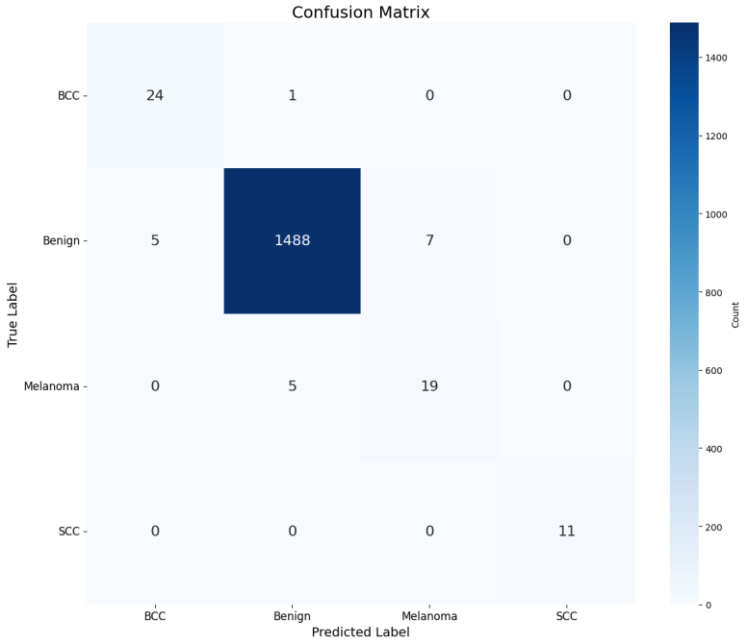

Figure 7. Confusion Matrix

3.4. ROC Curve Analysis

The examination of the ROC curve presented in Figure 8 illustrates excellent classification performance across all diagnostic categories, each achieving an AUC value of 1.00. This signifies that the model can flawlessly distinguish between positive and negative cases. The curves closely align with the top-left corner of the plot, demonstrating elevated true positive rates and minimal false positive rates, even in the presence of class imbalance within the dataset. Additionally, the ISIC 2024 pAUC metric visualization indicates a pAUC of 0.2000 in Figure 9, demonstrating high sensitivity while maintaining specificity—an essential aspect for skin cancer screenings, where failing to identify malignant cases

Vol. 6, No. 3, Juni 2025, Page. 1389-1404
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4546

poses significant risks, whereas misclassifying benign lesions can be effectively managed with follow-up assessments.

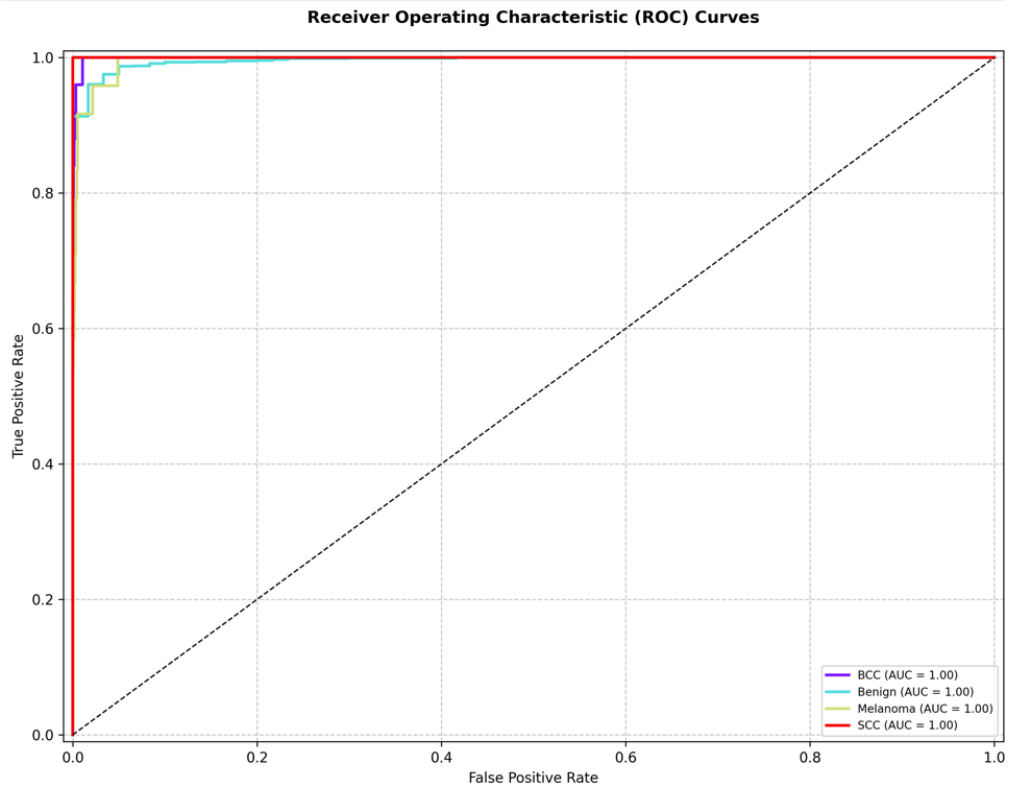

Figure 8. ROC Curve Plot

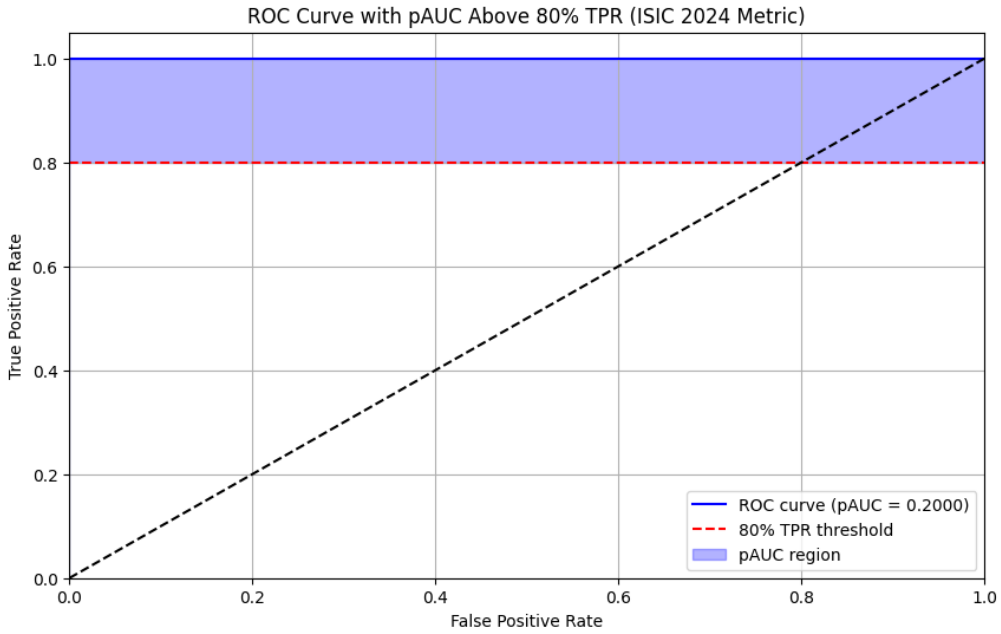

Figure 9 ROC Curve with pAUC

4. DISCUSSIONS

The proposed study MobileNetV3 with CBAM architecture augmented by Bayesian hyperparameter tuning reveals higher performance metrics compared to prior research techniques for

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4546

skin cancer classification. The comparative analysis reveals several significant findings that highlight the contributions of our methodology as stated in Table 2.

Table 2 Comparative analysis of the performance between the proposed model and prior research

References	Method	Dataset	Year	Accuracy↑	Precision↑	Recall↑	F1-Score↑
[31]	MobileNetV1	HAM10000	2023	92.45%	91.25%	89.18%	88.34%
[32]	Vision Transformer (VIT)	HAM10000	2022	94.30%	94.10%	-	94.10%
[33]	MobileNetV2	HAM10000	2023	94.47%	92.36%	90.78%	89.36%
[34]	Hybrid DenseNet + ResNet	HAM10000	2023	95 %	-	-	-
[35]	VGG19 + SVM	ISIC-2019	2023	96%	92%	92%	92%
[36]	InceptionV3	ISIC-2019	2023	96.40%	96.40%	96.38%	96.29%
[37]	CNN and VGG16	ISIC-2019	2022	96.91%	92.19%	92.18%	92.18%
[25]	DCNN	ISIC-2019	2023	97.11%	97.09%	97.12%	97.08%
Ours	MobilenetV3+CBAM	ISIC-2024	2024	**98.97%**	**98.99%**	**98.97%**	**98.98%**

The results show that the proposed model demonstrates outstanding performance, attaining 98.97% accuracy, 98.99% precision, 98.97% recall, and 98.98% F1-score on the ISIC2024 dataset. This is a significant improvement over prior leading models, including the DCNN model by Houssein et al. [25], which reached 97.11% accuracy on the ISIC-2019 dataset. The difference is even greater than other models, like InceptionV3 (96.40% accuracy) and CNNVGG16 (96.91% accuracy) on ISIC-2019.

Several reasons contribute to this improvement. First, employing CBAM with MobileNetV3 creates a strong architecture that effectively captures important details in dermatoscopic images. This helps the model focus on key areas within lesions while highlighting the most relevant features. Second, using Bayesian optimization for adjusting hyperparameters is a significant step forward in Compared to conventional approaches such as grid or random search, this strategy enables more efficient traversal of the hyperparameter space with a significantly reduced number of evaluations.

Furthermore, models evaluated on the HAM10000 dataset show low performance compared to those assessed on ISIC datasets. For example, the Hybrid DenseNet and residual network by Jasil et al. [34] achieved 95% accuracy, and MobileNetV2 by Ogundokun et al. [33] reached 94.47%. This difference could be due to variations in dataset characteristics, such as differences in lesions, image quality, and class distribution. Our model's strong performance on the ISIC2024 dataset shows it can handle modern dermatoscopic images with varied presentations.

When we compare similar architectural approaches, our MobileNetV3+CBAM setup significantly outperforms the original MobileNetV1 (92.45% accuracy) and MobileNetV2 (94.47% accuracy) tested on HAM10000. This improvement of 4.50-6.52% points in accuracy illustrates the benefits of our architectural upgrades, especially the attention mechanism and advanced regularization methods. The performance gains are clear across all parameters, including precision, recall, and F1-score, showing improvements of 6.62% - 7.73%, 8.19% - 9.79%, and 9.61% - 10.63% points, respectively.

Compared to transformer-based models like the Vision Transformer Network (VIT) by Xin et al. [32], which achieved 94.30% accuracy, our approach further confirms its advantages. While transformer architectures are increasingly popular in computer vision, our optimized convolutional neural network outperforms them while being computationally efficient. This shows that well-optimized CNNs with suitable attention mechanisms can effectively tackle medical image analysis tasks, especially where

deployment limits matter. The strong performance of our model has important implications for clinical use and future research.

First, the high recall of 98.97% across all classes indicates the model rarely misses malignant lesions, essential for the prompt identification and treatment of skin lesion. This is crucial for screening applications where missing malignancy is more serious than having false positives. The balanced results across different lesion types, shown by the confusion matrix analysis, suggest that the model effectively identifies key features for each diagnostic category despite challenges with class imbalance in dermatological datasets.

Second, combining CBAM with MobileNetV3 offers an accurate and less complex solution than larger models like VGG or DenseNet. This efficiency is essential for potential use in environments with limited resources or for mobile applications that could bring dermatological expertise to underserved areas. The lightweight design of MobileNetV3, paired with the focused feature enhancement from CBAM, strikes a good balance between performance and resource needs.

Third, our use of Bayesian optimization for hyperparameter tuning shows a systematic approach to model development that could apply to medical image analysis more broadly. Finding the best hyperparameters with fewer evaluations than older methods is critical when working with complex deep-learning models and limited computing power.

Melanoma classification underperformed relative to other categories, with precision of 78.26 % and recall of 75.17 %. This stems from two main issues: the melanoma subset's limited size, which restricts model exposure to lesion diversity, and the high visual heterogeneity of melanoma (variations in shape, color, and border irregularity). To improve performance, we recommend: (1) applying focal loss to emphasize underrepresented melanoma samples during training; (2) generating additional melanoma instances via intelligent oversampling methods such as SMOTE or generative adversarial networks; and (3) incorporating domain-specific augmentations pigment-based color shifts and micro-geometric distortions to better model melanoma's varied presentations. Integrating CBAM into MobileNetV3Small, alongside Bayesian hyperparameter optimization, establishes a systematic framework for efficient neural architecture search in medical imaging. This approach empowers computer vision researchers to develop edge-compatible models featuring adaptive attention mechanisms, ensuring optimal performance on resource-limited devices. It also validates automated hyperparameter tuning across extensive datasets.

5. CONCLUSION

This study introduces a new way to automatically classify skin cancer using the MobileNetV3 model. We improved this model with a CBAM and fine-tuned it using Bayesian hyperparameter tuning. Our tests on the ISIC 2024 dataset show outstanding results, achieving accuracy by 98.97%, precision by 98.99%, recall by 98.97%, and F1-score by 98.98%. This performance surpasses other leading methods for analyzing skin images.

CBAM enhancement really improves the model's capacity to identify key features by using both attention, spatial and channel. This helps the network highlight important areas and features that are important for diagnosis. This approach is especially useful for difficult cases like melanoma, where subtle differences are crucial for accurate identification. This research employs MobileNetV3 for its good consistency between productivity and capabilities, rendering the approach amenable to deployment in resource-constrained clinical settings.

This framework used Bayesian optimization for hyperparameter tuning, which is a better method than traditional ones. It allows for thorough exploration of complex parameter options with fewer tries, leading to better performance. This framework identifies the best settings for learning rate, regularization strength, network structure, and training methods, showing its effectiveness for creating

deep learning models in medical imaging. We implement targeted data augmentation to mitigate class imbalance, achieving balanced sensitivity in clinically critical high–true-positive–rate regions (pAUC = 0.2000). The optimal pAUC value in high-sensitivity regions verifies the model's efficacy for clinical screening, particularly in conditions where minimizing false negatives is crucial. The model revealed consistent performance across all diagnostic categories, despite the limitations presented by unbalanced dermatological data.

Our model demonstrates substantial enhancements compared to previous models using the ISIC and HAM10000 datasets. Our study results achieve enhancements of 1.86 to 6.52% points in accuracy relative to the optimal findings documented in the literature. These results prove that well-optimized convolutional neural networks with attention mechanisms can be very effective for medical image analysis, challenging the idea that transformer-based models are always the best choice for advanced computer vision tasks.

Future studies should focus on evaluate advanced loss functions such as focal loss and oversampling techniques based on generative adversarial networks to enhance melanoma detection. We will also validate our model against prospective clinical datasets acquired via smartphone-compatible dermatoscopes to confirm real-world utility.

In conclusion, our MobileNetV3+CBAM model with Bayesian hyperparameter optimization makes an important contribution to automated skin cancer classification. It offers a robust and efficient solution that could improve dermatology practices and enhance patient prognosis by facilitating earlier and more precise detection of malignant skin lesions.

REFERENCES

[1] Global Cancer Observatory, “Melanoma of Skin Fact Sheet,” 2022, *World Health Organization*. [Online]. Available: https://gco.iarc.who.int/media/globocan/factsheets/cancers/16-melanoma-of-skin-fact-sheet.pdf

[2] G. C. Observatory, “Melanoma of Skin Fact Sheet,” 2022, *World Health Organization*. [Online]. Available: https://gco.iarc.who.int/media/globocan/factsheets/cancers/16-melanoma-of-skin-fact-sheet.pdf

[3] N. Hasan *et al.*, “Skin cancer: understanding the journey of transformation from conventional to advanced treatment approaches,” *Mol Cancer*, vol. 22, no. 1, p. 168, Mar. 2023, doi: 10.1186/s12943-023-01854-3.

[4] B. Arouch and M. Elboim-Gabyzon, “The role of physical therapists in the early detection of skin melanoma: insights from an anonymous survey,” *Front Med (Lausanne)*, vol. 11, Mar. 2024, doi: 10.3389/fmed.2024.1436206.

[5] S. Kam and J. Zic, “42855 To ABCDE or not to ABCDE?: a Randomized Educational Intervention Study to Improve Patient Recognition of Atypical Nevi,” *J Am Acad Dermatol*, vol. 89, no. 3, p. AB98, Mar. 2023, doi: 10.1016/j.jaad.2023.07.394.

[6] M. A. Khan, T. Akram, Y.-D. Zhang, and M. Sharif, “Attributes based skin lesion detection and recognition: a mask RCNN and transfer learning-based deep learning framework,” *Pattern Recognit Lett*, vol. 143, pp. 58–66, 2021.

[7] P. P. Naik, “Cutaneous Malignant Melanoma: A Review of Early Diagnosis and Management,” *World J Oncol*, vol. 12, no. 1, pp. 7–19, 2021, doi: 10.14740/wjon1349.

[8] T. Tran *et al.*, “Expert Consensus Statement on Proficiency Standards for Dermoscopy Education in Primary Care,” *The Journal of the American Board of Family Medicine*, vol. 36, no. 1, pp. 25–38, Mar. 2023, doi: 10.3122/jabfm.2022.220143R1.

[9] A. Ennaji, E. H. Khoukhi, M. A. Sabri, and A. Aarab, “Malignant melanoma detection using multi-scale image decomposition and a new ensemble-learning scheme,” *Multimed Tools Appl*, vol. 83, no. 7, pp. 21213–21228, 2023.

[10] D. J. Derwin, O. J. Singh, B. P. Shan, K. U. Maheswari, and D. Lavanya, “An efficient multi-level pre-processing algorithm for the enhancement of dermoscopy images in Melanoma detection,” *Med Biol Eng Comput*, vol. 61, no. 11, pp. 2921–2938, 2023.

[11] A. Muh. F. and H. M. and D. H. Asfar, "COMBINATIONS OF FEATURE EXTRACTIONS AND MACHINE LEARNING ALGORITHMS FOR SKIN CANCER CLASSIFICATION," *Jurnal Teknik Informatika (JUTIF)*, vol. 5, pp. 1591–1598, Dec. 2024.

[12] N. Faujdar, R. Agrawal, and A. Agarwal, "Critical analysis of various supervised machine learning algorithms for detecting diabetic retinopathy in images," in *Artificial Intelligence and Machine Learning Techniques in Image Processing and Computer Vision*, Apple Academic Press, 2024, pp. 75–93.

[13] R. Kulkarni, S. Giri, S. Sanghvi, and R. Keskar, "Classification of Malignant Skin Cancer Lesion Using CNN, KNN, and SVM," 2023, pp. 731–742. doi: 10.1007/978-981-19-6631-6_50.

[14] A. Yilmaz, M. Kalebasi, Y. Samoylenko, M. E. Guvenilir, and H. Uvet, "Benchmarking of Lightweight Deep Learning Architectures for Skin Cancer Classification using ISIC 2017 Dataset," Mar. 2021.

[15] R. Agrawal, N. Gupta, and A. S. Jalal, "CACBL-Net: a lightweight skin cancer detection system for portable diagnostic devices using deep learning based channel attention and adaptive class balanced focal loss function," *Multimed Tools Appl*, Mar. 2024, doi: 10.1007/s11042-024-19485-1.

[16] A. Pandey, M. S. Teja, P. Sahare, V. Kamble, M. Parate, and M. F. Hashmi, "Skin cancer classification using non-local means denoising and sparse dictionary learning based CNN," *Journal of Electrical Systems and Information Technology*, vol. 11, no. 1, p. 36, Mar. 2024, doi: 10.1186/s43067-024-00162-0.

[17] K. DeVoe, G. Takahashi, E. Tarshizi, and A. Sacker, "Evaluation of the precision and accuracy in the classification of breast histopathology images using the MobileNetV3 model," *J Pathol Inform*, vol. 15, p. 100377, Dec. 2024, doi: 10.1016/j.jpi.2024.100377.

[18] U. K. Lilhore *et al.*, "A precise model for skin cancer diagnosis using hybrid U-Net and improved MobileNet-V3 with hyperparameters optimization," *Sci Rep*, vol. 14, no. 1, p. 4299, Mar. 2024, doi: 10.1038/s41598-024-54212-8.

[19] Y. Liu, P. Xiao, J. Fang, and D. Zhang, "A Survey on Image Classification of Lightweight Convolutional Neural Network," in *2023 19th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, 2023, pp. 1–10. doi: 10.1109/ICNC-FSKD59587.2023.10281072.

[20] B. Jiang *et al.*, "Research on facial expression recognition algorithm based on improved MobileNetV3," *EURASIP J Image Video Process*, vol. 2024, no. 1, p. 22, Mar. 2024, doi: 10.1186/s13640-024-00638-z.

[21] L. Jia *et al.*, "MobileNetV3 With CBAM for Bamboo Stick Counting," *IEEE Access*, vol. 10, pp. 53963–53971, 2022, doi: 10.1109/ACCESS.2022.3175818.

[22] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional Block Attention Module," 2018, pp. 3–19. doi: 10.1007/978-3-030-01234-2_1.

[23] L. Chen, H. Yao, J. Fu, and C. T. Ng, "The classification and localization of crack using lightweight convolutional neural network with CBAM," *Eng Struct*, vol. 275, p. 115291, Mar. 2023, doi: 10.1016/j.engstruct.2022.115291.

[24] P. Bhuyan, P. K. Singh, and S. K. Das, "Res4net-CBAM: a deep cnn with convolution block attention module for tea leaf disease diagnosis," *Multimed Tools Appl*, vol. 83, no. 16, pp. 48925–48947, Mar. 2023, doi: 10.1007/s11042-023-17472-6.

[25] E. H. Houssein, D. A. Abdelkareem, G. Hu, M. A. Hameed, I. A. Ibrahim, and M. Younan, "An effective multiclass skin cancer classification approach based on deep convolutional neural network," *Cluster Comput*, Jun. 2024, doi: 10.1007/s10586-024-04540-1.

[26] D. Wen *et al.*, "Characteristics of publicly available skin cancer image datasets: a systematic review," *Lancet Digit Health*, vol. 4, no. 1, pp. e64–e74, Mar. 2022, doi: 10.1016/S2589-7500(21)00252-1.

[27] N. R. Kurtansky *et al.*, "The SLICE-3D dataset: 400,000 skin lesion image crops extracted from 3D TBP for skin cancer detection," *Sci Data*, vol. 11, no. 1, p. 884, Aug. 2024, doi: 10.1038/s41597-024-03743-w.

[28] U. Kumar Lilhore *et al.*, "A precise model for skin cancer diagnosis using hybrid U-Net and improved MobileNet-V3 with hyperparameters optimization," *Sci Rep*, vol. 14, no. 1, p. 4299, Feb. 2024, doi: 10.1038/s41598-024-54212-8.

[29] L. Jia *et al.*, "MobileNetV3 With CBAM for Bamboo Stick Counting," *IEEE Access*, vol. 10, pp. 53963–53971, 2022, doi: 10.1109/ACCESS.2022.3175818.

[30] International Skin Imaging Collaboration (ISIC), "ISIC 2024 - Skin Cancer Detection with 3D-TBP."

[31] W. N. Ismail and H. A. Alsalamah, "Efficient Harris Hawk Optimization (HHO)-Based Framework for Accurate Skin Cancer Prediction," *Mathematics*, vol. 11, no. 16, p. 3601, Aug. 2023, doi: 10.3390/math11163601.

[32] C. Xin *et al.*, "An improved transformer network for skin cancer classification," *Comput Biol Med*, vol. 149, p. 105939, Oct. 2022, doi: 10.1016/j.compbiomed.2022.105939.

[33] R. O. Ogundokun *et al.*, "Enhancing Skin Cancer Detection and Classification in Dermoscopic Images through Concatenated MobileNetV2 and Xception Models," *Bioengineering*, vol. 10, no. 8, p. 979, Aug. 2023, doi: 10.3390/bioengineering10080979.

[34] S. P. G. Jasil and V. Ulagamuthalvi, "A hybrid CNN architecture for skin lesion classification using deep learning," *Soft comput*, Mar. 2023, doi: 10.1007/s00500-023-08035-w.

[35] M. Saeed, A. Naseer, H. Masood, S. U. Rehman, and V. Gruhn, "The Power of Generative AI to Augment for Enhanced Skin Cancer Classification: A Deep Learning Approach," *IEEE Access*, vol. 11, pp. 130330–130344, 2023, doi: 10.1109/ACCESS.2023.3332628.

[36] E. S. Nugroho, I. Ardiyanto, and H. A. Nugroho, "Boosting the performance of pretrained CNN architecture on dermoscopic pigmented skin lesion classification," *Skin Research and Technology*, vol. 29, no. 11, Nov. 2023, doi: 10.1111/srt.13505.

[37] A. Naeem, T. Anees, M. Fiza, R. A. Naqvi, and S.-W. Lee, "SCDNet: A Deep Learning-Based Framework for the Multiclassification of Skin Cancer Using Dermoscopy Images," *Sensors*, vol. 22, no. 15, p. 5652, Jul. 2022, doi: 10.3390/s22155652.