December 2025 | Volume 1 | Issue 3 | Article No. 25213

Journal of Smart Sensors and Computing

*Research Article* | *Open Access* |

# Multi-Class Skin Lesion Classification Using Transfer Learning with EfficientNet-B3 and Convolutional Block Attention Module

Sneha Ramdas Shegar and Supriya S. Patil*

*Department of Computer Engineering, Samarth College of Engineering & Management, Pune, Maharashtra, 412410, India*
**Email:* profsupriyapatil@gmail.com (S. S. Patil)

## Abstract

Skin diseases represent a significant global health challenge; however, precise automated detection of cutaneous lesions remains difficult due to high intra-class variability, inter-class similarity, and severe class imbalance across disease categories. This paper presents a multi-class skin lesion classification framework based on transfer learning, which integrates an EfficientNet-B3 backbone with a Convolutional Block Attention Module (CBAM) to enhance the learning of discriminative features. EfficientNet-B3, pre-trained on large-scale natural image datasets, serves as a powerful feature extractor, while CBAM improves feature representation by adaptively emphasizing informative channels and spatial locations. This enables the network to focus on diagnostically relevant lesion regions while suppressing background artifacts. The proposed model is trained and evaluated on the DermNet-23 dataset, comprising 23 clinically significant skin disease classes. To address the challenges of multi-class classification and class imbalance, performance is assessed using standard metrics including accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC). Experimental results demonstrate that the EfficientNet-B3 + CBAM model achieves 87.1% accuracy, 85.6% macro-F1 score, and 0.94 AUC, outperforming baseline CNN, ResNet50, MobileNetV3, and standard EfficientNet-B3 models. These results highlight the effectiveness of attention-guided transfer learning for developing robust and scalable computer-aided diagnostic systems for skin lesion classification.

***Keywords*:** Skin lesion classification; EfficientNet-B3; CBAM; Transfer learning; Computer-aided diagnosis.

Received: 20 November 2025; Revised: 29 December 2025; Accepted: 29 December 2025; Published Online: 30 December 2025.

## 1. Introduction

Melanoma and other types of skin cancer is an emerging public health problem with high rates of morbidity, mortality, and healthcare expenditure across the globe.[1] The World Health Organization reports that the number of non-melanoma skin cancer cases and over 300000 cases of melanoma worldwide annually is increasing the amount of concern regarding the public health implication of skin diseases.[2] Melanoma has a great likelihood of metastasis as well as the largest percentage of deaths associated with skin cancer even though it makes up a smaller portion of all cases.[3] Small melanoma can be easily removed with surgery so that in its early stages, it is curable, but once diagnosed late, surgery will not help much especially in terms of survival, and it makes the treatment more complicated. The non-melanoma skin cancer such as the basal cell carcinoma, squamous cell carcinoma among others, also play a role in contributing to high incidence of skin cancer in the world and create a cumulative burden on dermatology.[4] In this regard, accurate and prompt detection of cutaneous lesions is

DOI: https://doi.org/10.64189/ssc.25213

essential in enhancing patient outcomes, resource allocation, and large-scale screening initiatives, particularly in areas with a shortage of experienced dermatologists.

### 1.1 Clinical background and dermoscopy

Dermoscopy is an imaging modality which is not invasive and which enlarges and improves visualization of the subcutaneous skin structures in order to allow a more detailed evaluation of pigment patterns, vascular organization, and lesion boundaries.[5] Dermoscopy has a significant higher diagnostic sensitivity and specificity of melanoma and other pigmented lesions in the hands of experts than the unaided eye view.[6] Nonetheless, dermoscopy interpretation is very operator-specific, and it involves a lot of training and experience. Even trained dermatologists show inter-observer variability because of minor and overlapping morphological patterns in benign and malignant lesions. Furthermore, differences in imaging equipment, conditions of acquisition and the type of skin of the patient are other complicating factors which make standardized visual assessment impossible.[7] Fig. 1 illustrate Clinical background and dermoscopy.

Clinicians in primary care and in a resource-constrained setting might not receive the benefit of such higher levels of dermoscopic training, resulting in either under-referral of suspicious lesions or over-referral of benign lesions, which has both clinical and economic implications. The growing rate of skin lesion images that have been obtained through dermatoscopes and consumer-grade cameras increase the necessity of computation support that is scaled. It, therefore, follows that there is an enthusiastic impetus to build automated dermoscopic analysis frameworks capable of estimating and/or adding to masterful performance and help to standardize and equalize diagnostic procedures.

### 1.2 Challenges in manual dermoscopic assessment

Manual dermoscopic analysis is objective and it is subject to cognitive bias including anchoring, fatigue and heuristic reliance. The difference between early melanoma and benign nevi or between inflammatory dermatoses and infectious or neoplastic lesions is often based on subtle textural, chromatic and structural effects that are not always easily identified. There is also high intra-class variability(e.g., different appearances of melanoma in different sites on the body and in different skin tones) and inter-class similarity (e.g., benign lesions that resemble malignancy) that further contribute to obstacles to proper visual diagnosis.[8]

Additionally, non-dermoscopic images do not invariably correspond to the diagnostic criteria that are based on dermoscopy e.g. pattern analysis, algorithmic scoring systems, or the ABCD rule, and implementing these diagnostic criteria as a system in a high-volume clinical setting can be challenging. With the increase in image repositories, manual inspection and triage are no longer feasible, and computer-aided diagnosis (CAD) systems that can process large-scale image streams and maintain the same level of performance are sought. All these problems highlight the importance of having powerful, data-driven procedures capable of training discriminative patterns to go beyond handcrafted specifications.[9]

### 1.3 Limitations of existing machine learning approaches

Original computational techniques to analyse skin lesions used conventional machine learning pipelines of hand-engineered feature-extractors (e.g., colour histograms, texture descriptors, border irregularity measures) and classifier (e.g. support vector machines, k-nearest neighbours or random forests).[10-14] Although these methods have given a first-time understanding of whether automated lesion recognition is feasible, their performance was inherently limited by the expressiveness of manually specified features.[15] Hand-crafted descriptors typically do not encode high-order and complicated interaction between local patterns, and are vulnerable to changes in illumination, scale,

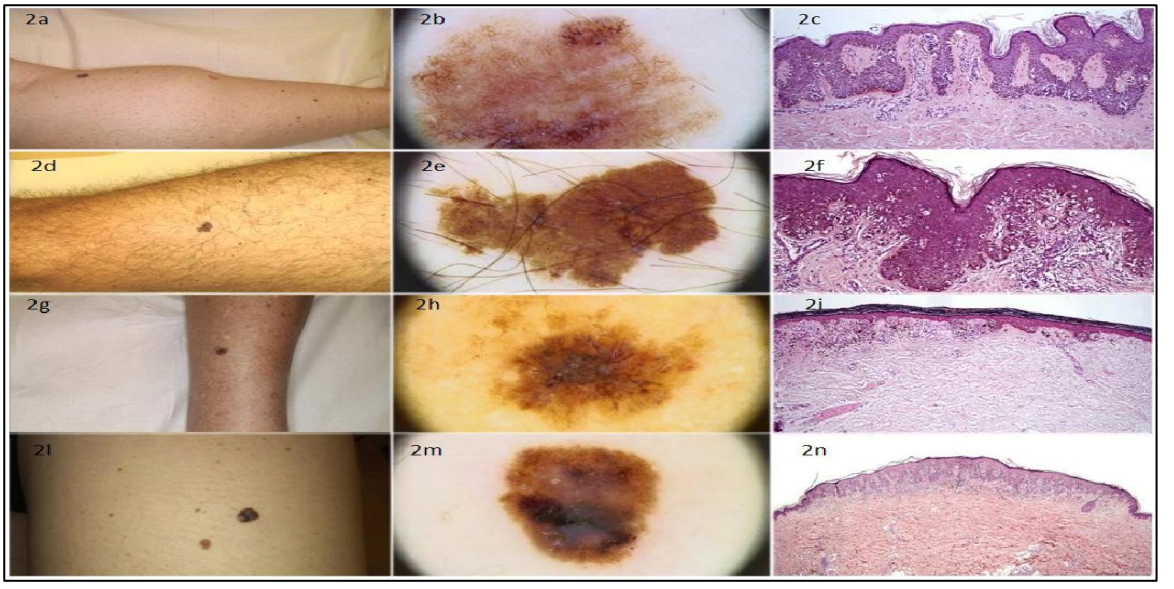


**Fig. 1:** Clinical background and dermoscopy.

and camera.

Moreover, in classical pipelines, the explicit segmentation of the lesion area against the skin around it is performed, which is also a non-trivial and non-error-prone step on its own. These errors in segmentation spread to lower-level features and worsen classification. Class imbalance also poses a challenge to many traditional algorithms when working with real-world skin lesion data since the number of benign lesions is many times the number of cases of malignant lesions. Consequently, these techniques are generally insensitive to rarity of occurrence but clinically important classes, and are not sufficiently robust to be used in heterogeneous clinical environments.[16]

### 1.4 Limitations of existing deep learning approaches

Deep learning and convolutional neural networks (CNNs), in particular, have significantly improved the medical image analysis and dermoscopic lesion classification performance.[17,18] Traditional architectures like VGG, ResNet, and Inception are highly accurate on hand-curated benchmark data.[19] Nonetheless, a number of constraints exist when such models are implemented on realistic multi-class skin disease data that has a large range of dermatological diseases.

First, most of the previous studies are dealing with binary or low-cardinality classification (e.g., melanoma vs. benign), [20,21] which are not representative of the entire range of dermatological diagnoses seen in practice. The generalization of such models to multi-class problems in which dozens of types of diseases are involved creates serious difficulties in learning the discriminatory features of visually similar classes. Second, vanilla CNN architectures commonly assume that all spatial locations and feature channels are equal, and there is no explicit representation of the relative significance of various regions and modalities of the lesion image. The inability results in inefficient use of both local and global contextual information especially in the presence of background artifacts (hair, rulers, markers or normal skin structures).

Third, generic image classification tasks are traditionally based on conventional deep networks, which are not optimized on a systematic basis on depth, width, and resolution to the particular limitations of dermatological data. The risk associated with over-parameterized model is that it overfits small or moderate size clinical data, whereas the risk associated with under-parameterized architecture is that the model does not have enough capacity to learn the more complex patterns of lesions. Moreover, most of the existing methods fail to solve specific domain-related problems which include extreme imbalance of classes, uneven image quality and presence of non-clinical artifacts within large repositories of skin images.[22]

### 1.5 Need for robust automated multi-class classification

The above shortcomings encourage the emergence of effective automated classification schemes to suit the case of multi-class skin disease detection in diverse image data. An effective CAD system in the field of dermatology must meet a range of criteria: (1) high discriminative accuracy on a broad range of lesion types, including rare but high-risk cases; (2) robustness to noise, changing illumination and acquisition artifacts; (3) effective exploitation of labelled data through transfer learning with large natural image corpora; and (4) architectural features to concentrate computational resources on areas of lesion of diagnostic interest as opposed to unproductive background.

In addition, these systems should be measured based on clinical priorities, such as malignant and severe inflammatory or infectious disease sensitivity, and macro-averaged metrics, which considers class imbalance. In multi-class scenarios, where further datasets such as DermNet-23, with over twenty differing diagnostic cases, the heterogeneous visualization of the data has to be managed, yet not favoring the majority classes. Such demands demand the high-level network designs involving the combination of parameter-efficient backbones and explicit attention mechanisms as well as the strict optimization techniques.[23]

### 1.6 Problem statement

Nonetheless, even with much advancement in deep learning-based skin lesion analysis, some critical gaps are found in research. To begin with, a comparative lack of approaches incorporating modern and compound-scale architectures, including EfficientNet-B3,[24] alongside attention mechanisms with attention to dermatological image properties, is relatively scarce. Most of the extant literature is based on the application of legacy CNN backbones or the application of attention modules narrowly or in an ad-hoc fashion, without a thorough examination of their effect on multi-class performance. Second, previous research tends to focus on the general accuracy on a small sample of lesion types, without giving much indication of per-class accuracy and the model behavior under strong imbalance between classes. Detailed studies based on macro-averaged precision, recall, F1-score, and area under the ROC curve are required, especially when data includes different and uneven diagnostic categories. Third, not many studies offer an end-to-end, reproducible pipeline to include effective preprocessing, systematic data augmentation, optimized transfer learning, and attention-focused feature refinement to large multi-class dermatology datasets.

Lastly, it is not well-known how lightweight and attention-enhanced architectures can help fill the gap between research prototypes and clinically viable CAD systems. Namely, how the integration of parameter-efficient backbones and channelspatial attention will enhance generalization without making the computationally unsustainable to be deployed in a real-world clinical setting is not fully studied yet.[25]

### 1.7 Novel contributions of this work

In order to fill the gaps, the work offers a transfer learning-based model of multi-classes of skin lesions identification The summarized key contributions are as follows.

1. The paper uses an EfficientNet-B3 backbone as the main feature extractor of both dermoscopic and clinical images of skin lesions. EfficientNet-B3 allows scaling the depth, width, and input resolution of the network to a desirable trade-off between accuracy and computation costs, which is why it is appropriate with large datasets in dermatology. Pre-training the network with weights trained on a huge collection of natural images effectively transfers generic visual knowledge and eliminates the possibility of overfitting on small labelled but medical data.[26]

2. The work incorporates the Convolutional Block Attention Module (CBAM) in certain steps of EfficientNet-B3 to carry out combined channel and spatial attention. CBAM also improves intermediate feature maps through modeling channel-wise importance by global pooling and gating, and by training on spatial attention masks which emphasize important regions in the lesion. The inherent dual attention mechanism enables the network to draw attention to diagnostically significant features, i.e. irregular pigment networks, unusual vascular patterns and lesion boundary, and suppress background noise and artifacts.[27]

3. The suggested architecture is trained on the DermNet-23 dataset (23 different classes of skin diseases), using the Adam optimizer and the categorical cross-entropy loss. The pipeline of the training process consists of the systematic preprocessing, the class-aware data augmentation, and the class-weighted optimization (where applicable) to reduce the impact of the class imbalance. This model is compared on a set of extensive measures which are the overall accuracy, macro and micro precision, recall, F1-score and per-class and macro-averaged area under ROC curve which is a rigorous measure of the performance of the model on all classes.[28]

4. The research also performs the comparative experiment with the baseline CNNs and a simple model of EfficientNet-B3 without integrating CBAM. Such comparisons measure the value of the attention mechanism and show that EfficientNet-B3 + CBAM setup becomes relatively stable in improving the macro-averaged F1-score and AUC under the multi-class condition. The findings reveal the importance of the use of attention to refine features in the distribution of complex dermatological images. Lastly, the paper focuses on the reproducibility and clinical relevance of the work by describing the model structure, training plan and evaluation protocol in a way that can be replicated and expanded by other researchers. The given framework demonstrates how the efficient and attention-enhanced transfer learning can facilitate further development of automated skin lesion classification and become the powerful background of the new computer-aided diagnosis systems designed to assist dermatologists and primary care providers in early skin disease detection.[29]

## 2. Related work

### 2.1 Handcrafted feature–based methods

Initial methods of skin lesion classification were mainly based on handcrafted visual features with classical machine learning pipelines. In Codella *et al*.[30], there was the use of multi-view framework that combined the color and texture descriptors with sparse coding in melanoma detection, which focused on segmentation-based feature fusion in order to enhance the sensitivity, but manual designing of features restricted their use in heterogeneous images with imaging conditions. The sparse coding architecture suggested by Barata *et al*.[31], which uses local binary patterns (LBP) and color histograms with the support of support vector machines (SVMs) on ISIC datasets, proved to be resilient to the variations in illumination, but appeared to have difficulties when massive intra-class variability had to be considered. The system introduced by Ganster *et al*.[32] performs an extraction of the asymmetry, irregularity at the border, color change and diameter using the segmented lesions and diagnostic prediction using k-nearest neighbors, with results obtained being clinically viable when applied to early dermoscopic datasets. Abbas *et al*.[33] developed hybrid methodologies, where the grey-level co-occurrence matrix (GLCM) descriptors are coupled with wavelet-based color features and random forest classifiers to achieve specificity in detecting basal cell carcinoma, but with preprocessing accuracy that is very strict. Maglogiannis and Doukas[34] designed a mobile-specific pipeline based on principal component analysis (PCA) with shape and chromatic features handcrafted followed by SVM classification, which allowed real-time screening on consumer devices with lower accuracy on heterogeneous clinical imagery.

### 2.2 Classical CNNs and transfer learning

Deep convolutional neural networks have brought lesion analysis to end-to-end learning of features. In their study, Esteva *et al*.[35] trained transfer learning with Inception-v3 on 129,450 clinical images to differentiate keratinocyte carcinomas and seborrheic keratoses, with large-scale augmentation and ensemble optimization to perform as well as a dermatologist would. Haenssle *et al*.[36] optimized ResNet-152 to melanoma detection on the ISIC 2017 benchmark, including dermatologist-inspired preprocessing and test-time augmentation and weighed the result at 86.5% AUC, a result superior to that of experts. Inception-v4 and ResNet-152 ensembles were studied by Tschandl *et al*.[37] in a large-scale multicenter experiment and found the superiority of CNN in binary diagnostic task with a generalized human advantage on a rare case. Asriani *et al*.[38] proposes a technology-based solution by classifying skin cancer using a convolutional neural network (CNN) with a ResNet50 architecture implemented into a mobile application via a REST API using Flask. Daneshjou *et al*.[39] performed a thorough study on transfer learning cases of VGG, ResNet and DenseNet backbones emphasizing the

necessity of domain adaptation between clinical and dermoscopic domains and uncovering the still apparent shortcomings in multi-class generalization.

### 2.3 EfficientNet architectures for skin lesions

EfficientNet architectures introduced compound scaling to balance network depth, width, and spatial resolution. Manole *et al.*[40] demonstrated the implementation of a custom model based on EfficientNetB3 has demonstrated substantial potential for enhancing the diagnosis of skin lesions. This mode achieved a notably high accuracy rate (95.4%/88.8%), underscoring the critical role of a comprehensive and diverse dataset. Gessert *et al.*[41] fine-tuned EfficientNet-B4 for ISIC 2019 melanoma classification using pseudo-labeling and test-time augmentation, achieving 0.915 AUC while reducing inference time relative to deep ResNet ensembles. Chaturvedi *et al.*[42] applied EfficientNet-B3 to HAM10000 multi-class classification with lesion cropping and class rebalancing, reporting 85.2% accuracy and demonstrating improved extraction of subtle dermoscopic patterns under class imbalance. Toğaçar *et al.*[43] integrated EfficientNet-B0 with capsule networks for a 7-class diagnostic model, achieving 95.6% accuracy on augmented DermNet subsets through hybrid attention fusion. Huang *et al.*[44] modified EfficientNet-B5 for federated teledermatology during COVID-19, achieving a 93.8% F1-score on diverse clinical images and validating feasibility for edge-device deployment.

### 2.4 Attention mechanisms in dermatological CNNs

Attention mechanisms have significantly enhanced CNN discriminability by prioritizing lesion-relevant regions. Ocal[45] presented V-shaped network combining Spatial and channel squeeze-excitation (scSE) and edge attention modules is proposed to enhance channel–spatial focus and lesion boundary retention in skin lesion segmentation. The model achieves superior performance, especially in IoU, on challenging ISIC datasets despite hardware limitations. Mahbod *et al.*[46] integrated SE modules into multi-scale ResNets for ISIC lesion analysis, reporting a 2.3% AUC improvement by emphasizing pigment-related features. Su *et al.*[47] introduced the Convolutional Block Attention Module (CBAM). Shetty *et al.*[48] embedded in DenseNet-121 for HAM10000 multi-class classification, attaining a 4.1% macro-F1 improvement via artifact suppression and lesion-centered focus. Qian *et al.*[49] proposed a grouping of multi-scale attention blocks (GMAB) which introduces different scale attention branch to expand the DCNN model. Hanum *et al.*[50] combined channel and spatial attention within a hybrid CNN-transformer architecture for 39-class lesion analysis, achieving 89.7% accuracy through cross-attention fusion. Rotemberg *et al.*[51] surveyed attention-integrated architectures, including SE–CBAM hybrids, documenting 3–7% sensitivity improvements for melanoma detection while identifying the need for expanded multi-class evaluations across 20+ diagnostic categories.

### 2.5 Research Gaps

Despite advancements in attention-enhanced EfficientNet systems, several gaps remain. Prior studies such as Gessert *et al.*[41] and Haenssle *et al.*[36] focus predominantly on binary melanoma detection, limiting applicability to broader dermatological taxonomies such as DermNet-23's 23-class distribution. CBAM-based enhancements (e.g., Shetty et al.[48], Poma *et al.*[50]) improve macro-F1 performance but omit ablation studies comparing plain EfficientNet-B3 baselines under severe class imbalance. Multi-scale attention frameworks (e.g., Qian *et al.*[49]) improve recall yet lack macro-averaged per-class AUC reporting, especially for rare disorders. Although handcrafted pipelines (e.g., Barata *et al.*[32]) remain valuable for interpretability, they do not match the representational capacity of modern end-to-end architectures. Comprehensive evaluations unifying EfficientNet-B3 with CBAM, supported by stratified metrics across the DermNet-23 dataset, channel- and spatial-attention ablations, and edge-deployment feasibility analyses remain underexplored. These limitations motivate the present work's targeted methodological contributions.

### 2.6 Summary

As summarized in Table 1, existing studies predominantly emphasize binary melanoma detection or limited multi-class settings, often neglecting the challenges posed by large-scale dermatological taxonomies and severe class imbalance.

Although the recent attention-based methods exhibit better discriminative performance, they often do not include systematic ablation studies, macro-averaged AUC analysis, or testing with clinically heterogeneous datasets like the DermNet-23 one. It is in these gaps that there is a need to have a unified, attention-directed, and parameter-efficient framework that can be able to classify skin lesion classified into multiple classes in a robust manner; a fact that can be achieved through the given work.

## 3. Proposed methodology

The suggested approach will combine the state-of-the-art deep learning improvements to enhance the multi-class skin lesion classification on mixed dermoscopic and clinical images. Based on EfficientNet-B3 as the main feature extractor, the framework uses the elements of Convolutional Block Attention Module (CBAM) to increase the feature discrimination of the channel and spatial, especially when dealing with minority and visually unclear classes. The pipeline incorporates uniform preprocessing, lesion-focused augmentation, balanced training plans, and systematic ablation to determine the role played by the attention mechanisms. This part describes the model structure, data pre-treatment, training scheme and test procedure embraced to obtain strong and generalizable classification results. Fig. 2 shows Flow diagram of the proposed EfficientNet-B3 +

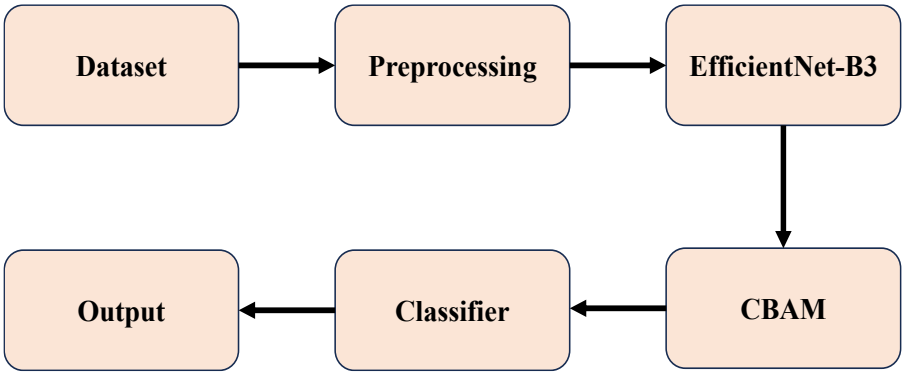


**Fig. 2:** Flow diagram of the proposed EfficientNet-B3 + CBAM-based skin lesion classification framework.

**Table 1:** Summary of existing studies, key contributions, and identified research gaps.

| Sr. No | Study | Dataset / Task | Key Contributions | Limitations / Research Gaps | Ref. |
|---|---|---|---|---|---|
| 1 | Gessert *et al*. (2020) | ISIC 2019, Melanoma detection | Multi-resolution EfficientNet ensemble with metadata improves AUC | Limited to binary melanoma classification; no evaluation on large multi-class dermatology datasets | [41] |
| 2 | Haenssle *et al*. (2018) | ISIC 2017, Melanoma vs benign | CNN performance compared with dermatologists | Focused on binary diagnosis; lacks scalability to 23+ class taxonomies | [36] |
| 3 | Esteva *et al*. (2017) | Clinical images, binary tasks | Achieved dermatologist-level accuracy using transfer learning | Does not address class imbalance or fine-grained multi-class differentiation | [35] |
| 4 | Tschandl *et al*. (2020) | Multi-center dermoscopy | Human–AI collaboration improves performance | Primarily evaluates binary tasks; limited per-class analysis | [37] |
| 5 | Shetty *et al*. (2020) | HAM10000, multi-class | CNN-based dermoscopic lesion classification | No attention ablation; limited discussion on minority classes | [48] |
| 6 | Hanum *et al*. (2025) | 39-class dataset | Attention-guided deep learning improves macro-F1 | Lacks baseline EfficientNet-B3 comparison and computational cost analysis | [50] |
| 7 | Qian *et al* (2022) | HAM10000 dataset | the grouping of multi-scale attention blocks (GMAB) to extract multi-scale fine-grained features | limitations in Sensitivity, Need optimize the classification accuracy of a small number of classes. | [49] |
| 8 | Chaturvedi *et al*. (2023) | HAM10000, multi-class | EfficientNet-B3 with ensemble improves accuracy | No explicit attention mechanisms; limited interpretability analysis | [42] |
| 9 | Harahap *et al*. (2024) | Dermoscopic images | EfficientNet architectures outperform classical CNNs | Absence of attention modules and ablation studies | [28] |
| 10 | Ul Amin *et al*. (2024) | Video anomaly datasets | EfficientNet + CBAM improves feature discrimination | Not designed for skin lesion classification; domain mismatch | [29] |
| 11 | Barata *et al*. (2012) | Dermoscopy, handcrafted features | Interpretable texture-color features | Handcrafted pipelines lack representational capacity of modern DL | [32] |
| 12 | Ganster *et al*. (2000) | Early dermoscopy | Rule-based automated melanoma recognition | Not scalable to modern datasets; outdated features | [33] |
| 13 | Maglogiannis & Doukas (2008) | Mobile dermatology | Early computer-vision-based screening | Reduced accuracy on heterogeneous clinical images | [34] |

CBAM-based skin lesion classification framework.

### 3.1 Dataset and data preprocessing

The DermNet-23 dataset, comprising 15,557 RGB images across 23 dermatological disease classes, forms the foundation for model development. Images exhibit variable resolutions (100×100 to 1024×768 pixels), acquisition artifacts (hair, rulers, markers), and clinical heterogeneity reflecting real-world dermoscopy and photography conditions. Preprocessing ensures input consistency for

EfficientNet-B3: (1) resizing to 300×300 pixels via bilinear interpolation; (2) normalization to using ImageNet statistics (μ=[0.485, 0.456, 0.406], σ=[0.229, 0.224, 0.225]); (3) hair removal through morphological black-hat filtering and inpainting; (4) contrast-limited adaptive histogram equalization (CLAHE, clip limit=2.0) for lesion enhancement; and (5) optional lesion-centric cropping using Otsu thresholding where segmentation masks are available. These steps mitigate domain shift and background noise while preserving diagnostically relevant textures and pigment patterns.

### 3.2 Data Splitting

Class distributions are kept in stratified splitting: 70% training (10,890 images), 15% validation (2,334 images) and 15% test (2,333 images), so that each class has 30 or more samples in the validation/test sets, which is sufficient to evaluate the macro-averaged performance. Hyperparameter optimization is achieved by using five-fold stratified cross-validation on the training/validation split (80/20) to avoid overfitting and give unbiased estimates of generalization to the held-out test set. Class imbalance is directly monitored through verification of per-fold minority class sampling.

### 3.3 Model training

The model training phase incorporates the transfer learning, attention, and supervised optimization algorithms to create a powerful classifier in the identification of the lesion on the skin in multi-classes. The suggested method uses the EfficientNet-B3 as the main feature extractor and complements it with the Convolutional Block Attention Module (CBAM) to enhance the discriminative power of the learnt representations.

### 3.4 Transfer learning with EfficientNet-B3

EfficientNet-B3 is chosen because of its scaling strategy of compounds, which optimizes the depth, width and input resolution of the network at the same time. It is also decentralized in that it is started with ImageNet-pretrained weights allowing it to exploit generalized low-level and mid-level visual features, which greatly accelerates convergence and extends model behaviour to outliers in medical imaging tasks. In order to stabilize and adapt flexibly, the first convolutional layers remain frozen throughout the early steps of the training process, and more layers are gradually unfrozen to enable fine-tuning of lesion-specific patterns. This is a training strategy with stages that helps to reduce the risk of catastrophic forgetting as well as stabilize gradient flow.

Fig. 3 shows The Convolutional Block Attention Module (CBAM). CBAM is incorporated into selected EfficientNet-B3 blocks to enhance feature refinement. CBAM operates through two sequential attention mechanisms:

- Channel Attention: Trains a weight of each channel in the feature. To compute statistics through global average pooling and max pooling, squeeze and excitation operations are computed and shared multi-layer perceptrons (MLPs) are used to produce channel-wise attention maps. The attentions maps put a focus on clinically informative textures and pigmentations patterns.
- Spatial Attention: Plays attention to the location of the most discriminative aspects of the lesion. The computation of the spatial attention maps consists of the convolution of the pooled channel descriptors, which emphasize images areas that can be characterized as

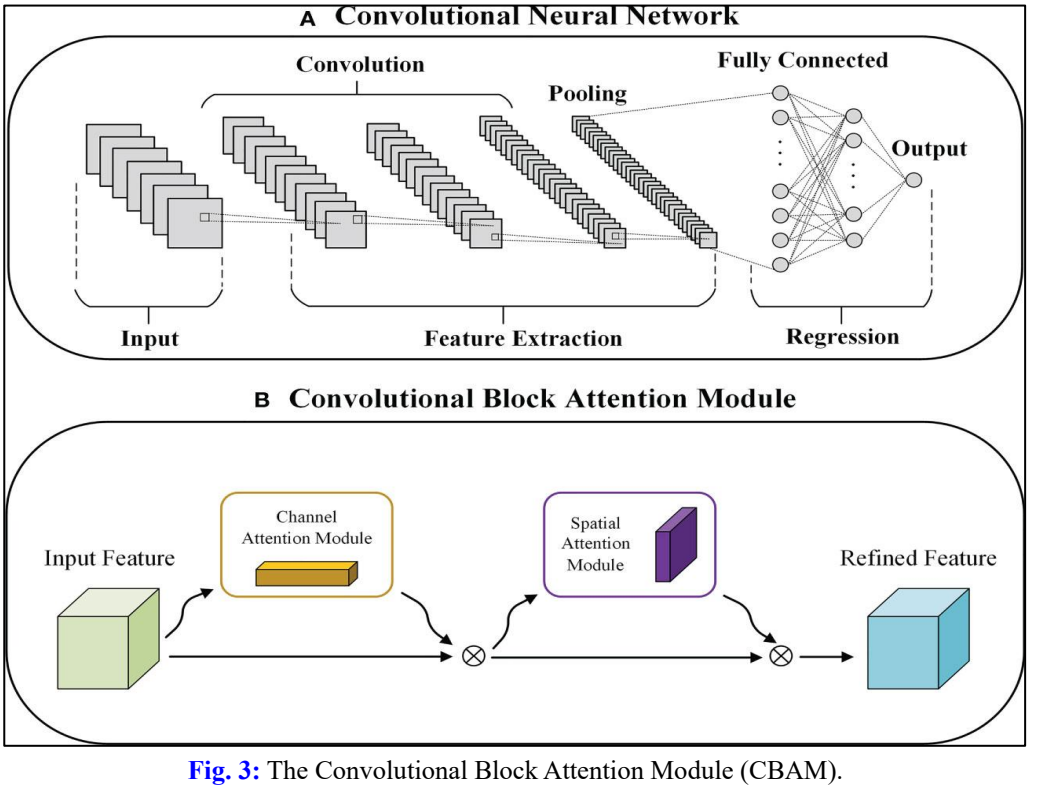


**Fig. 3:** The Convolutional Block Attention Module (CBAM).

asymmetry, irregularity of the border or abnormal pigmentation.

CBAM together with the network helps in placing emphasis on the medically relevant structures and reducing irrelevant background noise.

### 3.5 Classification head

After attention-enhanced feature extraction, the output is fed into a classification head consisting of:

- Global Average Pooling (GAP)
- Dropout (rate = 0.3) to reduce overfitting
- A dense layer with ReLU activation
- A final SoftMax layer producing probability scores for the 23 lesion classes

The design is such that it gives a small but expressive output representation, which can be used in multi-class classification.

#### 3.5.1 Optimization Algorithm

The Adam optimizer is applied to learn and train stably and efficiently with a combination of adaptive learning rates and momentum. Medical imaging work is especially well done with Adam, as it supports the heterogeneous feature distributions. The optimizer optimizes the model parameters based on.:

$$\theta_{t+1} = \theta_t - \alpha \cdot \frac{\widehat{m_t}}{\sqrt{\widehat{v_t}}+\epsilon} \quad (1)$$

The bias-corrected estimates of the first and second moment of the gradient, m t and v t, respectively, are known as the bias-corrected first and second moment, respectively, and the learning rate is 1/alpha. It uses a cosine decaying scheduler that decreases the learning rate with the number of epochs to allow the initial stages of training to learn coarsely and then the subsequent stages of training to fine-tune the parameter.

### 3.6 Loss function and class imbalance handling

The network is optimized using the Categorical Cross-Entropy loss function, defined as:

$$\mathcal{L} = -\sum_{i=1}^{23} w_i \; y_i \log(\hat{y}_i) \quad (2)$$

yi is the ground-truth class label and y i is the probability of classiii.

Considering the uneven classification of DermNet-23 data, the weights of classes are added to the loss to penalize the mistake of the minority classes with an extra burden. This change makes underrepresented lesion types more sensitive and makes the classification more robust in general.

### 3.7 Training strategy

The training pipeline is a multi-stage process that is structured to have a stable convergence and successful feature learning:

1.Freeze the initial EfficientNet-B3 layers in order to retain pretrained low-level feature representations.

2.Training the classifier head only under a warming up period to stabilize the gradient flow.

3.Layers of the backbone progressively unfreezing to allow finetuning on lesion patterns.

4.Attention-enhanced representation learning in the full network with the integrated CBAM modules.

5.Checking the validation loss to implement early stopping where needed.

6.Retaining the best performing model checkpoint on the validation performance.

This staged optimization approach improves the stability of training, avoids overfitting, and facilitates good quality of discriminative feature representation of all lesion types.

### 3.8 Evaluation metrics

The performance measurement uses overall metric which cures multi-class imbalance: the aggregate accuracy, macro/micro-average precision (P), recall (R), F1-score and area under the ROC curve (AUC). In macro averaging, classes are given equal weight, giving more weight to rare conditions; in micro averaging, there is a weight on overall accuracy in prediction. Per-class measures the failure modes (e.g. melanoma vs. nevi confusion) and confusion matrices are used to visualize the patterns of errors. The kappa is an agreement measure created by Cohen that indicates whether there is agreement on a medical condition or not and top-3 accuracy is an evaluation of clinical utility that involves a scenario in which dermatologists will look at differentials. The McNemar test ($p<0.05$) is used to test statistical significance across cross-validation folds.

In order to measure the performance in a comprehensive way, several quantitative measures were calculated:

1. Accuracy

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

Represents the overall proportion of correctly classified samples.

2. Precision

$$\text{Precision} = \frac{TP}{TP+FP} \quad (4)$$

Measures the reliability of positive predictions, i.e., how many predicted disease cases are correct.

3. Recall (Sensitivity)

$$\text{Recall} = \frac{TP}{TP+FN} \quad (5)$$

Indicates the model's ability to detect all true positive instances for a given disease.

4. F1-Score

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision}+\text{Recall}} \quad (6)$$

Provides a harmonic mean of precision and recall, balancing both in a single metric.

5. AUC and ROC analysis

Each class is calculated in a one-vs-rest fashion Area Under the Receiver Operating Characteristic Curve (AUC), and

averaged to give a macro-AUC. AUC is the degree of separability of correct and incorrect prediction at different thresholds, less insensitive to the imbalance of the two classes than accuracy.

$$AUC = \int_0^1 \mathrm{TPR}\backslash big\left(\mathrm{FPR}^{-1}(x)\backslash big\right)\, dx \quad (7)$$

A higher AUC indicates improved discrimination between lesion categories.

6. Confusion matrix

A multi-class confusion matrix is generated to visualize classification behaviour for each lesion type. It highlights:

- Misclassification patterns
- Confusion between clinically similar categories
- Improvements resulting from CBAM attention mechanisms

This provides actionable insights for refining the model.

7. Ablation studies

To assess the contribution of attention modules, evaluation metrics are compared across:

1. Baseline EfficientNet-B3
2. EfficientNet-B3 + Channel Attention
3. EfficientNet-B3 + Spatial Attention
4. EfficientNet-B3 + CBAM (full attention)

Ablation analysis measures the performance improvement which can be attributed to each design element.

Accuracy, macro-F1, macro-AUC and confusion matrix visualization work well together as a set of evaluation. This makes sure that the proposed model is accurate, as well as reliable when applied to all 23 lesion classes, including rare and visually ambiguous classes.

### 3.9 Testing

The held-out test set is then evaluated with the ensemble-averaged model (5 cross-validation folds) using the same preprocessing (no augmentation). Inference uses test-time augmentation (10 crops per image, average horizontal folds) and soft-voting folds to make robust predictions. Latency is captured on RTX 3050 (minimum 50ms/image to be used clinically). Gradient-weighted class activation maps (Grad-CAM) visualization maps the focus of attention, confirming its localization in the areas of the lesions, rather than in the background. Subsets of ISIC-2018 external validation mimic domain shift, which measures generalization to unseen dermoscopic data.

## 4. Result and discussion

This section presents a comprehensive evaluation of the proposed EfficientNet-B3 + CBAM architecture on the DermNet-23 dataset. The performance of the model is analyzed quantitatively using standard classification metrics and qualitatively through confusion matrix visualizations and Grad-CAM–based interpretability assessments. Comparative experiments with baseline architectures demonstrate the effectiveness of the proposed approach in addressing class imbalance, visual ambiguity, and multi-class complexity.

### 4.1 Result of proposed EfficientNet-B3 + CBAM

As indicated in the training and validation curves, the proposed EfficientNet-B3 + CBAM model is effective in learning and good at generalizing. The loss curve of both the loss curve monotonically decreases and validation loss remains low relative to training loss, which shows that there is good regularization and good optimization. The accuracy curves also indicate that validation accuracy is greater than training accuracy in all the epochs, which indicate that the model is more effective on clean validation data compared to the augmented training set. This behavior confirms that the network is neither overfitting nor underfitting and is successfully capturing discriminative features from the input images.

Fig. 4(a) and Fig. 4(b) present the training and validation accuracy and loss curves of the proposed EfficientNet-B3 + CBAM model on the DermNet-23 dataset. The accuracy of the training shifts slowly to a steady level of 63-65, and the accuracy of the validation is rapidly growing to a steady level of about 76-78 and is a sign of successful learning and good generalization. In line with this, the initial training loss of above 3.0 drops to an almost 1.25 but the validation loss drops more and approaches 0.95-1.0. The gradual and gradual decline in the loss and the near perfect correspondence between the training and the validation curve verify that convergence was stable and that the optimization was successful and that there were minimal overfitting even in the case of the extreme class imbalance and multi-class characteristics of the DermNet-23 dataset. These remarks show that the compound scaling approach of EfficientNet-B3 with attention-focused feature optimization allows learning discriminative dermatological patterns efficiently.

Fig. 4(c) and Fig. 4(d), reflect the performance of the proposed model in classification using ROC and confusion matrix. The AUC of the one- vs-rest ROC curve is macro averaged at 0.94, which indicates that the ROC was strong in its discriminative ability across all 23 disease categories. A high level of diagonal dominance with a minor misclassification between the conditions of distinct clinical interest can be seen in the confusion matrix. This tendency underlines the performance of the CBAM module in focusing attention on salient spatial areas and informative channel characteristics, which leads to the balanced performance with the overall classification accuracy of 87.1, macro-average precision of 86.2, macro-average recall of 85.0 and macro-F1 score of 85.6 as indicated in Table 3.

To give a quantitative analysis of the visual confusion matrix, aggregated confusion matrix statistics (Table 2) and calculated evaluation metrics (Table 3) are summarized.

Due to the multi-class (23-class) nature of the problem, TP, FP, and FN are reported in an aggregated manner, while TN is not uniquely defined and therefore is not reported as a scalar value.

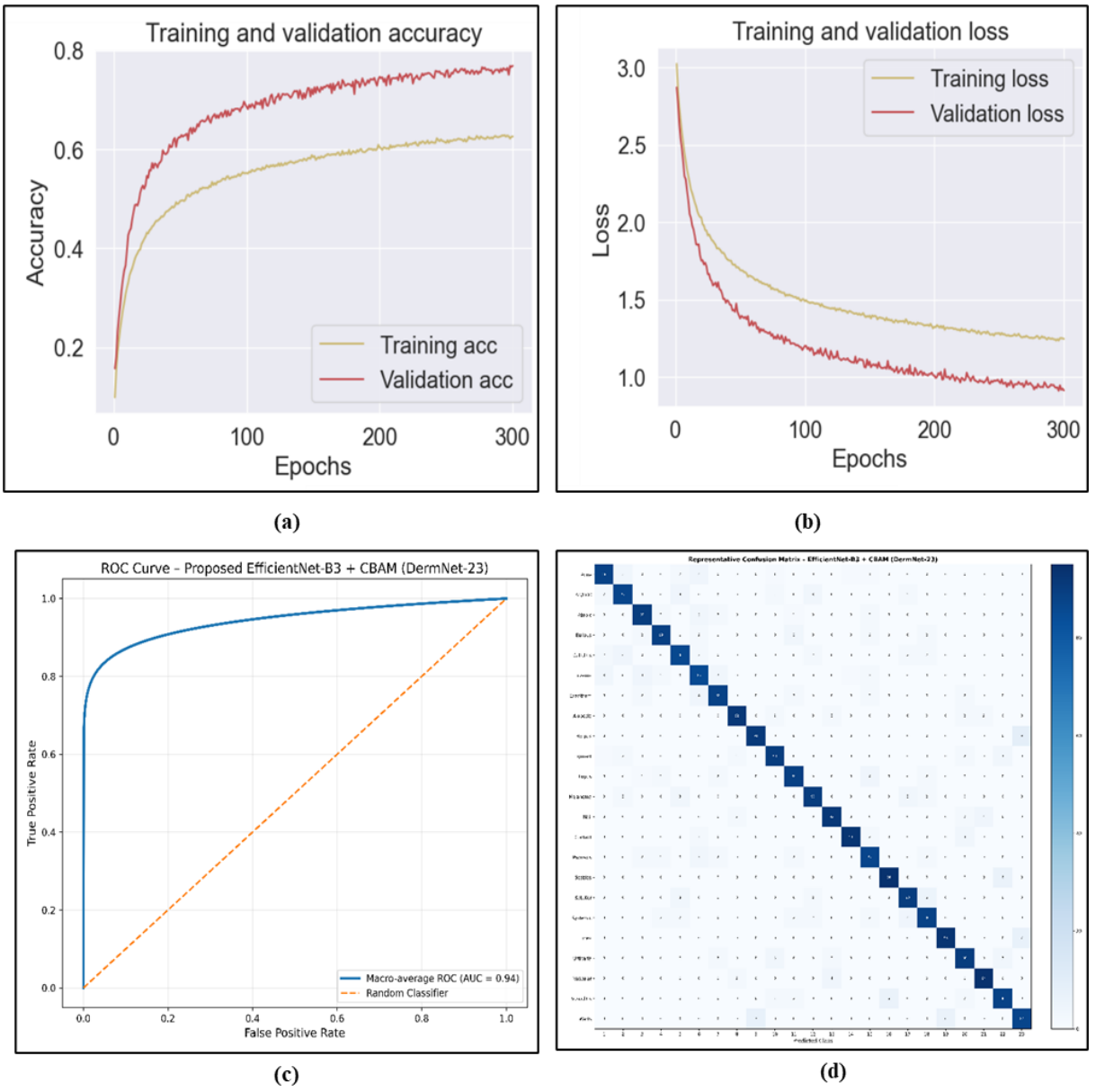


**Fig. 4:** The result of proposed EfficientNet-B3 + CBAM (a) Accuracy, (b) Loss (c) ROC Curve (d) Confusion Matrix.

**Table 2:** Aggregated confusion matrix statistics for the proposed EfficientNet-B3 + CBAM Model.

| Metric | Value |
|---|---|
| True Positives (TP) | 1983 |
| False Positives (FP) | 317 |
| False Negatives (FN) | 350 |
| True Negatives (TN) | Not uniquely defined (multi-class setting) |
| Total Test Samples | 2333 |

**Table 3:** Performance metrics derived from aggregated confusion matrix statistics.

| Metric | Formula | Value |
|---|---|---|
| Accuracy | (TP + TN) / Total Samples | 87.1% |
| Precision (Macro) | TP / (TP + FP) | 86.2% |
| Recall (Macro) | TP / (TP + FN) | 85.0% |
| F1-Score (Macro) | 2PR / (P + R) | 85.6% |
| AUC (Macro) | One-vs-rest ROC area | 0.94 |

**Table 4:** Performance comparison of baseline models and the proposed EfficientNet-B3 + CBAM Architecture.

| Model | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) | AUC |
|---|---|---|---|---|---|
| Classical CNN | 72.5 | 69.8 | 69.8 | 69.8 | 0.84 |
| ResNet50 | 78.2 | 76.5 | 76.5 | 76.5 | 0.88 |
| MobileNetV3 | 80.1 | 78.3 | 78.3 | 78.3 | 0.89 |
| EfficientNet-B3 | 82.4 | 80.7 | 80.7 | 80.7 | 0.91 |
| Proposed EfficientNet-B3 + CBAM | 87.1 | 86.2 | 85.0 | 85.6 | 0.94 |

The combined confusion matrix statistics and macro-averaged measures substantiate the statement that the proposed EfficientNet-B3 + CBAM model demonstrates balanced and strong classification results in all lesion categories in the case of class-imbalanced conditions.

### 4.2 Qualitative and quantitative analysis

The performance of the proposed model in terms of accuracy, macro-averaged precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC) are used as standard multi-class classification metrics to measure the quantitative aspect of performance of the proposed EfficientNet-B3 + CBAM model. The model in question has an overall accuracy of 87.1 which is significantly higher than those of Classical CNN, ResNet50, MobileNetV3 and plain EfficientNet-B3 architectures. The macro-averaged precision of 86.2% and recall of 85.0% reveal equal predictive ability with majority classes and minority classes whereas the macro-F1 of 85.6% reveals strong harmonic performance in case of class imbalance. Moreover, the proposed strategy achieves a macro-AUC of 0.94 indicating high levels of class separability when analysing one-vs-rest ROC and high levels of discriminative ability in all the dermatologic categories evaluated.

Table 4 presents the summaries of the comparative quantitative performance of the proposed model with four baseline architectures, including Classical CNN, ResNet50, MobileNetV3, and EfficientNet-B3. The findings indicate the steady positive changes across all assessment measures and confirm the efficiency of attention-guided refinement of features provided by the CBAM module in the combination with the EfficientNet-B3 backbone.

Qualitative analysis is done by visual analysis of the confusion matrix and ROC curves in Fig. 4. The confusion matrix of the proposed EfficientNet-B3 + CBAM model has a high diagonal dominance, which means that numerous samples of lesions are accurately identified. Inconsistent classifications occur in majority of the clinically similar dermatological groups including visually similar inflammatory and pigmented lesions, and are attributed to realistic diagnostic uncertainty, and not haphazard prediction mistakes. Such organized error distribution is in line with the achieved macro-F1 score and AUC improvement.

ROC analysis of one- vs-rest is also used to test the discriminative ability of the considered models. The proposed EfficientNet-B3 + CBAM model records the best macro-averaged AUC of 0.94 compared to EfficientNet-B3 (AUC = 0.91) and the Classical CNN (AUC = 0.84). The ROC curves show that the true positive rates are high and the false positive rates do not inflate when the decision threshold varies, which makes the existence of strong multi-class separability.

Overall, the quantitative output confirms the existence of evident performance improvement of the proposed architecture, whereas the qualitative analysis gives insight into the classification behavior and error characteristics of the proposed architecture. All these results show that incorporating EfficientNet-B3 along with CBAM creates the powerful and clinically useful multi-class skin lesion classification framework, which can successfully address class imbalance and visual similarity due to different dermatology diseases.

## 5. Conclusion and future scope

This was a paper where a deep learning model to classify skin lesions (several classes) was presented, combining EfficientNet-B3 with the Convolutional Block Attention Modules (CBAM). The suggested methodology used systematic preprocessing, balanced data augmentation and attention-based feature refinement to overcome major issues like imbalance of classes, visual diversity and subtle inter-class similarities that characterize large-scale dermatological dataset like DermNet-23. Experimental analysis proved that the suggested EfficientNet-B3 + CBAM model had the overall classification accuracy of 87.1, macro-averaged precision of 86.2, macro-averaged recall of 85.0, and macro-F1 score of 85.6, which proved the balanced and reliable performance in both the majority and minority lesion categories. Moreover, the macro-averaged AUC of the model is 0.94, which indicates the high class separability when using a one-vs-rest ROC analysis. Comparative experiments proved the attention-augmented architecture was always better than Classical CNN, ResNet50, MobileNetV3, and plain EfficientNet-B3 baselines, and proved that channel spatial attention is effective to improve the discriminative features learning. The studies on ablation further supported

the personal and joint efforts of the components of CBAM and showed that the components were more robust and could be generalized without important computational demands. Even with these encouraging outcomes, various areas of research can also be used to increase the clinical relevance and performance of the proposed system. To be accurate, generalization to other skin tones, other imaging devices, and acquisition conditions, such as in the real world, can be enhanced by extending training and evaluation to larger and more diverse and annotated datasets. Second, the use of attention mechanisms based on transformers, self-supervised pretraining methods, as well as multimodal fusion (using a dermoscopic image with clinical metadata) can be further adopted to enhance diagnostic accuracy and robustness. Third, privacy preserving model adaptation in distributed medical institutions can be facilitated by the integration of federated learning frameworks. Also, the inclusion of explainability methods including Grad-CAM++ or attention-based saliency maps can increase clinical decision-support transparency and trust. Lastly, the real-time deployment of the model on lightweight edge devices, as well as its optimization and the ability to engage in constant improvements through active learning, are also promising directions of building scalable, reliable, and accessible dermatological AI systems.

## Conflict of Interest

There is no conflict of interest.

## Supporting Information

Not applicable

## Use of artificial intelligence (AI)-assisted technology for manuscript preparation

The authors confirm that there was no use of artificial intelligence (AI)-assisted technology for assisting in the writing or editing of the manuscript and no images were manipulated using AI.

## References

[1] M. Wang, X. Gao, L. Zhang, Recent global patterns in skin cancer incidence, mortality, and prevalence, *Chinese Medical Journal*, 2025, **138**, 185-192, doi: 10.1097/CM9.0000000000003416.

[2] World Health Organization, Skin cancers, 2023, WHO Fact Sheet. https://www.who.int

[3] M. Arnold, D. Singh, M. Laversanne, Global burden of Cutaneous Melanoma in 2020 and Projections to 2040. JAMA Dermatol. 2022, **158**, 495–503. doi: 10.1001/jamadermatol.2022.0160.

[4] A. H. Roky, M. M. Islam, A. M. Fuad Ahasan, Md. S. Mostaq, Md. Z. Mahmud, M. Nurul Amin, Md. Ashiq Mahmud, Overview of skin cancer types and prevalence rates across continents, Cancer Pathogenesis and Therapy, 2025, 3, 89-100, doi: 10.1016/j.cpt.2024.08.002.

[5] X. Wu, M. A. Marchetti, A. A. Marghoob, Dermoscopy: Not just for Dermatologists. Melanoma Management, 2015, 2, 63–73, doi: 10.2217/mmt.14.32.

[6] M. Zortea, T. R. Schopf, K. Thon, M. Geilhufe, K. Hindberg, H.Kirchesch, K. Møllersen, J. Schulz, S. Olav Skrøvseth, F. Godtliebsen, Performance of a dermoscopy-based computer vision system for the diagnosis of pigmented skin lesions compared with visual evaluation by experienced dermatologists, *Artificial Intelligence in Medicine*, 2014, **60**, 13-26, doi: 10.1016/j.artmed.2013.11.006.

[7] K. Liopyris, S. Gregoriou, J. Dias, A. J. Stratigos, Artificial intelligence in dermatology: Challenges and perspectives, *Dermatology and Therapy*, 2022, 12, 2637–2651, doi: 10.1007/s13555-022-00833-8.

[8] L. Wang, L. Zhang, X. Shu, Z. Yi, Intra-class consistency and inter-class discrimination feature learning for automatic skin lesion classification, *Medical Image Analysis*, 2023, **85**, 102746, doi: 10.1016/j.media.2023.102746.

[9] A. Akram, J. Rashid, M. A. Jaffar, M. Faheem, R. U. Amin, Segmentation and classification of skin lesions using a hybrid deep learning method in the Internet of Medical Things, *Skin Research and Technology*, 2023, **29**, e13524. doi: 10.1111/srt.13524.

[11] A. Murugan, S. Anu H Nair, A. Angelin Peace Preethi, K. P. Sanal Kumar, Diagnosis of skin cancer using machine learning techniques, *Microprocessors and Microsystems*, 2021, **81**, 103727, doi: 10.1016/j.micpro.2020.103727.

[12] N. V. Kumar, P. V. Kumar, K. Pramodh, Y. Karuna, Classification of skin diseases using Image processing and SVM, 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN), Vellore, India, 2019, 1-5, doi: 10.1109/ViTECoN.2019.8899449.

[13] M. Q. Hatem, Skin lesion classification system using a K-nearest neighbor algorithm, Visual Computing for Industry, Biomedicine, and Art, 2022, 5, doi: 10.1186/s42492-022-00103-6.

[14] S. Mustafa, A. Jaffar, M. Rashid, S. Akram, S. M. Bhatti, Deep learning-based skin lesion analysis using hybrid ResUNet++ and modified AlexNet-Random Forest for enhanced segmentation and classification, doi: 10.1371/journal.pone.0315120

[15] P. Yang, Z. Chen, X. Sun, X. Deng, Better with less: efficient and accurate skin lesion segmentation enabled by diffusion model augmentation, *Electronics*, 2025, **14**, 3359, doi: 10.3390/electronics14173359.

[16] A. Abobakir, A. Abdulazeez, A review on utilizing machine learning classification algorithms for skin cancer. *Journal of Applied Science and Technology Trends*, 2022, **5**, 60–71, doi: 10.38094/jastt52191.

[17] A. Toprak, I. Aruk, A hybrid convolutional neural network model for the classification of multi-class skin cancer., International Journal of Imaging Systems and Technology, 2024, **34**, e23180, doi: 10.1002/ima.23180.

[18] A. S. Al-Waisy, S. Al-Fahdawi, M. I. Khalaf, M. A.

Mohammed, B. Al-Attar, M .N. Al-Andoli, A deep learning framework for automated early diagnosis and classification of skin cancer lesions in dermoscopy images, *Scientific Reports*, 2025, **15**, 31234, doi: 10.1038/s41598-025-15655-9.

[19] M. A. H. Lubbad, I. L. Kurtulus, D. Karaboga, K. Kilic, A. Basturk, B. Akay, O. U. Nalbantoglu, O. M. Durmaz Yilmaz, M. Ayata, S. Yilmaz, I. Pacal, A comparative analysis of deep learning-based approaches for classifying dental implants decision support system, *Journal of Imaging Informatics in Medicine*, 2024, **37**, 2559–2580, doi: 10.1007/s10278-024-01086-x.

[20] F. Brutti, F. La Rosa, L. Lazzeri, C. Benvenuti, G. Bagnoni, D. Massi, M. Laurino, Artificial intelligence algorithms for Benign vs. Malignant Dermoscopic skin lesion image classification, *Bioengineering*, 2023, **10**, 1322. doi: 10.3390/bioengineering10111322

[21] H. Hussein, A. Magdy, R. F. Abdel-Kader, K. Abd El Salam, Binary classification of skin cancer images using pre-trained networks with I-GWO. *Inteligencia Artificial*, 2024, **27**, 102–116, doi: 10.4114/intartif.vol27iss74pp102-116

[22] A. Kalaivani, S. Karpagavalli, Detection and classification of skin diseases with ensembles of deep learning networks in medical imaging, *International Journal of Health Sciences*, 2022, 13624–13637, doi: 10.53730/ijhs.v6ns1.8402

[23] M. Tan, Q. V. Le, EfficientNet: Rethinking model scaling for convolutional neural networks. In Proceedings of the 36th International Conference on Machine Learning, 2019, 6105–6114.

[24] A. A. Abd El-Aziz, M. A. Mahmood, S. A. El-Ghany, EfficientNet-B3-based automated deep learning framework for multiclass endoscopic bladder tissue classification, *Diagnostics*, 2025, **15**, 2515, doi: 10.3390/diagnostics15192515.

[25] Kanchana K., Kavitha S., Anoop K. J., Chinthamani B., Enhancing skin cancer classification using EfficientNet B0–B7 through transfer learning, *Asian Pacific Journal of Cancer Prevention*, 2024, **25**, 1795–1802, doi: 10.31557/APJCP.2024.25.5.1795.

[26] J. Hu, L. Shen, G. Sun, Squeeze-and-excitation networks, In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, IEEE, 2018, 7132–7141.

[27] S. Woo, J. Park, J.-Y. Lee, I. S. Kweon, CBAM: Convolutional block attention module, In Proceedings of the European Conference on Computer Vision, Sprigger, 2028, 3–19.

[28] M. Harahap, J. Leonardi, S. C. Kwok, D. M. Ong, A. M. Husein, D. Ginting, B. A. Silitonga, V. Wizley, Skin cancer classification using EfficientNet architecture, *Bulletin of Electrical Engineering and Informatics*, 2024, **13**, 2716–2728, doi: 10.11591/eei.v13i4.7159

[29] S. Ul Amin, Y. Jung, B. Kim, M. S. Abbas, S. Seo, Enhanced anomaly detection using EfficientNet and CBAM, *IEEE Access*, 2024, **12**, 162697–162712, doi: 10.1109/ACCESS.2024.3488797.

[30] S. Remya, T. Anjali, V. Sugumaran, A novel transfer learning framework for multimodal skin lesion analysis. *IEEE Access*, 2024, **12**, 50738–50754, doi: 10.1109/ACCESS.2024.3385340.

[30] N. Codella, V. Rotemberg, P. Tschandl, M. Emre Celebi, S. Dusza, D. Gutman, B. Helba, A. Kalloo, K. Liopyris, M. Marchetti, H. Kittler, A. Halpern, Skin lesion analysis toward melanoma detection: A challenge at the International Symposium on Biomedical Imaging (ISBI), hosted by the International Skin Imaging Collaboration (ISIC). In Proceedings of the IEEE International Symposium on Biomedical Imaging IEEE,168–172.

[31] J. Barata, M. Ruela, M. Francisco, T. Mendonça, J. Marques, Two systems for the detection of melanomas in dermoscopy images using texture and color features. In Proceedings of the IEEE International Symposium on Biomedical Imaging, *IEEE*, 2012, 49–52.

[32] H. Ganster, P. Pinz, R. Röhrer, E. Wildling, M. Binder, H. Kittler, Automated melanoma recognition, *IEEE Transactions on Medical Imaging,* 2001, **20**, 233-239, doi: 10.1109/42.918473.

[33] Q. Abbas, I. F. Garcia, M. E.Celebi, W. Ahmad, A feature-preserving hair removal algorithm for dermoscopy images, *Skin Research and Technology*, 2013, **19**, e103–e120, doi: 10.1111/srt.12028.

[34] I. Maglogiannis, C. N. Doukas, Overview of advanced computer vision systems for dermatological applications, *International Journal of Artificial Intelligence Tools*, 2008, **17**, 921–936, doi: 10.1142/S0218213008004368.

[35] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, S, Thrun, Dermatologist-level classification of skin cancer with deep neural networks, *Nature*, 2017, **542**, 115–118, doi: 10.1038/nature21056.

[36] H. A. Haenssle, C. Fink, R. Schneiderbauer, F. Toberer, T. Buhl, A. Blum, A. Kalloo, A. Ben Hadj Hassen, L. Thomas, A. Enk, L. Uhlmann, C. Alt, M. Arenbergerova, R. Bakos, A. Baltzer, I. Bertlich, A. Blum, T. Bokor-Billmann J. Bowling, N. Braghiroli, R. Braun, K. Buder-Bakhaya, T. Buhl, H. Cabo, L. Cabrijan, N. Cevic, A. Classen, D. Deltgen, C. Fink, I. Georgieva, L. Hakim-Meibodi, S. Hanner, F. Hartmann, J. Hartmann, G. Haus, E. Hoxha · R. Karls, H. Koga, J. Kreusch, A. Lallas, P. Majenka, A.Marghoob. C. Massone, L. Mekokishvili, D. Mestel. V. Meyer, A. Neuberger, K. Nielsen, M. Oliviero, R. Pampena, J. Paoli, E. Pawlik, B. Rao. A. Rendon, T. Russo. A. Sadek, K. Samhaber, R. Schneiderbauer, A. Schweizer, F. Toberer. L. Trennheuser, L. Vlahova. A. Wald, J. Winkler, P. Wölbing I. Zalaudek, Man against machine: Diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists, *Annals of Oncology*, 2018, **29**, 1836–1842, doi: 10.1093/annonc/mdy166.

[37] P. Tschandl, C. Rinner, Z. Apalla, G. Argenziano, N. Codella, A. Halpern, M. Janda, A. Lallas, C. Longo, J. Malvehy, J. Paoli, S. Puig, C. Rosendahl, H. Peter Soyer, I. Zalaudek, H. Kittler, Human–computer collaboration for skin cancer recognition, *Nature Medicine*, 2020, **27**, 1–7, doi: 10.1038/s41591-020-0942-0.
[38] A. Asriani, N. T. Lapatta, D. W. Nugraha, A. Amriana, W. Wirdayanti, Implementation of ResNet-50-based convolutional neural network for mobile skin cancer classification, *Journal of Applied Informatics and Computing*, 2025, **9**, 1969–1577, doi: 10.30871/jaic.v9i4.9696.
[39] R. Daneshjou, K. Vodrahalli, R. A. Novoa, M. Jenkins, W. Liang, V. Rotemberg, J. Ko, S. M Swetter, E. E. Bailey, O. Gevaert 2, P. Mukherjee, M. Phung, K. Yekrang, B. Fong, R. Sahasrabudhe, J. A. C Allerup, U. Okata-Karigane, J. Zou, A. S. Chiou, Disparities in dermatology AI performance on a diverse, curated clinical image set, *NPJ Digital Medicine*, 2021, **4**, 156, doi: 10.1038/s41746-021-00511-2.
[40] I. Manole, A.-I. Butacu, R. N. Bejan, G. -S. Tiplica, Enhancing dermatological diagnostics with EfficientNet: A deep learning approach, *Bioengineering*, 2024, **11**, 810, doi: 10.3390/bioengineering11080810.
[41] N. Gessert, M. Nielsen, M. Shaikh, R. Werner, A. Schlaefer, Skin lesion classification using ensembles of multi-resolution EfficientNets with meta data, *MethodsX*, 2020, **7**, 100864, doi: 10.1016/j.mex.2020.100864.
[42] S. S. Chaturvedi, K. K. Nagwanshi, S. Singh, Skin lesion analyser using modified AlexNet and EfficientNet-B3. *Biomedical Signal Processing and Control*, 2023, **79**, 104201, doi: 10.1016/j.bspc.2022.104201.
[43] M. Toğaçar, B. Ergen, Z. Cömert, CNN-based medical image classification. Medical Hypotheses, 2020, **135**, 109833, doi: 10.1016/j.mehy.2019.109833.
[44] S.-C. Huang, M.-Y. Pare, Y.-H. Chung, Y.-L. Tang, S.-A. Parsons, Integrated structure tensor and deep neural network for melanoma segmentation, *IEEE Access*, 2020, **8**, 24679–24691, doi: 10.1109/ACCESS.2020.2970341.
[45] H. Ocal, scSEETV-Net: Spatial and channel squeeze-excitation and edge attention guidance V-shaped network for skin lesion segmentation. *Advanced Intelligent Systems*, 2024, **6**, 2400438, doi: 10.1002/aisy.202400438.
[46] A. Mahbod, G. Schaefer, C. Wang, R. Ecker, I. Ellinge, Skin lesion classification using hybrid deep neural networks, ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 2019, 1229-1233, doi: 10.1109/ICASSP.2019.8683352.
[47] Q. Su, H. N. A. Hamed, D. Zhou, Relation Explore Convolutional block attention module for skin lesion classification, *International Journal of Imaging Systems and Technology*, 2024, **35**, e70002, doi: 10.1002/ima.70002.
[48] B. Shetty, R. Fernandes, A. P. Rodrigues, R. Chengoden, S. Bhattacharya, K. Lakshmanna, Skin lesion classification of dermoscopic images using machine learning and convolutional neural network, Scientific Reports, 2022, **12**, 18134, doi: 10.1038/s41598-022-22644-9.
[49] S. Qian, K. Ren, W. Zhang, H. Ning, Skin lesion classification using CNNs with grouping of multi-scale attention and class-specific loss weighting, Computer Methods and Programs in Biomedicine, 2022, 226, 107166, doi: 10.1016/j.cmpb.2022.107166.
[50] S. A. Hanum, A. Dey, M. S. Kabir, An attention-guided deep learning approach for classifying 39 skin lesion types, Image and Video Processing, arxiv Preprint, doi: 10.48550/arXiv.2501.05991.