

LightDermNet: A Lightweight Attention-Augmented CNN for Reproducible Multi-Class Dermoscopic Skin Lesion Classification

N.R. Rajalakshmi, ME, PhD¹ , Anupam Singh, PhD² , Sachi Shome¹, and Md Mojahidul Islam¹

¹Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala RD Institute of Science and Technology, Chennai, India; ²Department of Computer Science and Engineering, R. P. Shaha, University, Narayanganj, Bangladesh

DOI: <https://doi.org/10.30953/thmt.v1i1.687>

Corresponding Author: N. R. Rajalakshmi, Email: rajirajasekaran@gmail.com

Keywords: class imbalance, convolutional block attention module, focal loss, cross-validation, HAM10000

Abstract

Background: Automated dermoscopic skin lesion classification presents a persistent challenge in clinical artificial intelligence (AI): achieving reliable multi-class discrimination under severe class imbalance while remaining computationally viable for resource-constrained deployment. Existing approaches predominantly rely on large pretrained architectures evaluated under single train-test splits with global pre-augmentation, introducing data leakage and limiting reproducibility.

Objectives: This work presents LightDermNet, a purpose-built lightweight convolutional neural network that integrates depthwise separable convolutions and convolutional block attention modules across four progressive feature-extraction stages, resulting in a total of 104,840 trainable parameters.

Methods: Trained on the HAM10000 dataset (10,015 images, seven lesion classes) under a strictly leakage-free five-fold stratified cross-validation protocol with within-fold augmentation, inverse-frequency class weighting, and focal loss ($\gamma = 2.0$).

Results: LightDermNet achieves a mean cross-validation accuracy of 0.7820 (± 0.0050) and a held-out test accuracy of 0.7558, with a macro receiver operating characteristic – area under the curve of 0.9271. Gradient-weighted class activation mapping visualizations confirm that the model consistently localizes diagnostically relevant lesion regions, supporting its interpretability for clinical decision-support applications. Systematic ablation across four training configurations confirms that the joint application of augmentation, class weighting, and focal loss collectively drives performance gains.

Conclusions: Benchmark evaluation against MobileNetV2, MobileNetV3Small, DenseNet121, and ResNet50V2 under identical conditions demonstrates that LightDermNet surpasses all baselines in both mean accuracy and cross-fold stability while utilizing 97–99.6% fewer parameters. These findings establish that sub-200K-parameter models trained with rigorous cross-validation can match or exceed pretrained architectures on the HAM10000 benchmark, providing:

- a reproducible and computationally accessible foundation,
- for dermoscopic AI deployment,
- in telehealth and point-of-care settings.

Plain Language Abstract

Cutaneous malignancies constitute a prevalent and potentially lethal disorder arising from exposure to ultraviolet (UV) radiation through solar and artificial illumination sources.¹ People with light complexions, a history of sunburns, prolonged UV exposure, or frequent use of tanning beds face an elevated risk of skin cancer. Stratospheric ozone reduction represents a major environmental contributor to this pathology. Principal cutaneous malignancy categories encompass basal cell carcinoma, melanoma, squamous cell carcinoma (SCC), and actinic keratosis. Skin cancer is one of the most common cancers, representing roughly 33% of global malignancy cases,² and is expected to continue to be a prominent cancer type in the present decade. Melanoma represents the most lethal cutaneous malignancy, accounting for 75% of cutaneous cancer deaths.³ The International Agency for Research on Cancer⁴ predicts that the number of new cases of cutaneous melanoma is projected to rise by over 50% from 2020 to 2040. This highlights:

- the need for effective methods to diagnose and treat skin cancer,
- including preventive measures and
- early detection through screening.

Key Takeaways

- the mean validation accuracy and standard deviation across all five folds,
- LightDermNet achieves a mean five-fold validation accuracy,
- LightDermNet achieved a test accuracy of 0.7558, a macro F1-score of 0.56, a weighted F1-score of 0.73, and a macro receiver operating characteristic—area under the curve (AUC-ROC) of 0.9271 under one-versus-rest evaluation.

Dermoscopy, a noninvasive diagnostic imaging modality, can properly detect melanoma, which is more deadly than other kinds of skin cancer. Due to its benefits, including strong visual perception and a low incidence of scoring errors, it has attracted considerable interest. Established diagnostic protocols, including the ABCD rule, 7-point checklist, Menzies process, and CASH methodology, have enhanced clinicians' capacity to differentiate melanoma from benign lesions on imaging.⁵ Because of the similarities and differences among classes, even a highly trained dermatologist may have trouble making an accurate diagnosis of skin lesions. Consequently, there is a critical need to create automated techniques that improve the accuracy of melanoma detection.

Dermoscopic imaging enables visualization of the subsurface cutaneous architecture, revealing lesion characteristics through chromatic and textural patterns that are typically imperceptible to the naked eye.⁶ This modality is commonly employed within diagnostic frameworks and has emerged as a valuable supplementary tool for clinical dermatologists. Computer-aided diagnosis systems depend substantially on automated classification of lesion patterns in dermoscopic imagery.⁷ Cancer survival rates across multiple malignancy types have improved markedly in recent decades, owing to accelerated advances in computational imaging and artificial intelligence (AI) methodologies for disease identification and prognostic assessment.⁸

Multiple computational methodologies have been developed for automated evaluation of dermoscopic images. Traditional frameworks in the early developmental stages predominantly employed sequential processing pipelines comprising image quality enhancement, lesion perimeter identification, characteristic extraction, optimal feature selection, and pathological categorization.^{9–11} Such computational systems leveraged diverse attributes, including chromatic properties, surface texture characteristics, boundary morphology, geometric descriptors, and classification algorithms such as support vector machines, feed-forward neural networks, random forests, and naive Bayes.^{12,13} Nevertheless, identifying optimal feature sets, extraction methods, and classification frameworks remains a major challenge.

Earlier systems used manual feature extraction techniques, which were time-consuming and resulted in a minimal feature set. Additionally, these approaches exhibited suboptimal diagnostic performance due to morphological similarities across pathological categories and phenotypic variations within individual lesion types, rendering them inadequate for routine clinical use. Conversely, deep neural architectures, specifically convolutional neural networks (CNNs), have demonstrated exceptional performance across diverse computer vision applications.^{14–17} CNNs are better at obtaining complete visual feature representations from training datasets.¹⁸ Several recent studies employ convolutional neural networks pretrained on extensive image corpora to support dermoscopic image categorization via transfer learning.^{19–21} However, the combination of subtle inter-class distinctions and significant intra-class variability in skin lesion images can weaken the discriminative power of these unaltered deep feature sets, thereby reducing melanoma detection accuracy.^{22–24} In general, the field of dermoscopy image analysis has moved from traditional, hand-crafted feature extraction methods to deep learning methods, especially CNNs.

Despite progress in deep learning-driven dermoscopy analysis, most published models are computationally demanding, necessitating advanced GPU infrastructure and considerable memory resources that remain inaccessible in remote clinics, mobile screening units, and low-resource healthcare environments prevalent across South Asia, sub-Saharan Africa, and rural areas worldwide. Telehealth platforms increasingly require AI inference to operate on-device or at the network edge, where model size and response latency directly determine clinical viability. This work addresses the deployment gap by introducing LightDermNet, a purpose-built, attention-augmented CNN architecture that prioritizes the accuracy–efficiency trade-off over raw benchmark performance, evaluated under a reproducible, leakage-free protocol suitable for real-world clinical deployment.

Literature Review

The research included a review of the literature (Table 1) as well as an in-depth examination of the chosen subject. Extensive prior research has established that computer-aided diagnostic systems for the detection of skin pathology constitute a well-defined research domain.

Contemporary investigations by Gajera et al. conducted systematic evaluations across multiple benchmark repositories to examine dermoscopic image analysis for the diagnosis of melanoma by pre-trained convolutional network feature extraction.⁵ Their findings demonstrated the significance of various methodological parameters, including data preprocessing strategies, dataset partitioning approaches, feature normalization protocols, and hierarchical feature representation levels, for melanoma diagnostic performance.

An integrated deep learning methodology combining EfficientNetV2S and SwinTransformer networks was introduced by Shehzad et al.²⁵ for automated skin cancer identification across multiple diagnostic categories. Through custom alterations to the fifth stage of EfficientNetV2S, integration of Swin-Transformer feature streams, and fusion of outputs from both the original and adapted networks, the ensemble attained 99.10% accuracy on the held-out test set, with sensitivity of 99.27% and specificity of 99.80%.

In,² an efficient AlexNet architecture utilizing transfer learning was proposed for accurate classification and diagnostic assessment of melanoma. The proposed method utilized images with a specified region of interest to isolate only the distinguishing features.

Shrestha et al.²⁶ developed three deep-learning segmentation approaches, including the U-Net, ResUNet, and DeeplabV3+ architectures, for cutaneous malignancy classification. In these evaluations, DeeplabV3+ achieved the highest performance, achieving 96.21% overall accuracy, with precision and recall of 93.26 and 93%, respectively.²⁶

The DSCC net framework was introduced by Tahir et al. for automated classification of cutaneous malignancies across four pathological categories: SCC, basal cell carcinoma (bcc), melanoma (mel), and melanocytic nevi (mn).²⁷ The DSCC net architecture achieved 94.17% accuracy, 93.76% recall, and 93.93% F1-score through the systematic application of modified convolutional blocks designed for early-stage tumor identification, 94.28% precision, and 99.42% area under the curve (AUC).

Keerthana et al.²⁸ introduced dual hybrid convolutional architectures incorporating SVM classification at the terminal

Table 1. Comparative analysis of dermoscopic classification approaches on HAM10000 (7 classes).

Author Ref (Year)	Approach	Params (M)	Acc (%)	F1 (%)	Ref.
Ioannis Kousis et al. (2022)	DenseNet169	14.0	92.25	93.27	36
Srinivasu et al. (2021)	MobileNet + LSTM	~4.2	85.34	–	37
Gessert et al. (2019)	CNN + Attention	NR	–	–	38
Razzak et al. (2020)	Multistage Residual Net	NR	96.07	–	39
Chaturvedi et al. (2020)	CNN Ensemble	NR	92.83	84.00	40
Rajput et al. (2021)	Customized AlexNet	~60.0	98.20	–	41
Khan et al. (2021)	DL + Optimization	NR	90.67	–	24
Shehzad et al. (2023)	EfficientNetV2S + Swin	~28.0	99.10	–	25
Zia Rehman et al. (2022)	MobileNetV2 + DenseNet201	~23.0	85.34	–	42
Kadampur et al. (2020) LightDermNet: A Lightweight Attention-Augmented CNN for Reproducible Multi-Class Dermoscopic Skin Lesion Classification	Inception v3	23.9	99.77	95.74	35
Polat et al. (2020)	CNNs + OVA	NR	92.90	–	43
Mamun et al. (2025)	Lightweight Custom CNN	0.69	96.70	–	33
Sarker et al. (2026)	LGGC-Net	0.81	88.05	–	34
LightDermNet	CBAM + Depthwise Sep.	0.10	78.20 [†]	0.56*	–

All studies use the HAM10000 dataset (seven classes). *Macro F1 on held-out test set ($n = 1,503$). [†]Mean accuracy across stratified five-fold CV with within-fold augmentation only; all other entries report single train–test split accuracy. CBAM: convolutional block attention module; CNN: convolutional neural networks; DL: deep learning; LGGC: large-gap GC; LSTM: long short-term memory; NR: not reported; ~ = estimated from standard architecture.

processing stage; 1 EL identified dermoscopy images as benign or malignant tumors. The SVM classifier receives a combined set of the retrieved parameters from the first and second CNN models. The initial ensemble architecture, combining DenseNet-201 with MobileNet, achieved 88.02% classification accuracy, whereas the alternative hybrid configuration, integrating DenseNet-201 with ResNet-50, attained 87.43% accuracy.

In their study, Syed et al.²⁹ used a deep spiking neural network to assess 3,670 melanoma cases and 3,323 non-melanoma specimens from the ISIC 2019 collection. With far fewer trainable parameters, the proposed spiking VGG-13 model achieved 89.57% accuracy and 90.07% F1-score, outperforming the standard VGG-13 and AlexNet architectures.

Jinnai et al.³⁰ used a faster region-based CNN (FCRNN) instead of dermoscopy to identify melanoma from 5,846 clinical pictures to prepare the training dataset. Bounding boxes for lesion locations were manually constructed with enhanced precision, and the FCRNN model exceeded the diagnostic performance of 10 certified dermatologists and 10 resident physicians.

Alwakid et al.³¹ implemented CNN architectures, including a customized ResNet-50 model, to evaluate the HAM10000 dataset. The research was conducted using an asymmetrically distributed collection of skin cancer samples. ESRGAN was used to improve image quality before data augmentation was utilized to address class imbalance. The outcome was determined using the 86 and 85.3% accurate CNN and ResNet-50 models, respectively.

Rashid et al.³² proposed a transfer learning framework utilizing the MobileNet-V2 architecture for automated melanoma detection, differentiating malignant from benign cutaneous lesions. Their convolutional neural network approach was validated on the ISIC 2020 repository. To mitigate class distribution disparities, multiple synthetic data generation strategies were implemented, resulting in improved diagnostic performance of the deep learning framework.

Recent work has begun to address the computational efficiency gap in dermoscopic classification. Mamun et al.³³ showed that a custom CNN with only 692K parameters can match the accuracy of ResNet50 on HAM10000 while cutting FLOPs by more than 99%. They concluded that large pretrained models don't offer much of an accuracy boost at a high computational cost. Similarly, Sarker et al.³⁴ proposed LGGC-Net, a lightweight attention-based CNN with 0.81 million parameters, achieving 88.05% accuracy on HAM10000 and positioning it as a deployment-ready solution for resource-constrained clinical settings. These findings motivate the development of sub-200K parameter architectures evaluated under rigorous, leakage-free protocols.

While prior studies demonstrate high classification accuracy on HAM10000, the majority rely on single train–test splits without cross-validation and employ large pretrained architectures exceeding 20 million parameters.^{25,35} No published work in this benchmark evaluates a sub-200K-parameter model under stratified k -fold cross-validation with strictly within-fold augmentation. LightDermNet is positioned to address this efficiency–reproducibility gap, targeting deployment in telehealth-enabled, resource-constrained clinical environments where model size and inference latency directly determine clinical viability.

Methodology

This study develops LightDermNet, a purpose-built attention-augmented convolutional neural network designed for efficient and reproducible dermoscopic image classification. The diagnostic framework begins with dataset acquisition, followed by hair removal preprocessing, image normalization, and within-fold data augmentation. The extracted features are used to train the LightDermNet model for classification, which assigns predicted class labels and probabilities to the seven skin lesion categories. A stratified five-fold cross-validation protocol is employed throughout to ensure leakage-free evaluation appropriate for clinical application. The complete experimental pipeline is illustrated in Figure 1.

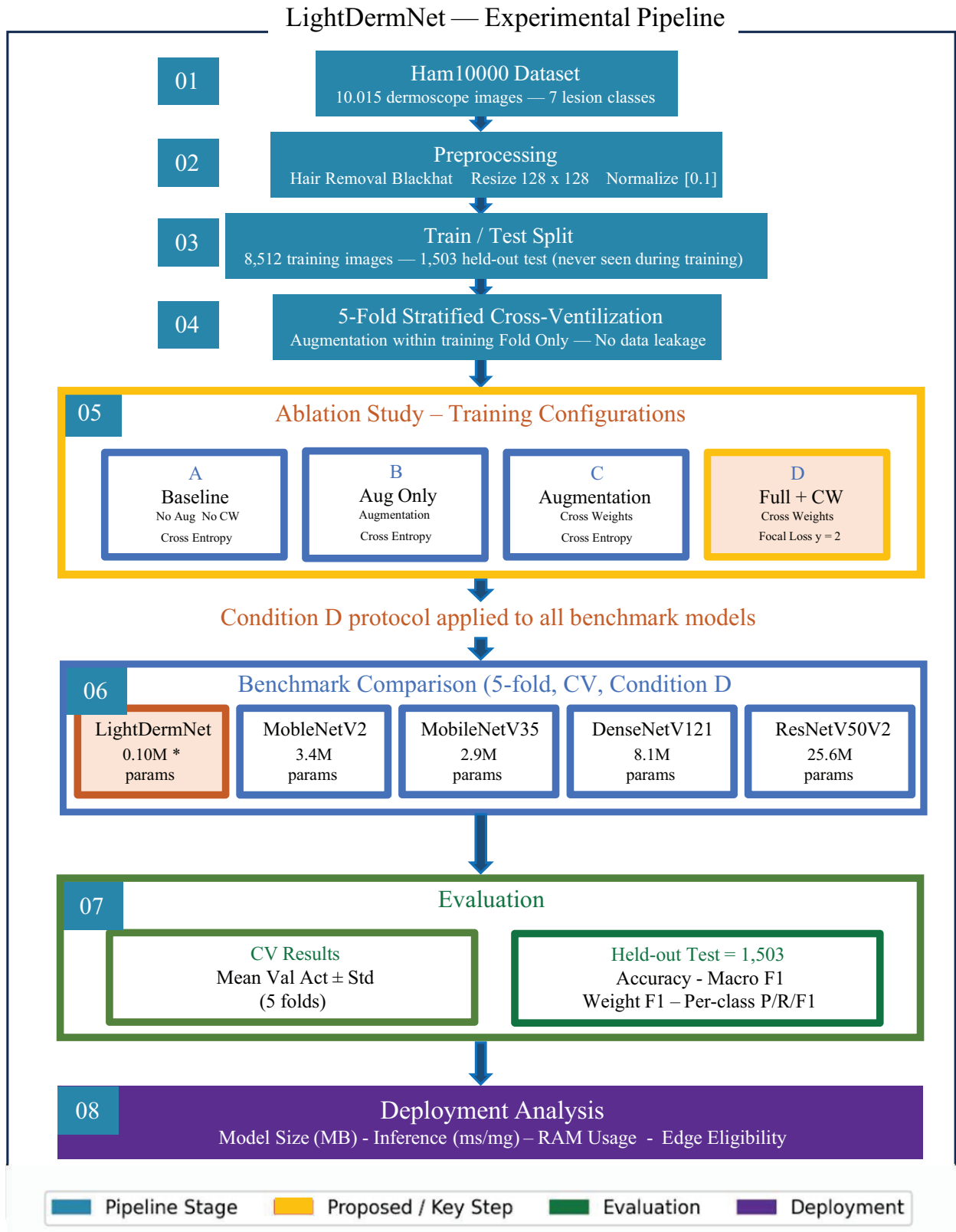


Fig. 1. Experimental pipeline of LightDermNet.

CBAM: convolutional block attention module, Conv2D: standard 2D convolution, CV: computer vision, DWS: depthwise separable conv, RAM: residual attention module.

Dataset

Our experimental framework employs the HAM10000 collection established by the International Skin Image Collaboration in 2018 as the foundational data resource.⁴⁴ This repository comprises 10,015 dermoscopic photographs obtained from

heterogeneous patient cohorts at two collection sites. The dataset includes seven categories of skin lesions: actinic keratosis along with intraepithelial carcinoma (akiec), vascular lesions (vasc), dermatofibroma (df), benign keratosis-like lesions (bkl), and melanocytic nevi (nv). Each diagnostic

category represents a distinct dermatological condition targeted for automated identification. The repository's development involved collaborative efforts between Cliff Rosendahl from Queensland University Medical School in Australia and the ViDIR Group at the University of Vienna's Dermatology School in Austria.

The complete collection of 10,015 images was partitioned into 8,512 training images and 1,503 held-out test images prior to any preprocessing or augmentation steps, ensuring the test set was completely independent across all experiments. A notable limitation was the substantial asymmetry in class-wise sample distribution: the most prevalent category (melanocytic nevi) contained 6,705 specimens, while the least represented class (dermatofibroma) had only 115 samples, accounting for 1.1% of the complete collection. The class distribution across the seven lesion categories is illustrated in Figure 2, with representative sample images shown in Figure 3.

Data Preprocessing

In dermoscopic image analysis, acquisition artifacts, most notably hair occlusion, represent a significant source of interference that can impair automated classification performance.

Skin Lesion Counts in the HAM10000 dataset

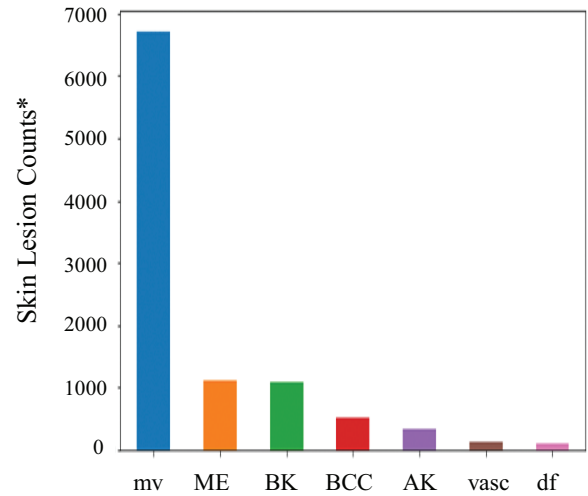


Fig. 2. Seven types of skin lesion counts in the HAM10000 dataset. *AK: actinic keratosis; BCC: basal cell carcinoma; BK: benign keratosis; df: dermatofibroma; ME: melanoma; mv: melanocytic nevi; vasc: vascular lesions.

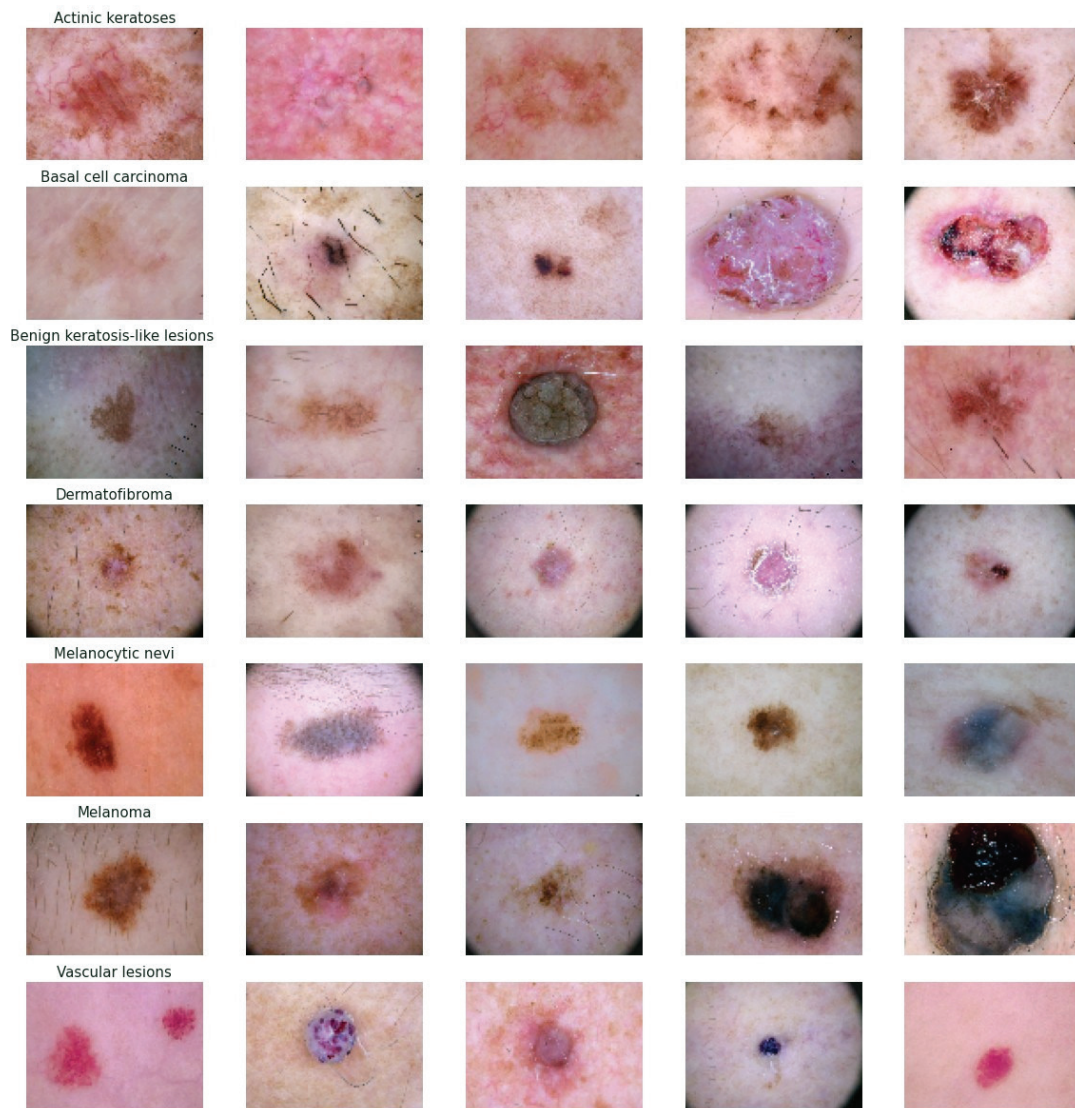


Fig. 3. Representative images from the HAM10000 dataset across all seven lesion categories.

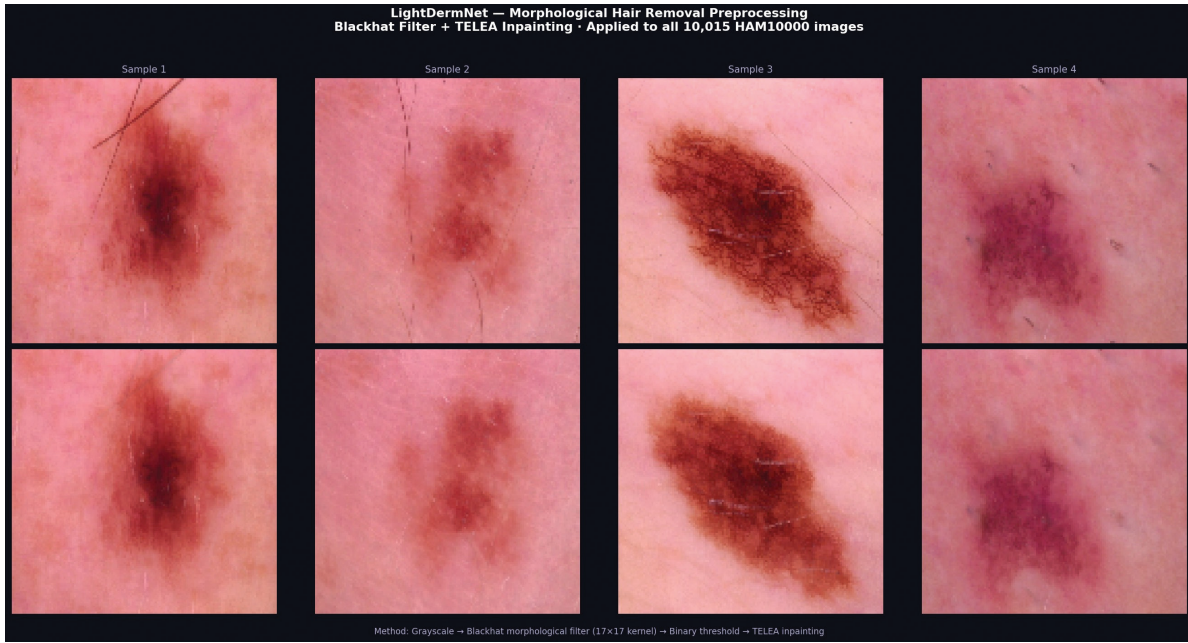


Fig. 4. Hair removal pipeline results: four sample images showing original (top) and processed (bottom) dermoscopic images after Black Hat morphological filtering and Telea inpainting.

To systematically address this issue, a structured hair removal pipeline was applied to all 10,015 images prior to the train-test partition, ensuring consistent preprocessing across all experimental folds.

The hair removal procedure operates in four sequential stages. First, each image is converted to grayscale to isolate luminance variation associated with hair structures. A morphological Black Hat filter with a 17×17 elliptical structuring kernel is then applied to enhance dark, elongated hair artifacts against the surrounding skin texture. To make a binary hair mask, the response map is binarized using a fixed threshold of $t = 10$. Finally, detected hair regions are reconstructed using Telea inpainting⁴⁵ with a 3-pixel radius, which fills masked areas by propagating surrounding skin texture inward. Representative before-and-after outputs of this pipeline are presented in Figure 4.

Following hair removal, all images are resized from their original resolution of 450×600 pixels to a standardized resolution of 128×128 pixels. Pixel values are subsequently normalized from uint8 integer representation to float32 in the range $[0,1]$ to ensure consistent input scaling across the network. This combination of artifact suppression, resizing, and normalization prepares the image corpus for augmentation and model training.

Data Augmentation

The HAM10000 repository exhibits significant class-wise sample disparities, with specific diagnostic groups having substantially fewer specimens than the dominant melanocytic nevi category. To address this imbalance without introducing data leakage, augmentation was applied exclusively within the training partition of each cross-validation fold, while validation and held-out test images remained unaugmented across all experiments.

Augmentation procedures incorporated random rotational transformations up to ± 20 degrees, horizontal and

vertical reflection-based augmentations, positional shifts of 10% in the height and width dimensions, and zoom operations with a 10% magnitude. These transformations were applied on the fly during training using automated augmentation pipelines, ensuring that no augmented images were shared between the training and validation partitions across any fold. By introducing controlled variations in rotation, zoom, shifts, and flips, the augmentation strategy enhances the model’s ability to generalize across diverse lesion orientations, scales, and spatial positions, thereby improving classification robustness without compromising evaluation integrity.

LightDermNet Architecture

LightDermNet is a purpose-built convolutional neural network designed to balance diagnostic accuracy with computational efficiency, making it suitable for deployment in resource-constrained telehealth environments. The architecture comprises four progressive convolutional stages that incorporate depthwise separable convolutions and the convolutional block attention module (CBAM), enabling efficient multi-scale feature extraction with a substantially reduced parameter footprint compared to standard pretrained architectures.

Network Architecture

The input layer accepts dermoscopic images with a shape of $128 \times 128 \times 3$ after hair-removal preprocessing and float32 normalization. Each of the four convolutional stages progressively increases the number of feature maps through a sequence of 32, 64, 128, and 256 filters, respectively, using 3×3 depthwise separable convolutional kernels with rectified linear unit (ReLU) activation and the same padding to preserve spatial dimensions. Following each convolutional stage, max-pooling with 2×2 kernels performs spatial downsampling, reducing feature map dimensions by 50% and concentrating the most discriminative activations.

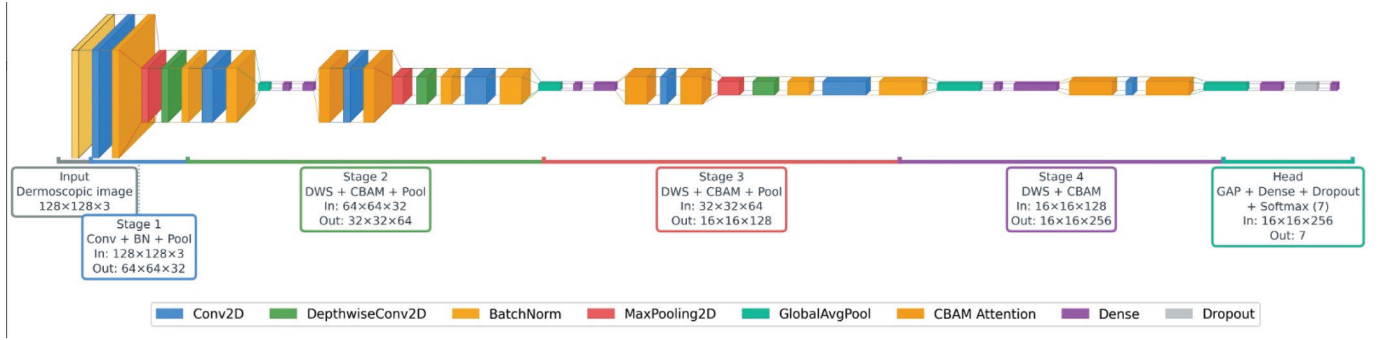


Fig. 5. LightDermNet architecture.

Conv2D: standard 2D convolution, DWS: depthwise separable conv, RAM: residual attention module.

CBAM attention gates are inserted after each pooling operation, applying sequential channel-wise and spatial attention to recalibrate feature responses and suppress uninformative regions. A flattening operation converts the final feature representations into a one-dimensional array, followed by two fully connected layers with ReLU activation, incorporating dropout regularization to mitigate overfitting. The output layer consists of seven neurons corresponding to the distinct lesion categories, with SoftMax activation applied to generate probability distributions over classes. The complete model encompasses 104,840 trainable parameters, representing a reduction of over 99% compared to standard pretrained architectures such as ResNet50V2 (25.6M parameters), as illustrated in Figure 5. Depthwise separable convolutions decompose a standard convolution into a depthwise spatial filtering step followed by a pointwise channel projection, reducing the computational cost by a factor expressed as:

$$\frac{Cost_{DWS}}{Cost_{Std}} = \frac{1}{N} + \frac{1}{DK_2} \quad (1)$$

where N denotes the number of output channels, and DK is the kernel size. For $DK = 3$ and $N = 256$, the configuration yields a reduction factor of approximately 0.115, meaning depthwise separable convolutions consume roughly 8.7 $\times$ fewer multiply-accumulate operations than their standard counterparts at equivalent receptive field size. CBAM augments each convolutional stage by sequentially recalibrating channel and spatial attention. The channel attention map $\mathbf{M}_c \in \mathbb{R}^{1 \times 1 \times C}$ is computed as:

$$\mathbf{M}_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) \quad (2)$$

where σ denotes the sigmoid activation, and MLP is a shared two-layer network with reduction ratio $r = 8$. The resulting spatial attention map $\mathbf{M}_s \in \mathbb{R}^{H \times W \times 1}$ is then derived as:

$$\mathbf{M}_s(F) = \sigma f^{7 \times 7}([\text{AvgPool}(F); \text{MaxPool}(F)]) \quad (3)$$

where $f^{7 \times 7}$ denotes a convolutional operation with a 7×7 kernel and $[\cdot]$ denotes channel-wise concatenation along the spatial dimensions. The refined feature map is obtained by

element-wise multiplication of the input with both attention maps sequentially: $F' = \mathbf{M}_s(\mathbf{M}_c(F) \otimes F) \otimes F$.

Training Strategy

The model is compiled using the Adam optimizer with an initial learning rate of 1×10^{-3} . To address the substantial class imbalance inherent in the HAM10000 dataset, two complementary strategies are employed: class weights inversely proportional to class frequency are incorporated into the loss computation, and focal loss⁴⁶ with $\gamma = 2.0$ and $\alpha = 0.25$ is used in place of standard categorical cross-entropy. Focal loss down-weights well-classified majority-class examples and concentrates gradient updates on difficult minority-class instances, improving sensitivity for underrepresented lesion categories such as dermatofibroma and vascular lesions. Focal loss is formally defined as:

$$L_{FL} = -\alpha_i (1 - p_i)^\gamma \log(p_i) \quad (4)$$

where p_i is the model's estimated probability for the ground-truth class, $\gamma = 2.0$ controls the down-weighting of easy examples, and $\alpha = 0.25$ balances positive and negative class contributions. Class weights are computed as:

$$w_c = \frac{N}{K \cdot n_c} \quad (5)$$

where N is the total number of training samples, $K = 7$ is the number of lesion categories, and n_c is the sample count for class c . For the dermatofibroma class ($n_c = 95$ in the training partition), this yields $w_c \approx 12.8$, ensuring that minority class gradients receive proportionally amplified updates during backpropagation.

The learning rate is adaptively reduced by a factor of 0.5 when validation accuracy stagnates for five consecutive epochs, via the ReduceLROnPlateau scheduler, thereby supporting stable convergence throughout the full training duration.

Evaluation Protocol

where N is the total number of training samples, $K = 7$ is the number of lesion categories, and n_c is the sample count for class c . For the dermatofibroma class ($n_c = 95$ in the training partition), this yields $w_c \approx 12.8$, ensuring that minority class gradients receive proportionally amplified updates during backpropagation.

The learning rate is adaptively reduced by a factor of 0.5 when validation accuracy stagnates for five consecutive epochs, via the ReduceLROnPlateau scheduler, thereby supporting stable convergence throughout the full training duration.

Results and Discussion

This section presents the experimental results for LightDermNet across four evaluation stages: an ablation study, cross-validation training behavior, benchmark model comparison, and final held-out test set performance. All results are reported under the stratified five-fold cross-validation protocol described in Section 3, with the 1,503 held-out test images evaluated exclusively with the best-performing fold (Fold 5; validation accuracy = 0.7882).

Ablation Study

To systematically identify the contribution of each training component, four experimental conditions were evaluated under identical five-fold stratified cross-validation. Condition A served as the baseline, training LightDermNet without augmentation, class weighting, or focal loss, relying solely on standard cross-entropy. Condition B introduced within-fold augmentation, Condition C applied class weights without augmentation, and Condition D combined all components—augmentation, class weights, and focal loss—representing the full proposed training strategy.

Table 2 summarizes the mean validation accuracy and standard deviation across all five folds for each condition. Condition D achieves the highest mean validation accuracy of 0.7820 (± 0.0050), confirming that the combination of augmentation, class weighting, and focal loss yields the most stable and accurate configuration. Notably, Condition C (class weights only, 0.7701) performed below both Condition A (0.7784) and Condition B (0.7777), suggesting that class weight amplification without complementary augmentation can introduce gradient instability on severely minority classes such as dermatofibroma ($n_c = 95$) when the training distribution remains unbalanced. The validation accuracy curves across all folds and conditions are presented in Figure 6, with a comparative summary shown in Figure 7.

Figure 8 presents per-class recall across all four conditions on the validation set, revealing that Condition D consistently improves minority class recall—particularly for melanoma (0.353 $\rightarrow$ 0.413), vascular lesion (0.667 $\rightarrow$ 0.810), and basal cell.

Carcinoma (0.468 $\rightarrow$ 0.506)—while maintaining dominant class (nv) recall above 0.93.

Figure 9 presents the normalized confusion matrix for Condition D on the validation set (Fold 5), confirming that the dominant misclassification pattern—minority classes predicted

as melanocytic nevi—persists consistently across training configurations and is attributable to the intrinsic visual similarity between early-stage malignant lesions and benign nevi in the HAM10000 collection.

Training Behavior—Best Fold

Condition D Fold 5 achieved the highest validation accuracy of 0.7882 and was selected as the representative model for test set evaluation. Figure 10 illustrates the training and validation accuracy and focal loss curves for this fold. Training accuracy increased steadily from 0.5207 at Epoch 1 to 0.8739 at the point of early stopping (Epoch 49), while validation accuracy converged to 0.7882 at Epoch 34 before stabilizing under the ReduceLROnPlateau schedule. The consistent narrowing of the train-validation accuracy gap across epochs confirms that the combination of within-fold augmentation, dropout regularization, and adaptive learning rate reduction effectively mitigated overfitting despite the dataset’s substantial class imbalance.

Benchmark Model Comparison

To contextualize LightDermNet’s performance relative to established pretrained architectures, four transfer learning baselines were evaluated under identical Condition D protocol—MobileNetV2 (3.4M parameters), MobileNetV3Small (2.9M parameters), DenseNet121 (8.1M parameters), and ResNet50V2 (25.6M parameters). All baselines employed two-phase fine-tuning: an initial head-only training phase of 10 epochs followed by partial unfreezing of the top 30% of base layers. Table 3 presents the per-fold and mean validation accuracy for all five models.

LightDermNet achieves a mean five-fold validation accuracy of 0.7820 (± 0.0024), surpassing MobileNetV2 by 1.46 percentage points, DenseNet121 by 1.77 points, MobileNetV3Small by 2.50 points, and ResNet50V2 by 3.32 points—while utilizing 104,840 parameters compared to 3.4M–25.6M for the baselines, representing a parameter reduction of 97–99.6%. Critically, LightDermNet also exhibits the lowest cross-fold standard deviation (0.0024) among all evaluated models, indicating that the combination of CBAM attention and depthwise separable convolutions produces more consistent generalization across data partitions than pretrained feature extractors under equivalent training conditions. The benchmark comparison and efficiency frontier are illustrated in Figures 11 and 12.

Test Set Performance

Final evaluation was conducted on the 1,503 held-out test images using the Condition D Fold 5 model (validation accuracy = 0.7882). This evaluation was performed once, after all training, validation, and hyperparameter decisions were finalized, ensuring complete independence of the reported metrics. LightDermNet achieved a test accuracy of 0.7558, a macro F1-score of 0.56, a weighted F1-score of 0.73, and a macro receiver operating characteristic—area under the curve (AUC-ROC) of 0.9271 under one-versus-rest evaluation. The per-class classification report is presented in Table 4, with per-class metrics and AUC-ROC curves shown in Figures 13–15.

Table 2. Ablation study results—five-fold mean validation accuracy.

Condition	Aug	CW	Loss	Mean Val Acc	$\pm$ Std
A—Baseline	✗	✗	Cross-Entropy	0.7784	0.0072
B—Aug Only	✓	✗	Cross-Entropy	0.7777	0.0078
C—CW Only	✗	✓	Cross-Entropy	0.7701	0.0028
D—Full (Proposed)	✓	✓	Focal ($\gamma = 2$)	0.7820	0.0050

Aug: Within-fold augmentation; CW: inverse-frequency class weights.

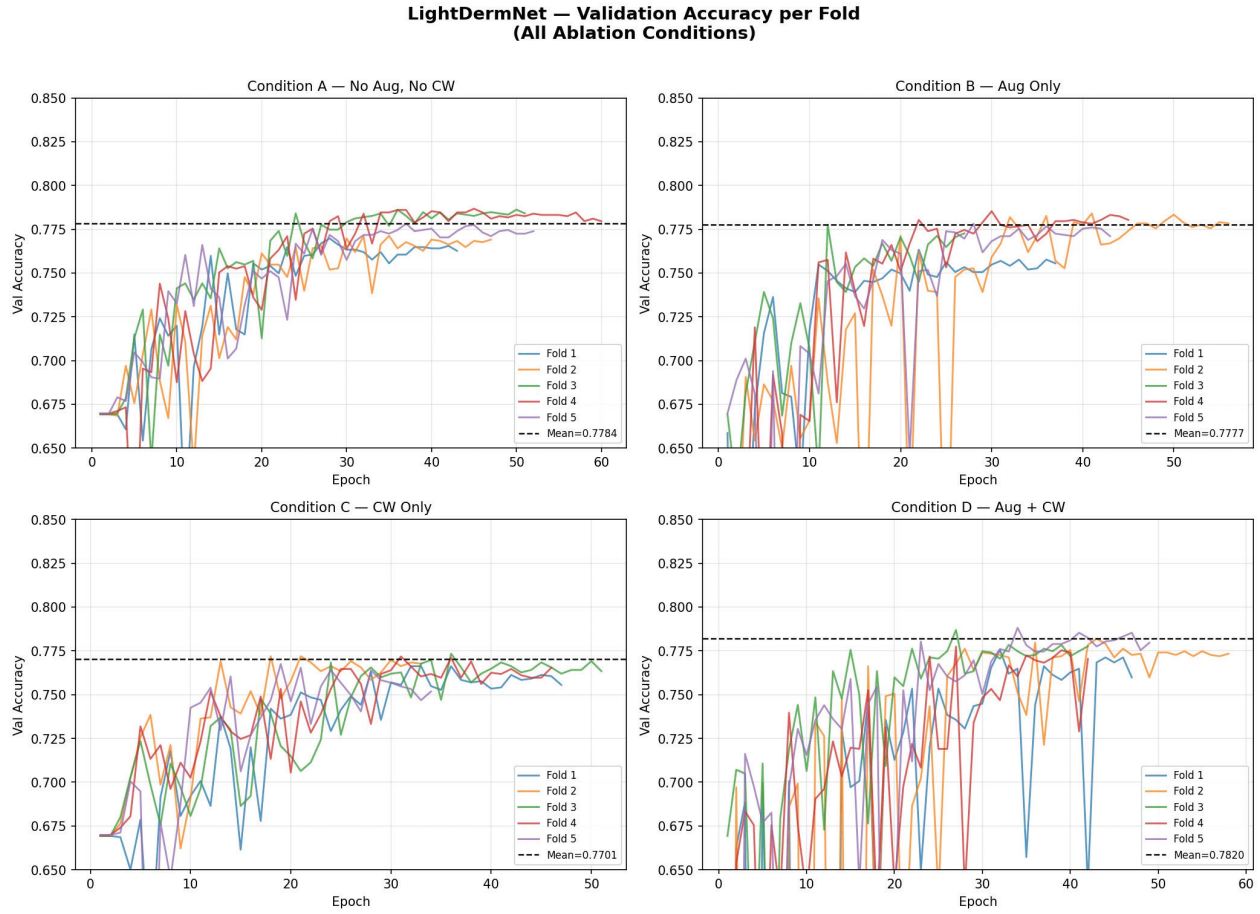


Fig. 6. Validation accuracy curves across all five folds for Conditions A–D. Condition D demonstrates the highest and most consistent convergence across folds.

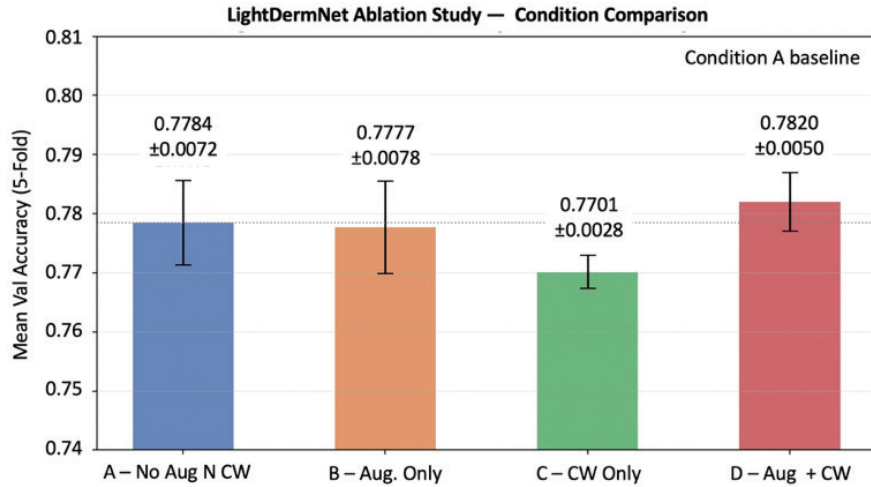


Fig. 7. Mean validation accuracy ($\pm$ std) per ablation condition. Condition D (Aug + CW + Focal Loss) achieves the best performance at 0.7820 ± 0.0050 .

Class-wise performance reflects the inherent distributional asymmetry of the HAM10000 dataset. The dominant melanocytic nevi class achieves an F1 of 0.88, consistent with its support of 1,006 test samples (66.9% of the test set). Vascular lesion attains the highest minority class F1 of 0.83 and AUC of 0.9981, attributable to its morphologically distinctive appearance relative to other lesion categories.

Conversely, melanoma recall of 0.31 represents the most clinically significant limitation: 49.1% of melanoma instances were misclassified as melanocytic nevi—a common failure pattern in dermoscopic classifiers arising from the visual similarity between early-stage melanoma and benign nevi.³⁸ Actinic keratosis precision of 0.75 with recall of 0.37 indicates high specificity but low sensitivity, suggesting the model

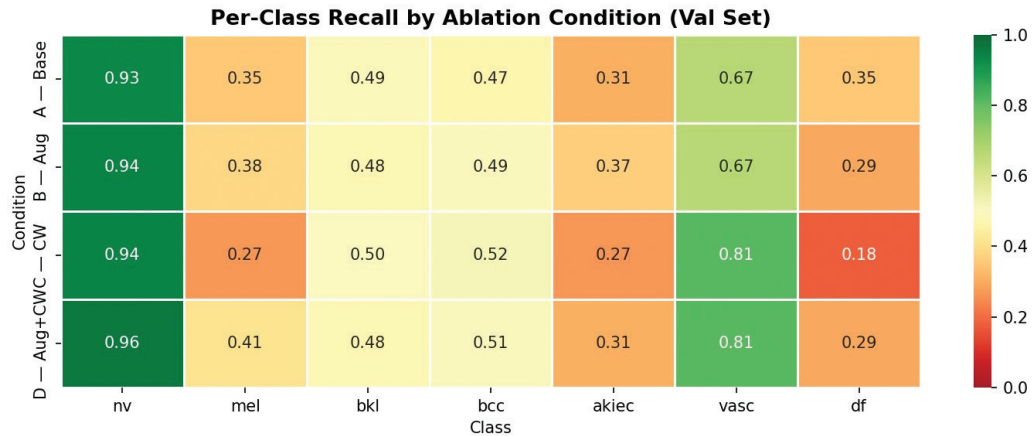


Fig. 8. Per-class recall heatmap across ablation Conditions A–D (validation set). Condition D achieves the most balanced recall profile across all seven lesion categories.

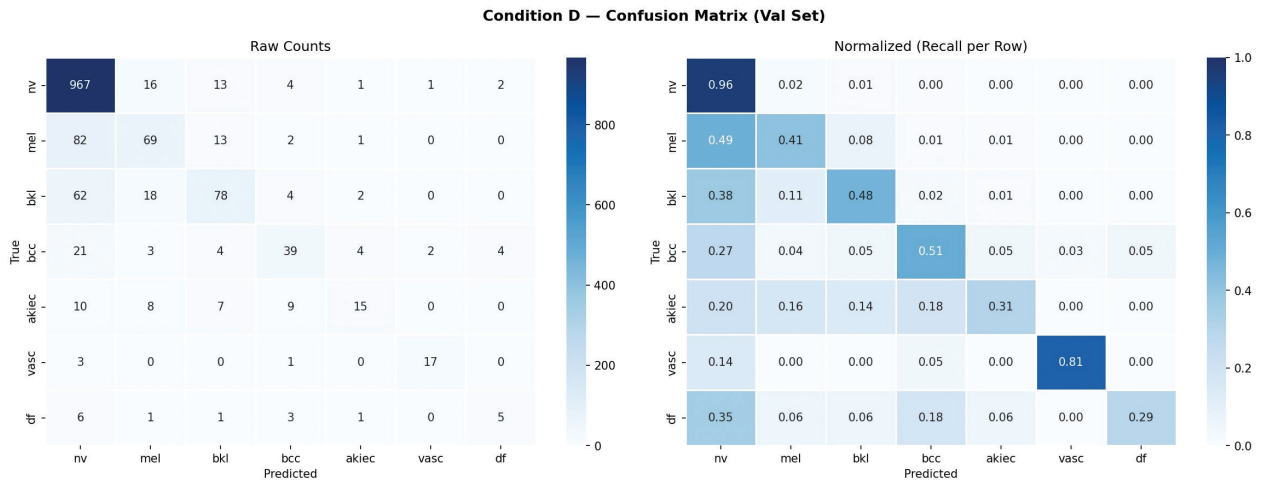


Fig. 9. Confusion matrix for Condition D, Fold 5 on the validation set ($n = 1,499$). Left: raw counts; right: normalized recall per row. The dominant off-diagonal pattern reflects minority class confusion with melanocytic nevi (nv).

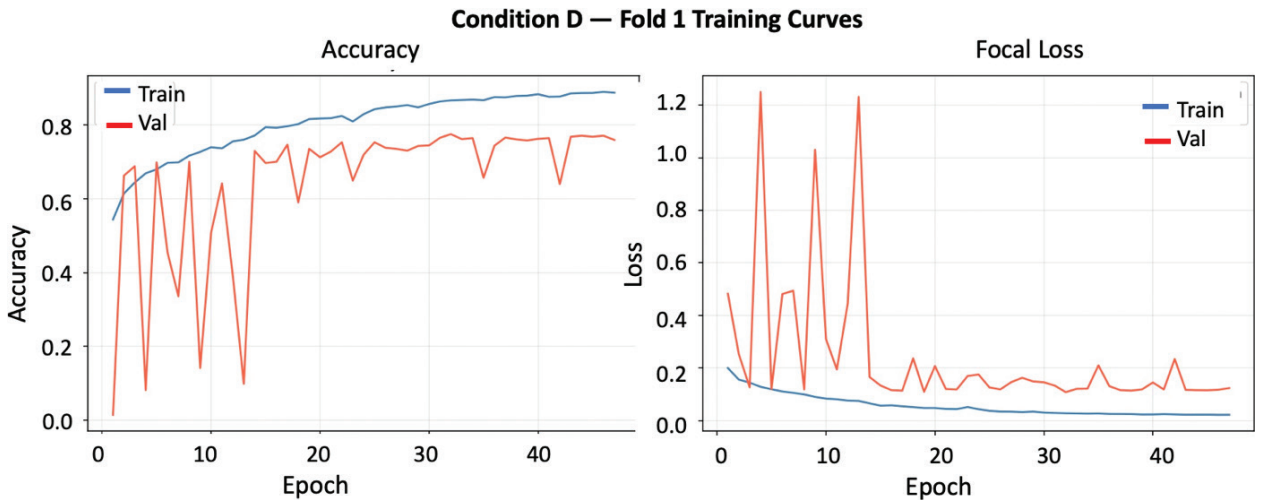


Fig. 10. Training and validation accuracy and focal loss curves for Condition D, Fold 5 (best fold, val acc = 0.7882). Early stopping triggered at Epoch 49; best weights restored from Epoch 34.

is conservative in flagging akiec lesions. Dermatofibroma F1 of 0.40 reflects the combined challenge of minimal support ($n = 17$) and morphological overlap with benign keratosis. These class-wise disparities motivate future investigation

into lesion-specific attention mechanisms and class-balanced sampling strategies beyond uniform augmentation.

The macro AUC-ROC of 0.9271 demonstrates strong discriminative capacity across all seven lesion categories

Table 3. Benchmark comparison—five-fold mean validation accuracy (Condition D protocol).

Model	Params	F1	F2	F3	F4	F5	Mean	±Std
LightDermNet	104K	0.7812	0.7843	0.7785	0.7852	0.7808	0.7820	0.0024
MobileNetV2	3.4M	0.7612	0.7684	0.7627	0.7689	0.7760	0.7674	0.0052
DenseNet121	8.1M	0.7676	0.7684	0.7719	0.7589	0.7546	0.7643	0.0065
MobileNetV3S	2.9M	0.7698	0.7548	0.7562	0.7511	0.7532	0.7570	0.0066
ResNet50V2	25.6M	0.7527	0.7477	0.7498	0.7382	0.7553	0.7488	0.0059

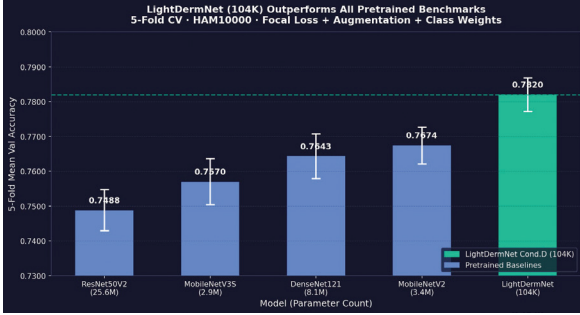


Fig. 11. Five-fold mean validation accuracy comparison across all benchmark models under Condition D protocol. LightDermNet achieves the highest accuracy with the lowest parameter count.

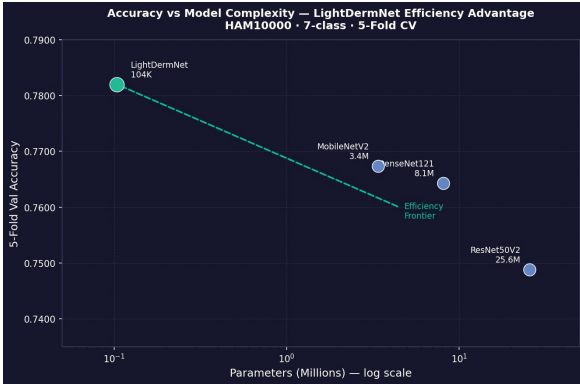


Fig. 12. Model accuracy versus parameter count (log scale). LightDermNet occupies the efficiency frontier, achieving the highest accuracy at 97–99.6% fewer parameters than all evaluated baselines.

Table 4. Per-class classification report—held-out test set ($n = 1,503$), Condition D, Fold 5.

Class	Precision	Recall	F1	Support	AUC
Melanocytic nevi (nv)	0.82	0.94	0.88	1,006	0.9128
Melanoma (mel)	0.47	0.31	0.37	167	0.8768
Benign keratosis (bkl)	0.54	0.36	0.43	165	0.8807
Basal cell carcinoma (bcc)	0.50	0.49	0.50	77	0.9631
Actinic keratosis (akiec)	0.75	0.37	0.49	49	0.9717
Vascular lesion (vasc)	0.89	0.77	0.83	22	0.9981
Dermatofibroma (df)	0.39	0.41	0.40	17	0.8864
Macro average	0.62	0.52	0.56	1,503	0.9271
Weighted average	0.73	0.76	0.73	1,503	-

Overall accuracy: 0.7558. AUC: one-vs-rest macro AUC-ROC; AUC-ROC: receiver operating characteristic—area under the curve.

at the ranking level, suggesting that while the model’s hard classification accuracy is constrained by minority class confusion, the underlying probability estimates remain well calibrated for clinical triage support in resource-constrained environments where ranked confidence outputs are actionable.

Explainability Analysis—Gradient-Weighted Class Activation Mapping Visualization

To interpret the spatial attention behavior of LightDermNet, gradient-weighted class activation mapping (Grad-CAM) was applied to one correctly classified sample per lesion class using the final convolutional layer of the proposed architecture. Figure 16 presents the original dermoscopic images alongside their corresponding Grad-CAM activation maps for all seven HAM10000 categories.

The activation maps confirm that LightDermNet consistently localizes diagnostically relevant regions rather than background skin texture. Melanoma activations concentrate on the irregular pigmented core and asymmetric border regions, consistent with established ABCD clinical criteria. Basal cell carcinoma maps highlight the central lesion structure, while vascular lesion activations converge tightly on the vascular nodule—corroborating its high AUC of 0.9981 and F1 of 0.83. Actinic keratosis activations are notably diffuse across the erythematous field, reflecting the morphologically heterogeneous nature of this class and partially explaining its lower recall of 0.37. These visualizations provide qualitative evidence that the CBAM attention mechanism directs gradient flow toward lesion-specific features, supporting the model’s suitability for clinical decision-support applications where interpretability is a regulatory requirement.⁴⁷

Conclusion

This study presented LightDermNet, a lightweight attention-augmented convolutional neural network for automated multiclass dermoscopic image classification on the HAM10000 benchmark. By integrating depthwise separable convolutions with CBAMs, the proposed architecture achieves competitive diagnostic performance with only 104,840 trainable parameters—a reduction of 97–99.6% relative to standard pretrained architectures—demonstrating that efficient, deployment-ready models need not sacrifice classification accuracy.

A strictly leakage-free five-fold stratified cross-validation protocol was adopted throughout, ensuring reproducible and unbiased performance estimation. The systematic

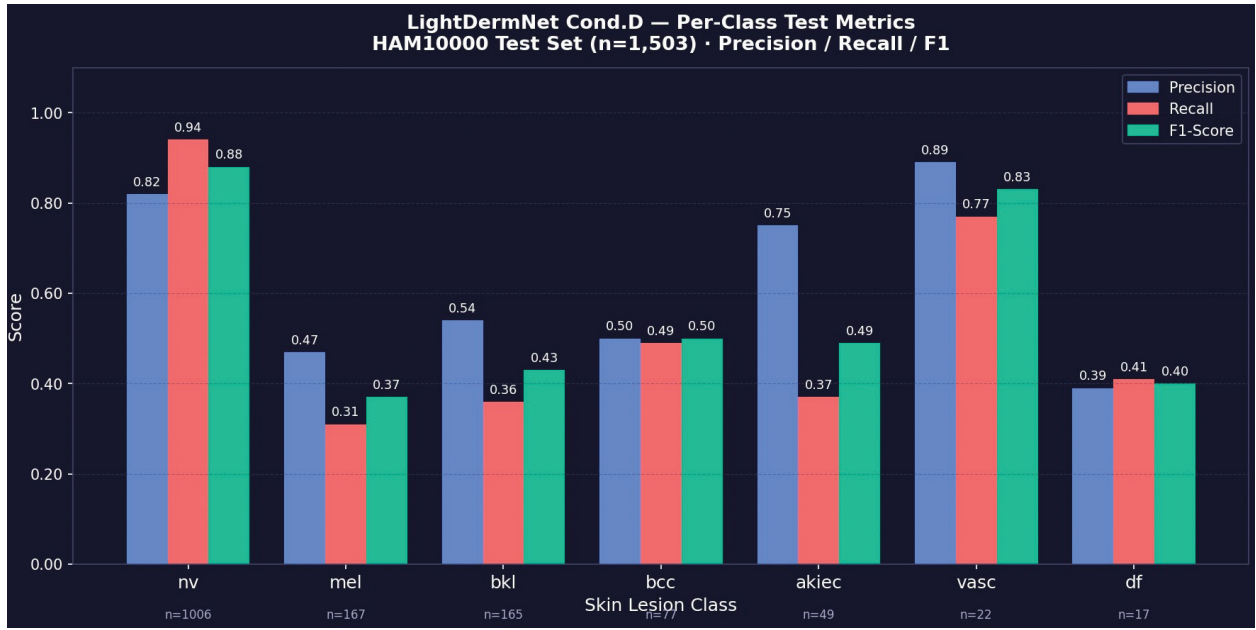


Fig. 13. Per-class precision, recall, and F1-score on the held-out test set ($n = 1,503$). Support counts are annotated below each class label.

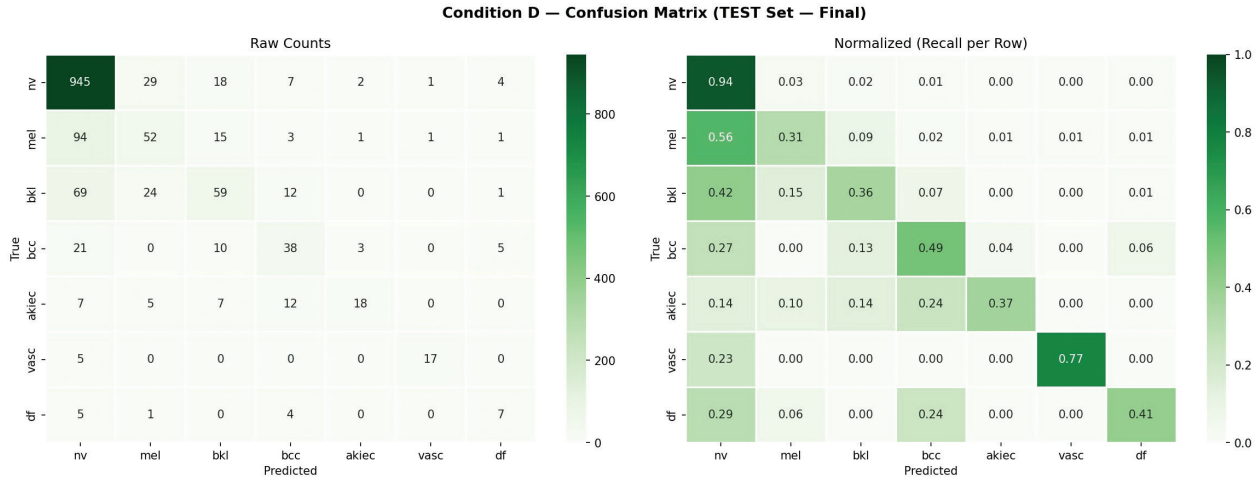


Fig. 14. Confusion matrix on held-out test set (Condition D, Fold 5). Left: raw counts; right: normalized recall per row. The dominant misclassification pattern is minority classes predicted as melanocytic nevi (nv).

ablation study confirmed that combining within-fold augmentation, inverse frequency class weighting, and focal loss yields the most stable training configuration, while benchmark evaluation demonstrated that LightDermNet surpasses all four pretrained baseline architectures under identical training conditions. The macro AUC-ROC of 0.9271 across seven lesion categories further confirms the model's strong discriminative capacity at the probability ranking level.

Future research directions include incorporating lesion-aware spatial attention mechanisms to improve minority class sensitivity, exploring advanced rebalancing strategies such as mixup augmentation and knowledge distillation, and conducting prospective deployment evaluation on embedded hardware platforms for real-world telehealth integration. These enhancements aim to further optimize LightDermNet for practical clinical application in resource-constrained dermatological screening environments.

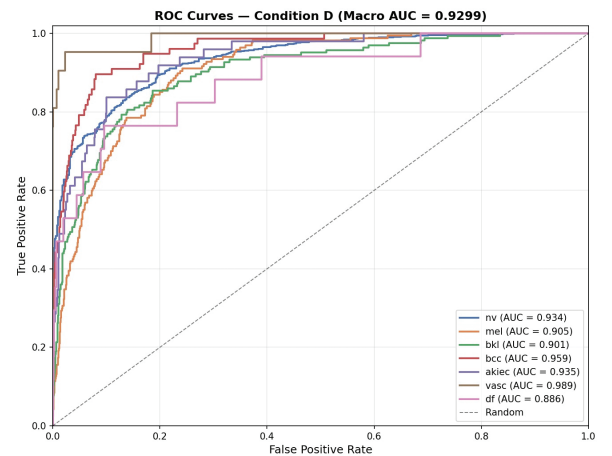


Fig. 15. One-versus-rest ROC curves for all seven lesion classes. Macro AUC = 0.9271. Vascular lesion achieves near-perfect discrimination (AUC = 0.9981). ROC: receiver operating characteristic; AUC: area under the curve.

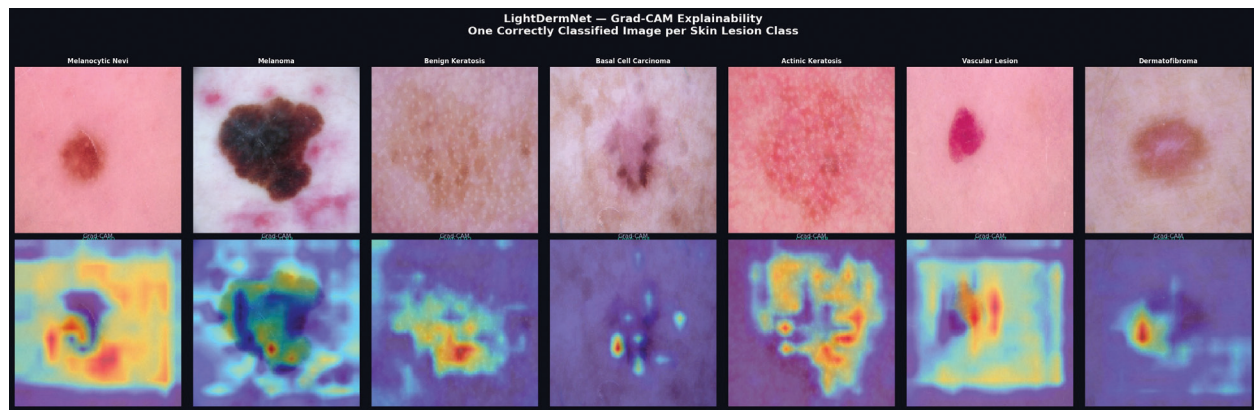


Fig. 16. Grad-CAM explainability visualizations for LightDermNet across all seven HAM10000 lesion classes. Top row: original dermoscopic images (one correctly classified sample per class). Bottom row: corresponding Grad-CAM activation maps overlaid on the final convolutional feature layer. Warm regions (red/orange) indicate discriminative areas used for classification decisions. Grad-CAM: gradient-weighted class activation mapping.

References

- Karimkhani C, Green A, Nijsten T, Weinstock M, Dellavalle R, Naghavi M, et al. The global burden of melanoma: results from global burden of disease study 2015. *Br J Dermatol*. 2017;177(1): 134–40. <https://doi.org/10.1111/bjd.15510>
- Ashraf R, Afzal S, Rehman AU, Gul S, Baber J, Bakhtyar M, et al. Region-of-interest based transfer learning assisted framework for skin cancer detection. *IEEE Access*. 2020;8:147858–71. <https://doi.org/10.1109/ACCESS.2020.3014701>
- Skin cancer statistics—world cancer research fund international [Internet]. Available from: <https://www.wcrf.org/dietandcancer/skin-cancer-statistics/>
- Cancer [Internet]. Available from: <https://www.who.int/news-room/fact-sheets/detail/cancer>
- Gajera HK, Nayak DR, Zaveri MA. A comprehensive analysis of dermoscopy images for melanoma detection via deep CNN features. *Biomed Signal Process Control*. 2023;79:104186. <https://doi.org/10.1016/j.bspc.2022.104186>
- Argenziano G, Puig S, Zalaudek I, Sera F, Corona R, Alsina M, et al. Dermoscopy improves accuracy of primary care physicians to triage lesions suggestive of skin cancer. *J Clin Oncol*. 2006;24:1877–82. <https://doi.org/10.1200/JCO.2005.05.0864>
- Tang P, Liang Q, Yan X, Xiang S, Zhang D. GP-CNN-DTEL: global-part CNN model with data-transformed ensemble learning for skin lesion classification. *IEEE J Biomed Health Inform*. 2020;24:2870–82. <https://doi.org/10.1109/JBHI.2020.2977013>
- Iqbal I, Younus M, Walayat K, Kakar MU, Ma J. Automated multi-class classification of skin lesions through deep convolutional neural network with dermoscopic images. *Comput Med Imaging Graph*. 2021;88:101843. <https://doi.org/10.1016/J.COMPMEIMAG.2020.101843>
- Barata C, Celebi ME, Marques JS. A survey of feature extraction in dermoscopy image analysis of skin cancer. *IEEE J Biomed Health Inform*. 2019;23:1096–109. <https://doi.org/10.1109/JBHI.2018.2845939>
- Adegun A, Viriri S. Deep learning techniques for skin lesion analysis and melanoma cancer detection: a survey of state-of-the-art. *Artific Intellig Rev*. 2021;54:811–41. <https://doi.org/10.1007/S10462-020-09865-Y/METRICS>
- Goswami T, Dabhi VK, Prajapati HB. Skin disease classification from image—a survey. In: 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS). 2020; pp. 599–605. <https://doi.org/10.1109/ICACCS48705.2020.9074232>
- Afza F, Khan MA, Sharif M, Rehman A. Microscopic skin laceration segmentation and classification: a framework of statistical normal distribution and optimal feature selection. *Microsc Res Tech*. 2019;82:1471–88. <https://doi.org/10.1002/JEMT.23301>
- Khan MA, Akram T, Sharif M, Saba T, Javed K, Lali IU, et al. Construction of saliency map and hybrid set of features for efficient segmentation and classification of skin lesion. *Microsc Res Tech*. 2019;82:741–63. <https://doi.org/10.1002/JEMT.23220>
- Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. In: 3rd International Conference on Learning Representations, ICLR 2015—Conference Track Proceedings [Internet]. 2014. Available from: <https://arxiv.org/abs/1409.1556v6>
- He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2016; December 2016; pp. 770–8. <https://doi.org/10.1109/CVPR.2016.90>
- Huang G, Liu Z, Maaten LVD, Weinberger KQ. Densely connected convolutional networks. In: Proceedings—30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017; 2017; pp. 2261–69. <https://doi.org/10.1109/CVPR.2017.243>
- Nayak DR, Dash R, Majhi B. Automated diagnosis of multi-class brain abnormalities using mri images: a deep convolutional neural network based method. *Pattern Recogn Lett*. 2020;138:385–91. <https://doi.org/10.1016/J.PATREC.2020.04.018>
- Lecun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521:436–44. <https://doi.org/10.1038/nature14539>
- Menegola A, Fornaciali M, Pires R, Bittencourt FV, Avila S, Valle E. Knowledge transfer for melanoma screening with deep learning. In: Proceedings—International Symposium on Biomedical Imaging [Internet]. 2017; pp. 297–300. Available from: <http://arxiv.org/abs/1703.07479>
- Hosny KM, Kassem MA, Foad MM. Classification of skin lesions using transfer learning and augmentation with alex-net. *PLoS One*. 2019;14:e0217293. <https://doi.org/10.1371/JOURNAL.PONE.0217293>
- Mahbod A, Schaefer G, Wang C, Dorffner G, Ecker R, Ellinger I. Transfer learning using a multi-scale and multi-network ensemble for skin lesion classification. *Comput Methods Prog Biomed*. 2020;193:105475. <https://doi.org/10.1016/J.CMPB.2020.105475>
- Khan MA, Zhang YD, Sharif M, Akram T. Pixels to classes: intelligent learning framework for multiclass skin lesion localization and classification. *Comput Electr Eng*. 2021;90:106956. <https://doi.org/10.1016/J.COMPELECENG.2020.106956>
- Khan MA, Akram T, Sharif M, Kadry S, Nam Y. Computer decision support system for skin cancer localization and classification. *Comput Mater Contin*. 2021;68:1041–64. <https://doi.org/10.32604/CMC.2021.016307>
- Khan MA, Akram T, Zhang YD, Sharif M. Attributes based skin lesion detection and recognition: a mask RCNN and transfer learning-based deep learning framework. *Pattern Recogn Lett*. 2021;143:58–66. <https://doi.org/10.1016/J.PATREC.2020.12.015>

25. Shehzad K, Zhenhua T, Shoukat S, Saeed A, Ahmad I, Bhatti S, et al. A deep-ensemble-learning-based approach for skin cancer diagnosis. *Electronics*. 2023;12(6):1342. <https://doi.org/10.3390/electronics12061342>
26. Shrestha H, Jaganathan SCB, Dhasarathan C, Suriyan K. Detection and classification of dermatoscopic images using segmentation and transfer learning. *Multimedia Tools Appl*. 2023;82:23817–31. <https://doi.org/10.1007/s11042-023-14752-z>
27. Tahir M, Naeem A, Malik H, Tanveer J, Naqvi RA, Lee SW. DSCC_Net: multi-classification deep learning models for diagnosing of skin cancer using dermoscopic images. *Cancers*. 2023;15(7):2179. <https://doi.org/10.3390/cancers15072179>
28. Keerthana D, Venugopal V, Nath MK, Mishra M. Hybrid convolutional neural networks with SVM classifier for classification of skin cancer. *Biomed Eng Adv*. 2023;5:100069. <https://doi.org/10.1016/j.bea.2022.100069>
29. Gilani SQ, Syed T, Umair M, Marques O. Skin cancer classification using deep spiking neural network. *J Digital Imaging*. 2023;36(3):1137–47. <https://doi.org/10.1007/S10278-023-00776-2/METRICS>
30. Jinnai S, Yamazaki N, Hirano Y, Sugawara Y, Ohe Y, Hamamoto R. The development of a skin cancer classification system for pigmented skin lesions using deep learning. *Biomolecules*. 2020;10:1–13. <https://doi.org/10.3390/biom10081123>
31. Alwakid G, Gouda W, Humayun M, Sama NU. Melanoma detection using deep learning-based classifications. *Healthcare*. 2022;10:2481. <https://doi.org/10.3390/HEALTHCARE10122481>
32. Rashid J, Ishfaq M, Ali G, Saeed MR, Hussain M, Alkhalifah T, et al. Skin cancer disease detection using transfer learning technique. *Appl Sci*. 2022;12(11):5714. <https://doi.org/10.3390/app12115714>
33. Mamun AA, Ray PC, Nasib MRU, Das A, Uddin J, Absur MN. Optimizing deep learning for skin cancer classification: a computationally efficient cnn with minimal accuracy trade-off [Internet]. 2025. arXiv:2505.21597. Available from: <https://arxiv.org/abs/2505.21597>
34. Sarker MAR, Kabir MA, Hossain MS. LGGC_Net: a local-global graph and color attention-based lightweight CNN for skin cancer classification. *Sci Rep*. 2026;16(1):17790. <https://doi.org/10.1038/s41598-026-48724-8>
35. Kadampur MA, Riyae SA. Skin cancer detection: applying a deep learning based model driven architecture in the cloud for classifying dermal cell images. *Inform Med Unlock*. 2020;18:100282. <https://doi.org/10.1016/j.imu.2019.100282>
36. Kousis I, Perikos I, Hatzilygeroudis I, Virvou M. Deep learning methods for accurate skin cancer recognition and mobile application. *Electronics*. 2022;119:1294. <https://doi.org/10.3390/electronics11091294>
37. Srinivasu PN, SivaSai JG, Ijaz MF, Bhoi AK, Kim W, Kang JJ. Classification of skin disease using deep learning neural networks with MobileNet V2 and LSTM. *Sensors*. 2021;21(8):2852. <https://doi.org/10.3390/s21082852>
38. Gessert N, Sentker M, Madesta F, Schmitz R, Kahl H, Iszatt I, et al. Skin lesion classification using CNNs with patch-based attention and diagnosis-guided loss weighting. *IEEE Trans Biomed Eng*. 2020;67(2):495–503. <https://doi.org/10.1109/TBME.2019.2915839>
39. Razzak MI, Naz S, Zaib A. Deep learning for medical image processing: overview, challenges and future. In: *Classification in BioApps*. 2018; pp. 323–50. <https://doi.org/10.1007/978-3-31965981-712>
40. Chaturvedi SS, Tembhurne JV, Diwan T. A multi-class skin cancer classification using deep convolutional neural networks. *Multimedia Tools Appl*. 2020;79:28477–98. <https://doi.org/10.1007/s11042-020-09388-2>
41. Rajput G, Agrawal S, Raut G, Vishvakarma SK. An accurate and non-invasive skin cancer screening based on imaging technique. *Int J Imaging Syst Technol*. 2022;32(1):354–68. <https://doi.org/10.1002/ima.22616>
42. Rehman MZ, Zahid A, Tahir R, Hassan B, Asim M, Ali MU. Skin lesion classification using NASNet mobile and DenseNet-201 on HAM10000 dataset. *Computers*. 2022;11(6):80. <https://doi.org/10.3390/computers11060080>
43. Polat K, Koc R, Alcin OF. Skin lesion classification using machine learning algorithms. *Int J Appl Math Electr Comput*. 2020;8(4):20–5. <https://doi.org/10.18100/ijamec.839564>
44. Tschandl P, Rosendahl C, Kittler H. The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci Data*. 2018;5:180161. <https://doi.org/10.1038/sdata.2018.161>
45. Telea A. An image inpainting technique based on the fast marching method. *J Graph Tools*. 2004;9(1):23–34. <https://doi.org/10.1080/10867651.2004.10487596>
46. Lin T-Y, Goyal P, Girshick R, He K, Dollár P. Focal loss for dense object detection. *IEEE Trans Pattern Anal Machine Intellig*. 2020;42(2):318–27. <https://doi.org/10.1109/TPAMI.2018.2858826>
47. Woo S, Park J, Lee J-Y, Kweon IS. CBAM: convolutional block attention module. In: *Proceedings of the European Conference on Computer Vision (ECCV)*. Springer; 2018, pp. 3–19. <https://doi.org/10.1007/978-3-030-01234-2>

Copyright Ownership: This is an open-access article distributed in accordance with the Creative Commons Attribution Non-Commercial (CC BY-NC 4.0) license, which permits others to distribute, adapt, enhance this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See <http://creativecommons.org/licenses/by-nc/4.0>. The authors of this article own the copyright.