



## Enhanced Skin Disease Classification Through Convolutional Neural Networks: A Performance Advantage Study

Subha T<sup>1</sup>, Narendra Kumar Rao<sup>2\*</sup>, Ranjana R<sup>3</sup>, Indurekaa S R<sup>3</sup>

<sup>1</sup> Department of Educational Media and Technology, NITTTR, Chennai 600113, India

<sup>2</sup> AI&ML, Sri Ramachandra Institute of Higher Education and Research, Chennai 600116, India

<sup>3</sup> AI&DS, Dr.Mahalingam College of Engineering and Technology, Pollachi 642003, India

Corresponding Author Email: [narendrakumarraob@gmail.com](mailto:narendrakumarraob@gmail.com)

Copyright: ©2026 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/isi.310330>

### ABSTRACT

**Received:** 10 August 2025

**Revised:** 1 November 2025

**Accepted:** 20 March 2026

**Available online:** 31 March 2026

#### Keywords:

*skin lesion classification, Convolutional Neural Network, HAM10000, EfficientNet, DenseNet, TensorFlow Lite*

Skin cancer is among the most prevalent cancers worldwide, and early detection significantly improves survival outcomes. This study presents a deep learning approach for classifying seven types of skin lesions using the HAM10000 dataset. We propose a Convolutional Neural Network (CNN) architecture enhanced with data augmentation (rotation, zoom, shift) and RandomOverSampler to address class imbalance. The proposed CNN was compared against two state-of-the-art architectures—EfficientNetB0 and DenseNet121—under identical preprocessing and training conditions. The CNN model achieves a classification accuracy of 98.6%, substantially outperforming EfficientNetB0 (14.4%) and DenseNet121 (24.9%). Precision, recall, and F1-score metrics (98.5%, 98.4%, and 98.4%, respectively) further confirm its superior performance. The lower accuracy of the pretrained models is attributed to their reliance on higher input resolution and sensitivity to class imbalance when not sufficiently fine-tuned. To enable practical deployment, the trained CNN model was optimized and converted to TensorFlow Lite format, facilitating cross-platform execution on mobile and embedded systems. This work demonstrates that a well-designed CNN, even without pretrained weights, can outperform complex architectures on moderate-sized medical image datasets when coupled with appropriate preprocessing and imbalance handling. The proposed system offers an effective, scalable, and accessible solution for skin cancer screening, particularly suitable for resource-limited clinical environments.

## 1. INTRODUCTION

Skin cancer is one of the most popular cancer types across the globe and the increases in the number of people diagnosed with skin cancer can largely be ascribed to the increased exposure to the sun and to ultraviolet (UV) radiation. The formation of skin cancers occurs due to the excessive growth of the skin cells on the surface of our bodies and, subsequently, lead to the formation of malignant tumors. It is critical that skin cancers be diagnosed at an early stage, because early diagnosis and treatment will significantly decrease the potential of metastasis and provide a much better chance of survival (i.e. long-term survivor). Currently accepted methods for diagnosing skin cancer (e.g., visual examination by dermatologists) can produce variable results because they are subjective and incomplete when differentiating between skin conditions that appear to be the same. Advances in deep learning (DL) and machine learning (ML) have yielded potential options for computer-aided automated skin cancer detection in the recent past. Convolutional Neural Networks (CNNs), more than other DL algorithms, have exhibited extraordinary efficacy in classification on the basis of images. DL models possess the capability to automatically learn raw image data hierarchies in feature, making them exceedingly appropriate for skin cancer as well as general medical image

analysis.

In recent years, advances in DL and ML have provided possible solutions for automated computer-assisted skin cancer detection. CNNs have far surpassed all the other types of DL algorithms in their ability to classify using images. DL models are able to automatically learn hierarchical structures of feature from the raw image data; therefore, they are ideal for skin cancer detection and also for general medical image analysis.

Training deep learning models to classify medical images is often hindered by class imbalance among the skin conditions. Such class imbalances can lead to biased predictions of the trained model, as bias is typically towards majority classes. There are various strategies available to mitigate class imbalance such as RandomOverSampler and the use of data augmentation techniques (rotation, zooming, and shifting) to produce equal representation of each skin lesion class while also enhancing the robustness of the trained deep learning model.

The present study proposes a deep learning approach using CNNs for classifying skin lesions from the HAM10000 dataset. To enhance performance and generalization of the model, various augmentation/transformation strategies were utilized such as up-sampling and down-sampling. Additionally, the model has been implemented in TensorFlow

lite for mobile/wearable application so that it can be utilized to assist with the diagnosis of skin cancers in low to moderate resource areas. By improving the identification of skin cancers through detection algorithms, this solution can potentially improve the manner in which healthcare professionals diagnose skin cancers.

## 2. RELATED WORK

Deep learning network systems have greatly advanced in recent years for skin cancer diagnosis through use of Convolutional Neural Network (CNN) technology. Haque et al. [1] proposed using CNN-based image analysis for classifying skin lesions based on dermato-pathological images; their work produced effective results. The technique involved designing an architecture with multiple convolutional and pooling layers to improve skin lesion classification performance. Batch normalization and parallel processing strategies were also employed to accelerate computation and enhance convergence speed.

Hasan et al. [2] studied using deep learning technologies to diagnose liver disease with results demonstrating that CNN models work successfully-- even within the field of medical imaging. Yadav and Bhat [3] also analyzed many different types of deep learning algorithms as applied to dermatology, providing evidence for CNN models outperforming traditional approaches such as SVM & k-NN. The results show substantial differences concerning ability to extract features, giving an advantage to deep network models when faced with unique problems (e.g - classifying skin cancer).

In their study, Prasanna Lakshmi and Vidhyashree [4] evaluated the use of CNN and SVM to classify skin lesions using machine learning techniques. They found that CNN-based techniques performed significantly better than previous algorithms due to their ability to learn hierarchical features from raw images directly. They also identified the issue of class imbalance as a limiting factor to the performance of the models, which was then addressed using methods such as oversampling and data augmentation.

Tan and Le [5] utilized EfficientNetB0 for the detection of skin cancer and proved that it could outperform previous CNNs, predominantly after being fine-tuned on small datasets. However, they also found that EfficientNet requires high-resolution images to perform at an ideal level and many hyperparameters for tuning to get the most favourable result from this architecture.

Rahman et al. [6] have implemented Self-Augmented Multistage Deep Learning Network along with Coot Optimization Algorithms to optimize Deep Belief Networks in order to analyze dermoscopic images. Ashtagi et al. [7] discussed an EfficientNetB0-based deep learning model for identifying and classifying skin cancer, which automates the process of analysing skin lesions.

## 3. EXISTING AND PROPOSED WORK

### 3.1 Existing work

Research on skin cancer detection using machine learning and deep learning techniques is a very active area of research within the medical imaging community, and there has been great progress in this area. Traditional methods for detecting

skin cancer were primarily based on subjective visual assessments by dermatologists, and are therefore influenced by human error.

CNNs have recently become the benchmark for measuring performance on tasks involving image classification. CNNs are capable of automatically learning how to extract relevant features from raw pixel data, resulting in a very efficient method for performing image classification tasks (such as those found in the medical imaging domain). Numerous research studies have evaluated CNN's ability to classify skin lesions and detected various forms of skin cancer at a high level of accuracy using CNN-based models. However, current CNNs are often adversely affected by multiple problems, including the problem of class imbalance (certain types of skin cancer, i.e. melanoma, are poorly represented within the training dataset). Class imbalance results in CNN performance suffering from generalization issues and biased predictions. In addition, some of the key factors that influence CNN performance are the number of training samples available and the degree to which the training data are diverse.

Also, pre-trained models such as EfficientNet and DenseNet have been investigated for skin cancer detection. The pre-trained models, grounded in transfer learning, have demonstrated that employing pre-trained weights on large datasets can enhance accuracy while decreasing training time. Nevertheless, such models tend to be overly complicated or not optimized for low resolutions in images, thus restricting their functionality in some real-world use cases. Current efforts in this area also underscore the requirement for cross-platform portability for real-time deployment in healthcare environments.

There has been significant interest in solutions for assisting early diagnosis of skin cancer as well as reducing the workload of dermatologists, with automated skin cancer detection systems being developed for this purpose. Automated skin cancer detection systems initially relied heavily on traditional image processing and machine learning techniques based on handcrafted feature extraction such as colour, texture, shape etc. Although these techniques provided an initial understanding of the potential for skin cancer detection, performance is heavily dependent on accurate segmentation of lesions and proper handcrafted feature extraction of lesions by trained experts. As such, many practical implementations of these techniques are affected by a variety of factors including:

1. Variations in lighting and how lesions appear.
2. The presence of artefacts (for example hair) and air bubbles that can reduce the quality of the image and lead to classification errors.
3. Interference from lighting, etc. when attempting to classify lesions at a distance without the presence of a trained professional examining the image directly (via dermatoscopy).

Conventional machine learning classifiers such as support vector machines and k-nearest neighbour classifiers are highly sensitive to misclassification errors in the segmentation step and any changes in lighting, etc. that can result in poor image quality. Since lesion boundaries are influenced by noise and can be difficult to determine, classification inaccuracies can have a strong negative impact on the accuracy of the classifiers used for skin cancer detection, thereby significantly reducing the effectiveness of automated skin cancer detection systems in real world scenarios. Deep Learning-based approaches (CNN's etc.) have been shown to learn the features necessary to classify dermoscopic images more effectively than conventional machine learning approaches, however, there are

still a number of challenges that exist with Deep Learning based skin cancer detection systems.

High-resolution images are often necessary for many architectures to maintain the fine-grained details of visual features such as irregular borders, asymmetry, and subtle variations in color. However, the increased cost associated with computing such high-resolution images makes them impractical for real-time or resource-limited clinical environments.

A major barrier in skin cancer research is the lack of adequate labeled medical datasets and their inherent imbalance. The number of malignant lesion samples tends to be much less than the number of benign cases, causing a bias in learning which results in decreased sensitivity towards important classes. Models that are trained using small or imbalanced datasets tend to have a higher tendency to overfit, resulting in a high level of accuracy when trained, but very limited generalization capabilities to unseen data. Although techniques such as data augmentation and resampling have been utilized, the effectiveness of these techniques may not always be consistent across lesion type categories.

Generalization across a variety of imaging conditions is another unsolved problem. The acquisition device and magnification level are only two variables among many, including skin tone and clinical location, that can produce large variations in dermoscopic image quality. As evident by the results of many models trained on a small number of benchmark datasets, studies by several researchers demonstrate that when these models are utilized with images collected from other sources, there is often a performance decline due to the domain shift and the dataset bias.

Interpretability has also been a significant issue in automated diagnosis systems due largely to the fact that many deep learning models act as "black boxes" with minimum information about how they come to their conclusions when making predictions. This limited view of the model's performance has been an obstacle for clinical acceptance and confidence because physicians and other healthcare professionals require understandable explanations of their decision support tool (DST) to assist in making diagnoses.

While a large number of studies have reported the overall accuracy of classification, they often neglected to assess the metrics used to evaluate class-wise performance (sensitivity, specificity, and recall). In medical diagnosis, the misclassification of a malignant lesion as benign can result in severe consequences; thus, it is essential to report on performance for each class individually to provide a complete picture.

The lack of standardisation appears to be another limitation of the literature reviewed. The vast differences in methodology for preprocessing, training and testing sets, the balances of training and testing data, and the selected performance metrics hinder the ability to conduct a proper fair comparison between competing models; therefore, the improvements reported may be a result of the limitations of those experiments as opposed to superiority of methodology.

### 3.2 Proposed work

This work presents a state-of-the-art deep learning-based method for the classification of seven classes of skin lesions employing a CNN architecture trained on the HAM10000 dataset.

One of the crucial elements of the suggested work includes

the implementation of deeper CNN architecture with BatchNormalization and Dropout layers to increase the feature extraction strength and avoid overfitting. Moreover, proposed model comparison against EfficientNetB0 and DenseNet121 would be conducted because these are cutting-edge deep models with proven performance and efficiency to carry out classification tasks in the field of image classification.

By converting this application to run as a native app on both mobile devices and embedded systems, real time skin cancer detection can be accomplished in places where there are not enough resources available. The method proposed here is to provide an effective, easily deployable and scalable solution for skin cancer detection allowing healthcare practitioners to provide more accurate and timely diagnoses.

Skin cancer is a type of cancer that can be fatal and is still one of the most common types of cancer and one of the most deadly forms of cancer in the world today. The early detection of skin cancer is extremely important to being able to provide effective treatment. Recently researchers have focused a lot of their attention on using a CNN as the main type of deep-learning architecture for classifying skin lesions automatically. One study, Rashad et al. [8] described developing a fully automated CNN-based screening pipeline that utilizes various advanced techniques for preprocessing the data and augmenting the images. This work was shown to improve significantly the accuracy of classification for dermoscopic images using CNNs to over 90% on large dermoscopic data sets, thus illustrating the effectiveness of preparing the images for use in the model prior to using them to train the model.

In addition, the researchers S et al. [9] used explainable AI methods (e.g., Grad-CAM) and ensemble learning to improve the accuracy of the model and to address the need for explainability in clinical AI systems so that dermatologists can better understand the models' predictions.

Natha and Rajeswari [10] also investigated ensemble methods and integrated several different types of CNNs together to get better classification of melanomas and non-melanomas versus using individual models. AlSadhan et al. [11] created a single CNN framework that can classify different lesion types (e.g., melanoma, basal cell carcinoma) with the same architecture for improved generalizability of their classification model using the ISIC dataset. Therefore, this approach would be beneficial for implementation in a variety of real-world settings, as the same system must be used for classifying lesions of multiple types (e.g., melanoma, basal cell carcinoma, etc.) in the same test.

The optimization of speed and accuracy was a major focus of many researchers working on developing CNN architectures as well. The BMC Med. Imaging paper [12] focused heavily on the systematic hyperparameter tuning of CNNs through proper design methodologies, thereby accelerating convergence while still maintaining classification accuracy via enhanced model checkpointing. Similarly, Shah et al. [13] applied PSO to both the hyperparameter tuning of their CNN architecture as well as the application of explainability to validate CNN outputs through expert human verification of model output.

Jeyalakshmi et al. [14] have developed an efficient and quick-to-deploy CNN that can be operated on both mobile and low-resource platforms, to facilitate Point of Care diagnostics within a resource-constrained environment. This objective aligns with Applied Sciences [15] evaluation of CNN performance with edge computing using Raspberry Pi and

Jetson Nano, including the trade-offs in inference speed vs. accuracy of diagnosis.

Additionally, Architectural Studies include comparison at an architectural level. As an example, Vieira et al. [16] looked at skin lesion segmentation and classification algorithms based using different architectural bags composed of CNN, Transformers and Hybrid Deep Learning. One of the conclusions drawn within their research was that Hybrid Models provide greater performance than Conventional CNN's for some segmentation or classification tasks. Lastly,

Deshmukh et al. [17] focused on CNN's used for skin cancer detection by integrating their built-in model optimization techniques with the properties of interpretability into 'Smart Healthcare' so that doctors can use them within a real-time environment in an appropriate way.

A CNN has been developed to classify dermoscopic images and focuses on the colours, textures, and structural features of lesions. The architecture of the neural network has three convolutional blocks that are feed-forward to fully connected layers.

**Table 1.** Stages in Convolutional Neural Network (CNN) architecture

Layer	Type	Filters	Kernel Size	Stride	Output Shape
Input	Input	–	–	–	224 × 224 × 3
Conv1	Convolution	32	3 × 3	1	224 × 224 × 32
Pool1	Max Pooling	–	2 × 2	2	112 × 112 × 32
Conv2	Convolution	64	3 × 3	1	112 × 112 × 64
Pool2	Max Pooling	–	2 × 2	2	56 × 56 × 64
Conv3	Convolution	128	3 × 3	1	56 × 56 × 128
Pool3	Max Pooling	–	2 × 2	2	28 × 28 × 128
Flatten	Flatten	–	–	–	100,352
FC1	Fully Connected	256	–	–	256
Dropout	Dropout	0.5	–	–	256
Output	Fully Connected	7	–	–	7

Each convolutional block has a convolutional layer, a Rectified Linear Unit (ReLU) activation function, and a max-pooling layer. Convolutional layers use a kernel size of three, a stride of one, and an additional layer of bye to ensure that spatial data is stored that is required to classify or separate boundaries of lesions, with the number of filters continually increased throughout the network (32, 64 and 128), to extract features that become increasingly abstract, but more discriminating.

Max-pooling layer size two reduces the spatial resolution of the image while preserving the important characteristics of the lesion. Once all of the features in the image have been extracted by the convolutional layers, the feature map is flattened before being transferred into a fully connected layer with 256 neurons. A dropout of 50% is incorporated to reduce overfitting of the dataset due to class imbalance, while the output layer contains seven neurons (one for each of the seven types of lesions) and a softmax activation function. The sequential order of the CNN layers is summarised in Table 1.

Overall, the advances made in CNN-based skin cancer detection, including increases in computer vision precision, explanation ability, and performance efficiency, demonstrate how rapidly this research area has progressed. Still, researchers are struggling with issues like finding a trade-off between the performance of edge devices and the CPUs required for processing images; achieving good performance when training on one data set and testing on another; and developing models that can both accurately classify images and provide explanations so clinicians can effectively use those models to make their treatment decisions. Addressing these issues will require continued research to close the gap between laboratory results and the successful implementation of these models in clinical practice.

Specifically, a number of these studies [8-20] demonstrate advances in skin cancer detection based on CNNs including work related to improving accuracy, understanding, use of ensemble methods, and optimally deploying on edge devices. Nevertheless, each of these studies has several aspects that require improvement:

1. **Generalizing across different populations:** A

majority of the models used in this review have been designed using publicly available datasets such as ISIC and HAM10000 [8, 10, 11, 16] which have been shown to be biased towards specific skin types; furthermore, there are no validation studies using regional, real world clinical data to show whether the models would work consistently across all relevant populations.

2. **Balanced speed vs. accuracy trade-off:** Many of the models developed so far tend to either be geared toward providing very accurate predictions [8, 9, 10] or are designed for faster inference using lighter weight architecture [14, 15]; however, there have not yet been any hybrid approaches developed that meet both criteria without making significant compromises.
3. **Clinical Interpretability and Trust:** Many of the studies reviewing AI applications in dermatology have included the use of explainable AI [9, 13, 17]; however, these studies have not included rigorous user studies to evaluate the usefulness of explanation methods in real-world diagnostic workflows by dermatologists.
4. **Integration into Point-of-Care Systems.** Some studies have been conducted examining deploying AI model on the edge [14, 15]; however, no implementations of AI models have been developed for full integration within a comprehensive tele-dermatology or mobile health system, creating barriers to practical implementation in rural areas and resource-limited countries.
4. **Multi-Modal Data Use:** Most models use only dermoscopic or clinical images. There are very few attempts to integrate multiple modalities of input, including patient history or metadata as well as other imaging modalities, to create more accurate predictions.
5. **Longitudinal and Predictive Ability:** Newly developed models mainly focus on identifying static images. There has been very little work down to date monitoring change in lesions over time, in order to predict the development of malignancy. This information could provide support for the clinician to

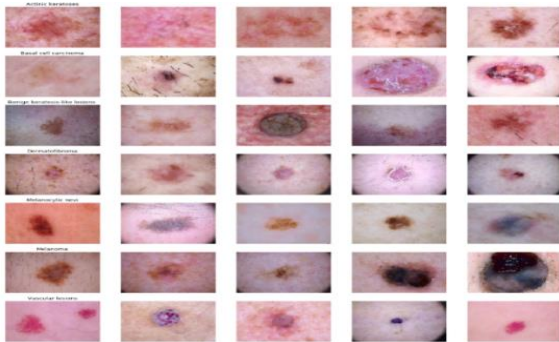
make proactive clinical decisions.

Even though skin cancer detection using CNNs has shown high performance when tested on data sets, there is an urgent requirement for clinically validated lightweight to highly accurate models that are explainable and from multiple types of data sources that will work in different population and settings. It is important to address these gaps in order to facilitate the transition from research prototype to actual clinical implementation.

## 4. METHODS AND MATERIALS

### 4.1 Dataset and preprocessing

This study used the HAM10000 dataset (N = 10,015) that contains dermoscopic images of seven different types of skin lesions. The images are 28 × 28 pixels in size and have been classified into seven different types of benign or malignant skin condition (melanoma, benign keratosis, dermatofibroma, etc.). Figure 1 illustrates sample images from the HAM10000 dataset, showing examples of the seven different skin lesion types. Because the dataset is imbalanced, the RandomOverSampler was employed to re-balance the distribution of classes. After the oversampling process, there will be 6705 cases in each of the seven classes of images.



**Figure 1.** Sample images from the HAM10000 dataset



**Figure 2.** Sample images after preprocessing and augmentation

Pre-processing included data rescaling, where the pixel values are rescaled from the integers 0-255 to float numbers in the range [0,1] and data augmentation methods such as rotation, zooming and shifting were applied. Figure 2 shows the sample images after preprocessing and augmentation. Data augmentation methods were randomly applied each training epoch for increased generalization (to reduce overfitting) and to allow for good performance of the model. A total of 75% (8,681) of the data will be used for training, and the remaining

25% (2,893) will be used for testing purposes.

### 4.2 Model architecture

In this work, the main model used is a CNN, which has been designed for image classification purposes. Each Convolutional layer of the CNN contains MaxPooling Layers that reduce the size of the feature maps produced by each convolutional layer in order to decrease the spatial dimensions of the feature maps. The Activations of each convolutional layer are done using the ReLU activation function, which will also add non-linearities to the model, and then Batch Normalization will be applied to help stabilize training and improve the performance of the models. In addition to the convolutional layers and pooling layers, there are Fully Connected layers which will be used to perform the final classification. A Softmax activation function will be used in the output layer to provide the probability of each of the 7 classes.

To make a comparison, EfficientNetB0 and DenseNet121 pretrained models were also used. EfficientNetB0 was selected as it can efficiently scale with less parameterization and still preserve performance. DenseNet121, recognized for its ability to reuse features efficiently, was used to study its potential in enhancing accuracy for skin lesion classification.

### 4.3 Training process

The deep learning model to be used is trained with a set of hyperparameters that are chosen to guarantee stable convergence, efficient optimization and good generalization performance is shown in Table 2. The reasons for each parameter are explained below:

#### (A) Learning Rate Selection:

The learning rate determines the step size used to update the model parameters during gradient-based optimization. Let  $\theta$  denote the trainable parameters and  $L(\theta)$  the loss function. The parameter update rule is:

$$\theta(t+1) = \theta(t) - \eta \nabla_{\theta} L(\theta(t)) \quad (1)$$

A learning rate of  $\eta = 0.001$  is selected to achieve a balance between convergence speed and numerical stability.

#### (B) Optimization Algorithm:

The Adam (Adaptive Moment Estimation) optimizer is used to adjust the parameters. Adam calculates the first and second moments of the gradients with an exponentially decaying average as follows:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (2)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (3)$$

The above adaptive strategy makes the convergence faster and the training more stable. Bias-corrected estimates are obtained using:

$$\hat{m}(t) = \frac{m(t)}{1 - \beta_1^t} \quad (4)$$

$$\hat{v}(t) = \frac{v(t)}{1 - \beta 2^t} \quad (5)$$

The final parameter update is expressed as

$$\theta(t+1) = \theta(t) - \eta \cdot \frac{\hat{m}(t)}{\sqrt{\hat{v}(t) + \epsilon}} \quad (6)$$

(C) *Training Epoch:*

An epoch is a single pass through the whole training dataset. The model goes through the training 25 epochs, thus the empirical risk is gradually minimized. Let the dataset size be  $N$  and batch size be  $B$ .

$$U = \frac{N}{B} \quad (7)$$

(D) *Mini-Batch Size:*

Mini batch stochastic optimization is performed with a batch size of 128 which is a good compromise between gradient variance reduction and computational efficiency. Stochastic gradient descent approximates the true gradient by

$$\nabla L(\theta) \approx \left(\frac{1}{B}\right) \Sigma \nabla Li(\theta) \quad (8)$$

With  $B = 128$ :

- Reduced gradient variance  $\rightarrow$  smoother convergence
- Maintains stochasticity needed to escape saddle points
- Efficient parallel computation on GPUs

(E) *Loss Function:*

For multiclass classification, the categorical cross entropy loss function is adopted. This expression corresponds to maximum likelihood estimation and yields well conditioned gradients. The cross entropy loss is given by:

$$L = - \Sigma y(k) \cdot \log(\hat{y}(k)) \quad (9)$$

(F) *Evaluation Metrics:*

The quality of the model is measured by classification accuracy on training and validation datasets to estimate generalization performance.

$$Accuracy = \left(\frac{1}{N}\right) \Sigma I(\hat{y}(i) = y(i)) \quad (10)$$

(G) *Data Augmentation:*

In order to enhance generalization, data augmentation methods like rotation, zooming, and spatial shifting are implemented. This helps the network to become invariant to the transformation and also serves as a regularization method.

$$x' = T(x) \quad (11)$$

The Augmented Objective Function is given by

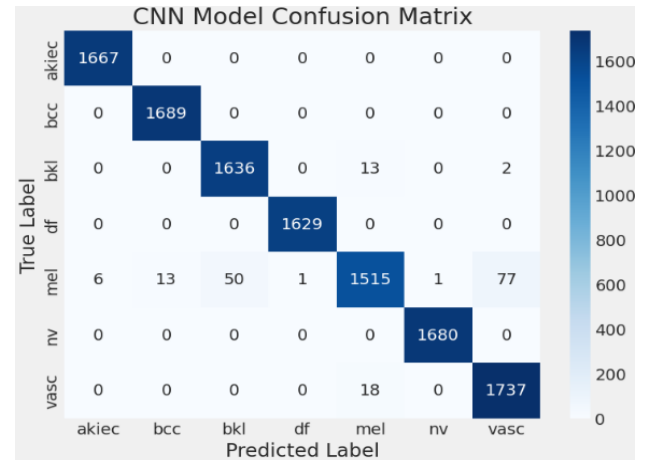
$$\mathcal{L}_{aug}(\theta) = \mathbb{E}_{(x,y),T} [L(f_{\theta}(T(x)), y)] \quad (12)$$

#### 4.4 Evaluation and results

Upon training, the models were tested on the test set of 2,893 images. The confusion matrix for the proposed CNN model is presented in Figure 3, illustrating the classification performance across the seven lesion types. The CNN model achieved an accuracy of 98.6%, substantially outperforming EfficientNetB0 (14.4%) and DenseNet121 (24.9%).

**Table 2.** Training parameters table

Parameter	Value
Learning Rate	0.001
Optimizer	Adam
Epochs	25
Batch Size	128
Loss Function	Categorical Cross-Entropy
Evaluation Metrics	Accuracy, Validation Accuracy
Augmentation	Rotation, Zoom, Shifting

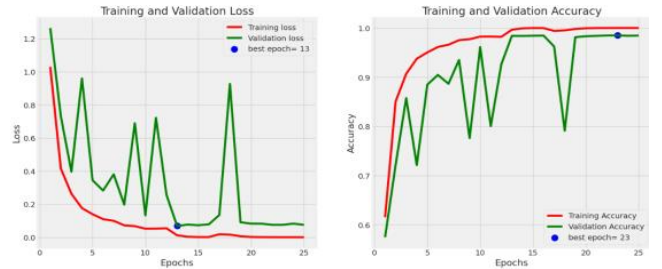


**Figure 3.** Confusion matrix for proposed Convolutional Neural Network (CNN)

## 5. ANALYSIS OF RESULTS

### 5.1 Confusion matrix

The confusion matrix is shown in Figure 4, illustrating the performance of the CNN model on the test set.



**Figure 4.** Training and validation accuracy plot

Figure 4 shows the comparing the training and validation accuracy Training and validation accuracy plot over 25 epochs for the CNN model. This plot helps visualize the model's learning curve and identifies potential overfitting issues.

Figure 5 visualizes the F1-score for CNN model. Detailed CNN classification report, including precision, recall, and F1-score for each lesion class, is provided in Figure 6.

Figure 7 shows the t-SNE projection of the CNN's learned features, demonstrating good feature separability among the seven lesion classes.

The model performs stably throughout training, and validation accuracy rises smoothly without drastic oscillations, indicating the robust learning process of the model.

Extensive performance comparison is performed among CNN, EfficientNet, and DenseNet for the HAM10000 dataset. Table 3 summarizes the Performance comparison of CNN, EfficientNetB0, and DenseNet121. Experimental results show that the proposed CNN has the highest overall accuracy and F1-score, with a stronger classification capability for skin lesion analysis against EfficientNet and DenseNet. Although EfficientNet and DenseNet benefit from deep and dense feature extraction, their performance is more affected by class imbalance, resulting in reduced effectiveness for minority lesion categories.

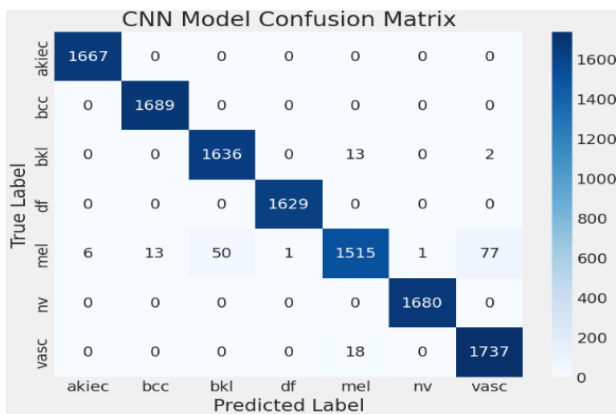


Figure 5. F1-score for Convolutional Neural Network (CNN) model

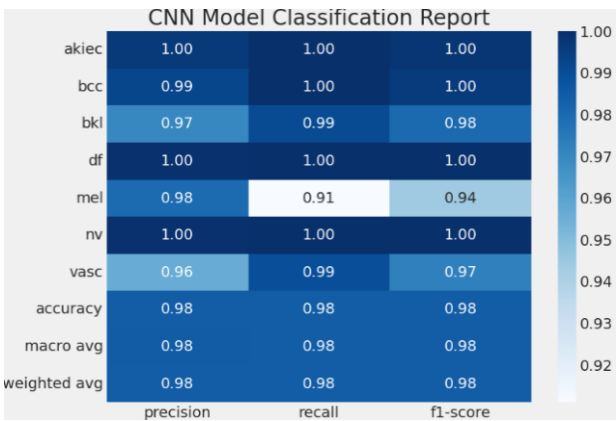


Figure 6. Detailed Convolutional Neural Network (CNN) classification report

In contrast, the proposed CNN has more consistent class-wise performance for all types of lesions, making it more robust to the imbalance in dermatological data. The CNN model also has a lesser number of parameters and lower computational complexity compared with EfficientNet and DenseNet, respectively; thus, it is more suitable for practical deployment. These results prove the advantages brought by this proposed CNN model regarding accuracy, stability, and efficiency, and the abstract has been updated to properly set up and reflect these comparative findings and their significance.

CNN Model: The CNN model was excellent, with 98.6%

accuracy. The high accuracy can be explained by the fact that deeper layers of CNN were used, batch normalization, and dropout, which allowed the model to learn complex features and generalize to the unseen test data well.

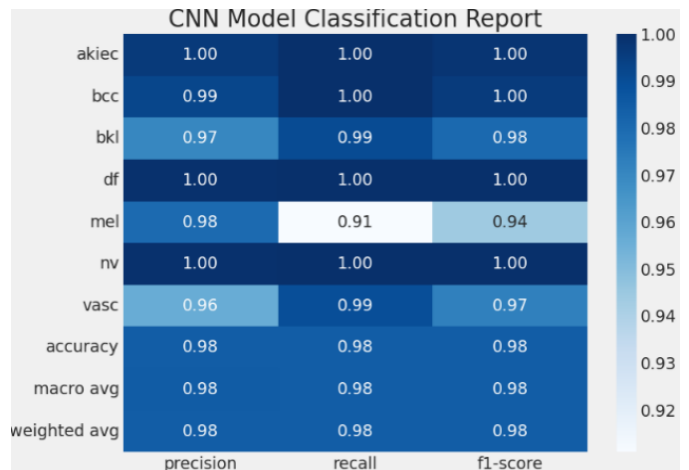


Figure 7. T-SNE projection of the Convolutional Neural Network (CNN)'s learned features

EfficientNetB0: Although EfficientNetB0 is efficient in other assignments, it performed poorly in this research because its input image resolution is low ( $28 \times 28$ ). In addition, the model is also demanding in terms of fine-tuning and requires higher resolution for images to perform well. Therefore, it only recorded 14.4% accuracy and is therefore not very appropriate for this particular task.

Table 3. Performance comparison of Convolutional Neural Network (CNN), EfficientNetB0, and DenseNet121

Model	Accuracy	Precision	Recall	F1-Score
CNN	98.6%	98.5%	98.4%	98.4%
EfficientNet	14.4%	14.2%	14.6%	14.4%
DenseNet	24.9%	24.2%	25.4%	24.8%

DenseNet121: DenseNet121 outperformed EfficientNetB0 but also underperformed compared to the CNN model. The model recorded 24.9% accuracy, presumably due to the architectural complexity and insufficient input resolution. DenseNet121 is superior in cases where higher-resolution images are utilized and fine-tuning is conducted on a larger dataset.

## 6. CONCLUSIONS

In this paper, it is demonstrated that a model based on CNN is efficient enough to identify skin cancer with an accuracy of 98.6% on the HAM10000 dataset. The suggested model outperformed EfficientNet and DenseNet, which did not perform well because of low resolution input and poor fine-tuning. Applying data augmentation, Random Over Sampler, and batch normalization allowed the CNN model to handle class imbalance effectively and improve generalization. The results highlight that CNNs offer a scalable and effective solution to skin cancer detection, which makes them suitable for real-time deployment, especially in mobile and embedded health systems. This work can be used for early diagnosis and assist physicians to devise treatment plans for better patient outcomes.

## 7. DISCUSSION AND FUTURE WORK

The CNN model also worked better for the detection of skin cancer than EfficientNet and DenseNet since it was able to properly extract the features of the HAM10000 dataset. Despite high accuracy, class imbalance and low image resolution constraints still remained hurdles, particularly for more advanced models. More research will further tune EfficientNet and DenseNet from images with a higher resolution and larger datasets. Moreover, the integration of ensemble learning algorithms with real-time implementation in mobile applications can enhance the system's performance and accessibility in resource-constrained environments.

## REFERENCES

- [1] Haque, S., Ahmad, F., Singh, V., Mathkor, D.M., Babegi, A. (2025). Skin cancer detection using deep learning approaches. *Cancer Biotherapy and Radiopharmaceuticals*, 40(5): 301-312. <https://doi.org/10.1089/cbr.2024.0161>
- [2] Hasan, M.K., Elahi, M.T.E., Alam, M.A., Jawad, M.T., Martí, R. (2022). DermoExpert: Skin lesion classification using a hybrid convolutional neural network through segmentation, transfer learning, and augmentation. *Informatics in Medicine Unlocked*, 28: 100819. <https://doi.org/10.1016/j.imu.2021.100819>
- [3] Yadav, R., Bhat, A. (2024). A systematic literature survey on skin disease detection and classification using machine learning and deep learning. *Multimedia Tools and Applications*, 83(32): 78093-78124. <https://doi.org/10.1007/s11042-024-18119-w>
- [4] Prasanna Lakshmi, G.S., Vidhyashree, B (2024). A comprehensive survey on skin lesion detection using deep learning. *Frontiers in Health Informatics*, 13(3): 169.
- [5] Tan, M.X., Le, Q. (2019). EfficientNet: Rethinking model scaling for Convolutional Neural Networks. In *Proceedings of the 36th International Conference on Machine Learning*, pp. 6105-6114. <http://proceedings.mlr.press/v97/tan19a/tan19a.pdf>
- [6] Rahman, S.Z., Singasani, T.R., Shaik, K.S. (2023). Segmentation and classification of skin cancer in dermoscopy images using SAM-based deep belief networks. *Healthcraft Frontiers*, 1(1): 15-32. <https://doi.org/10.56578/hf010102>
- [7] Ashtagi, R., Kharat, P.V., Sarmalkar, V., Hosmani, S., Patil, A.R., Akkalkot, A.I., Padthe, A. (2024). Enhancing melanoma skin cancer diagnosis through transfer learning: An EfficientNetB0 approach. *Acadlore Transactions on AI and Machine Learning*, 3(1): 57-69. <https://doi.org/10.56578/ataiml030105>
- [8] Rashad, N.M., Abdelnabi, N.M., Seddik, A.F., Sayedelahl, M.A. (2025). Automating skin cancer screening: A deep learning. *Journal of Engineering and Applied Science*, 72(1): 6. <https://doi.org/10.1186/s44147-024-00573-w>
- [9] S, A., R, S., Nambiar, S.K., Cherian, A.K. (2025). A comprehensive deep learning approach for skin cancer diagnosis: Integrating interpretability and advanced techniques. *Multimedia Tools and Applications*, 84(35): 43643-43663. <https://doi.org/10.1007/s11042-025-20854-7>
- [10] Natha, P., Rajeswari, P. (2024). Advancing skin cancer prediction using ensemble models. *Computers*, 13(7): 157. <https://doi.org/10.3390/computers13070157>
- [11] AlSadhan, N.A., Alamri, S.A., Ben Ismail, M.M., Bchir, O. (2024). Skin cancer recognition using unified deep convolutional neural networks. *Cancers*, 16(7): 1246. <https://doi.org/10.3390/cancers16071246>
- [12] Musthafa, M.M., TR, M., V, V.K., Guluwadi, S. (2024). Enhanced skin cancer diagnosis using optimized CNN architecture and checkpoints for automated dermatological lesion classification. *BMC Medical Imaging*, 24(1): 201. <https://doi.org/10.1186/s12880-024-01356-8>
- [13] Shah, S.A.H., Shah, S.T.H., Khaled, R., Buccoliero, A., Shah, S.B.H., Di Terlizzi, A., Di Benedetto, G., Deriu, M.A. (2024). Explainable AI-based skin cancer detection using CNN, Particle Swarm Optimization and Machine Learning. *Journal of Imaging*, 10(12): 332. <https://doi.org/10.3390/jimaging10120332>
- [14] Jeyalakshmi, V.S., Bala Shunmugam, N., Kavitha, M., Dani, D.P.D. (2025). Automized quick prediction of skin cancer diagnosis by enhanced deep convolutional neural network. *Advances in Artificial Intelligence and Machine Learning in Big Data Processing*, 2202: 292–302. [https://doi.org/10.1007/978-3-031-73065-8\\_24](https://doi.org/10.1007/978-3-031-73065-8_24)
- [15] Vincent, Darian, G., Surantha, N. (2025). Performance evaluation of convolutional neural network (CNN) for sskin cancer detection on edge computing devices. *Applied Sciences*, 15(6): 3077. <https://doi.org/10.3390/app15063077>
- [16] Vieira, J., Mendonça, F., Morgado-Dias, F. (2025). Deep learning approaches for skin lesion detection. *Electronics*, 14(14): 2785. <https://doi.org/10.3390/electronics14142785>
- [17] Deshmukh, G.K., Jadhav, S., Kadam, A., Aware, S., Shukla, T. (2025). Interpretable skin cancer detection via optimized CNN models for smart healthcare solutions. *Research & Reviews: A Journal of Bioinformatics*, 12(1): 41-45. <https://journals.stmjournals.com/rrjobi/article=2025/view=196911>
- [18] Jalaboi, R., Winther, O., Galimzianova, A. (2023). Dermatological diagnosis explainability benchmark for convolutional neural networks. *arXiv preprint arXiv:2302.12084*. <https://doi.org/10.48550/arXiv.2302.12084>
- [19] Sakthiradha, M., Sharvanthika, K.S., Priyanka, V., T, S., Sriya, G. (2025). Skin disease detection using deep learning. In *2025 5th International Conference on Intelligent Technologies (CONIT), HUBBALI, India*, pp. 1-8. <https://doi.org/10.1109/CONIT65521.2025.11166843>
- [20] Sasikala, S., Kumar, S.A., Shivappriya, S.N., T, P. (2020). Towards improving skin cancer detection using transfer learning. *Bioscience Biotechnology Research Communications*, 13(11): 55-60. <https://doi.org/10.21786/bbrc/13.11/13>