AVE Trends in Intelligent Health Letters

Intelligent Skin Disease Detection and Classification Using Convolutional Neural Networks on Dermoscopic Images

A. Muthukumaravel[1,*], S. Silvia Priscila[2], B. M. Praveen[3]

[1]Faculty of Arts and Science, Bharath Institute of Higher Education and Research, Chennai, Tamil Nadu, India.
[1,3]Institute of Engineering and Technology, Srinivas University, Dakshina Kannada, Karnataka, India.
[2]Department of Computer Science, Bharath Institute of Higher Education and Research, Chennai, Tamil Nadu, India.
dean.arts@bharathuniv.ac.in[1], silviaprisila.cbcs.cs@bharathuniv.ac.in[2], bm.praveen@yahoo.co.in[3]

*Corresponding author

Abstract: Skin disorders are a leading global health problem. The exact diagnosis of melanoma will help to improve patient outcomes. Methods of diagnosis, such as visual inspection by a dermatologist, are widely used for diagnosing various types of skin disease. However, they tend to suffer from experts' subjectivity. It is also time-consuming and limited by the availability of medical expertise. To overcome the problems, this research paper proposes an automated skin illness classification system using DL (Deep Learning) methods, which can assist in reliable and efficient diagnosis. The system suggested using a CNN to classify skin diseases from dermoscopic images. The HAM10000 dataset is an established, clinically verified dataset containing images of seven common pigmented skin lesion categories, used to train and test the model. Before training models, images are preprocessed using steps such as resizing, normalisation, colour normalisation, and noise reduction. A pre-trained CNN architecture has been used to extract visual features via TL (Transfer Learning), achieving good performance with limited medical data. The Adam optimiser trains the model with categorical cross-entropy loss. To evaluate robustness across all disease classes, accuracy, precision, recall, F1-score, and confusion matrix are used. The skin lesions are accurately classified, proving that the method works. The approach has great diagnostic potential for dermatologists. This can be expanded to online or mobile apps for early identification of skin disease and greater accessibility.

Keywords: Skin Disease; Deep Learning (DL); CNN Architecture; Medical Expertise; Noise Reduction; Dermoscopic Images; Colour Normalisation; Transfer Learning; Adam Optimiser.

Cite as: A. Muthukumaravel, S. S. Priscila, and B. M. Praveen, "Intelligent Skin Disease Detection and Classification Using Convolutional Neural Networks on Dermoscopic Images," *AVE Trends in Intelligent Health Letters*, vol. 3, no. 1, pp. 40–52, 2026.

Journal Homepage: https://avepubs.com/user/journals/details/ATIHL

Received on: 13/02/2025, **Revised on:** 07/06/2025, **Accepted on:** 12/09/2025, **Published on:** 05/01/2026

DOI: https://doi.org/10.64091/ATIHL.2026.000264

1. Introduction

The major organ of the human frame is the skin. It shields the entire body [17]. It is the body's first line of defence against external factors [16]. One of the most prevalent cancers in the US (United States) is skin tumours. This tumour happens due to the rapid multiplication and excessive division of skin cells in the body. Skin tumours are harmful, but it is better to detect them early. Because of the difficulty of identification, specialists sometimes reach an incorrect conclusion about this cancer [11]. Millions of people across the globe suffer from one or more skin diseases every year. An appropriate and precise diagnosis is

very important for the active treatment of any medical condition. This holds particularly true for serious conditions that can turn fatal if not diagnosed on time, such as melanoma, which is a type of skin cancer. Conventionally, dermatologists have relied on visual examination and dermoscopy to identify skin lesions. Unfortunately, these processes can be quite time-consuming and prone to human error. Furthermore, the nature of the differences between skin lesions makes examination more difficult. The growing need for reliable diagnostic tools has spurred interest in the design of automated systems to aid in the identification and classification of skin diseases [25].

1.1. Importance of Skin Disease Classification

Classifying skin diseases is highly valuable, as it enables the healthcare system to identify, diagnose, and treat skin conditions with greater accuracy. Manual diagnosis often leads to mistakes, given that several skin conditions, such as skin cancer, acne, psoriasis, and eczema, share a variety of symptoms. With automated disease classification and the latest DL and ML technologies, disease patterns and image analysis of affected skin can be performed with precision. With such technologies, assisting dermatologists can lead to better resolution of their clinical problems and to the refinement of the diagnosis and treatment of skin diseases. For remote services, classification of skin diseases greatly benefits them, as dermatological services can be rendered to patients in remote regions to provide initial diagnoses and classifications using technological devices. The workload of medical personnel can also be lightened by using such services, as large numbers of people can be screened using the classifications.

1.2. Challenges of Skin Lesion Classification

The classification of skin lesions is quite difficult due to the high visual resemblance among diseases, the diverse types of skin, and the various lesion styles. The signs of skin diseases are varied, and changes in symptoms are a long-term process [21]. For the untrained eye, it can be difficult to determine the type of skin condition. A lot of laypersons also ignore the symptoms of skin conditions, which can lead to serious issues such as irreversible skin damage and even skin cancer. Also, early treatment for skin cancer can reduce mortality [20]. Skin lesion diagnosis poses unique difficulties, stemming from the overlap of visual characteristics such as lesion size, shape, texture, colour, and the various types of skin on which these lesions appear. Many dermatological conditions visually appear similar, and lesions characterised as benign or malignant can often change appearance over time, concealing the potential for either an acute or chronic medical condition. Factors such as lighting, environmental conditions, changes in lesion size, and lesion position can complicate a dermatological diagnosis. For those without medical training, an accurate and objective skin analysis can be difficult, often leading to a more severe dermatological concern, as skin lesions may worsen over time and potentially progress to skin cancer or irreversible skin damage. Dermatological practitioners must have access to accurate and reliable methods for analysing and classifying skin lesions.

1.3. Traditional Skin Cancer Diagnostic Procedures

Traditional diagnostic procedures primarily rely on dermatologists' visual inspection. These procedures, usually supported by a dermoscope, can often be subjective and rely on clinical skill. These factors make it difficult to obtain an accurate diagnosis on time, given the growing number of skin diseases. The accuracy of skin cancer diagnosis depends on the specificity and sensitivity of the physical examination method used. While several large studies have demonstrated inconclusive benefits of screening the general adult population, the specificity and sensitivity of the physician's speciality were also inconclusive due to aggregated accuracy reporting across several practitioner types [22]. It has suggested a new approach to classifying melanocytic tumours as benign or malignant by analysing digital dermoscopic images [29]. In skin cancer, traditional diagnostic methods rely heavily on the clinician's experience to examine skin lesions and assess their suspiciousness. For example, when visualising the skin, a clinician evaluates a lesion's size, colour, shape, and texture to formulate a diagnosis. One method they use to evaluate the potential severity of a lesion is the ABCDE rule, in which each letter corresponds to a different critical factor in determining its potential to be a melanoma.

For example, A - asymmetry, B - border defection, C - colour, D - derivate size of the lesion, and E - how the lesion has changed over time. Physicians also examine lesions using dermoscopy, a skin-imaging technique. When dermoscopy is unavailable to the dermatologist, a lesion biopsy is performed, and the samples are sent for microscopic examination. Traditional methods of diagnosis are time-consuming, and for many lesions, diagnosis can be delayed by the numerous steps that require the dermatologist's experience. To reduce time to diagnosis and reliance on the dermatologist's experience, many researchers are investigating the use of computer-assisted diagnostic systems in conjunction with lesion imaging analysis. Poor contrast and undesirable noise are common problems across many image modalities, including medical, satellite, aerial, and general photos [30]. A parametric model that can be used to generate high-quality samples of natural images. The suggested method uses a series of convolutional networks integrated into a Laplacian pyramid to generate images via a coarse-to-fine refinement procedure [28].

1.4. Modern Skin Cancer Diagnostic Procedures

Recent developments in AI (Artificial Intelligence), especially DL, have sparked new prospects for automating medical image analysis. Skin conditions can be identified at an earlier stage, thereby reducing healthcare costs, and the application of automated deep CNN techniques reduces misdiagnosis of skin illnesses and enhances forecast accuracy [13]. CNNs are DL algorithms designed to recognise visual patterns in images and videos. They can provide the best classification accuracy. Automatic skin disease classification systems can provide consistent support to clinicians whilst improving overall outcomes. Similarly, they will continue to become smarter and more efficient for better diagnosis. Advanced techniques for evaluating skin cancer have been introduced, such as artificial intelligence image analysis, tape sampling for messenger RNA (mRNA) on the skin surface, reflectance confocal microscopy, and impedance spectroscopy [23]; [24]. A class imbalance occurs when one or the other pathology is significantly overrepresented [26]. Neural networks are therefore likely to overfit and exhibit poor generalisation to new data [27].

1.5. Importance of Skin Cancer Datasets

In this regard, large-scale annotated datasets such as HAM10000 are becoming a necessary tool for research, providing dermoscopic images. Systems that automatically process these images could assist dermatologists in clinical practice, improve detection rates, and reduce workload in some cases. It will help improve public health by enabling faster, more consistent identification of skin disorders, thereby enhancing patient health outcomes.

2. Related Work

Dermatoscopic analysis, biopsy, and histopathological investigation are always challenging for automatic classification of dermatoscopic images of skin lesions due to fine-grained variations. The CNN is a powerful DL technique that outperforms traditional procedures. Udriștoiu et al. [1] proposed a CNN design that classifies skin lesions using only image pixels and analysis labels as input. DL models effectively learn features that accurately interpret complex patterns. Srinivasu et al. [2] proposed a process for categorising skin sickness using DL-based methods. MobileNet V2 is effective, achieving higher accuracy while using lightweight computational strategies. The model can preserve sufficient information to make accurate predictions. The results are compared with state-of-the-art models and a CNN design that was expanded with a few changes. The proposed technique has achieved greater than 85% accuracy on the HAM10000 dataset. The speed at which it identifies the affected area is essentially leading the area to be used up by computations by almost 2× compared to the conventional MobileNet model. An app is designed to enable prompt, proper action. The image of the affected area helps the patient and dermatologist identify the type of skin disease at an early stage. The recommended system can assist general practitioners in efficiently diagnosing skin abnormalities, thereby reducing additional difficulties and sickness. Skin issues are a worldwide health problem.

The danger of the toxicities is invisible, which creates physical pain and initiates mental depression. Also, in extreme circumstances, it may cause skin cancer. After that, identifying skin disease from clinical images is one of the most challenging tasks in healthcare image analysis. Furthermore, doctors using their expertise still take time, and it's a subjective process. As a result, both affected individuals and dermatologists need automated skin disease prognostication to enable earlier action. Ahammed et al. [3] proposed a digital hair-removal method based on morphological filtering and an inpainting procedure, followed by Gaussian filtering to deblur the images. They automatically apply the Grabcut segmentation algorithm to isolate infected lesions. To obtain hidden inputs, statistical feature techniques and GLCM have been applied to skin images. For efficient classification, three ML techniques are applied to the extracted features. Two standard datasets are being evaluated to validate the models. The SVM classifier is marginally better than the others. The outcome has been compared with conventional models. Because external factors across crowds and patient populations can influence similar skin lesions, diagnosing skin diseases becomes difficult. The objective of the study by Alruwaili and Mohamed [4] was to explain a fusion-level DL model that enhances stability and performance in skin disease categorisation. The model's architecture combines three independent CNNs across separate divisions.

The NN (Neural Network) framework leverages its ability to extract intricate features from diverse robust architectures to achieve accurate results and tight classification precision. Through dense and dropout layers, the framework effectively generalises the features it learns by mitigating overfitting. This study demonstrates the application of DL for skin disease classification. Therefore, the method has clinical potential to help automate dermatological diagnosis. More than one-third of the world's population will suffer from a skin disease at some point, but their actual impact is far less acknowledged than physical ailments. They must automate the classification of these diseases to enable accurate, rapid diagnosis. Mohan et al. [5] explained the use of Vision Transformers, Swin Transformers, and DinoV2 in dermatology tasks, introducing the latter. The effectiveness of DinoV2 is confirmed using standard datasets, where it outperforms previous models. They will also compare ConvNeXt and further CNN designs and highlight the advantages of transformer models. The applications of this particular

transformer, specifically DinoV2, suggest it may be useful in dermatology. Skin ailments are among the most prevalent health concerns worldwide, affecting people of all ages and significantly lowering their quality of life.

It is necessary to accurately and quickly diagnose and classify diseases. According to studies, soft-tissue injuries pose challenges in both practical and operative treatment. According to Sarı and Keser [6], fungal skin diseases include 693 cases of eczema. The classifier was built using the Relief algorithm, which prioritises classification success. To achieve better classification success and higher-quality selection, Relief was used in this study, which focused on three different skin diseases. An SVM classifier using Relief's approach achieves higher accuracy than classification with other techniques. The suggested model makes an original contribution to the literature, notably by combining feature selection with a simplified design. This success rate shows that DL is an effective method for classifying skin diseases, reducing cancer mortality through early, effective action, and enabling us to differentiate between them easily. Skin diseases are an important category of disease in medicine, and they are often difficult to diagnose and are often misdiagnosed. The use of DL for skin disease classification is highly valuable for clinical analysis. According to Liu et al. [7], the skin disease classification model is based on multiscale channel attention. To begin with, an improved pyramid segmentation attention module has been added to the model for extracting multi-scale features from the entire image. To validate the proposed model's performance, this study used common skin disease datasets. In dermatological research, the ability to accurately identify different types of skin lesions, such as nodules, can facilitate early diagnosis and significantly ease treatment. The research paper by Mubeen and Dulhare [8] described an innovative approach in which a unified attention (UA) network is paired with a deep CNN for feature extraction of skin lesions and nodules.

The UA network processes sequential patient histories, while LSTM networks monitor nodule progression. This technology has better precision and sensitivity than other methods. This can improve dermatology diagnoses. The timely identification and management of several skin illnesses depend significantly on skin lesions. Computer vision and ML techniques have improved significantly, and therefore learning-based methods have attracted much attention. According to Debelee [9], the importance of analysing skin lesions in medicine was explained, along with the challenges posed by physical inspection. The in-depth review of earlier research papers on skin lesion classification aims to accurately classify skin disease types from dermoscopic, macroscopic, and other lesion images. The study papers selected for the present study were analysed for their contributions and limitations regarding the techniques used, DL architectures, or conventional ML methods. Aksoy et al. [10] evaluated the performance of various DL models for lesion classification, identified the most suitable architecture, and developed a web-based portal to enhance accessibility for diagnostic and educational purposes. Several DL models were compared for classification accuracy. They incorporated the top model into a web application that allows users to upload images for automatic classification, with confidence scores indicating the model's confidence in its prediction. The tool's enhanced visualisation feature helps explore feature maps produced by convolutional layers for better interpretability. Techniques of web scraping and summarisation were also used to provide dermatological information from reputable sources.

3. Proposed Methodology

DL techniques provide faster, more precise processing of large-scale datasets than classical ML methods, enabling more effective feature extraction. Simultaneously, DL algorithms can assist clinicians with more comprehensive data analysis and evaluation of test results [26]. The design of the projected system includes a DL pipeline for classifying skin disease from dermoscopic images. The process follows the pre-test phase to control the flow and organise the data. Figure 1 summarises the first steps in building a DL system for classifying skin-related diseases. Images of dermoscopic skin lesions from the HAM10000 dataset serve as the system's input. Each dermoscopy image in the dataset shows a different skin lesion from a different diagnostic category. Raw dermoscopic images captured under various lighting conditions can introduce noise. Images can additionally vary in size and colour. Consistency and quality must therefore be improved using pre-processing techniques. These techniques will enable further analysis while removing less relevant features from dermoscopic images. The preprocessing steps before model training improve the quality of dermoscopic images. These steps include colour correction, noise reduction, and consistent image resizing. Consistent resizing and colour correction are also forms of normalisation. Normalisation is the process of controlling the scaled pixel values. Noise reduction and other forms of reduction are intended to remove various unwanted artefacts.

These techniques allow the DL model to learn features that are relevant from the data. The relevance of the features stems from increased image clarity and quality. Once preprocessing is finished, the system undergoes transfer learning employing pre-trained models from CNN architectures such as ResNet50 and VGG16. Here, the models are used to recognise strong patterns, textures, and structures, as well as to detail the skin lesions from the learned images. Each neural network operates through multiple layers. Models such as ResNet, VGGNet, and EfficientNet CNNs are used for lower-level image processing tasks and can learn highly complex features through pre-training on large image datasets. Therefore, transfer learning can provide high overall model performance. Once feature extraction is complete, classification is performed using the EfficientNetB0 architecture. Each learned feature is learned from the model's output layer to recognise and categorise images of skin lesions.

EfficientNetB0 is opted as it provides high performance with low descent complexity. This is ideal for healthcare models, as it solves the core dilemma of engineering: the trade-off between ease and efficiency.

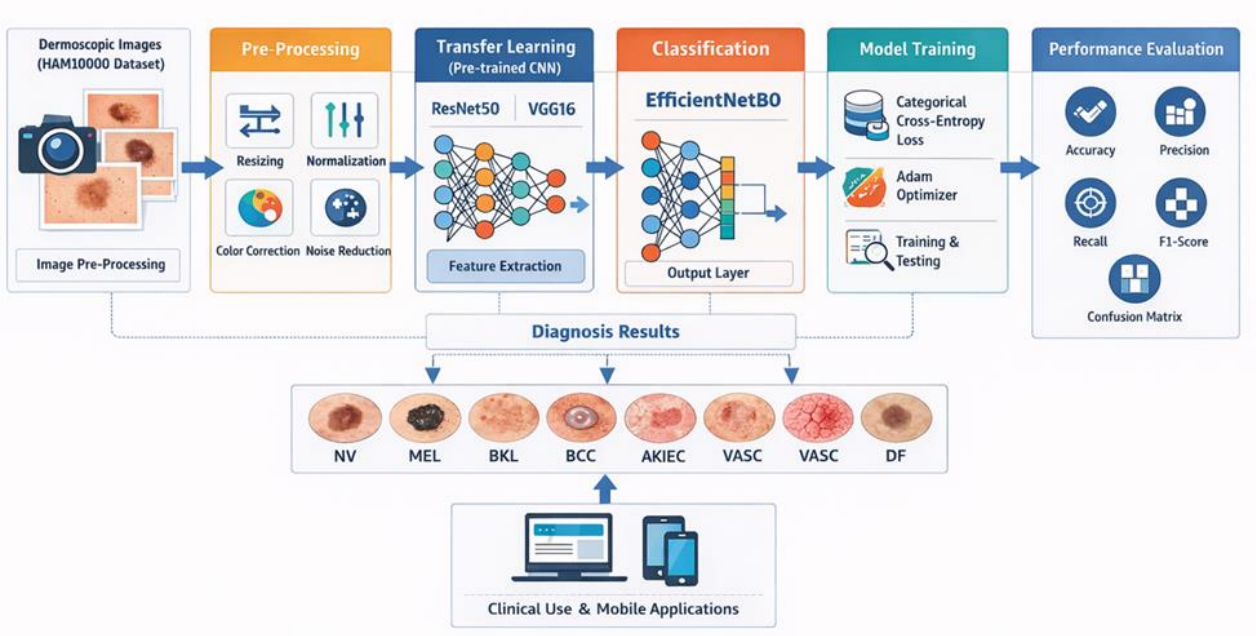

Figure 1: Automatic skin disease classification using DL models

The next step is to train the model, which learns to classify skin lesions using labelled data. During training, the model is trained using the categorical cross-entropy loss function to quantify the divergence between the predicted and actual class labels. The model's learning rate is improved by using the Adam optimiser to update its weights. To determine how the model generalises to new data, the dataset is split into training and test sets. The final part of the process involves evaluating the system's classification and the classification system using multiple evaluation parameters. Other metrics for evaluating model classification performance include precision, recall, and the F1-score. In addition, a confusion matrix can be used better to understand the predictive results for each skin lesion category. The system's final output is a diagnosis, with images of skin lesions categorised into 7 classes: NV, ME, BK, BC, AK, VL, and DF. With these types of predictions, dermatologists can be assisted in detecting skin disease. Moreover, the suggested method can be implemented in clinical settings and in web- or mobile-based applications. This offers healthcare providers and patients the opportunity to conduct remote dermatological screenings, obtain initial diagnostic support, and, most importantly, improve access to dermatological care in underserved regions.

3.1. Image Preprocessing

In DL-based skin disease classification, it is necessary to preprocess raw dermoscopic images to provide suitable inputs for the CNN. Normalising and standardising input data can help accelerate model convergence, reduce computational complexity, and increase classification accuracy.

3.1.1. Image Resizing Using Spatial Normalisation

To use the ResNet50 or EfficientNet CNN architectures, all dermoscopy images are resized to a static input dimension of 224 × 224 pixels. Uniformity is achieved by resizing images, which reduces computational cost. If $I_{orig} \in R^{H \times W \times 3}$ is the original input image and $I_{resized} \in R^{224 \times 224 \times 3}$ It represents the image after the resizing task can be illustrated as:

$$I_{resized} = Resize\left(I_{orig}, 224{,}224\right) \quad (1)$$

3.1.2. Pixel Normalisation Using Min-Max Scaling

The pixel intensities in raw images range from 0 to 255. By scaling the pixel values of raw images from (0, 255) to (0, 1), numerical stability improves, and model training converges faster:

$$I_{norm} = \frac{I - I_{min}}{I_{max} - I_{min}} \quad (2)$$

Where:

- It signifies the actual pixel intensity value.
- Initialize $I_{min} = 0$ and $I_{max} = 255$
- I_{norm} represents the normalized value of a pixel.

Post-normalisation, each image has a uniform intensity value that can be used in DL.

3.1.3. Noise and Artefact Removal Using Gaussian Smoothing

Hair, air bubbles, or reflections contained in dermoscopic images may interfere with model learning. Gaussian smoothing can be used to reduce noise. Related noise models, formulated based on these environmental and ambient variables, include the Gaussian distribution or a distribution with a characteristic function close to it. It is well known that in digital image processing, spatial filters (mean, median, and Gaussian smoothing) are applied to remove Gaussian noise [15]:

$$I_{smooth}(x, y) = \sum_{i=-k}^{k} \sum_{j=-k}^{k} G(i, j).I(x + i, y + j) \quad (3)$$

$$G(i, j) = \frac{1}{2\pi\sigma^2} e^{-\frac{i^2+j^2}{2\sigma^2}} \quad (4)$$

Where:

- $G(i, j)$ indicates the Gaussian kernel data.
- σ denotes the SD (Standard Deviation) controlling smoothing.
- k defines the kernel's radius.

This removes high-frequency noise while preserving lesion boundaries.

3.1.4. Colour Normalisation

Some assessments indicate that colour loyalty significantly enhances the accuracy of analysing malicious lesions in images. A comparison of local and global methods revealed that colour features are more useful than texture features for melanoma identification [14]. The strategy based on local characteristics among those evaluated shows better performance, as colour adjustment enhances the accuracy of spontaneous diagnosis [19]. This indicates that lighting correction procedures have favourable effects on achieving colour constancy [18]. Images usually vary in colour depending on the type of dermatoscope and lighting. Colour normalisation or standardisation is used to minimise this variation:

$$I_c^{norm} = \frac{I_c - \mu_c}{\sigma_c} \quad (5)$$

Where:

- I_c represents pixel intensity in colour channel c (R, G, B).
- μ_c and σc denotes the mean value and SD of channel c.

By maintaining a consistent colour distribution, CNNs extract features more efficiently.

3.2. Dataset Splitting: Train–Validation–Test Split

The classification dataset was divided into three subsamples: a training subset (70%), a validation subset (15%), and a test subset (15%).

3.3. Feature Extraction Using CNN

CNNs utilise multiple convolutional and pooling layers to automatically learn features from images. It fuses spatial features from the illness probability maps with people-specific clinical information, thereby refining the overall model's performance.

3.3.1. Convolution Operation

$$(F * K)(x, y) = \sum_{i=1}^{m} \sum_{j=1}^{n} F(x + i, y + j)K(i, j) \quad (6)$$

F is denoting the original image, while K is denoting the Convolution Kernel.

3.3.2. Activation Function: Rectified Linear Unit (ReLU)

$$f(x) = \max(0, x) \quad (7)$$

ReLU introduces nonlinearity and accelerates convergence during training.

3.3.3. Pooling Using Max Pooling

$$P(x, y) = \max_{(i,j) \in R} F(i, j) \quad (8)$$

Dimensionality & complexity are reduced by pooling.

3.4. Transfer Learning: Pre-Trained CNN (ResNet50 and VGG 16)

Transfer learning initialises the model using weights trained on ImageNet. The last classification layers are replaced and optimised on the HAM10000 dataset. Classification Layer by Softmax Function:

$$P(y_i) = \frac{e^{z_i}}{\sum_{j=1}^{C} e^{z_j}} \quad (9)$$

C indicates the number of disease classes and $\boldsymbol{z_i}$ is the output logit.

3.4.1. Loss Function: Categorical Cross-Entropy

$$L = -\sum_{i=1}^{C} y_i \log(\hat{y}_i) \quad (10)$$

3.4.2. Optimisation Using Adam Optimiser

Adam combines momentum and adaptive learning rates for efficient training. Using CNNs for feature extraction has shown great promise for image-based classification, particularly in medical imaging analysis. One advantage of CNNs is their ability to learn features such as lines, shapes, and textures from images, eliminating the need for custom-built feature extraction. The model's ability to learn from visual images enables it to discern image features at both low-level (undiscerning) and high-level (discerning) levels, ultimately leading to significant improvements in classification. In analysing images for the detection of skin diseases, the subtleties of the images can at times go unnoticed to the naked eye.

$$\theta_{t+1} = \theta_t - \alpha \frac{m_t}{\sqrt{v_t} + \epsilon} \quad (11)$$

As a result, these features may go unnoticed by the human eye, but a CNN automatically compensates for them and highlights them. This automatic feature extraction tends to improve the effectiveness of the analysis. CNNs reduce the amount of data the model must analyse by providing features that represent the data; this is a method for reusing information. It improves their effectiveness through information recycling, making them ultimately more effective in the long run. This method of hierarchical learning also provides the CNN with some measure of generalisation when applied to feature analysis, which is effective for automated image analysis. Feature analysis will be effective for automated disease diagnosis.

3.5 Model Evaluation

Assessing model performance is important for reliability. This is especially true in the medical field, such as skin or other disease classification. Model evaluation is fundamental to assessing a predictive system's effectiveness and reliability. It also helps evaluate a model's generalisation ability and assures the quality and utility of its predictions. In particular, in medical applications such as disease detection, model evaluation becomes more critical, as inaccurate predictions can lead to incorrect diagnoses and/or treatment delays. Evaluation metrics are comparison points to differentiate models. By studying metrics such as accuracy, precision, and recall, among others, researchers can articulate the merits and flaws of models. Furthermore, model

evaluation also articulates how to develop the model by suggesting where enhancements or more sophisticated training are needed. In conclusion, extensive evaluation is vital to ensure a model is reliable for real-world applications. Various metrics help measure the performance of the multi-class model shown below.

3.5.1. Accuracy

Accuracy is the proportion of correctly classified samples to the total number of samples. It delivers a general indication of model performance but may be distorted in unbalanced datasets:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (12)$$

- TP denotes True Positives.
- TN refers to True Negatives.
- False positives are represented as FP.
- FN denotes False Negatives.

3.5.2. Precision

Precision, also known as Positive Predictive Value, is the ratio of properly forecast positive cases to the total forecast positive cases. A high precision means fewer FPs:

$$Precision = \frac{TP}{TP+FP} \quad (13)$$

3.5.3. Recall (Sensitivity)

Recall, also called Sensitivity, assesses the proportion of actual positives the model correctly identifies. It demonstrates how successfully the model captures TPs:

$$Recall = \frac{TP}{TP+FN} \quad (14)$$

3.5.4. F1-Score

The F1 score, the harmonic mean of precision and recall, provides an overall assessment that reflects both false positives and negatives. It is especially helpful for unbalanced datasets:

$$F1 - Score = 2.\frac{Precision.Recall}{Precision+Recall} \quad (15)$$

3.5.5. Confusion Matrix

A confusion matrix is a Table 1 used to visualise the performance of a classification system. It does so by comparing the actual label and the predicted label. The counts of TPs, TNs, FPs, and FNs for each class are illustrated. The performance of the skin disease classification system on seven classes (NV, ME, BK, BC, AK, VL, DF) using the EfficientNetB0 Model is provided in the confusion matrix.

Table 1: Arrangement for a 7-class problem (HAM10000 Dataset)

Actual/Predicted	NV	ME	BK	BC	AK	VL	DF
NV	6599	45	28	8	4	4	4
ME	40	1045	18	9	4	9	7
BK	29	19	1025	9	6	2	1
BC	17	4	5	481	4	1	2
AK	4	2	3	4	302	4	4
VL	4	4	5	2	4	135	0
DF	1	2	2	1	2	0	101

Most of the samples are classified correctly, as indicated by the higher values along the diagonal. This shows a good degree of accuracy in the classification model, especially for NV, ME, and BK. The proposed deep learning model performs robustly, with minimal misclassifications, mostly between visually similar lesions (Figure 2).

Figure 2: Confusion matrix for classification of skin disease

Where:

- **NV:** Melanocytic nevi
- **ME:** Melanoma
- **BK:** Benign keratosis
- **BC:** Basal cell carcinoma
- **AK:** Actinic keratosis
- **VL:** Vascular lesions
- **DF:** Dermatofibroma

EfficientNetB0 achieved the highest Accuracy, Precision, Recall, and F1-score for multi-class skin disease classification on the given dataset, making it the best choice.

Table 2: Comparison of DL models on HAM10000 dataset

Class Name	Output Metric	ResNet50 Model	VGG16 Model	EfficientNetB0 Model
NV	Accuracy	0.972	0.964	0.987
	Precision	0.977	0.979	0.991
	Recall	0.978	0.969	0.975
	F1-Score	0.975	0.969	0.988
ME	Accuracy	0.977	0.979	0.990
	Precision	0.919	0.909	0.935
	Recall	0.905	0.910	0.925
	F1-Score	0.919	0.910	0.930
BK	Accuracy	0.989	0.979	0.993
	Precision	0.934	0.924	0.940
	Recall	0.936	0.918	0.942
	F1-Score	0.940	0.919	0.941
BC	Accuracy	0.990	0.986	0.991
	Precision	0.939	0.933	0.950
	Recall	0.933	0.919	0.940

	F1-Score	0.939	0.929	0.945
AK	Accuracy	0.987	0.987	0.990
	Precision	0.931	0.910	0.938
	Recall	0.917	0.910	0.925
	F1-Score	0.926	0.910	0.931
VL	Accuracy	0.988	0.979	0.989
	Precision	0.837	0.819	0.850
	Recall	0.879	0.858	0.890
	F1-Score	0.860	0.838	0.870
DF	Accuracy	0.988	0.982	0.992
	Precision	0.882	0.869	0.900
	Recall	0.929	0.909	0.940
	F1-Score	0.908	0.891	0.920

4. Results and Discussion

4.1. Dataset Description

The proposed methodology utilises the HAM10000 dataset. This is the most famous public dataset for skin lesion analysis. The dataset comprises 10,015 dermoscopic images of pigmented skin lesions, grouped into 7 classes. The attached images are labelled and verified by expert dermatologists. Many diagnoses are confirmed on histopathology for clinical reliability. The images in the dataset were acquired under standardised conditions but vary in lesion size, colour, shape, and texture. This helps the models generalise to cases they are likely to meet in the real world. The dataset shows a natural class imbalance between benign and malignant lesions, likely due to a higher incidence of benign lesions in the population. Hence, it is sufficient to train, validate, and test DL models for automatic skin disease classification [12].

4.2. Findings and Discussion

The consequences of the experiment showed that three DL models, ResNet50, VGG16, and EfficientNetB0, attained high classification performance for all seven skin diseases (NV, ME, BK, BC, AK, VL, and DF). Among the models tested, EfficientNetB0 shows the best results across most classes, with higher accuracy, precision, recall, and F1-score. The model achieved very high accuracy (>0.98) on the dominant classes NV and BK, indicating strong learning from dermoscopic images. Still, for the difficult and minority classes (VL and DF), EfficientNetB0 maintained performance above that of ResNet50 and VGG16, as measured by precision and recall. The confusion matrix further shows that most predictions lie on the diagonal, indicating that the model correctly classifies the samples while minimising misclassification between similar-looking lesions. At the same time, confusion arises between certain pigmented lesions, ML and NV, because they are so similar in appearance. The results confirm that TL with EfficientNetB0 has superior performance for automated skin disease classification. The proposed system shows strong potential as a reliable clinical decision-support tool. Moreover, it can be extended to real-time web- or mobile-based investigative applications to enable timely detection and improve patient outcomes. Table 2 provides a comparison of the performance of the three DL models, ResNet50, VGG16, and EfficientNetB0, when used to identify and classify the various categories of skin lesions present in the HAM10000 dataset. The models are evaluated using standard industry performance measures: accuracy, precision, recall, and F1-score.

These values help analyse the model's overall performance in recognising and classifying skin lesions. The study is based on seven classes of skin lesions, namely, NV, ME, BK, BC, AK, VL, and DF. The analyses show that EfficientNetB0 performs better than ResNet50 and VGG16 across almost all skin lesion classes. This demonstrates the model's superior capacity to extract relevant features from dermoscopic images, thereby improving its classification performance. The EfficientNetB0 design optimally balances depth, width, and resolution, facilitating the capture of intricate details in a collection of images of skin lesions. EfficientNetB0 is the top performer for classifying the NV class, one of the skin lesion types, with an accuracy of 0.987, the highest among all models. Furthermore, it delivers the best precision (0.991) and an F1-score of 0.988. ResNet50 achieves an accuracy of 0.972 and an F1-score of 0.975, while VGG16 performs slightly worse, with an accuracy of 0.964 and 0.969 and an F1-score of 0.969, which is significantly lower than the F1-score of EfficientNetB0. The greater precision of EfficientNetB0 also means the model misclassifies fewer lesions of this class. In ME class, EfficientNetB0 has 0.990 accuracy with lower precision (0.935) and recall (0.925) and an F1-score of 0.930, which indicates that its performance is better compared to ResNet50 and VGG16, which had accuracy of 0.977 and 0.979, respectively. Furthermore, the precision and recall remain lower than those of EfficientNetB0, suggesting a lower ability to identify melanoma correctly.

Among the BK classes, the EfficientNetB0 model achieves the highest accuracy of 0.993, making it the best-performing model. It has also recorded precision and recall of 0.940 and 0.942, with an F1-score of 0.941. The ResNet50 model has also exhibited reasonably strong performance, with an accuracy of 0.989 and an F1-score of 0.940, while the VGG16 model recorded slightly lower values of 0.979 and 0.919, respectively. These results show that although all the models perform impressively well at detecting benign keratosis cases, the EfficientNetB0 model still achieves slightly higher accuracy and F1 scores, which is worth considering. In the BC class, EfficientNetB0 has once again shown better classification performance with an accuracy of 0.991 and an F1-score of 0.945. ResNet50 also exhibited strong performance, with an accuracy of 0.990 and an F1-score of 0.939, while VGG16 demonstrated slightly lower performance, with an accuracy of 0.986 and an F1-score of 0.929. The improved recall of 0.940 achieved by EfficientNetB0 indicates that it correctly identifies the majority of actual basal cell carcinoma cases without missing many cases. EfficientNetB0 demonstrates consistent performance on the AK class, achieving 0.990 accuracy, 0.938 precision, 0.925 recall, and 0.931 F1-score. While ResNet50 and VGG16 have the same accuracy (0.987), they have lower precision and recall. This results in smaller F1-scores.

This means that EfficientNetB0 is better at balancing the identification of positive cases with the correct classification of negatives. All models exhibit slightly poorer performance in the VL class, likely due to the overlap in appearance between vascular lesions and some benign skin diseases. Nonetheless, EfficientNetB0 still records the best performance with 0.989 accuracy, 0.850 precision, 0.890 recall, and 0.870 F1-score. VGG16 and ResNet50 have lower precision and recall, meaning that they are poorer at classifying vascular lesions. Lastly, for the DF class, the best result was achieved by EfficientNetB0, with an accuracy of 0.992, precision of 0.900, recall of 0.940, and F1 score of 0.920. As for ResNet50, it performed reasonably with an accuracy of 0.988 and an F1 score of 0.908, while VGG16 had a score of 0.982 (accuracy) and an F1 score of 0.891, meaning it performed the worst of the three. The highest recall value for EfficientNetB0 indicates that the model correctly identified the majority of dermatofibroma cases. The results in Table 2 show that EfficientNetB0 outperforms ResNet50 and VGG16 across all skin lesion classes in the HAM10000 dataset. It consistently performs best in terms of accuracy, precision, recall, and F1 score across all runs, indicating that the model excels at identifying discriminative features between classes in dermoscopic images. ResNet50 achieves the second-best results and is competitive at the class level in a select few cases, while VGG16 performs the worst most of the time. This shows that EfficientNetB0 is the best-performing DL model for automated skin lesion classification and for initial screening of skin diseases.

5. Conclusion and Future Work

This paper presents an automated skin illness classification system based on DL techniques. The study uses dermoscopic images from the HAM10000 dataset. The proposed method involves image pre-processing and transfer learning, using and comparing three pre-trained CNN architectures: ResNet50, VGG16, and EfficientNetB0. Experimental results demonstrated that all models achieved high classification performance across all seven skin lesion classes: NV, ME, BK, BC, AK, VL, and DF. Among them, EfficientNetB0 achieved the highest accuracy, precision, recall, and F1-score, thanks to its effective feature extraction and lower computational load. As the findings show, automated diagnosis via DL can minimise subjectivity. It may reduce the need for expert assessment and dermatologist consultation. The proposed system has strong potential as a clinical decision-support system for diseases such as melanoma. Deploy the EfficientNetB0 model in web- and mobile-based applications due to its lightweight, efficient nature. As a result, this model can make it more accessible and allow for remote screening and timely medical assistance.

Increased accessibility to more comprehensive dermoscopic imaging datasets is essential to enhancing s dermoscopic imaging-based skin disease classification in future work. The model would greatly benefit from dermoscopic images obtained in various healthcare settings and imaging situations, including diverse dermal pigmentation. With these modifications and the use of advanced data augmentation and preprocessing strategies, the classification process's robustness could be enhanced while also addressing class imbalance. There has also been little work on using hybrid DL methods and/or ensemble methods that allow multiple classifications. Explainable AI (XAI) techniques could also build the greatly needed trust in automated dermatologic systems by analysing a model's decision processes. Finally, for countries with increasing populations and little to no qualified healthcare personnel, especially dermal oncologists, web- and smartphone-based healthcare systems utilising state-of-the-art EfficientNetB0s enable users to screen skin lesions and, eventually, facilitate remote diagnostics and early recognition of serious dermal neoplasias, such as melanoma.

Acknowledgement: The authors sincerely acknowledge Srinivas University and Bharath Institute of Higher Education and Research for their continuous support, guidance, and provision of essential research facilities.

Data Availability Statement: The datasets supporting the findings of this study are available from the corresponding author upon reasonable request, ensuring openness, verification, and reproducibility of the research.

Funding Statement: This research was not funded by any organization/institution.

Conflicts of Interest Statement: There is no conflict of interest related to this manuscript.

Ethics and Consent Statement: The authors affirm their full consent for this work to be openly accessible, allowing readers to engage with, learn from, and benefit from the content without restriction.

Reference

1. A. L. Udriștoiu, A. E. Stanca, A. E. Ghenea, C. M. Vasile, M. Popescu, Ș. C. Udriștoiu, A. V. Iacob, S. Castravete, L. G. Gruionu, and G. Gruionu, "Skin Diseases Classification Using Deep Learning Methods," *Curr. Health Sci. J.*, vol. 46, no. 2, pp. 136–140, 2020.
2. P. N. Srinivasu, J. G. S. Sai, M. F. Ijaz, A. K. Bhoi, W. Kim, and J. J. Kang, "Classification of Skin Disease Using Deep Learning Neural Networks with MobileNet V2 and LSTM," *Sensors (Basel)*, vol. 21, no. 8, p. 2852, 2021.
3. M. Ahammed, M. A. Mamun, and M. S. Uddin, "A machine learning approach for skin disease detection and classification using image segmentation," *Healthcare Analytics*, vol. 2, no. 11, p. 100122, 2022.
4. M. Alruwaili and M. Mohamed, "An Integrated Deep Learning Model with EfficientNet and ResNet for Accurate Multi-Class Skin Disease Classification," *Diagnostics*, vol. 15, no. 5, p. 551, 2025.
5. J. Mohan, A. Sivasubramanian, V. Sowmya, and V. Ravi, "Enhancing skin disease classification leveraging transformer-based deep learning architectures and explainable AI," *Comput. Biol. Med.*, vol. 190, no. 5, p. 110007, 2025.
6. M. O. Sarı and K. Keser, "Classification of skin diseases with deep learning based approaches," *Sci. Rep.*, vol. 15, no. 1, p. 27506, 2025.
7. H. Liu, Y. Dou, K. Wang, Y. Zou, G. Sen, X. Liu, and H. Li, "A skin disease classification model based on multi scale combined efficient channel attention module," *Sci. Rep.*, vol. 15, no. 1, p. 6116, 2025.
8. A. Mubeen and U. N. Dulhare, "Enhanced Skin Lesion Classification Using Deep Learning, Integrating with Sequential Data Analysis: A Multiclass Approach," *Eng. Proc.*, vol. 78, no. 1, p. 6, 2024.
9. T. G. Debelee, "Skin Lesion Classification and Detection Using Machine Learning Techniques: A Systematic Review," *Diagnostics*, vol. 13, no. 19, p. 3147, 2023.
10. S. Aksoy, P. Demircioglu, and I. Bogrekci, "Deep Learning-Based Web Application for Automated Skin Lesion Classification and Analysis," *Dermato*, vol. 5, no. 2, p. 7, 2025.
11. Q. Huang, H. Ding, and F. R. Sheykhahmad, "A skin cancer diagnosis system for dermoscopy images according to deep training and metaheuristics," *Biomed. Signal Process. Control*, vol. 83, no. 5, p. 104705, 2023.
12. K. S. Mader, "Skin Cancer MNIST: HAM10000," *Kaggle Dataset*, 2020. [Accessed by 12/12/2024].
13. R. A. Sulthana, V. Chamola, Z. Hussain, F. Albalwy, and A. Hussain, "A novel end-to-end deep convolutional neural network-based skin lesion classification framework," *Expert Syst. Appl.*, vol. 246, no. 7, p. 123056, 2024.
14. X. Zhao, H. Zhang, Q. Zheng, and C. Jing, "A hybrid deep learning framework for skin disease localization and classification using wearable sensors," *PeerJ Comput. Sci.*, vol. 11, no. 7, p. e3002, 2025.
15. Z. Kuş and M. Aydin, "Removal of Gaussian Distributed Noise in Images with Deep Neural Network Models," in *30th Signal Processing and Communications Applications Conference (SIU)*, Safranbolu, Turkey, 2022.
16. E. N. H. Kırğıl and Ç. B. Erdaş, "Enhancing Skin Disease Diagnosis Through Deep Learning: A Comprehensive Study on Dermoscopic Image Preprocessing and Classification," *Int. J. Imaging Technol.*, vol. 34, no. 4, p. e23148, 2024.
17. N. M. S. Kumar, K. Hariprasath, S. Tamilselvi, A. Kavinya, and N. Kaviyavarshini, "Detection of Stages of Melanoma Using Deep Learning," *Multimedia Tools Appl.*, vol. 80, no. 12, pp. 18677–18692, 2021.
18. A. Adegun and S. Viriri, "Deep learning techniques for skin lesion analysis and melanoma cancer detection: A survey of state-of-the-art," *Artif. Intell. Rev.*, vol. 54, no. 2, pp. 811–841, 2021.
19. A. R. F. Santos, K. R. T. Aires, and R. M. S. Veras, "Aspects of Lighting and Color in Classifying Malignant Skin Cancer with Deep Learning," *Appl. Sci.*, vol. 14, no. 8, p. 3297, 2024.
20. B. C. Baumann, K. M. MacArthur, J. D. Brewer, W. M. Mendenhall, C. A. Barker, J. R. Etzkorn, N. J. Jellinek, J. F. Scott, H. A. Gay, J. C. Baumann, F. A. Manian, P. M. Devlin, J. M. Michalski, N. Y. Lee, W. L. Thorstad, L. D. Wilson, C. A. Perez, and C. J. Miller, "Management of primary skin cancer during a pandemic: Multidisciplinary recommendations," *Cancer*, vol. 126, no. 17, pp. 3900–3906, 2020.
21. M. Wei, Q. Wu, H. Ji, J. Wang, T. Lyu, J. Liu, and L. Zhao, "A Skin Disease Classification Model Based on DenseNet and ConvNeXt Fusion," *Electronics*, vol. 12, no. 2, p. 438, 2023.
22. T. Datzmann, O. Schoffer, F. Meier, A. Seidler, and J. Schmitt, "Are patients benefiting from participation in the German skin cancer screening programme? A large cohort study based on administrative data," *Br. J. Dermatol.*, vol. 186, no. 1, pp. 69–77, 2022.

23. B. Brouha, L. K. Ferris, and M. K. Skelsey, "Real-world utility of a non-invasive gene expression test to rule out primary cutaneous melanoma: A large US registry study," *Journal of the American Academy of Dermatology*, vol. 83, no. 6, p. AB147, 2020.
24. A. T. Young, N. B. Vora, J. Cortez, A. Tam, Y. Yeniay, L. Afifi, D. Yan, A. Nosrati, A. Wong, A. Johal, and M. L. Wei, "The role of technology in melanoma screening and diagnosis," *Pigment Cell Melanoma Res.*, vol. 34, no. 2, pp. 288–300, 2021.
25. I. Castiglioni, L. Rundo, M. Codari, G. D. Leo, C. Salvatore, M. Interlenghi, M. Interlenghi, F. Gallivanone, A. Cozzi, N. C. D. Amico, and F. Sardanelli, "AI Applications to Medical Images: From Machine Learning to Deep Learning," *Phys. Med.*, vol. 83, no. 3, pp. 9–24, 2021.
26. D. Stutz, M. Hein, and B. Schiele, "Disentangling Adversarial Robustness and Generalization," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, California, United States of America, 2019.
27. A. Bissoto, F. Perez, E. Valle, and S. Avila, "Skin Lesion Synthesis with Generative Adversarial Networks," In International Workshop on Computer-Assisted and Robotic Endoscopy, *Springer International Publishing*, Cham, Switzerland, 2018.
28. E. L. Denton, S. Chintala, A. Szlam, and R. Fergus, "Deep generative image models using a Laplacian pyramid of adversarial networks," in *Proc. 28th Int. Conf. Neural Inf. Process. Syst. (NIPS)*, Montreal, Canada, 2015.
29. F. Xie, H. Fan, Y. Li, Z. Jiang, R. Meng, and A. Bovik, "Melanoma Classification on Dermoscopy Images Using a Neural Network Ensemble Model," *IEEE Trans. Med. Imaging*, vol. 36, no. 3, pp. 849–858, 2016.
30. P. Janani, J. Premaladha, and K. S. Ravichandran, "Image Enhancement Techniques: A Study," *Indian J. Sci. Technol.*, vol. 8, no. 22, pp. 1–12, 2015.