
Journal of Informatics and Web Engineering

Vol. 5 No. 3 (October 2026)

eISSN: 2821-370X

Deep Learning for Skin Cancer Classification Using Data Augmentation and Pre-Trained CNN Models

Pa Pa Min^{1*}, Fang Yong Lim², Jashila Nair Mogan³, Astie Darmayantie^{4}**

^{1,3}Centre for Intelligent Cloud Computing, Centre of Excellence for Advanced Cloud, Ayer Keroh 75450, Malaysia

^{1,2,3}Faculty of Information Science and Technology, Multimedia University, Ayer Keroh 75450, Malaysia

⁴Dept. Information Technology, Gunadarma University, Depok, Jawa Barat 16424, Indonesia

*corresponding author: (papamin@mmu.edu.my; ORCID: 0000-0002-3306-4073)

**corresponding author: (astie@staff.gunadarma.ac.id; ORCID: 0009-0001-1622-4741)

Abstract - Skin cancer is one of the most common forms of cancer worldwide, and its early detection is critical for effective treatment and improved survival rates. In recent years, advancements in Artificial Intelligence (AI), particularly deep learning, have shown great promise in enhancing the accuracy and efficiency of medical image analysis. This study aims to develop a non-invasive, AI-powered diagnostic system for the early detection and classification of skin cancer using dermoscopic images. The proposed approach leverages deep learning techniques to automatically recognize and differentiate between various types of skin lesions, including both malignant and benign cases. The project focuses on the end-to-end development of this system, beginning with the acquisition and preprocessing of a publicly available dataset of annotated skin lesion images. The images undergo enhancement and normalization processes to improve model performance. Pre-trained Convolutional Neural Network (CNN) architecture is designed and trained to classify the lesions based on visual features. The model's performance is evaluated using an independent validation set. By providing a reliable and accessible screening tool, this work aims to support dermatologists in early diagnosis and decision-making, potentially reducing the burden on healthcare systems and improving patient care. The study demonstrates the potential of integrating deep learning into medical diagnostics for skin cancer, paving the way for future research and clinical implementation.

Keywords—Dermoscopic Images, Skin Cancer, Pre-trained Model, Data Augmentation, Convolutional Neural Network

Received: 15 July 2025; Accepted: 26 March 2026; Published: 16 October 2026

This is an open access article under the [CC BY-NC-ND 4.0](#) license.



1. INTRODUCTION

Skin cancer is the most commonly diagnosed cancer worldwide, with millions of new cases reported each year. Its incidence continues to rise due to increased exposure to ultraviolet radiation and lifestyle changes. Despite being preventable and highly treatable when detected early, skin cancer can become life-threatening if diagnosis is delayed. [1]. Early detection significantly increases the chances of successful treatment, reduces the need for invasive procedures, and improves survival rates. However, conventional diagnostic methods are heavily reliant on the

experience and subjective judgment of dermatologists, which may lead to inconsistent outcomes, particularly in early-stage lesions that often appear visually benign or lack distinctive morphological features.

In clinical practice, visual examination is typically followed by dermoscopic evaluation and, if necessary, a biopsy. While dermoscopy enhances the diagnostic capabilities of dermatologists, it still requires considerable expertise, and diagnostic accuracy varies among practitioners. Misdiagnosis or delayed diagnosis can result in either unnecessary excision of benign lesions or failure to treat malignant ones promptly. These challenges highlight the urgent need for intelligent, objective, and consistent diagnostic tools to support clinical decision-making and ensure timely intervention [2], [3].

In this context, Artificial Intelligence (AI), particularly deep learning, has emerged as a powerful tool in medical image analysis. Deep learning models, especially Convolutional Neural Networks (CNNs), have demonstrated remarkable success in tasks such as object detection, image classification, and medical diagnostics. These models have the potential to extract and learn complex features from medical images, enabling the automated classification of skin lesions with accuracy comparable to, and sometimes surpassing, that of human experts [4], [5], [6], [7].

This study aims to develop an AI-based, non-invasive diagnostic system for the early detection and classification of skin cancer using dermoscopic images. The primary objective is to design a deep learning model capable of classifying skin lesions into multiple categories with high accuracy, thus aiding in early diagnosis and reducing the dependency on specialist interpretation. To achieve this, two publicly available and widely used datasets—HAM10000 and ISIC—are employed. These datasets contain a diverse set of dermoscopic images labelled with clinically verified diagnoses, encompassing a variety of skin lesion types such as Actinic Keratosis, Basal Cell Carcinoma, Dermatofibroma, Melanoma, Nevus, Pigmented Benign Keratosis, Seborrheic Keratosis, Squamous Cell Carcinoma, and Vascular Lesions.

The study begins with preprocessing and augmentation of the image datasets to address class imbalance and improve model generalization. Various pre-trained deep learning architectures are explored and implemented, including VGG16, VGG19, and ResNet50. These models are chosen for their proven effectiveness in image classification tasks and are trained to classify the dermoscopic images accordingly. The performance of each model is evaluated using standard metrics such as accuracy, sensitivity, specificity, precision, recall, and F1-score.

Beyond accuracy, this study also considers the feasibility of deploying pretrained models for real-time clinical use, particularly in remote or under-resourced settings. While the current work does not implement compression strategies, future research could explore techniques such as pruning or quantization to further reduce computational requirements and enhance deployment of mobile or edge devices.

The key contributions of this study are fourfold. First, transfer learning is applied with tailored preprocessing and augmentation strategies to enhance the classification of skin lesions. Second, a comparative evaluation of multiple pretrained CNN architectures is conducted on a benchmark dermoscopic dataset. Third, the study provides a comprehensive assessment of performance using accuracy, precision, recall, F1-score and confusion matrices to address the challenge of class imbalance. Finally, the feasibility of deploying the proposed approach in under-resourced clinical settings is discussed, highlighting its potential for real-world impact. In summary, this research contributes to the development of robust, interpretable, and accessible AI-driven diagnostic tools for skin cancer detection. By improving diagnostic consistency and enabling timely intervention, these tools have the potential to revolutionize dermatological care and significantly enhance patient outcomes on a global scale.

2. LITERATURE REVIEW

CNNs have demonstrated exceptional effectiveness in processing medical images for skin cancer detection, leveraging their capacity to extract complex features for accurate classification. These architectures enable enhanced diagnostic accuracy and reliability by automating the detection pipeline and supporting ensemble techniques for rapid processing.

Shah et al. [3] introduced LRNet, a deep CNN tailored for low-resolution skin cancer classification. The model utilized resized dermatoscopic images and focused on reducing computational overhead. Preprocessing included normalization and data augmentation such as flipping, jittering, rotation, and brightness/contrast adjustment. The

architecture comprised convolutional, dropout, batch normalization, and pooling layers to support efficient learning on low-resolution inputs.

Pillai et al. [4] proposed a 26-layer CNN model for the diagnosis and classification of skin cancer into seven distinct classes using the HAM10000 dataset. To mitigate the problem of class imbalance, the authors applied data augmentation techniques such as rotation, zooming, and flipping. Alshalman et al. proposed a cost-effective CNN model for the early detection and classification of skin cancer. Their model classifies dermoscopic images into seven distinct skin cancer categories. Images from the HAM10000 dataset, comprising 10,015 annotated images of various skin lesions, were utilized for both training and testing. Preprocessing involved resizing the images to 100×75 pixels and applying Z-score normalization to minimize bias and enhance training efficiency [5].

Zhu et al. [6] conducted an analysis of three CNN architectures—VGG-16, VGG-19, and a custom-designed model—for the task of nine-class skin cancer classification. The objective was to identify the most effective model and integrate it into a web-based application for early diagnosis support. The study utilized a publicly available dataset from Kaggle comprising 6,857 dermoscopic images of nine distinct skin cancer types. Preprocessing steps included resizing the images to 128×128 pixels and normalizing pixel values to a $[0,1]$ range by dividing by 255. The dataset was subsequently split into training and testing subsets. All models were trained using the Adam optimizer with sparse categorical cross-entropy as the loss function [4].

Dhir et al. [7] proposed a robust CNN-based framework for the classification of skin lesions into seven categories to facilitate early detection of skin cancer. The model was trained on the HAM10000 dataset, and comprehensive data preprocessing steps were applied. These included converting the images to JPG format, resizing them to 28×28 pixels, and performing data augmentation techniques such as rotation and flipping to address class imbalance. Additionally, missing values (e.g., null ages) were replaced with the mean age to ensure data completeness. The CNN model was trained using the Adam optimizer with an initial learning rate of 0.001, a batch size of 64, and 25 epochs. Hyperparameter tuning was conducted to optimize the architecture, including the number of dense neurons, activation functions, and learning rates [6].

In a related study, Subramanian et al. [8] introduced a 16-layer CNN model to detect and classify skin cancer using clinical images from the HAM10000 dataset. The goal was to achieve high accuracy with minimal false negatives (below 10%) and high precision (above 80%). To address the issue of class imbalance, data augmentation techniques such as random cropping, rotation, and both horizontal and vertical shifts were applied. The input images were resized to 75×100 pixels and normalized. The CNN architecture included convolutional layers, max-pooling layers, and fully connected layers, and was trained in TensorFlow with a batch size of 32 over 50 epochs. The proposed model attained a training accuracy of 83.11% and a testing accuracy of 83.04% on the HAM10000 dataset [8].

Transfer learning has emerged as a powerful technique in skin cancer detection by enabling the fine-tuning of pre-trained models on domain-specific datasets. This facilitates effective learning even in data-scarce scenarios and supports the development of efficient models suitable for clinical integration. Several studies have explored and compared different deep learning techniques based on speed of inference, classification capability, and overall effectiveness to identify optimal approaches for real-world medical applications. Pattanaik et al. [9] proposed a cascaded framework combining U-Net for lesion segmentation with CNN-based classifiers to enhance the reliability of skin cancer diagnosis. The ISIC dataset comprising 2,357 images from nine skin cancer classes was utilized, and data augmentation methods such as rotation, zooming, and flipping were applied to address class imbalance [10]. Three CNN models—MobileNet, DenseNet121, and ResNet50 were trained on segmented images using the Adam optimizer for 50 epochs, with a batch size of 32.

Pilania et al. [11] introduced a hybrid deep learning framework that sequentially integrates a binary classifier, ResNet for feature extraction, and a CNN for final classification. The dataset used contained 10,015 dermoscopic images obtained from Kaggle. ResNet was initialized with pre-trained weights and fine-tuned for the skin cancer task. Patel et al. [12] conducted a comparative study using multiple pre-trained deep learning models VGG16, EfficientNetB7, ResNet50, InceptionResNetV2, and DenseNet201 as well as a custom CNN architecture for classifying skin lesions into seven categories. The HAM10000 dataset was used, and data balancing was performed through augmentation techniques including rotation, zooming, and flipping. These models were trained using the Adam optimizer and categorical cross-entropy loss, and performance was evaluated on both balanced and imbalanced datasets [12].

Kondaveeti et al. [13] developed an enhanced classification framework for early skin cancer diagnosis using transfer learning with several pre-trained CNNs (ResNet50, InceptionV3, MobileNet, and Xception). The HAM10000 dataset,

consisting of 10,015 images, was pre-processed by removing duplicates, resizing to 224×224 pixels, applying augmentation (e.g., flipping and cropping), and stratified sampling. Pre-trained models were modified by replacing top layers with global average pooling, dropout, and dense layers and Class weights were also introduced to counteract class imbalance during training.

Jasil et al. [14] proposed a deep neural network for skin cancer classification using the pre-trained DenseNet201 architecture through transfer learning. The model was developed for classifying seven types of dermoscopic images. The ISIC 2018 dataset was used, and preprocessing included resizing images to 224×224 pixels, applying augmentation techniques such as zooming, flipping, shifting, and shearing, along with normalization to scale pixel values

Rajesh et al. [15] introduced a methodology utilizing Deep Convolutional Neural Networks (DCNNs) for skin cancer detection and lesion intensity analysis. Their work involved fine-tuning two architectures—Inception V3 and DenseNet201—and further proposed an ensemble model combining both. The HAM10000 dataset was used with images resized to 192×256 pixels and normalized. Data augmentation strategies included rotation, shear, zoom, and shifting. The models were enhanced through transfer learning and additional architectural modifications.

Segmentation and classification play a key role in skin cancer diagnosis. Kumar et al. [16] conducted a comparative study involving Artificial Neural Networks (ANNs), CNNs, and transfer learning methods. The study used a dataset of dermoscopic images from Harvard's "Human Against Machine" project. Preprocessing steps involved cleaning, resizing images to 450×600 pixels, and removing artifacts such as hair using the Dull Razor technique. Models were developed using standard activation functions and fine-tuned pre-trained networks for classification tasks.

AI-powered skin cancer diagnostic tools aid in reducing the risk of late-stage cancers by enabling remote diagnosis through teledermatology and assisting dermatologists in clinical decision-making. These tools can increase access to care, reduce unnecessary treatments, and support personalized medicine in a more cost-effective manner.

Barbadekar et al. [17] proposed a two-stage system for skin cancer detection. The first stage involved lesion detection using the BCDU-Net architecture, while the second stage used VGG-19 and DenseNet for classifying the lesions. The system utilized dermoscopic images from the HAM10000 and ISIC-19 datasets. Preprocessing involved normalization and various augmentation techniques such as flipping, cropping, brightness adjustment, and contrast enhancement

Dubal et al. [18] developed a method for skin lesion classification using general camera images. Their system incorporated preprocessing, lesion segmentation, and feature extraction based on the ABCD rule—Asymmetry, Border, Colour, and Diameter. Feature-based classification was carried out using a feedforward artificial neural network. Preprocessing steps included grayscale conversion, lighting enhancement, and hair removal using bottom-hat filtering and morphological operations [13].

Pavan et al. [19] proposed a classification methodology using an enhanced AlexNet architecture combined with transfer learning. The HAM10000 dataset was used, and preprocessing involved resizing, normalization, and a variety of augmentation methods including cropping, scaling, flipping, and contrast adjustment. This methodology aimed to optimize classification by refining pre-trained networks for the task of skin cancer diagnosis [13].

Monajemi et al. [20] presented a lightweight diagnostic approach using the MobileNet model for the classification of multi-type lesions. Their approach involved preprocessing, standardization, and model construction. They applied model fusion techniques and used datasets such as ISIC2019 and HAM10000. The method emphasized efficiency and addressed challenges of data imbalance and limited labelled samples using augmentation and pair-based sampling.

More recent studies have continued to refine these directions. Shakya et al. [21] carried out a comprehensive comparison of fine-tuned pre-trained networks (VGG19, ResNet18, and MobileNetV2) employed both as end-to-end classifiers and as feature extractors coupled with conventional machine learning classifiers, reporting a peak accuracy of 92.87% on the ISIC 2018 dataset when ResNet18 and MobileNet features were concatenated and classified using an SVM. Alotaibi and AlSaeed [22] integrated self-, soft-, and hard-attention mechanisms into a pre-trained Xception backbone for binary lesion classification on HAM10000, raising accuracy from 91.05% for the plain Xception model to 94.11% with self-attention. Alrabai et al. [23] combined image augmentation with four pre-trained architectures (InceptionV3, Xception, ResNet50V2, and DenseNet121) and coupled the best-performing Xception model with explainable AI techniques, reporting 90.15% accuracy while emphasising interpretability as a prerequisite for clinical trust. Addressing class imbalance more directly, Kılıç and Dogan [24] applied SMOTE and data augmentation

alongside a layer-based pruning strategy and Avg-TopK pooling to an Xception backbone, attaining 91.52% accuracy on HAM10000 while reducing the parameter count from 20.9 million to 13.5 million. Hybrid architectures have also gained traction: Yousaf et al. [25] proposed a transformer-CNN framework in which a global contextual vision transformer performs classification and a companion network performs segmentation, reporting accuracies between 97% and 100% across the PH2, ISIC-2019, ISIC-2020, and HAM10000 datasets, while Murali and Mazumder [26] fused dual attention with patient metadata and Grad-CAM explanations to achieve 94.8% accuracy and a macro F1-score of 91.9% on HAM10000. From a deployment perspective, Tjong et al. [27] benchmarked a CNN trained on 10,015 dermoscopic images on single-board computers such as the Raspberry Pi 5 and Jetson Nano, reporting 98.25% accuracy with an inference time of 0.01 s and demonstrating that on-device screening is feasible without continuous network connectivity. Taken together, these recent works confirm that transfer learning remains the dominant paradigm, that class imbalance and augmentation continue to govern achievable accuracy, and that attention mechanisms, hybrid transformer designs, and model compression are the principal directions along which the field is currently advancing.

3. RESEARCH METHODOLOGY

The methodology of this study involves the exploration and deployment of various pre-trained deep learning architectures tailored for the classification and detection of skin cancer. The pre-trained model based on CNNs, such as VGG16, VGG19, and ResNet50 are used for this study. The underlying rationale is to harness their architectural strengths to enhance the precision and robustness of lesion classification tasks. CNN falls under the class of deep learning, which finds many applications in the image domain because these kinds of models process data in gridded format by convolutional layers autonomously and hierarchically learning features. The architecture of the three pre-trained model are illustrated in Figure 1, 2 and 3.

3.1 Pre-trained Model

3.1.1 VGG16

The VGG16 model applied in this study employs a pre-trained convolutional base trained on the ImageNet dataset (Figure 1). It comprises five convolutional blocks with 3x3 filters and ReLU activations. Each block ends with a MaxPooling2D layer to reduce feature map dimensions. The number of filters in each block increases progressively from 64 to 512. The top fully connected layers are excluded (i.e., include_top=False), and the convolutional base outputs feature maps. These are passed to a MaxPooling2D layer, followed by a Dense layer with 1024 neurons (ReLU activation) and an output layer with seven softmax-activated neurons. The pre-trained base layers are frozen during fine-tuning.

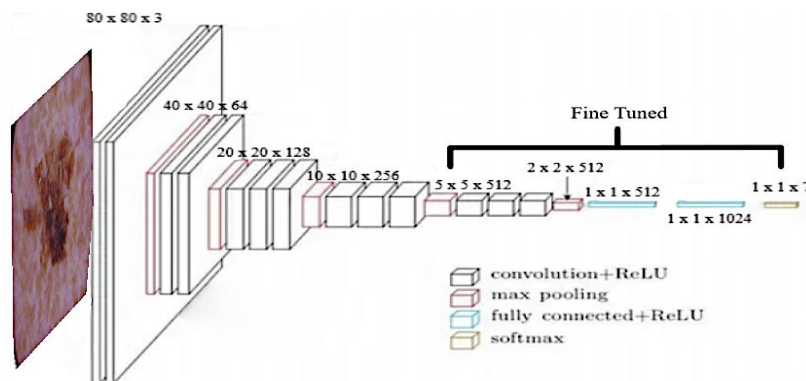


Figure 1. The Architecture of VGG-16 for Skin Cancer Classification

3.1.2 VGG19

The VGG19 architecture shares structural similarities with VGG16 but includes additional convolutional layers for deeper representation learning (Figure 2). It is organized into five convolutional blocks, the first two having two

Conv2D layers each and the last three having four Conv2D layers each, all with 3x3 kernels and ReLU activation. MaxPooling2D layers with a 2x2 pool size follow each block. The feature maps generated by the convolutional base (with include_top=False) are processed using a GlobalAveragePooling2D layer. Two Dense layers—one with 1024 neurons and another with seven neurons (softmax activation)—constitute the classification head. Pre-trained weights are preserved by freezing the base layers during adaptation.

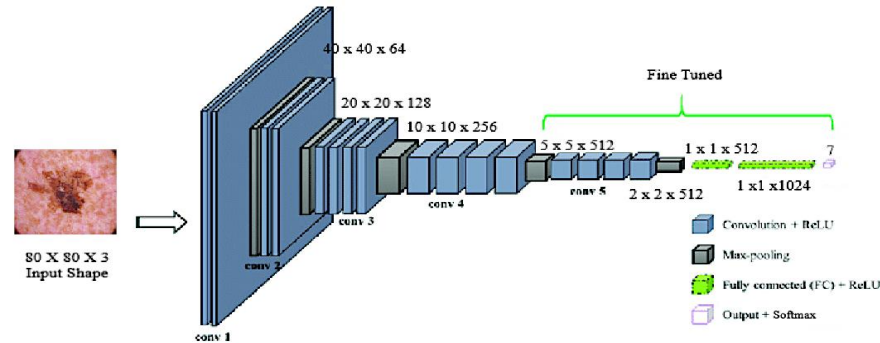


Figure 2. The Architecture of VGG-19 for Skin Cancer Classification

3.1.3 ResNet50

ResNet50, a deep residual network consisting of 50 layers, leverages identity shortcuts or residual connections to mitigate vanishing gradient problems in deep architectures (Figure 3). In this work, the pre-trained ResNet50 model (trained on ImageNet) is repurposed by removing its top layers and replacing them with a GlobalAveragePooling2D layer followed by two Dense layers: one with 1024 neurons (ReLU activation) and another with seven softmax-activated neurons for multiclass classification. The residual connections allow effective gradient flow, making the architecture suitable for deep and complex image classification tasks.

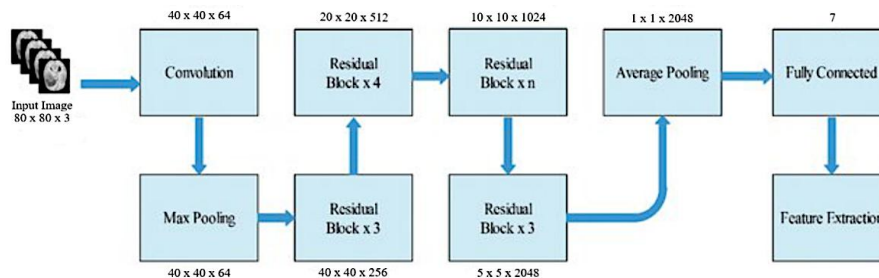


Figure 3. The Architecture of ResNet50 for Skin Cancer Classification

The models implemented in this study demonstrate a progression from basic to highly sophisticated deep learning architectures. Each model integrates critical components—such as convolutional feature extractors, pooling mechanisms, and dense classification layers—tailored to efficiently process dermoscopic images and distinguish between skin lesion categories. The selection and configuration of layers in each architecture reflect a deliberate strategy to balance computational complexity with classification accuracy in a multiclass setting.

3.2 Datasets

3.2.1 Skin Cancer MNIST: HAM10000

The HAM10000 dataset contains 10,015 dermoscopic images representing seven types of pigmented skin lesions: dermatofibroma, melanoma, nevus, seborrheic keratosis, squamous cell carcinoma, vascular lesions, and basal cell carcinoma. The images were collected from a diverse population using varied methods. Accompanying the dataset is a CSV file with metadata, including patient age, sex, lesion type, skin type, image, and image ID. Widely used in dermatology and machine learning research, HAM10000 serves as a benchmark for model evaluation and comparison.

with expert diagnoses. It is publicly available on platforms such as Kaggle and Papers with Code. The details of the HAM10000 dataset can be found in Table 1 and sample images can be found in Figure 4.

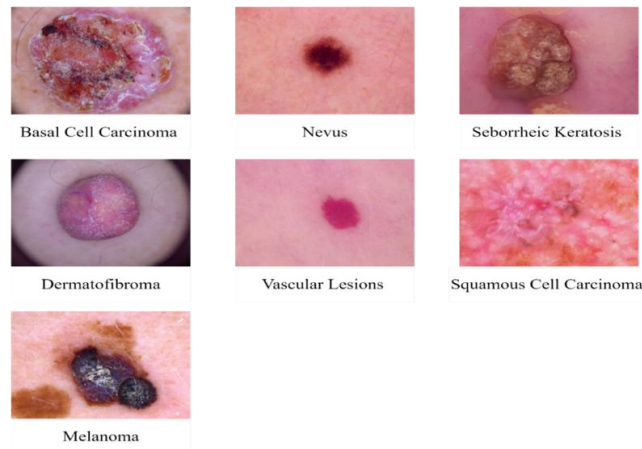


Figure 4. Sample Images from HAM-10000 Dataset

Table 1. Number of Images for Each Class in HAM-10000 Dataset

Skin Cancer	Total Number of Images
Basal Cell Carcinoma	514
Dermatofibroma	115
Melanoma	1113
Nevus	6705
Squamous Cell Carcinoma	327
Vascular Lesions	142
Seborrheic Keratosis	1099
Total	10015

3.2.2 Skin Cancer International Skin Imaging Collaboration (ISIC)

The International Skin Imaging Collaboration, or ISIC dataset has images of both malignant and benign oncological diseases from 2,357 pictures. The images were categorized according to ISIC and split into sets with an equivalent number of images, except in the case of melanomas and moles, where there were a little more. The dataset includes actinic keratosis, basal cell carcinoma, dermatofibroma, melanoma, nevus, pigmented benign keratosis, seborrheic keratosis, squamous cell carcinoma, and vascular lesions. The details of the ISIC dataset can be found in Table 2 and sample images can be found in Figure 5.

Table 2. Number of Images for Each Class in ISIC Dataset

Skin Cancer	Total Number of Images
Basal Cell Carcinoma	392
Dermatofibroma	111
Melanoma	454
Nevus	373
Squamous Cell Carcinoma	197
Vascular Lesions	142
Seborrheic Keratosis	80
Actinic Keratosis	130
Pigmented Benign Keratosis	478
Total	2357

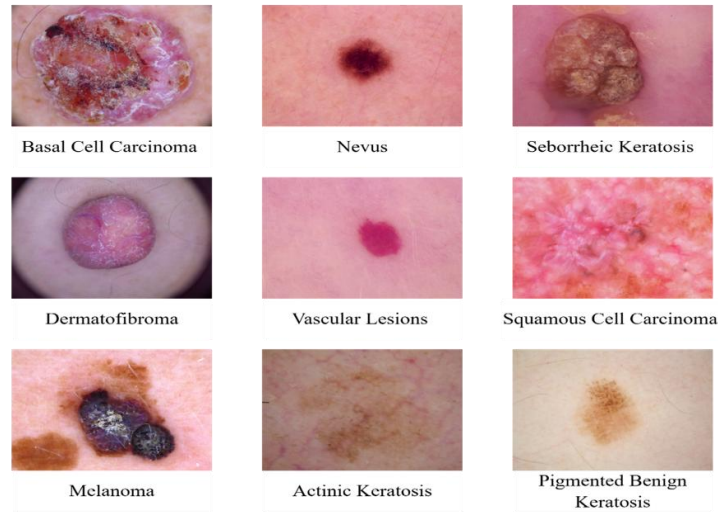


Figure 5. Sample Images from ISIC Dataset

3.3 Data Augmentation, Preprocessing and Splitting

To address class imbalance and enhance dataset diversity, both HAM10000 and ISIC datasets were augmented using a Python script. Techniques included random cropping, zooming (scaling from $1.1\times$ to $1.5\times$ followed by re-cropping), rotation (-45° to 45°), Gaussian blur (kernel sizes 3, 5, or 7), and image flipping (horizontal, vertical, or both). Augmentation continued until reaching a maximum of 6,700 images for MNIST HAM10000 and 550 for ISIC, based on the size of the largest class.

For preprocessing, metadata from HAM10000_metadata.csv was parsed using the pandas library. Images were organized into class-specific subdirectories, resized to 80×80 pixels in RGB format, and their labels were binarized using LabelBinarizer and converted to one-hot vectors via `tf.keras.utils.to_categorical`. This format is compatible pre-trained models like VGG16, VGG19, and ResNet50. Error handling was implemented to ensure a smooth workflow.

Finally, the datasets were split into training (60%), validation (20%), and test (20%) sets. The validation set facilitated hyperparameter tuning and overfitting control, while the test set evaluated model generalization. This ensured a comprehensive and reliable machine learning pipeline.

3.4 Model Optimization and Fine-Tuning

Model optimization in machine learning often involves hyperparameter tuning, such as adjusting the learning rate, batch size, or layer configurations. In this study, fine-tuning was performed by unfreezing the last eight layers of pre-trained models (VGG16, VGG19, and ResNet50) and training them for 10 epochs. This strategy retains general features learned from large datasets while allowing the final layers to adapt to the specific characteristics of the target dataset. By focusing on deeper task-specific patterns without overfitting, this fine-tuning approach enhances classification performance and accuracy.

4. RESULTS AND DISCUSSIONS

To evaluate the effectiveness of the proposed models and preprocessing strategies, experiments were conducted on four datasets: HAM10000 original, HAM10000 augmented, ISIC original, and ISIC augmented. The augmented datasets were generated to address class imbalance and improve generalization. Each dataset was used to train and test three pre-trained models (VGG16, VGG19, and ResNet50) with fine-tuning applied to their final layers.

Performance metrics such as accuracy, precision, recall, and F1-score were computed to assess model effectiveness across original and augmented datasets. The results reveal the impact of data augmentation and transfer learning on classification performance for skin lesion detection on Table 3 and Table 4 respectively.

Table 3. Accuracy of Proposed Pre-trained Models for Original Datasets

	HAM-10000 Original Dataset	ISIC Original Dataset
VGG16	80 %	56 %
VGG19	72 %	57 %
RESNET50	76 %	54 %

Table 4. Accuracy of Proposed Pre-trained Models for Augmented Datasets

	HAM-10000 Augmented Dataset	ISIC Augmented Dataset
VGG16	90 %	71 %
VGG19	89 %	71 %
RESNET50	89 %	74 %

Among the models tested, VGG16 consistently outperformed others on the HAM10000 dataset, particularly when data augmentation was applied. It achieved an impressive accuracy of 90% on the augmented dataset, compared to 80% on the original. This suggests that VGG16 benefits significantly from the increased diversity and balance provided by augmentation. The confusion matrix for VGG16 on the augmented dataset also revealed improved classification across all lesion classes, indicating better generalization and reduced class bias. The performance drop on the original dataset suggests that VGG16 may be sensitive to class imbalance and limited sample variability.

VGG19 and ResNet50 also demonstrated strong performance, particularly on the HAM10000 and ISIC augmented datasets. VGG19 achieved 89% accuracy on the HAM10000 augmented set and 71% on ISIC augmented, showing similar behaviour to VGG16 but with slightly lower performance on the HAM10000 original dataset (72%). ResNet50 showed the best results for the ISIC dataset, with 74% accuracy on the augmented version and 54% on the original. This indicates that ResNet50's deeper architecture and residual connections are more effective in capturing the complex patterns in ISIC images, especially after augmentation. Across all models, performance was consistently higher on augmented datasets, confirming the critical role of data augmentation in improving model robustness and accuracy.

The comparative analysis revealed that VGG16 is the most suitable model for HAM10000, while ResNet50 is more effective for the ISIC dataset. This distinction can be attributed to differences in dataset characteristics, where HAM10000 benefits more from simpler feature hierarchies captured by VGG16, and ISIC requires deeper, more abstract representations offered by ResNet50. These findings underscore the value of model selection based on dataset-specific attributes and the impact of augmentation in addressing data limitations common in medical image classification tasks. The confusion matrix got highest accuracy with VGG16 and VGG19 are illustrated in Figure 6 and 7.

5. CONCLUSION

In conclusion, this study demonstrates the significant potential of deep learning techniques in enhancing the accuracy and efficiency of skin cancer diagnosis through dermoscopic image analysis. Pre-trained models such as VGG16, VGG19, and ResNet50, when fine-tuned and supported by proper data preprocessing and augmentation, achieved diagnostic accuracy comparable to or surpassing that of experienced dermatologists. These improvements support early detection, enabling timely medical intervention and improved patient outcomes. Furthermore, the scalability and automation capabilities of such AI-based diagnostic systems open pathways for deployment in remote and underserved regions, addressing healthcare disparities and extending high-quality care to broader populations. The integration of deep learning in dermatological diagnostics presents a transformative opportunity not only to improve individual health outcomes but also to enhance the overall effectiveness of healthcare delivery systems. By promoting equitable

access and early detection, AI-driven image analysis tools hold the promise to revolutionize skin cancer diagnosis and treatment across diverse settings. In summary, this work contributes not only by applying deep learning for skin cancer detection, but also by rigorously evaluating pretrained models with multiple metrics, addressing dataset imbalance, and situating the findings in the context of clinical feasibility. These aspects extend the study beyond a straightforward application of existing techniques and provide a foundation for future research and deployment in real-world healthcare environments. Building on the recent literature, future work will focus on hybrid transformer-CNN designs and attention-based feature refinement [25], [26], and on compressing the resulting models for on-device inference so that screening remains feasible in settings without reliable connectivity [27].

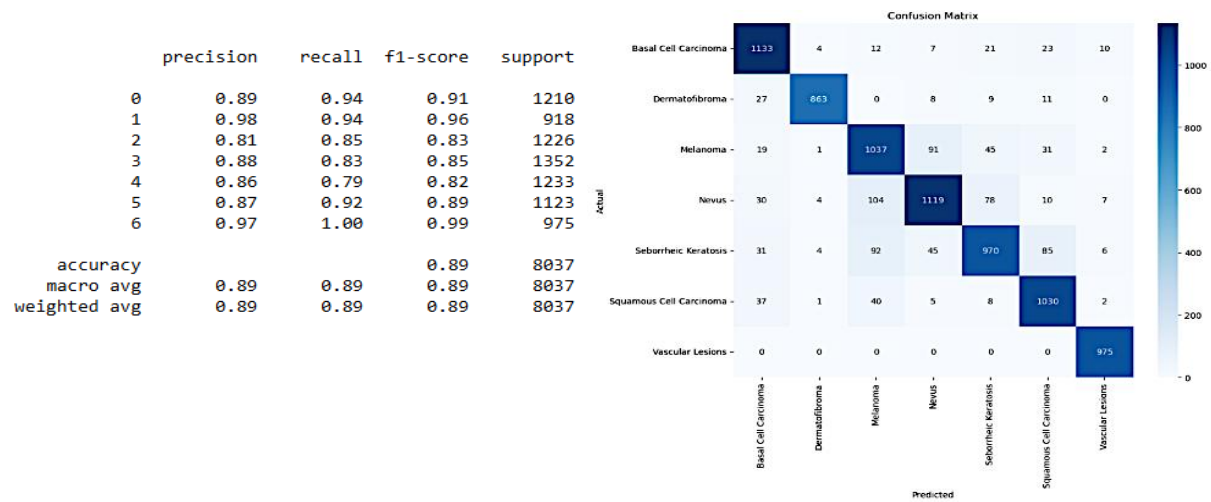


Figure 6. Confusion Matrix of VGG16 Model Training on HAM-10000 Augmented Dataset

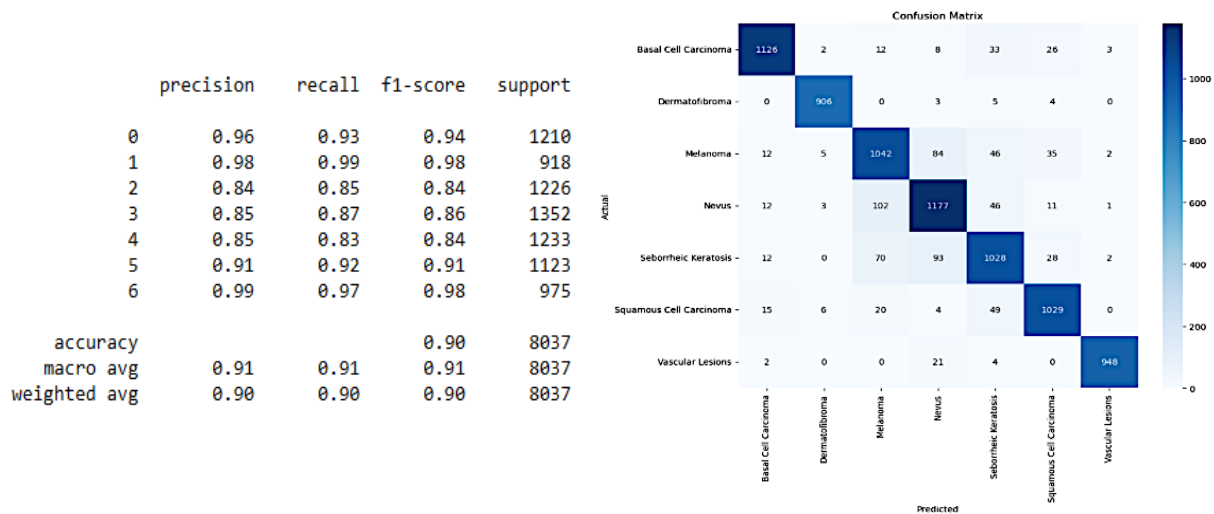


Figure 7. Confusion Matrix of VGG19 Model training on HAM-10000 augmented dataset

ACKNOWLEDGEMENT

The authors gratefully acknowledge the support and guidance from Multimedia University. Appreciation is also extended to the creators of the HAM10000 and ISIC datasets, which were essential for this research.

FUNDING STATEMENT

The authors received no funding from any party for the research and publication of this article.

AUTHOR CONTRIBUTIONS

Pa Pa Min: Methodology, Supervision, Editing, Review, Final Draft Preparation;
Fang Yong Lim: Conceptualization, Data Curation, Methodology, Validation, Writing – Original Draft Preparation;
Jashila Nair Mogan: Review & Editing;
Astie Darmayantie: Review & Editing.

CONFLICT OF INTERESTS

There is no conflict of interests.

ETHICS STATEMENTS

Our publication ethics follow The Committee of Publication Ethics (COPE) guideline. <https://publicationethics.org/>

DATA AVAILABILITY

The data supporting this study's findings are openly available in public repositories. The HAM10000 dataset is available via the ISIC archive, and the ISIC dataset can be accessed at <https://www.isic-archive.com>. These data were derived from sources in the public domain and are widely used for skin lesion analysis research.

REFERENCES

- [1] D. Praneetha, M.L. Pandala, and G. Nithin Sai, "Computer vision-based deep learning for comprehensive skin cancer detection: precise diagnosis and early intervention in dermatology," *2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSSES)*, 1–8, pp. 1–8, Dec. 2023, doi: 10.1109/ICSSES60034.2023.10465363.
- [2] A. S. Al-Waisy, S. Al-Fahdawi, M. I. Khalaf, M. A. Mohammed, B. Al-Attar, and M. N. Al-Andoli, "A deep learning framework for automated early diagnosis and classification of skin cancer lesions in dermoscopy images," *Scientific Reports*, vol. 15, art. no. 31234, Aug. 2025, doi: 10.1038/s41598-025-15655-9.
- [3] M. Shah, "LRNet: Skin cancer classification using low-resolution images," in *2021 International Conference on Communication information and Computing Technology (ICCICT)*, pp. 1–5, Jun. 2021, doi: 10.1109/ICCICT50803.2021.9510138
- [4] R. Pillai, N. Sharma, and R. Gupta, "Proposed convolution neural network for skin cancer diagnosis and classification," *2023 Int. Conf. Recent Advances in Electrical, Electronics & Digital Healthcare Technologies (REEDCON)*, pp. 308–312, May 2023, doi: 10.1109/REEDCON57544.2023.10151029.
- [5] M. Alshalman, B. F. Gargoum, T. Nagem, and K. A. Bozed, "Skin cancer detection by using deep learning approach," *2023 IEEE 11th International Conference on Systems and Control (ICSC)*, pp. 1–7, Dec. 2023, doi: 10.1109/ICSC58660.2023.10449804.
- [6] P. Zhu, "Convolutional neural networks based study and application for multicategory skin cancer detection," *2022 3rd International Conference on Electronic Communication and Artificial Intelligence (IWECAI)*, Jan. 2022, doi: 10.1109/IWECAI55315.2022.00114.

- [7] S. Dhir, A. Sinha, C. Bansal, R. Rani, N. Goel, and P. Bhutani, "A robust CNN-based approach for skin lesion detection and classification," *2023 3rd International Conference on Intelligent Technologies (CONIT)*, pp. 1–6, Jun. 2023, doi: 10.1109/CONIT59222.2023.10205907.
- [8] R. R. Subramanian, D. Achuth, P. S. Kumar, K. N. Kumar Reddy, S. Amara, and A. S. Chowdary, "Skin Cancer Classification Using Convolutional Neural Networks," *2021 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pp. 13–19, Jan. 2021, doi: 10.1109/Confluence51648.2021.9377155.
- [9] A. Pattanaik, L. Das, and S. Banerjee, "Cascaded Approach for Image Segmentation and Classification for Skin Cancer Detection," in *2024 5th International Conference for Emerging Technology (INCET)*, Karnataka, India, May 24–26, 2024, pp. 1–5, doi: 10.1109/INCET61516.2024.10592933
- [10] S. Sharmiladevi, S. Venkatraman, S. Malarvannan, A. Shriyans, and S. Swathisree, "A comprehensive study on skin cancer detection using deep learning," *2024 IEEE 3rd World Conference on Applied Intelligence and Computing (AIC)*, pp. 749–754, Jul. 2024, doi: 10.1109/AIC61668.2024.10731010.
- [11] U. Pilania, M. Kumar, P. Garg, and R. Kaur, "Detection and classification of skin cancer using binary classifier, residual network, and convolutional neural network," *2024 2nd International Conference on Sustainable Computing and Smart Systems (ICSCSS)*, pp. 1115–1122, Jul. 2024, doi: 10.1109/ICSCSS60660.2024.10625032.
- [12] H. K. Kondaveeti and P. Edupuganti, "Skin cancer classification using transfer learning," *2020 IEEE International Conference on Advent Trends in Multidisciplinary Research and Innovation (ICATMRI)*, pp. 1–4, Dec. 2020, doi: 10.1109/ICATMRI51801.2020.9398388.
- [13] H. Patel, D. Thakore, and N. Patel, "Skin lesion classification using deep learning techniques," *2024 Parul University International Conference on Engineering and Technology (PICET)*, pp. 1–6, May 2024, doi: 10.1109/PICET60765.2024.10716067.
- [14] S. P. G. Jasil and V. Ulagamuthalvi, "Skin lesion classification using pre-trained DenseNet201 deep neural network," *2021 3rd International Conference on Signal Processing and Communication (ICPSC)*, pp. 393–396, May 2021, doi: 10.1109/ICSPC51351.2021.9451818.
- [15] A. V. P. Rajesh, K. N. Rao, G. N. V. Sai, K. D. Kumar, and K. R. S. Karthik, "skin cancer detection and intensity analysis using deep learning," *2024 International Conference on Emerging Systems and Intelligent Computing (ESIC)*, pp. 376–381, Feb. 2024, doi: 10.1109/ESIC60604.2024.10481663.
- [16] A. A. Kumar, M. V. K. Reddy, V. M. Sanjay, V. Kothuru, K. P. Shah, and A. Kalucha, "Comparative analysis of skin cancer prediction using neural networks and transfer learning," *2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Mar. 2022, doi: 10.1109/ICACCS54159.2022.9785079.
- [17] A. Barbadekar, V. Ashtekar, and A. Chaudhari, "Skin cancer classification and detection using VGG-19 and DesNet," *2023 International Conference on Computational Intelligence, Networks and Security (ICCINS)*, pp. 1–6, Dec. 2023, doi: 10.1109/ICCINS58907.2023.10449996.
- [18] P. Dubal, S. Bhatt, C. Joglekar, and S. Patil, "Skin Cancer Detection and Classification," *2017 6th International Conference on Electrical Engineering and Informatics (ICEEI)*, pp. 1–6, Nov. 2017, doi: 10.1109/ICEEI.2017.8312419.
- [19] M. E. N. Pavan, U. B. Kumar Reddy, C. Mohan Ghantasala, M. Y. Devi, and A. Nesarani, "Enhanced skin cancer classification with alexnet and transfer learning," *2023 International Conference on Sustainable Communication Networks and Application (ICSCNA)*, pp. 1020–1024, Nov. 2023, doi: 10.1109/ICSCNA58489.2023.10370110.
- [20] M. Monajemi, S. O. Shahdi, S. S. H. Nasiri, and S. V. Shojaedini, "Efficiency analysis of convolutional neural networks in multiclass recognition of skin lesions via dermoscopy images," *2024 10th International Conference on Artificial Intelligence and Robotics (QICAR)*, pp. 119–124, Feb. 2024, doi: 10.1109/QICAR61538.2024.10496625.

- [21] M. Shakya, R. Patel, and S. Joshi, "A comprehensive analysis of deep learning and transfer learning techniques for skin cancer classification," *Scientific Reports*, vol. 15, art. no. 4633, Feb. 2025, doi: 10.1038/s41598-024-82241-w.
- [22] A. Alotaibi and D. AlSaeed, "Skin cancer detection using transfer learning and deep attention mechanisms," *Diagnostics*, vol. 15, no. 1, art. no. 99, Jan. 2025, doi: 10.3390/diagnostics15010099.
- [23] A. Alrabai, A. Echtioui, and F. Kallel, "Explainable deep learning approaches for skin cancer diagnosis," *Network Modeling Analysis in Health Informatics and Bioinformatics*, vol. 14, no. 1, art. no. 57, Dec. 2025, doi: 10.1007/s13721-025-00554-w.
- [24] S. Kilic and Y. Dogan, "A pruned and parameter-efficient Xception framework for skin cancer classification," *PLOS ONE*, vol. 21, no. 3, art. no. e0341227, Mar. 2026, doi: 10.1371/journal.pone.0341227.
- [25] N. Yousaf, J. Amin, W. H. Butt, A. Zafar, and S. Kim, "Advanced hybrid transformer CNN framework for improved skin lesion classification and segmentation," *Scientific Reports*, vol. 16, art. no. 13592, 2026, doi: 10.1038/s41598-026-43376-0.
- [26] P. Murali and D. H. Mazumder, "DermaScanAI: an explainable hybrid deep learning framework for automated skin lesion classification using dual attention and metadata fusion," *Scientific Reports*, vol. 16, art. no. 15130, 2026, doi: 10.1038/s41598-026-46011-0.
- [27] V. Tjong, G. Darian, and N. Surantha, "Performance evaluation of convolutional neural network (CNN) for skin cancer detection on edge computing devices," *Applied Sciences*, vol. 15, no. 6, art. no. 3077, Mar. 2025, doi: 10.3390/app15063077.

BIOGRAPHIES OF AUTHORS

	Pa Pa Min is a lecturer at Multimedia University, Malaysia, and a PhD researcher specializing in explainable artificial intelligence and socioeconomic data analysis. Her research interests include artificial intelligence, computer vision, deep learning, data analytics, and socioeconomic data science. She is actively involved in teaching and research in data science-related fields and has contributed to several academic publications and research projects. She can be contacted at email: papamin@mmu.edu.my.
	Fang Yong Lim graduated from the Faculty of Information Science and Technology (FIST) from Multimedia University with a specialization in Artificial Intelligence. His research interests include artificial intelligence, machine learning, and data analytics. He is particularly interested in applying AI techniques to solve real-world problems and develop intelligent data-driven systems. He can be contacted at email: 1201201544@student.mmu.edu.my.
	Jashila Nair Mogan is currently a Senior Lecturer in the Faculty of Information Science and Technology (FIST), Multimedia University. She received her Bachelor's Degree (Hons) majoring in Artificial Intelligence, Master of Science (I.T.), and Ph. D (I.T) from Multimedia University. Her research interests are in artificial intelligence, computer vision, deep learning, and gait recognition. She can be contacted at email: jashilanair@mmu.edu.my.
	Astie Darmayantie is a lecturer at Gunadarma University, Indonesia. Her research interests include artificial intelligence, machine learning, software engineering, and information systems. She has published several papers in international journals and conferences related to data analytics and intelligent systems. She can be contacted at email: astie@staff.gunadarma.ac.id.