Journal of Soft Computing and Computer Applications

Manuscript 1026

Skin Lesion Classification Using CNN Model and Augmented Dataset

Mohammed Nawzad Mohammed-Ramzi

Aso M. Aladdin

JOURNAL OF SOFT COMPUTING AND COMPUTER APPLICATIONS 2026;3:1026

ORIGINAL STUDY

Skin Lesion Classification Using CNN Model and Augmented Dataset

Mohammed Nawzad Mohammed-Ramzi [a], **Aso M. Aladdin** [b,*]

[a] Tishk International University, College of Applied Science, Information Technology Department, Qularaisi, 46001 Sulaymaniyah, Kurdistan Region, Iraq
[b] Charmo University, College of Science, Computer Science Department, Peshawa Street, Chamchamal, 46023 Sulaymaniyah, Kurdistan Region, Iraq

ABSTRACT

Skin cancer is a deadly disease. Skin lesion classification is a critical challenge due to its prevalent and deadly nature. Skin lesions are difficult for dermatologists to detect using eye examination, which is time-consuming and variable. A deep learning model of skin lesions classification has been proposed using a Convolutional Neural Network (CNN) trained on the HAM10000 dataset of 10,015 dermatoscopies. To improve resilience and address the dataset's extreme class imbalance, data augmentation techniques such as geometric transformations, brightness/contrast adjustments, blurring, noise addition, histogram equalization, color space alterations, and elastic deformations are used. With a carefully balanced 10% test set, the model can accurately distinguish seven skin lesions with an overall accuracy of 90.75%, a macro-average F1-score of 0.92, and a good overall Area Under the Curve - Receiver Operating Characteristics (AUC - ROC) score of 0.9911. To show its effectiveness, the model is compared to a trained ResNet50 model. The custom CNN performs better than ResNet50 (87.64% accuracy despite early overfitting). Deep learning models may be highly useful in skin cancer screening, according to the findings. This study could simplify worrisome lesion identification in primary care, mobile health apps, and underserved areas.

Keywords: Convolutional neural network, Skin lesion, Image augmentation, Classification

1. Introduction

Artificial Intelligence (AI) has become a powerful tool in the healthcare field, playing a significant role in developing models that assist in identifying and addressing critical medical problems [1]. These AI-driven models are applied across various domains, including heart diseases [2], kidney disorders [3], and psychological conditions [4]. Accurate diagnosis in these areas relies on diverse types of data, which are essential for evaluating

Received 13 January 2026; revised 28 April 2026; accepted 4 May 2026.
Available online 16 June 2026

* Corresponding author.
E-mail addresses: mohammed.nawzad@tiu.edu.iq (M. N. Mohammed-Ramzi), aso.aladdin@chu.edu.iq (A. M. Aladdin).

https://doi.org/10.70403/3008-1084.1026

the performance and precision of newly developed models [5]. AI integration improves diagnostic accuracy, efficiency, and decision-making in today's healthcare systems [6]. As a result, detecting skin cancer is largely based on the visual analysis of dermoscopic pictures, which is both time-consuming and extremely dependent on dermatologists' subjective knowledge. There are several barriers to identifying a skin lesion, including low contrast, visual noise, hair, fibers, and bubbles that mask the border of a lesion [7]. Plus, as skin lesions may look and feel alike in appearance, location, color, and size, early identification is important for effective treatment [8].

To date, most medical imaging tools cannot classify skin lesions accurately as they are complicated and have high variability. Even interpreting an image can be subjective and incongruent, thereby leading to the need for an accurate, automated classification system to help detect skin cancer at the earliest stage. While several studies have tried deep learning to detect skin cancer, there is still much to be done on compute power and dataset imbalance. For example, large, computationally expensive pre-trained networks such as ResNet or DenseNet have been developed that are subject to overfitting and require large amounts of memory. Lightweight, optimized CNNs are required to correctly class imbalances of datasets like HAM10000 without taking up unnecessary computational overhead of sophisticated ensemble or deep networks.

The purpose of this study is to develop a deep learning-based strategy for automatically classifying skin lesions to aid in the early diagnosis of probable skin malignancies. A modified CNN architecture is built and enhanced utilizing thorough data augmentation approaches to overcome dataset imbalance and increase model generalization.

This study is guided by the following research question:

RQ1. Can a Convolutional Neural Network (CNN) model, enhanced with data augmentation, accurately classify seven types of skin lesions from dermoscopic images?

RQ2. How effective is preprocessing, including hair removal and augmentation, in improving dermoscopic image quality and diversity for deep learning?

RQ3. How well does the proposed CNN model perform in terms of accuracy, precision, recall, F1-score, and Area Under the Curve – Receiver Operating Characteristics (AUC – ROC) when evaluated on a real-world dermoscopic image dataset?

Several studies have looked at CNN-based classification of skin lesions. For example, [9] noted the complexity in understanding skin lesion images and the need for large-scale training data sets; [10] suggested a deep learning model for classification and emphasizes the importance of precise categorization. Also, image distortions such as color, texture, and lighting may have a major impact on model accuracy [11].

This study has three main contributions. First introduces a custom CNN architecture for multiclass skin lesion categorization. Second, it illustrates a picture enhancement (augmentation) process of changing the geometric properties, brightness and color, noise addition, and elastic deformation of the image. Third, it compares its results with the results in the HAM10000 dataset.

The remainder of this study consists of Section 2 reviewing papers on skin lesion classification; Section 3 presenting the generated method; and Section 4 summarizing the results and performance. This study concludes with a summary of the results, implications, and opportunities for future research, which are presented in Section 5.

2. Related works

But the accuracy and speed of skin lesion identification have increased enormously with CNNs. CNNs are robust, consistent, and easier to grade than manual assessment, and

are well-suited for clinical and distant diagnosis. In particular study of [12], researchers proposed integrating ResNet and EfficientNet deep learning to learn features from skin lesions. Utilizing transfer learning on the HAM10000 data set, researchers were able to correctly identify seven types of skin lesions: melanocytic and non-melanocytic. Using EfficientNet's scaling ability and ResNet's depth, they achieved a top-of-modern accuracy of 99.14%, demonstrating the superiority of a hybrid CNN model over single architecture models like SqueezeNet for general dermatological diagnosis. Additionally, [8] also focused on CNN's uses in dermatology, which included dermoscopic and photographic illness classification. The authors enhanced these findings by combining and optimizing three existing deep learning architectures, correcting for outliers and class imbalance, and achieving encouraging results.

A study focuses on data augmentation (e.g., rotation and flipping) and ResNet and MobileNet to improve generalization. So, the model generalization performance in this study was 90.5% using ResNet with transfer learning [13]. Analytically, the YOLOv8x-cls model was 86.2% accurate in making clinical inference from a single model in a current clinical setting, according to a study on yolov8 variations of the HAM10000 dataset [14].

In article [15], a new MobileNet-based classifier has demonstrated an ability to be used easily as a tool for the early detection of skin cancer by providing an 85.0% category accuracy on the HAM10000 data set using transfer learning in resource-constrained settings. Likewise, [16] provided a hybrid deep learning approach, combining DenseNet-121 and AlexNet architecture already trained. They evaluated the efficiency of feature fusion to classify skin lesions by combining the strengths of both networks to produce an accuracy of 90.65% using the ISIC 2016–17 dataset.

[17] Enhanced the level of clarity and consistency in diagnostics by integrating Explainable AI (XAI) algorithms such as Grad-CAM and Grad-CAM++ into an Android-based healthcare system. Furthermore, [18] benchmarked 38 architectures in ten families (ResNet, Xception, and DenseNet) and came up with 87.6% of the HAM10000 accuracy. [19] Developed a framework that yielded 88.5% of the HAM10000 accuracy when using DenseNet for feature extraction and mask R-CNN for lesion segmentation. According to [20], 90.3% accuracy obtained by image augmentation. For more explaining, Table 1 lists the algorithms and scores of accuracies in these studies.

Previously, deep learning for skin lesion classification was studied in the context of deciding the best models to be used to diagnose certain diseases. The CNN used ranged from simple network such as SqueezeNet or MobileNet to more complex networks such as ResNet, DenseNet, and Xception. In addition, research has used methods such as transfer learning, image augmentation, and hybrid modeling, all of which have shown a significant improvement in classification accuracy. Ensemble learning and advanced data augmentation have shown to yield models that are accurate and generalizable. The numerous

Table 1. Related works summarization.

Reference	Methods	Dataset	Accuracy (%)
[8]	Hybridized CNN-DenseNet model	HAM10000 and PAD-UFES-20	95.5
[12]	EfficientNet and ResNet (Transfer Learning)	HAM10000	99.14
[13]	ResNet, data augmentation	HAM10000	90.5
[14]	YOLOv8x-cls (Single-model architecture	HAM10000	86.2
[15]	MobileNet (Transfer Learning)	HAM10000	85
[16]	Hybrid (AlexNet + DenseNet-121)	ISIC 2016–17	90.6
[17]	AI educational system	HAM10000	82
[18]	ResNet, Xception, DenseNet	HAM10000	87.6
[19]	DenseNet, Mask R-CNN	HAM10000	88.5
[20]	CNN with image augmentation	HAM10000	90.3

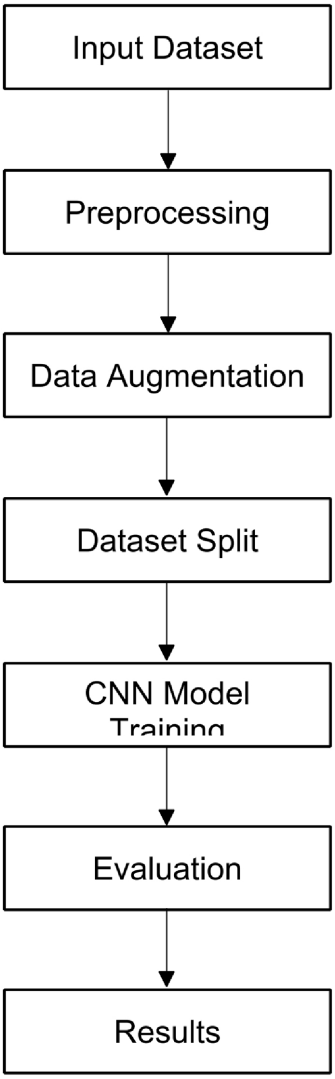

Fig. 1. The methodology.

methods presented here are indicative of ongoing efforts to improve and adapting deep learning models to the unique problem of medical image processing. These examples point to additional research that needs to be done to provide better diagnostic tools for dermatology.

As a result, combined with data augmentation and transfer learning, CNN models have been demonstrated to be highly effective in skin lesion classification. In this study, these results are built upon by employing a novel CNN design and using multi-channel augmentation to further improve the classification performance of skin lesions.

3. Methodology

According to these gaps, this study focuses on a proposed method to address these problems. Fig. 1 depicts the overall flow of the proposed model. The model relies on loading

thermoscopic image data and metadata, either in raw form or as a CSV file. The dataset is then processed using a range of image processing operations to enhance image quality, such as scaling and hair removal, to ensure the image has the same size and resolution as the input image. Afterwards, the data are augmented with various data manipulations such as flipping and turning to increase the diversity of the dataset and correct for class imbalance.

The processed dataset is divided into three subsets: testing, validation, and training. A model is trained using the training subset and performance is monitored and underfitting avoided by means of the validation subset.

When the model is ready, the test data has been analyzed. Then, using the analyzed data, the CNN model has been built and trained. Several performance parameters such as accuracy, precision, recall, F1-score, and AUC-ROC are used to assess the model after training. Lastly, the proposed model suggests outputs in the form of mismatch analysis, graphical visualizations, and quantitative evaluation metrics that will provide a complete view of the model's effectiveness and performance in categorization.

3.1. Dataset description

The HAM10000 dataset was used in this study to classify skin lesions. Despite being created for a different purpose, it was a useful tool for creating and evaluating a skin lesion classification model in the field of medical image analysis. The dataset comprises approximately 10,000 dermoscopic images categorized into seven classes: melanocytic nevi (nv), melanoma (mel), basal cell carcinoma (bcc), vascular lesions (vasc), actinic keratoses (akiec), dermatofibroma (df), and benign keratosis-like lesions (bkl). A numeric ID, ranging from 0 to 6, was assigned to each class for implementation purposes [21].

3.2. Data preprocessing

The loading and cleaning of images marked the beginning of the preprocessing process. All images were checked for integrity, and any corrupted or mislabeled files were removed. Each image was resized to a fixed resolution suitable for the CNN input dimensions. A custom hair removal technique was implemented using black hat filtering, Gaussian blurring, and inpainting to reduce occlusions caused by hair artifacts. As a result, the experiments revealed that CNN models coupled with data augmentation and transfer learning can help distinguish skin lesions. In line with these findings, the current study has combined a new CNN design with multi-channel augmentation to further classify skin lesions.

3.3. Data augmentation

The addition of data to the existing data set raised the available numbers and made it more resilient to changes in images. To do this, we used color jittering to mimic changes in illumination, brightness, and contrast to increase variability, and random rotations to adjust for different orientations of lesions. Also, horizontal and vertical flipping was used to correct for spatial bias and normalization was used to normalize the intensity distribution of pixels. All these changes in the model prevented overfitting and improved generalization.

3.4. Dataset split

It is prospective that there is a significant difference between the two sets which have been used. The dataset was divided into three subsets: testing (10%), validation (10%), and training (80%). While K-fold cross-validation is often used in smaller scale machine learning, it is incompatible with the computational overhead of deep neural networks on the significantly larger HAM10000 dataset. Therefore, the model was tested on a class balanced set of trials along with the early stopping and weight renewal features to verify the reliability of our model.

Additionally, the robustness to data augmentation (geometric, HSV, and elastic transformations) was an ablation threshold; without these augmentation parameters, the proposed models were strongly biased toward the majority class (Nevus) during initial training and were less generalizable than our model.

The batch size of the optimization step of Adam was set to 34 to allow for a balance between the limit of GPU memory and a consistent gradient estimate. Training was limited to 150 epochs, and the training process was automatically stopped by the loss of validation. Lastly, a benchmark experiment was run using a pre-trained ResNet50 model to measure the structural efficiency and diagnostics of the suggested CNN against the prevailing standards. For a baseline for classification accuracy and computing cost, the ResNet50 model was trained on the same balance of data and hyperparameters as the real-world examples.

3.5. Convolutional neural network model training

The model for image classification uses a modified CNN to extract hierarchical features from dermoscopic images. The model is composed of 32 convolutional layers using 33 kernels, followed by the Max pooling and ReLU activation layers. These initial layers are intended to extract fundamental visual features. To improve generalization and reduce overfitting, batch normalization and dropout at 0.2 have been used. To realize multi-scale pattern recognition, more convolutional blocks with 3×3 and 5×5 kernels and 64 filters are added as the model gets deeper. The feature abstraction is performed at these layers. Prior to arriving at a softmax output layer of seven units (corresponding to the seven skin lesion types), the final convolutional output is flattened and passed through dense layers of 300 and 100 neurons, respectively.

The Adam optimizer is used to train the model. An adaptive learning rate optimizer that works well for image classification problems is the Adam optimizer. Because the problem is multi-class, categorical cross-entropy is used as the loss function. A dynamic early halting strategy was used to avoid overfitting while giving the model enough time to learn intricate minority-class features. A patience of 15 epochs was used to monitor the validation loss, and the maximum training limit was set at 150 epochs. The ideal model weights from the best-performing era were automatically restored upon early termination. A dropout rate of 0.2 was selected to provide sufficient regularization without starving the network of critical hierarchical features.

It can be referred to the pseudocode below, which depicts training processing in each range for a better understanding of this model.

3.6. Evaluation metrics

For this proposed model, using several classification measures to assess the classification performance of the suggested CNN model, namely the F1-score, recall (sensitivity),

Pseudocode 1: Training Process

```
for epoch ← 1 to 150 do
    for each batch in training_data do
        predictions ← CNN.forward(batch.images)
        loss ← CategoricalCrossEntropy(predictions, batch.labels)
        Backpropagate loss gradients
        Update network parameters using optimizer
    end for
    if validation loss does not improve for 15 consecutive epochs then
        terminate training
    end if
end for
```

accuracy, and precision. Each of these measures is important in assessing the prediction power of the model, especially in multi-class categorization.

3.6.1. Accuracy

Accuracy measures the overall accuracy of predictions, indicating the model correctness for all classes. But if the set of data is uneven, and some classes dominate, then only accuracy may not be sufficient to describe the correctness of the models. Hence, consider other metrics as well to have a complete picture of the performance of models. Accuracy is calculated using Eq. (1) [22].

$$Accuracy = \frac{Tp}{TS} \tag{1}$$

Where TP is the True Positive, which indicates the number of instances in which the model correctly identifies the positive class, and TS is the Total Samples, which represents the total number of instances that have been evaluated by the model.

3.6.2. Precision

Precision measures the True Positive TP predictions from all the positive predictions. This precision is a key indication of the model's ability to avoid False Positives FP. FP occurs when a model incorrectly classifies a negative sample as positive, such as when a benign skin lesion is interpreted as malignant during medical imaging. Precision is computed using Eq. (2) [23].

$$Precision = \frac{Tp}{Tp + Fp} \tag{2}$$

3.6.3. Recall

Recall, or sensitivity, is the ability of the model to correctly identify all actual positive cases. This parameter indicates the extent to which the model successfully captures the relevant instances without missing them. This is particularly important for this specific area especially the case of medical diagnosis, where the failure to identify a true positive, such as a malignant lesion, can have grave consequences. Recall is computed by Eq. (3) [23].

$$Recall = \frac{Tp}{Tp + FN} \tag{3}$$

Table 2. Performance metrics.

Metric	Value
Accuracy	0.9075
Precision	0.92
Recall	0.92
F1-Score	0.92
AUC-ROC	0.9911

Where *FN* is the False Negative.

Likewise, high recall suggests that the model only missed very few actual positive cases. This is important when early detection of disease is needed, such as skin lesions classification.

3.6.4. F1-score

It is a measure of the performance that considers both recall and precision. It is a fair measure which does not penalize those classes in which there is a high disparity in recall and precision. The F1-score is especially useful in classes with unbalanced distribution, where it would be deceptive to rely solely on recall or precision. F1-Score is computed using Eq. (4) [23]:

$$F1 - Score = \frac{2 \times (Precision \times Recall)}{Precision + Recall} \tag{4}$$

3.6.5. Area under the curve – receiver operating characteristics

AUC-ROC compares the model's ability to correctly differentiate between two classes at any threshold of classification. Since the models can evaluate different bases of classification, the AUC-ROC is a general measure of the discriminative power of the model. A higher AUC indicates that the model is more effective in distinguishing classes, which is significant when evaluating CNN model in this situation [24].

4. Experimental results

The model worked well and will be considered in the next section. This section will sketchily discuss its effectiveness, with illustrations to tolerate these outcomes. It is discussed here with insights into the strengths and weaknesses of the model as well as possible gaps in the model's ability to properly classify skin lesions.

The key performance metrics of the model on an independent set of data are summarized in Table 2. The model's performance in prediction in real-world situations can be easily checked in this summary.

4.1. Benchmark comparison against ResNet50

The achieved 0.9075 accuracy indicates a high level of confidence in the model to correctly diagnose skin lesions in the entire dataset. To confirm the model and ensure that the model was not overdone by the class imbalance in the HAM10000 images (mostly Nevus images), per class metrics have been assessed. Overall, the model had an AUC-ROC of 0.9911 and is highly discriminative in all seven categories.

Importantly, the model performed exceptionally well on severe minority classes, which are hard to classify. bcc was classified with a high recall of 0.94 while mel, a highly

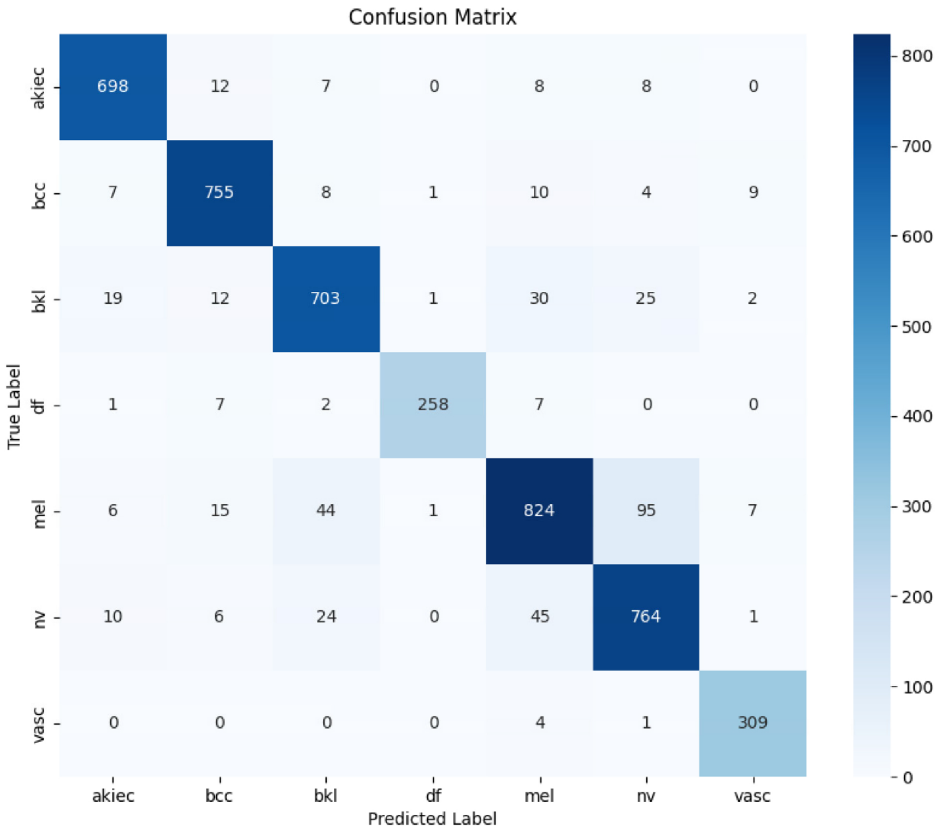

Fig. 2. Confusion matrix of the proposed model.

critical class for early detection, was classified with a recall of 0.83 and an F1-score of 0.86. A normalized confusion matrix was constructed to visually represent the TP rate and class misclassifications, indicating the effectiveness of the offered increase in the data in restoring the imbalance in the data and the bias of the majority class.

To evaluate the strength and efficiency of the proposed architecture, a benchmark test was performed against a powerful, current pretrained ResNet50 model under identical balanced data set. ResNet50 model was computationally costly, taking over 22 minutes for each training epoch but produced an overall accuracy of 87.64% and an F1-score of 0.89.

The custom lightweight CNN performed a higher 90.75% accuracy and 0.92 F1-score in a fraction of the time (around 2 minutes per epoch) at a fraction of the computational cost. This result shows that for specialized and augmented dermatological data, a lightweight built specifically for dermatology can perform better than heavy pretrained networks in diagnosing and computing tasks.

For illustration, Fig. 2 proves that the proposed model demonstrates strong *TP* rates across all classes, particularly excelling in distinguishing (*mel*) and (*nv*) despite the dataset's inherent complexities.

Visibly, Fig. 3 shows that this pretrained network produces more confusion between classes since it misclassifies a higher number of (*nv*) and (*mel*) occurrences than the custom architecture.

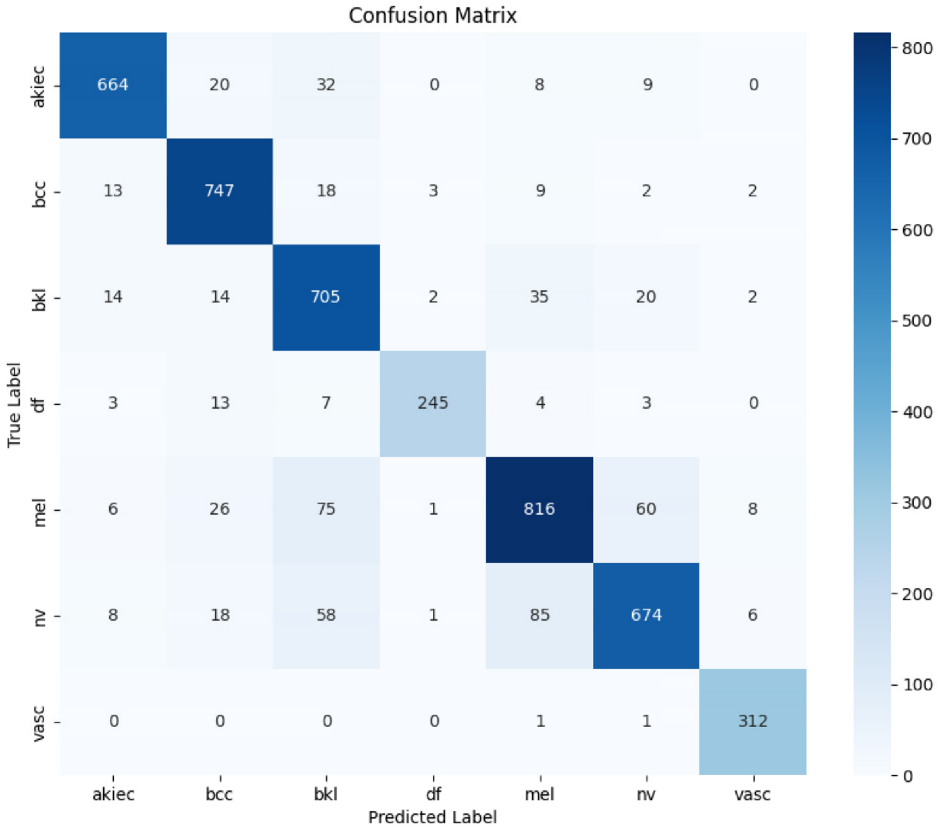

Fig. 3. Confusion matrix of the ResNet50.

5. Discussion

For further testing on the learning process, the growth in accuracy and loss over time of the training and validation data have also been evaluated in Figs. 4 and 5. This gives insight into the ability of the model to learn and generalize. As seen in Fig. 4, the growing accuracy and decreasing loss of the two sets indicate that the model is correctly representing the patterning of the dermatological data.

In the model, showing the validation and training curves are well aligned. The validity accuracy closely fits the training accuracy across epochs as pointed out in Fig. 4, while the validation loss is very similar to the training loss without any significant variation as pointed out in Fig. 5. Note that the gap between the curves is not as large as it usually is because of overfitting. It is an indication of the high effectiveness of the dropout rate (0.2) and the early stopping mechanism (15-epoch patience). The model also shows that it works perfectly with unseen data without any memorization. Fig. 4. Custom model training and validation accuracy. The close-to-center curves confirm robust learning and good generalization without overfitting.

Correspondingly, Fig. 5 plots the custom model training and validation loss. The slowed-down decline in training and validation losses indicates that the model is stable and that the regularization strategies used are effective.

Additionally, this study builds upon the results of prior studies indicating positive results from using CNNs in skin classification. The dataset is specifically designed to identify nv,

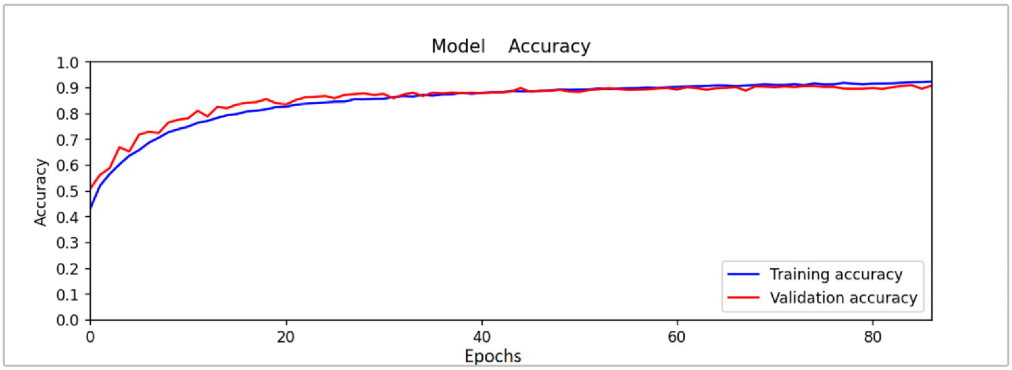

Fig. 4. Model accuracy.

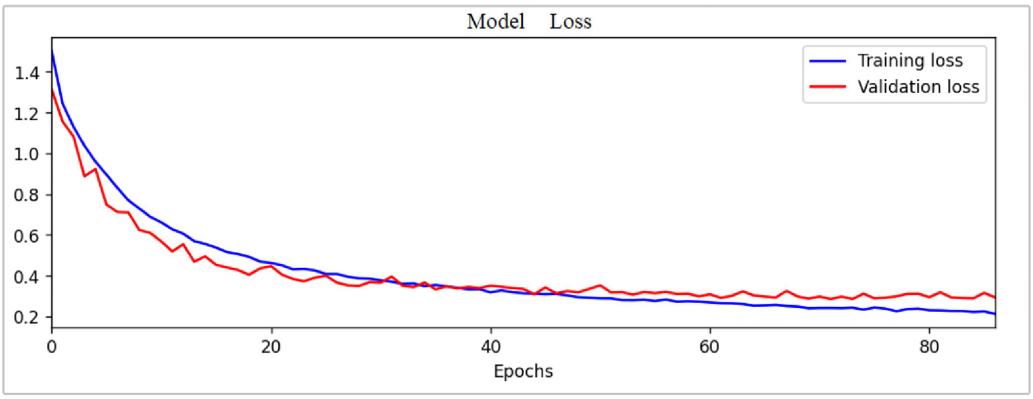

Fig. 5. Model loss.

mel, bkl, bcc, akiec, vasc, and df for the specific skin condition referred to in [25] regarding vitiligo. Unlike the previous work that focused on binary classification [25, 26]. Thus, the current study addresses the more complex task of multi-class classification for 7 types of skin lesions.

As mentioned in previous studies [13, 20], data augmentation has been employed to fill out the full size of the dataset and increase the resilience of the model. In this case, the augmentation included a large number of image transformations including flipping, rotating, and shifting, scaling, rotating, altering brightness and contrast by random brightness and random contrast, blurring with motion blur, median blur, Gaussian blur, adding noise with Gaussian noise, equalizing the histogram using clahe, changing colors with hue saturation value, and elastic deformations through elastic transform.

The extensive number of improvements is designed to show how the model is affected by real image appearance differences and can more effectively generalize unfamiliar images. These improvements, in combination with the proposed CNN design, explain the superior performance compared to some existing models. In addition, Table 3 shows the comparison of model accuracy to that of other studies. The current analysis is accompanied by an increasing number of highly complex architectures, such as Vision Transformers (ViTs) [27] and heavy ensemble deep learning models [28]. While they have excellent diagnostic capabilities, they require massive computational resources, high memory usage, and lengthy training times. For this reason, the proposed model has very competitive classifier

Table 3. Comparison with the state-of-the-art method.

Studies	Accuracy (%)
[13]	90.5
[14]	86.2
[15]	85
[19]	88.5
[20]	90.3
The Proposed Model	**90.75**

metrics that surpass those of several other models in the literature, and it avoids all the computational bottlenecks associated with these heavier state-of-the-art architectures. Because of this combination of high accuracy and low computational costs, the proposed model is very suitable for real-world clinical and mobile use.

Although it is difficult to directly compare accuracy since the data and the type of skin lesions varied, the results indicated that the proposed model could be as good, if not better, than the previous model in identifying skin lesion. Despite these difficulties, the model performed well. It has potential to be applied to more dermatological diagnostics and can be used in a wide range of therapeutic environments. The flexibility of custom CNNs extends beyond imaging applications to medicine. Similar deep learning systems have demonstrated that a convolutional feature extracted from the façade data can be sufficiently adaptable [29].

6. Conclusion

The results from this exploratory modelling study propose promising outcomes for the suggested mode in the detection of skin lesions. The proposed lightweight architecture has an overall accuracy of 90.75% with a precision of 0.92, a recall of 0.92, an F1-score of 0.92 and an AUC-ROC of 0.9911 on a strictly balanced set of data. The proposed model is more efficient and more accurate than ResNet50 and can discriminate between benign and malignant lesions in the difficult alternative class, which can lead to early detection of skin cancers.

The results of this study suggest that the model may help doctors in dermatology and primary care identify lesions that warrant further investigation. It could be used in mobile health apps that enable users to take a first-time assessment and then check themselves for injuries or illnesses and seek medical attention immediately. It could also be used to run remote skin cancer screenings at remote locations where patients cannot access a dermatologist and help them identify skin cancer earlier.

Future work could use a larger and more varied set of training data to allow for better generalizability across types of lesions, surface presentation of lesions, and skin tone. Integrating this model with other tools of diagnosis, such as dermoscopy or histology, could increase the accuracy of evaluation. Human trials are important for validating real-world performance and safety. Including interpretability tools such as SHapley Additive exPlanations (SHAP) or Grad-CAM will increase clinical trust. Parameter tuning or metaheuristics optimizations such as Lagrange Elementary Optimization (LEO) and Fitness Dependent Optimizer (FDO) could also improve the model performance. Overall, the deep learning model in this study offers significant promise to improve skin lesion screening and early detection, which would enable early diagnosis and prompt intervention as well as better patient outcomes.

Acknowledgment

None.

Authors' contributions

Writing—original draft, writing—review and editing, formal analysis, software, methodology, visualization, and data curation: Mohammed Nawzad Mohammed-Ramzi; writing—original draft, writing—review and editing, methodology, data curation, supervision, and funding acquisition: Aso M. Aladdin.

Conflict of interest

The authors declare no conflict of interest to any party.

Data availability

The HAM10000 dataset utilized in this study is publicly available through the Harvard Dataverse and Kaggle repositories.

References

1. M. Wang, X. Gao, and L. Zhang, "Recent global patterns in skin canser incidence, mortality, and prevalence," *Chinese Medical Journal*, vol. 138, no. 2, pp. 185–192, Jan. 2025, doi: 10.1097/CM9.0000000000003416.
2. L. Zhou, Y. Zhong, L. Han, Y. Xie, and M. Wan, "Global, regional, and national trends in the burden of melanoma and non-melanoma skin cancer: Insights from the global burden of disease study 1990–2021," *Scientific Reports*, vol. 15, Feb. 2025, Art. no. 5996, doi: 10.1038/s41598-025-90485-3.
3. F. Sanmarchi, C. Fanconi, D. Golinelli, D. Gori, T. Hernandez-Boussard, and A. Capodici, "Predict, diagnose, and treat chronic kidney disease with machine learning: A systematic literature review," *Journal of Nephrology*, vol. 36, no. 4, pp. 1101–1117, Feb. 2023, doi: 10.1007/s40620-023-01573-4.
4. C. Hickman, L. Pridgen, D. Hughes, L. S. Pair, A. C. Holland, "The role of artificial intelligence in increasing efficiency, reducing errors, and improving patient outcomes in clinical practice," *Clinical Journal for Nurse Practitioners in Women's Health*, vol. 2, no. 2, pp. 101–108, Apr. 2025, doi: 10.1891/CJNPWH-2515.
5. S. Asif *et al.*, "Advancements and prospects of machine learning in medical diagnostics: Unveiling the future of diagnostic precision," *Archives of Computational Methods in Engineering*, vol. 32, no. 2, pp. 853–883, Jun. 2024, doi: 10.1007/s11831-024-10148-w.
6. S. Arige, L. R. Atmakuri, R. Shaik, M. Gude, V. K. Ghanta, and R. Alluri, "Digital dermoscopy: Advancements in skin cancer diagnosis and monitoring," *Biomedical and Pharmacology Journal*, vol. 18, pp. 33–43, Mar. 2025, doi: 10.13005/bpj/3071.
7. M. S. and S. Gupta, "Deep learning applications in cardiovascular disease detection: A comprehensive systematic analysis," *International Journal for Research in Applied Science and Engineering Technology*, vol. 14, no. 1, pp. 411–418, Jan. 2026, doi: 10.22214/ijraset.2026.76859.
8. J. A. Rios-Duarte, A. C. Diaz-Valencia, G. Combariza, M. Feles, and R. A. Peña-Silva, "Comprehensive analysis of clinical images contributions for melanoma classification using convolutional neural networks," *Skin Research Technology*, vol. 30, no. 5, May 2024, Art. no. e13607, doi: 10.1111/srt.13607.
9. R. Thinakaran, J. Somasekar, V. Neerugatti, and K. Ganga, "Advancements in skin cancer detection: A comprehensive review of convolutional neural network approaches," in *Proc. 2024 14th Int. Conf. on Software Technology and Engineering (ICSTE)*, Macau, China, pp. 232–236, doi: 10.1109/ICSTE63875.2024.00047.
10. I. Ahmad, J. Amin, M. I. Lali, F. Abbas, and M. I. Sharif, "A novel deeplabv3+ and vision-based transformer model for segmentation and classification of skin lesions," *Biomedical Signal Processing and Control*, vol. 92, Jun. 2024, Art. no. 106084, doi: 10.1016/j.bspc.2024.106084.

11. H. M. L. Delower *et al.*, "Skin lesion detection and classification using machine learning: A comprehensive approach for accurate diagnosis and treatment," in *Proc. 2023 26th Int. Conf. on Computer and Information Technology (ICCIT)*, Cox's Bazar, Bangladesh, pp. 1–6, doi: 10.1109/ICCIT60459.2023.10441420.
12. M. Alruwaili and M. Mohamed, "An integrated deep learning model with EfficientNet and ResNet for accurate multi-class skin disease classification," *Diagnostic*, vol. 15, no. 5, Feb. 2025, Art. no. 551, doi: 10.3390/diagnostics15050551.
13. P. Mekala, S. B, G. Vishnu, G. J. Sai, and P. Subash, "Classification of skin lesions using deep learning models with transfer learning techniques," in *Proc. 2023 IEEE Int. Conf. on Recent Advances in Systems Science and Engineering (RASSE)*, Kerala, India, pp. 1–10, doi: 10.1109/RASSE60029.2023.10363548.
14. U. Saha, I. U. Ahamed, M. A. Imran, I. U. Ahamed, A.-A. Hossain, and U. D. Gupta, "YOLOv8-based deep learning approach for real-time skin lesion classification using the HAM10000 dataset," in *Proc. 2024 IEEE Int. Conf. on E-health Networking, Application & Services (HealthCom)*, Nara, Japan, pp. 1–4, doi: 10.1109/HealthCom60970.2024.10880715.
15. S. Khandizod, T. Patil, A. Dode, V. Banale, and C. D. Bawankar, "Deep learning based skin cancer classifier using MobileNet," *Internation Journal for Research in Applied Science and Engineering Technology*, vol. 10, no. 5, pp. 629–633, May 2022, doi: 10.22214/ijraset.2022.42260.
16. P. Gupta and S. Mesram, "AlexNet and DenseNet-121-based Hybrid CNN Architecture for Skin Cancer Prediction from Dermoscopic Images," *International Journal for Research in Applied Science and Engineering Technology*, vol. 10, no. 7, pp. 540–548, Jul. 2022, doi: 10.22214/ijraset.2022.45325.
17. K. Mridha, M. M. Uddin, J. Shin, S. Khadka, and M. F. Mridha, "An interpretable skin cancer classification using optimized convolutional neural network for a smart healthcare system," *IEEE Access*, vol. 11, pp. 41003–41018, 2023, doi: 10.1109/ACCESS.2023.3269694.
18. J. Vieira, F. Mendonça, and F. Morgado-Dias, "Deep learning approaches for skin lesion detection," *Electronics*, vol. 14, no. 14, Jul. 2025, Art. no. 2785, doi: 10.3390/electronics14142785.
19. M. A. Khan, T. Akram, Y.-D. Zhang, and M. Sharif, "Attributes based skin lesion detection and recognition: A mask RCNN and transfer learning-based deep learning framework," *Pattern Recognition Letters*, vol. 143, pp. 58–66, Mar. 2021, doi: 10.1016/j.patrec.2020.12.015.
20. G. Sharma, V. Anand, and V. Kumar, "A suitable approach for classifying skin disease using deep convolutional neural network," in *Proc. 2023 World Conf. on Communication & Computing (WCONF)*, Raipur, India, pp. 1–6, doi: 10.1109/WCONF58270.2023.10235079.
21. P. Tschandle, C. Rosendahl, and H. Kittler, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Scientific Data*, vol. 5, Aug. 2018, Art. no. 180161, doi: 10.1038/sdata.2018.161.
22. M. Badawi *et al.*, "Skin cancer classification and segmentation using deep learning," *International Journal of Telecommunications*, vol. 4, no. 1, pp. 1–23, Feb. 2024, doi: 10.21608/ijt.2024.280957.1045.
23. N. Gajare and P. Mane, "A robust CNN approach for multi-class skin disease detection and classification," in *Proc. 2025 Int. Conf. on Intelligent and Cloud Computing (ICoICC)*, Bhubaneswar, India, pp. 1–5, doi: 10.1109/ICoICC64033.2025.11052018.
24. O. Iparraguirre-Villanueva and M. Cabanillas-Carbonell, "Application of convolutional neural networks in skin disease prediction: accuracy and efficiency in dermatological image analysis," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 21, no. 2, pp. 18–37, Feb. 2025, doi: 10.3991/ijoe.v21i02.52871.
25. M. Arshad *et al.*, "A computer-aided diagnosis system using deep learning for multiclass skin lesion classification," *Computational Intelligence and Neuroscience*, vol. 2021, no. 1, Dec. 2021, Art. no. 9619079, doi: 10.1155/2021/9619079.
26. E. Göçeri, "Convolutional neural network based desktop applications to classify dermatological diseases," in *Proc. 2020 IEEE 4th Int. Conf. on Image Processing, Applications and Systems (IPAS)*, Genova, Italy, pp. 138–143, doi: 10.1109/IPAS50080.2020.9334956.
27. Furqan, N. Katuk, and D. Hartama, "Multiclass skin lesion classification algorithm using attention-based vision transformer with metadata fusion," *Journal of Applied Data Sciences*, vol. 7, no. 1, pp. 203–217, Jan. 2026, doi: 10.47738/jads.v7i1.1017.
28. M. A. Khan, S. Alam, and W. Ahmed, "Enhanced skin cancer diagnosis via deep convolutional neural networks with ensemble learning," *SN Computer Science*, vol. 6, no. 2, Jan. 2025, Art. no. 124, doi: 10.1007/s42979-024-03581-y.
29. S. Wang, S. Park, S. Park, and J. Kim, "Building façade datasets for analyzing building characteristics using deep learning," *Data in Brief*, vol. 57, Dec. 2024, Art. no. 110885, doi: 10.1016/j.dib.2024.110885.