

# Machine learning methods for automating skin lesion classification

Mohand Lokman Al Dabag<sup>1</sup>, Shaima Miqdad Mohamed Najeeb<sup>2</sup>, Razan Abdulhammed<sup>3</sup>, Haider TH. Salim ALRikabi<sup>4</sup>

<sup>1</sup> Department of Artificial Intelligence Technology Engineering, Northern Technical University, Mosul, Iraq

<sup>2,3</sup> Department of Cyber Security and Cloud Computing Techniques Engineering, Northern Technical University, Mosul, Iraq

<sup>4</sup> Electrical Engineering Department, College of Engineering, Wasit University, Wasit, Iraq

Corresponding author E-mail: [mohandaldabag@ntu.edu.iq](mailto:mohandaldabag@ntu.edu.iq)

Received Dec. 22, 2025

Revised Feb. 25, 2026

Accepted Apr. 13, 2026

Online May 6, 2026

## Abstract

The skin serves as the body's primary protective and most visible layer, making it particularly susceptible to lesions that may cause physical discomfort, psychological distress, and potential health risks. Addressing skin diseases is therefore both a medical necessity and a matter of social sustainability, as untreated conditions can significantly reduce quality of life and limit social participation. This paper presents an artificial intelligence-based approach for distinguishing between Acne and Actinic Keratosis using lightweight image-processing techniques that minimize computational requirements and support resource-efficient healthcare technologies. To address the lack of a unified dataset, lesion images were collected from multiple repositories, standardized in size, and enhanced through grayscale illumination. Local Binary Patterns (LBP) were employed for feature extraction and dimensionality reduction, followed by classification using machine learning models. The results show that the Multilayer Perceptron (MLP) achieved higher accuracy (99.5%) than the Support Vector Machine (SVM) (97.5%). Overall, the proposed approach combines diagnostic accuracy with computational efficiency, contributing to accessible and sustainable healthcare solutions that improve patient outcomes while reducing technological resource demands.

© The Author 2026.

Published by ARDA.

*Keywords:* Sustainability, Healthcare technology, Machine learning, Skin disease, Acne and Actinic Keratosis, Local binary pattern, SVM, MLP

## 1. Introduction

The skin is the outermost layer covering the human body and accounts for 16% of total body mass. This layer is considered the body's first line of defense against external diseases such as bacteria, parasites, and viruses. Therefore, maintaining healthy skin is essential for a healthy quality of life [1]. Skin diseases significantly affect the patient's psychological, physical, and social well-being. Common symptoms of skin diseases, such as itching and skin discoloration, often lead to a decline in quality of life. Depression, introversion, and low self-esteem may be among the consequences of skin diseases that affect an individual's social and productive performance. In addition, the long-term safe treatment of skin diseases imposes economic burdens on the patient, exacerbating their suffering [2].

Patients often try not to seek medical attention until the disease has progressed and its symptoms become clearly visible. Many patients, and even some unprofessional doctors, fail to identify the type of disease, which complicates treatment in serious cases like skin cancer. Therefore, there arose a need for a quick and effective machine-learning-based disease detection technique [3]. Several studies have proposed methods for identifying skin diseases using image processing and analysis techniques that, in turn, feed into machine learning models. Machine learning models are divided into two groups: the first, such as SVM and MLP, is fed with distinctive features extracted from image processing; the second relies on deep learning, which extracts the main characteristics of the disease and classifies it [4] [5].

Samir et al. outline a hybrid methodology that integrates deep learning and machine learning techniques for detecting skin diseases. Notably, the combination of ResNet50 for feature extraction and a Support Vector Machine for classification has proven effective for distinguishing skin diseases. It achieved 99.11% accuracy, which exceeds other tested combinations. The authors conduct a comprehensive comparative analysis of different feature extraction and classification strategies, reinforcing the efficacy of their hybrid technique [6].

Thangakani and Sornam [7] proposed a method for detecting skin cancer using SVM and Bayesian classifiers. They used GLCM (Gray-Level Co-Occurrence Matrix) to extract features of the underlying shape, color, and skin texture. The two classifiers are used to predict the skin lesions using the extracted feature. Zehra K. S., Recep G., et al. evaluate the performance of various ML algorithms, including SVM, ANN, and decision trees, for classifying skin lesions.

The proposed method preprocessed the skin images by resizing, filtering, normalizing, and enhancing them before feeding them to the ML algorithms. The article found that both Random Forest (RF) and K-Nearest Neighbors (K-NN) algorithms achieved the best performance for classifying the skin lesions [8]. Akasha et al. investigated the performance of three machine learning models: SVM, Random Forest (RF), and Decision Trees (DTs). These ML models are combined with state-of-the-art (SOTA) pre-trained deep learning models, EfficientNet, MobileNetV2, and DenseNet121. For feature reduction, Principal Component Analysis (PCA) was used. The RF with DenseNet121 achieved promising results compared with the other models [9], supporting sustainable healthcare innovation.

Sagheer et al. introduce an innovative deep learning framework that leverages the robust capabilities of transfer learning using a pre-trained VGG16 architecture, meticulously crafted to accurately identify and predict a variety of skin diseases, including chickenpox, measles, and monkeypox. The dataset comprises images depicting these conditions alongside images of healthy skin, systematically arranged into training and testing subsets. Notably, the model achieves testing accuracy of 93.29% [10]. Huang et al. suggest an optimal skin cancer detection system based on deep learning. They provide an improved metaheuristic algorithm for a Deep network to diagnose skin cancer from images. The performance of the proposed method is demonstrated through comparison with related work [11].

Sohaib employed various DL models to classify skin lesions. DenseNet201 and ResNet52V2) and Convolutional Neural Networks (CNNs) are used to find out which model provides the best performance. The combination of two DL models (DenseNet201 and ResNet152V2) has been shown to achieve the best performance [12]. Alaa et al. proposed a DL model for early prediction of skin cancer. Their model employed multiple schemes for preprocessing skin images to enhance image contrast and construct the ROI image. Then, a deep belief network was used to classify skin lesions [13]. Peng et al. trained various DL models to classify an imbalanced skin lesion dataset. They used an augmentation strategy to enlarge the small dataset. The proposed combination system shows acceptable performance for classifying skin lesions [14].

This article is organized into the following sections: the next section outlines the theoretical background for the methodological approaches relevant to this work, establishing a preliminary step for the proposed method, which is arranged in subsections. Section 3 illustrates the proposed method that serves for sustainability in healthcare technology. The next section presents intermediate experimental results for each component of the

proposed method, followed by its classification rate. Finally, Section 5 presents a discussion and conclusions derived from a comparison with related work.

## 2. Methods and materials

The crucial step in the image preprocessing is the contrast stretching. It is usually termed as image normalization. The objective of this technique is to improve image contrast by utilizing the available image intensity range, so that it occupies the full range (e.g., 0 to 255 for 8-bit images) [15]. The fundamental method involves applying a linear scaling function for each pixel. This is done based on the minimum and maximum intensity values  $a$  and  $b$  of an image  $I_{in}$ . The stretched image  $I_{out}$  is evaluated according to Equation 1.

$$I_{out}(x, y) = \left( \frac{I_{in}(x, y) - a}{b - a} \right) \times (L - 1) \dots \dots \dots 1$$

Where  $L$  is the maximum intensity level of the digital system, the original image often has a narrow histogram due to improper exposure or sensor limitations. After applying this process, the pixel density across the image spans the full range allowed by the system, enabling it to reveal subtle details in shadows and highlights [16]. However, the effectiveness of this method degraded in the presence of image outliers, even with a single extreme pixel value. This leads to suboptimal contrast enhancement in the image. To overcome this situation, additional operations, such as histogram clipping, can be employed, in which a small number of pixels at both ends of the histogram are excluded before the stretching operation is applied [17].

Improving the visual acuity of an image's edges and enhancing image details is done by a technique called image sharpening. It emphasizes the high frequency of an image in the spatial domain by convolving it with specific kernels [15]. The most popular approach uses the Laplacian kernel, a second-order derivative filter. The mathematical representation of the sharpening process is illustrated by Equation 2.

$$I_{sh}(x, y) = I(x, y) + c * \nabla^2 I(x, y) \dots \dots \dots 2$$

Where  $I(x, y)$  is the original image,  $\nabla^2 I(x, y)$  is the filtered image,  $c$  is the scaling constant that controls the sharpening strength, and  $I_{sh}(x, y)$  is the sharpened image [18].

The texture of an image can be evaluated using this method. It is an effective and straightforward technique for assessing the weight of a given point relative to surrounding points in a given window, as in Figure 1.

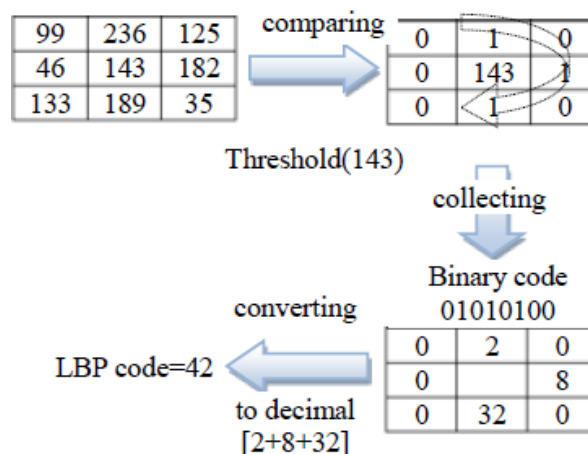


Figure 1. Evaluating the LBP [19]

This is done by dividing the image into 3x3 windows, and the center pixel value is used as a threshold for evaluating the surrounding pixels, starting from the top-left corner with a clockwise direction to cover all the window pixels. Each pixel below the threshold is replaced with 0; otherwise, it is replaced with 1. The center pixel weight is equal to the summation of 2 to the power  $N$  of only the nonzero pixels (ones pixels). Figure 1 illustrates the LBP technique [19] [20].

An artificial neural network has a foundational and widely used class called the Multilayer Perceptron (MLP), which is a feedforward network as in Figure 2. It consists of at least three layers. The first is called the input layer, which receives the incoming data. The next layer may comprise one or more hidden layers. These layers are responsible for learning complex patterns. The output layer is the last layer of the model. This layer produces the final prediction or classification. All these layers consist of nodes that are fully connected to the previous and next layer through adjustable weights. For a given neuron, the output is evaluated as the summation of its inputs multiplied by the corresponding weights plus a bias term. Then the computed value is mapped by an activation function. Equation 3 is the mathematical representation of a single neuron.

$$y = \sum_{i=1}^n w_i x_i + b, z = f(y) \dots \dots \dots (3)$$

where  $x_i$  is the input features,  $w_i$  are the associated weights,  $b$  is the bias, and  $f(\cdot)$  is an activation function such as ReLU or sigmoid. Figure 2 shows the MLP structure. During training, these weights are iteratively adjusted until reaching a minimum error, thereby improving model performance [21].

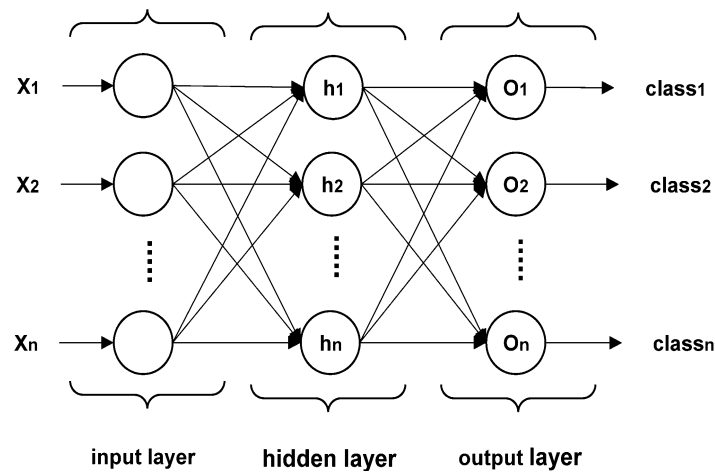


Figure 2. MLP architecture with one hidden layer

SVM is a supervised machine learning algorithm that uses target labels to learn the model. The main goal of SVM is to determine the best boundary (hyperplane) that splits data into different classes with the maximum possible margin. It searches for a small number of data points called support vectors. These points must be the closest to the decision boundary. The separation line or plane is determined via these points. SVMs maximize the margin between classes to improve the classification performance. This model can handle linear and nonlinear data by using kernel functions such as radial basis functions (RBFs) [21]. Figure 3 illustrates the SVM method for choosing the optimal hyperplane in binary classification. Due to the lack of a dataset that encompasses both diseases, images were collected from two different online websites, namely Kaggle and Roboflow.

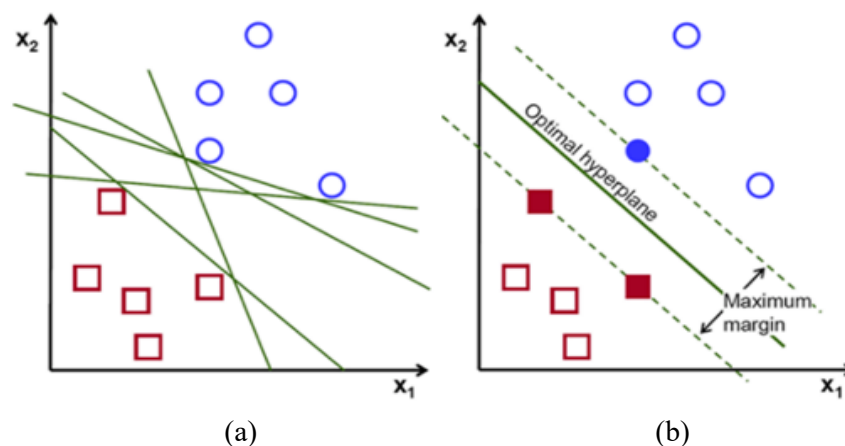


Figure 3. Classifying two classes using the SVM: (a) before the classification and (b) after the classification

*Actinic keratosis* images provided by the Kaggle dataset HAM10000. This dataset has 10015 dermatoscopic images. These images are collected from different populations, which include only skin lesions: basal cell carcinoma (bcc), Actinic keratoses, benign keratosis-like lesions (solar lentigines / seborrheic keratoses and lichen-planus-like keratoses, BKL), melanocytic nevi (NV), and vascular lesions (angiomas, angiokeratomas, pyogenic granulomas and hemorrhage, dermatofibroma (DF), melanoma (mel). These images are acquired using diverse modalities [22].

*Acne* images are downloaded from the Roboflow website. Unfortunately, the website provides limited metadata for this dataset [23]. It includes 745 images, but only 601 images are used in this work, depending on their quality. Finally, the composed dataset from both datasets consists of 601 *Acne* images and 460 *Actinic Keratosis* images. This dataset is split 70% for training, 15% for validation, and 15% for testing.

### 3. The proposed method

The proposed method is based on the characteristic appearance of the skin for each disease, based on Figure 4. Actinic keratosis typically appears as reddish patches that are noticeably larger than acne lesions, while acne presents as pimples in the form of blackheads, whiteheads, or larger pus-filled blisters. For these reasons, the edge extraction plays a key role in capturing distinctive features that help differentiate between the two diseases [24, 25]. Therefore, the edge extraction plays as a backbone of our proposed method, which consists of the following stages:

The preprocessing step contains two sub-operations. The first step converts the images to grayscale and resizes them to ensure all images in the dataset have the same size (283x640, the smallest image size in the dataset). The next sub-operation stretches the image contrast to have the full gray level. The sharpening stage used the Laplacian filter to extract only image object edges, then added them to the original image. This enlarges the image details.

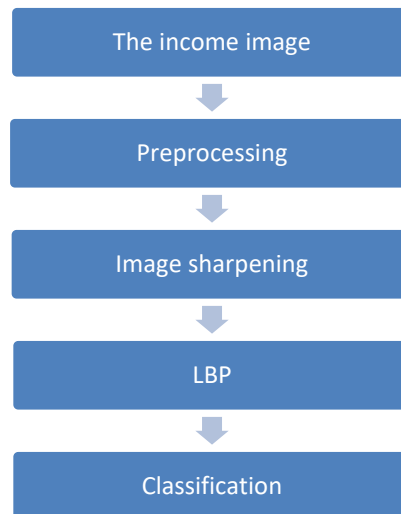


Figure 4. The proposed method block diagram

In addition to the disease's visual appearance, it is important to know how information is distributed at a particular point, which is captured using the Local Binary Pattern (LBP) technique as part of the proposed method for differentiating between the two diseases. The last stage classifies the skin disease. Figure illustrates the flowchart of the proposed method [26, 27].

### 4. Results and discussion

This section explores the outcome of all stages in the proposed method, as well as its classification performance. The first-stage operation is used to reduce computation time by converting the images to grayscale and resizing

them to a common size. It is also used to expand the original image spectrum into a full gray-level spectrum by stretching its contrast to the full gray-level range. Figure shows the effect of the preprocessing stage.

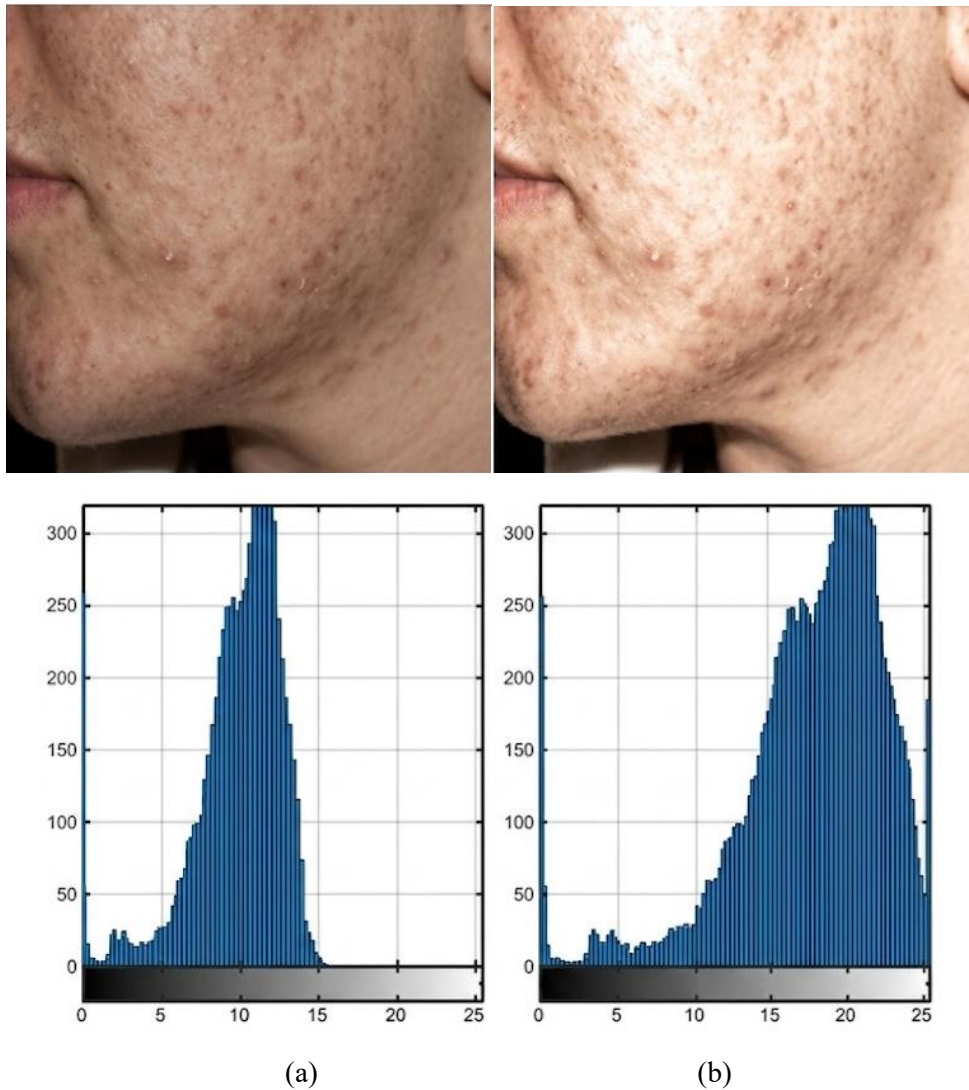


Figure 5. The effect of the preprocessing stage, (a) the original image, (b) the outcome of this stage

In the sharpening stage, the Laplacian filter is applied to the incoming image for extracting edge information, which is then added back to the output of the previous stage to enhance image details. Figure shows the outcome of this stage.

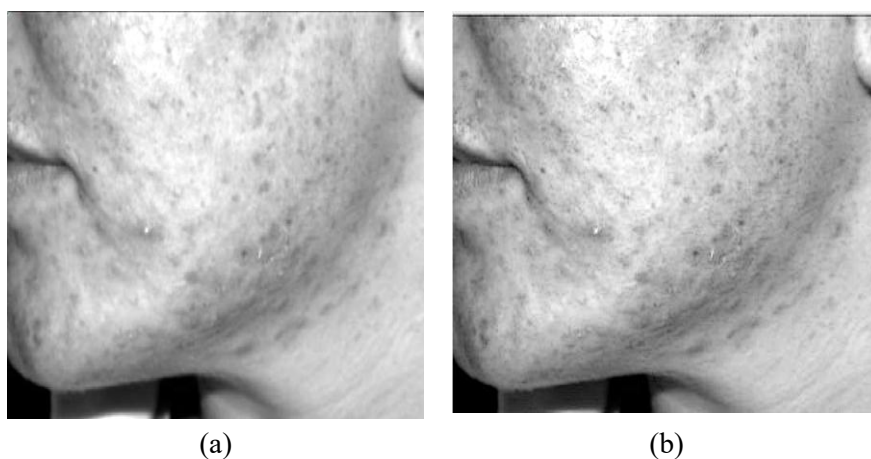


Figure 6. Sharpening stage: (a) before sharpening, (b) after sharpening

The next stage extracts the most effective features for categorizing the two skin diseases. It also reduces the processed data down to 0.03%, and it is done for each image (59/181120). This stage significantly reduces the classifier's computation time.

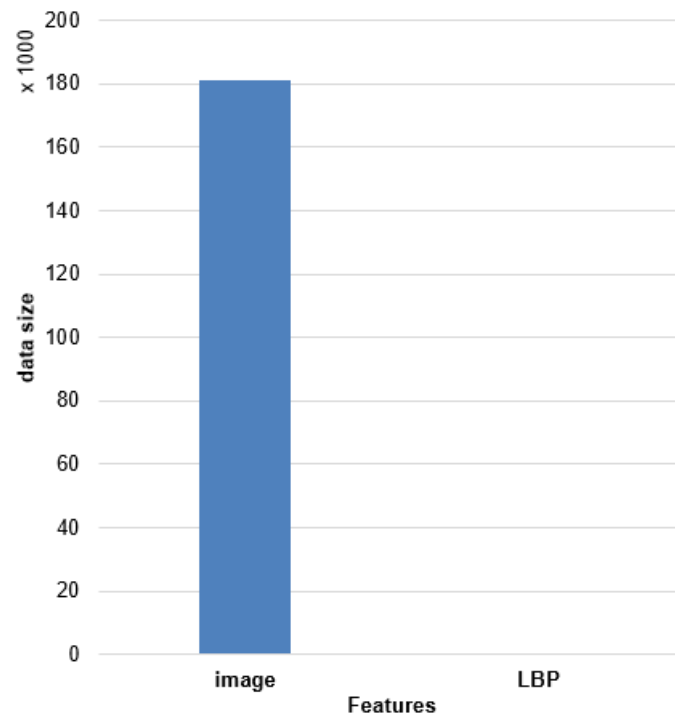


Figure 7. The amount of data reduction for the LBP stage

Two classifiers (SVM and MLP) were used to distinguish skin diseases. The MLP has one hidden layer with 10 nodes, and hyperbolic tangent and linear activation functions for the hidden and output layers, respectively. The backpropagation optimizer is also used in this classifier. Figure illustrates the performance of this classifier. This figure shows that the MLP needs only 10 epochs to train the classifier, indicating that the chosen features help the network learn more effectively during training. It also shows that the classifier doesn't suffer from overfitting, since the test curve is consistent with the training curve.

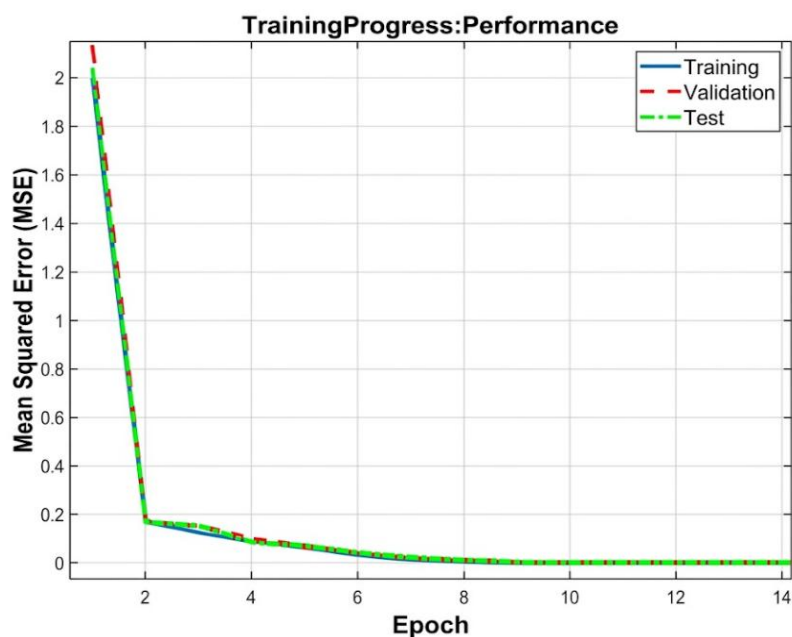


Figure 8. MLP training performance

The SVM is the second classifier used in the classification stage, with a radial basis function as the kernel function. The two classifiers show approximately the same performance; see Figure 9 for details. This figure shows that the MLP achieved 99.5% accuracy, while the SVM achieved 97.1% accuracy, indicating that the proposed method performs well at distinguishing between the two skin diseases [25, 26]. Moreover, the proposed method achieves performance comparable to, and in some cases surpassing, that of several recent studies [2, 28]. Future works can be considered to employ cloud computing [29], using machine learning [30] and optical techniques [31] to enhance the outcomes of this study.

Actual	TN	151	1
	TP	0	56
		TN	TP
		Predicted	

(a)

Actual	TN	147	5
	TP	1	55
		TN	TP
		Predicted	

(b)

Figure 9. The confusion matrices of the two classifiers: (a) MLP confusion matrix, (b) SVM confusion matrix

## 5. Conclusion

The skin is the body's outer layer that protects it and is prone to lesions. These can hurt, make you feel horrible about yourself, and even put your health at risk. It is not only necessary for medical reasons to treat skin disorders quickly, but it is also important for social sustainability since untreated problems can make patients' lives worse and make it harder for them to be involved in their communities. This research shows how to utilize AI to determine the difference between acne and actinic keratosis using image-processing approaches that don't need as many computational resources. This helps healthcare technologies use fewer resources.

Because there wasn't a single dataset, lesion photos were taken from different sources, resized to the same size, and improved by using grayscale light. We employed Local Binary Patterns (LBP) to extract and simplify features, and then we applied machine learning models to sort them. The results showed that the Multilayer Perceptron (MLP) was more accurate (99.5%) than the Support Vector Machine (SVM) (97.5%). This strategy combines accurate diagnosis with efficient computing to help make long-term healthcare solutions that are easier to find, can be utilized by more people, and can enhance patients' health while using fewer technology resources.

## Declaration of competing interest

The authors declare that they have no known financial or non-financial competing interests in any material discussed in this paper.

## Funding information

The author declares that they have received no funding from any financial organization to conduct this research.

## Author contribution

Mohand Lokman Al Dabag conceptualized the study, developed the methodology, and supervised the overall research. Shaima Miqdad Mohamed Najeeb contributed to data processing, implementation of machine learning models, and experimental analysis. Razan Abdulhammed assisted with validation, interpretation of results, and manuscript drafting. Haider T. H. Salim AlRikabi provided critical review, technical guidance, and final manuscript editing. All authors read and approved the final version of the manuscript.

## References

- [1] D. N. Tsedenya, M. S. Tilahun, and M. G. Berihun, “Fungal Skin Disease Classification Using the Convolutional Neural Network,” *Journal of Healthcare Engineering*, 2023. <https://doi.org/10.1155/2023/6370416>
- [2] Y.-C. Yang et al., “Skin disease classification using transfer learning model and fusion strategy,” *Intelligence-Based Medicine*, vol. 12, 2025. <https://doi.org/10.1016/j.ibmed.2025.100271>
- [3] M. Ahammed et al., “A machine learning approach for skin disease detection and classification using image segmentation,” *Healthcare Analytics*, vol. 2, 2022. <https://doi.org/10.1016/j.health.2022.100122>
- [4] R. Sadik et al., “An in-depth analysis of CNN architectures with transfer learning for skin disease diagnosis,” *Healthcare Analytics*, vol. 3, 2023. <https://doi.org/10.1016/j.health.2023.100143>
- [5] A. J. Mohammed et al., “Multi-objective Optimization in Satellite-Assisted UAVs,” *Journal of Intelligent Systems and Internet of Things*, vol. 16, no. 1, pp. 189–198, 2025. <https://doi.org/10.54216/JISIoT.160116>
- [6] K. B. Samir et al., “Machine Learning and Deep Learning Integration for Skin Diseases Prediction,” *IJETT*, vol. 70, no. 2, pp. 11–18, 2022.
- [7] T. S. Anisha and M. Sornam, “Skin Lesion Detection and Segmentation using SVM and Bayesian Classifier,” *IJCA*, vol. 9, no. 1, pp. 40–44, 2020.
- [8] K. S. Zehra et al., “Automated Classification of Skin Diseases Using Microscopic Images,” *Concurrency Computation Practice Experience*, 2025. <https://doi.org/10.1002/cpe.70220>
- [9] A. Akasha et al., “Early Detection of Skin Diseases Using Hybrid ML/DL Models,” *Information*, vol. 16, no. 2, 2025. <https://doi.org/10.3390/info16020152>
- [10] A. Sagheer et al., “Intelligent skin disease prediction using explainable AI,” *Scientific Reports*, vol. 15, 2025. <https://doi.org/10.1038/s41598-024-83966-4>
- [11] Q. Huang et al., “Skin cancer diagnosis using deep training and metaheuristics,” *Biomedical Signal Processing and Control*, vol. 83, 2023. <https://doi.org/10.1016/j.bspc.2023.104705>
- [12] N. H. Sohaib, “Accurate Deep Learning Algorithms for Skin Lesion Classification,” *ISI*, vol. 29, no. 4, pp. 1529–1539, 2024. <https://doi.org/10.18280/isi.290426>
- [13] S. A.-W. Alaa et al., “Deep learning framework for early diagnosis of skin cancer,” *Scientific Reports*, 2025. <https://doi.org/10.1038/s41598-025-04931-3>
- [14] Y. Peng et al., “Deep Learning on Imbalanced Datasets for Skin Lesion Classification,” *IEEE TMI*, vol. 41, no. 5, 2022. <https://doi.org/10.1109/TMI.2021.3136682>
- [15] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 4th ed., Pearson, 2018.
- [16] M. Tekalp, *Digital Video Processing*, 2nd ed., 2015.
- [17] J. C. Russ, *The Image Processing Handbook*, 7th ed., CRC Press, 2015.
- [18] R. Giovanni, “Sharpening by polynomial operators,” *IEEE Signal Processing*, vol. 23, no. 9, 2016.
- [19] R. R. O. Al-Nima, *Signal Processing and Machine Learning Techniques...*, thesis, 2017.
- [20] M. M. N. Shaima et al., “RNA Cancer Clustering Using SOM,” *JISIS*, vol. 15, no. 2, 2025. <https://doi.org/10.58346/JISIS.2025.I2.009>
- [21] L. Mohand and O. Nalan, “EEG classification based on cross-correlation,” *Signal Image Video Processing*, vol. 13, no. 3, 2019. <https://doi.org/10.1007/s11760-018-1383-9>

- [22] Kaggle Dataset, accessed Jan. 2026.
- [23] Roboflow Dataset, accessed Jan. 2026.
- [24] M. S. E. Ghada et al., "Privacy protection in healthcare using ML models," *Scientific Reports*, 2025. <https://doi.org/10.1038/s41598-025-04931-3>
- [25] I. B. Hasan et al., "SkinIncept ensemble model," *Discover Applied Sciences*, 2025. <https://doi.org/10.1007/s42452-025-06596-z>
- [26] A. Amina et al., "SkinHealthMate app," *Systems and Soft Computing*, vol. 6, 2024. <https://doi.org/10.1016/j.sasc.2024.200166>
- [27] M. Z. Uddin, M. A. Shahriar, B. W. Schuller, M. N. Mahamood, and M. A. R. Ahad, "Skin disease diagnosis using decision and feature level fusion of deep features," *Frontiers in Digital Health*, vol. 7, art. no. 1478688, 2025. <https://doi.org/10.3389/fdgth.2025.1478688>
- [28] Y. Ji et al., "ECG Classification using CNN," *Sensors*, vol. 19, no. 11, 2019. <https://doi.org/10.3390/s19112558>
- [29] Y. S. Mezaal, H. H. Madhi, T. Abd, S. K. Khaleel, "Cloud computing investigation for cloud computer networks using cloudanalyst," *Journal of Theoretical and Applied Information Technology*, vol. 96, no. 20, 2018.
- [30] M. S. Shareef, T. Abd, and Y. S. Mezaal, "Gender voice classification with huge accuracy rate," *TELKOMNIKA*, vol. 18, no. 5, p. 2612, 2020. <https://doi.org/10.12928/TELKOMNIKA.v18i5.13717>
- [31] N. Ahmed Al-Shareefi, J. A. Aldhaibaini, S. Adil Abbas, and H. S. Obaid, "Towards 5G millimeter-wave wireless networks: a comparative study on electro-optical upconversion techniques," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 20, no. 3, p. 1471, 2020. <https://doi.org/10.11591/ijeecs.v20.i3.pp1471-1478>