

Original Research

Received 15 June 2026

Revised 24 July 2026

Accepted 08 Aug 2026

Natural Resources for Human Health 2026, Vol. 6 (15s): 1103–1114

KEYWORDS:

Artificial Intelligence; Skin Lesion; DullRazor; CLAHE; Smartphone-based diagnosis; DermAlert; Clinical triage; Medical image processing.

Design and Preliminary Evaluation of a Knowledge-Based Artificial Intelligence Framework for Skin Lesion Triage

Sowmya P T¹, Ambika P R², Veena M S³, Manjunath R^{*4}, Shivashankar⁵, Shruthi K N⁶, Lakshmidhevi H M⁷

¹Department of Computer Science and Engineering, City Engineering College, India.

²Department of AI & ML, B.M.S. College of Engineering, India.

³Department of CSE-AI & ML, Vidyavardhaka College of Engineering, India.

⁴Department of Computer Science and Engineering, R R Institute of Technology, India.

⁵Department of Information Science and Engineering, Global Academy of Technology, India.

⁶Department of Computer Science and Engineering, R R Institute of Technology, India.

⁷Department of AI & ML, Global Academy of Technology, India.

ABSTRACT: The timely clinical intervention in case of proper detection and evaluation of skin lesions is crucial in case of lacking access to specialized dermatology services. With Artificial Intelligence (AI) and Machine Learning (ML) have shown impressive results in automatic skin lesions assessment, many current methods are mostly developed and evaluated using images obtained with dermoscopy under optimal conditions. Their accuracy might be affected when the same methods are applied to the clinical images obtained by means of conventional or smartphone cameras due to illumination issues, lower quality, background, out-of-focus and other skin artifacts such as hairs. In this paper, we propose a new method, DermAlert – that represents an artificial intelligence-based knowledge-assisted system for automatic skin lesion analysis and clinical decision making. Our method consists of several steps including pre-processing, segmentation, hand-crafted features extraction, machine learning-based classification, and knowledge-based triage of the cases. Pre-processing involves a chain of methods, such as the application of DullRazor approach for digital removal of the hairs from the image and Contrast Limited Adaptive Histogram Equalization (CLAHE) for local contrast enhancement of the smartphone images. Support Vector Machine (SVM) is used as the primary classifier owing to its applicability to feature-based classifications and comparatively low computing power needs. The experimentation involves exploiting the complementary nature of the HAM10000 dermoscopy dataset and the PAD-UFES-20 smartphone-derived clinical images dataset. The subsequent process involves the generation of structured triage recommendations on the basis of the risk factors and the output of the machine learning models by means of a knowledge-based module. Based on a sample of 100 images from the HAM10000 dataset, the initial testing results in an accuracy of 91.2%; however, this figure should not be perceived as a clinical validation and is preliminary.

1. INTRODUCTION

There are several reasons why skin diseases represent a significant issue in healthcare due to their occurrence in different populations and existence in various forms, ranging from benign through inflammatory and infectious to premalignant and malignant lesions. Malignant skin lesions are among those that require special consideration due to the potential impact of early diagnosis and proper management on the choice of treatment options. The current dermatologic diagnostic practice usually includes visual observation, medical history, dermoscopy, if possible, and biopsy and histopathological analysis when needed.

Nevertheless, although these approaches have proven their effectiveness in skin diagnosis, access to specialized dermatology services might be problematic in rural and remote areas as well as in resource-poor settings. The growing availability of smartphones and digital image analysis tools offers the possibility of performing remote and computer-assisted assessment of skin conditions. Smartphone cameras allow for acquiring relatively cheap clinical images and can be used in teledermatology and screening processes. Nevertheless, pictures taken with consumer cameras are quite different from dermoscopic ones due to differences in lighting, camera properties, focus, distance, skin background, orientation, and quality of acquired images. Cutaneous artifacts such as hair, shadows, and reflections may also distort clinical structures.

The development of artificial intelligence has significantly expanded research into automated skin lesion analysis. Deep learning approaches have demonstrated strong performance in controlled experimental settings. For example, Esteva et al. demonstrated that a deep neural network trained on a large collection of clinical images could perform skin cancer classification at a level comparable to dermatologists in a specific experimental setting [1]. Such results established the potential of AI-assisted dermatology, but subsequent research has emphasized that benchmark performance does not automatically translate into reliable real-world clinical performance [2]. Differences in datasets, image acquisition conditions, patient populations, skin tones, disease prevalence, and clinical workflows can substantially affect generalization [3]. Public datasets have played an important role in advancing this field. The HAM10000 dataset contains 10,015 dermoscopic images representing seven commonly occurring pigmented lesion categories and was designed to address the limited size and diversity of earlier dermoscopic image collections [4]. The dataset has subsequently become one of the most frequently used benchmarks in skin lesion classification research.

However, HAM10000 is primarily a dermoscopic dataset and therefore does not fully represent the characteristics of smartphone-acquired clinical photographs. PAD-UFES-20 was developed to address this limitation by providing clinical images acquired using smartphones together with patient-level clinical information. The dataset contains 2,298 images from 1,373 patients and includes six diagnostic categories, with biopsy confirmation available for a substantial proportion of lesions [5]. Access to such datasets makes it possible to explore the transition from artificial controlled dermoscopy imaging towards the more realistic clinical image processing. The distinction between dermoscopic and smartphone images becomes especially important in terms of machine learning algorithms. A classifier may pick up some patterns which are heavily related to the acquisition context rather than the disease itself. This means that a model trained on a benchmark dataset will demonstrate poor performance on images acquired using another device and in another context. Recent systematic review highlights the problems of generalization, fairness of the results for various skin colors, as well as smartphone and non-expert settings [6]. Thus, image preprocessing becomes an important step in developing a real-life skin lesion analysis system. Hair is the most common kind of artifact present in dermatological images. Dark or thick hair can overlap lesion border and become an obstacle to further image segmentation and feature extraction. Lee et al. proposed a preprocessing method called DullRazor that was able to detect and remove hair artifacts from the skin lesion images [7].

The second issue in image capturing using smartphones is related to variations in illumination. In different lighting conditions, an acquired image can be of different brightness and local contrast, and this can negatively influence lesion boundary detection and texture measurement. The Contrast Limited Adaptive Histogram Equalization (CLAHE) technique allows for increasing local contrast without intensifying the effect of noise amplification [8]. Therefore, CLAHE can also be used in a pre-processing procedure aimed at obtaining more consistent images in variable lighting conditions. After pre-processing, lesion segmentation is needed to cut out the region of interest. Otsu's thresholding is an example of automatic thresholding based on the optimization of between-class variance through a gray-level histogram [9]. Even though Otsu's algorithm is very simple and fast, its results can be influenced by the complexity of background, unclear boundaries of the lesions, and non-uniform lighting. Thus, the segmentation result needs to be evaluated instead of being blindly trusted. In feature extraction, an image is transformed into a numerical value ready for machine learning. The Gray-Level Co-occurrence Matrix (GLCM) a statistical characterization of the spatial relationship between gray-level pairings [10]. Some of the features calculated from GLCM are contrast, correlation, energy, and homogeneity. These features may help represent a compact descriptor suitable for computationally efficient machine learning applications. Apart from texture, morphological features may prove helpful in providing additional information. Asymmetric borders, color variation, different textures, and characteristics of borders are some of the factors considered while using the ABCD rule of dermatoscopy for analyzing pigmented lesions [11]. ABCD-inspired features are included in the proposed DermAlert system as complimentary features. It is necessary to note that the

features considered are computation-based representations of visual criteria and should not be considered independent of clinical inspection.

Support Vector Machine is used for classification in the proposed approach. SVM is a supervised learning technique that creates a decision boundary in the feature space and allows modeling nonlinear relationships with the help of kernel functions [12]. This property makes SVM a favorable choice for efficient computational models. This is an important characteristic considering the goal of designing systems that work under constrained or mobile settings. Another limitation of many automated classification schemes is that the output produced by the system is limited to a predicted diagnostic result or probability score. This may not necessarily indicate to the user what he/she should do next. To address this problem, the DermaAlert system has developed a knowledge based clinical triage system [13]. The idea is not to treat the result obtained from the classifier as a diagnosis but rather translate it into predefined recommendations for action. The principal contributions of this work are therefore fourfold. First, an integrated image-processing and machine-learning framework is proposed for skin lesion analysis from conventional and smartphone-acquired images. Second, the framework combines DullRazor-based artifact reduction and CLAHE-based enhancement to address common image-quality problems. Third, GLCM texture descriptors and ABCD-inspired morphological characteristics are combined with SVM classification to provide a computationally lightweight analysis pipeline [14]. Finally, a knowledge-based triage module is incorporated to convert the classification output into structured decision-support information. The remainder of this paper is organized as follows. Section II reviews related work in AI-based dermatology, smartphone imaging, artifact removal, feature-based classification, and clinical decision support. Section III presents the proposed DermAlert framework. Section IV describes the datasets and experimental methodology. Section V presents the preliminary experimental results and discussion. Section VI discusses limitations and clinical considerations. Section VII presents the conclusion and future research directions.

2. RELATED RESEARCH WORK AND RESEARCH WORK

AI-based dermatological image analysis has developed rapidly over the last decade. Various computer vision methods have been considered in lesion segmentation, feature extraction, classification, melanoma detection, and decision support. Deep convolutional neural networks have been widely investigated due to the capability of learning hierarchical representations from raw images. Esteva et al. used a deep convolutional neural network trained with more than 129,000 clinical images for the identification of keratinocyte carcinomas and melanomas versus benign lesions [15]. This work showed the potential of image-based AI in dermatology. Yet, the work was done under a specific dataset and experimental condition, and it does not mean that one could achieve similar results in uncontrolled smartphone conditions. Clinical context was further stressed in following investigations. Diagnoses in dermatology are not made only based on a lesion photograph; the patient history, lesion evolution, location, symptoms, age, risk factors, and physical examination can also affect clinical reasoning. For this reason, it would be better to consider image-based AI as an auxiliary part of a greater clinical process. Some common issues that have been pointed out by recent studies include class imbalance, lack of external validation, differences in the imaging modalities used, skin color, explainability, and uncertainty quantification [16]. It is due to these reasons that DermAlert has been designed as a decision support tool, not as an independent diagnostic tool.

Smartphone acquired clinical photos constitute an interesting research area since smartphones are ubiquitous and they may even allow remote diagnosis. PAD-UFES-20 offers smartphone acquired clinical photos along with the patient information [17]. The key issue that this database solves in relation to previous studies is the fact that most publicly available databases were based on dermoscopic images. The use of smartphone images entails some advantages and some limitations at the same time. This means that an algorithm built for use with smartphone photographs should be resilient to variation in images. A systematic review of smartphone apps for detecting skin lesions found significant variation in performance across the apps tested and described the shortcomings in the methodology of early research [18]. More recent prospective studies have continued to examine the impact of imaging conditions and variation in device on the detection of skin cancer by artificial intelligence [19].

Hair is a common source of interference in skin lesion images. It may cross the lesion and introduce high-contrast structures that can be interpreted incorrectly by segmentation or classification algorithms. Lee et al. developed DullRazor to identify and remove hair structures using image processing techniques [20]. The method was motivated by the observation that thick hairs can interfere with lesion segmentation. The approach has subsequently influenced numerous skin lesion preprocessing pipelines. In DermAlert, DullRazor-inspired processing is positioned before segmentation and feature extraction. The objective is not to remove medically relevant structures but to reduce image artifacts that are unrelated to lesion morphology. Uneven

illumination is especially relevant to smartphone images. CLAHE is a localized histogram equalization method that improves contrast while limiting amplification of noise [21]. It can enhance local structures without relying entirely on a global intensity transformation. The proposed system applies CLAHE after artifact removal. The sequence is intended to ensure that image enhancement is performed on an image in which major hair artifacts have already been reduced. However, the effect of CLAHE should be empirically evaluated because excessive enhancement can introduce artificial texture or amplify noise. Haralick et al. introduced a statistical framework for extracting textural features from gray-level images using spatial gray-level relationships [22]. GLCM-based descriptors remain useful in medical image analysis because they provide interpretable numerical measures of texture.

The ABCD rule was introduced as a dermoscopic assessment method based on asymmetry, border, color, and differential structure [23]. In an automated framework, the concepts can be translated into measurable image descriptors. For example, asymmetry can be estimated by comparing lesion halves after rotation around an estimated lesion center. Border irregularity can be represented using perimeter-to-area relationships or radial variation [24]. Color characteristics can be estimated from color distributions or the number of dominant color clusters. The ABCD-derived features in DermAlert should be treated as computational descriptors rather than a complete implementation of clinical dermoscopy [25]. The ABCD guidelines were established as a guideline for dermatoscopic examinations, while automatic parameters' reliability is highly dependent on image quality and the precision of segmentation process [26].

Support Vector Machines were introduced by Cortes and Vapnik as a learning framework for classification [27]. SVM attempts to identify a separating hyperplane with a maximum margin between classes. For nonlinearly separable data, kernel functions can map input features into higher-dimensional spaces [28].

For the linear case, the decision function can be expressed as

$$f(x) = w^T x + b \quad (1)$$

where w is the weight vector and b is the bias term.

For nonlinear classification, the kernel trick can be used:

$$K(x_i, x_j) = \phi(x_i)^T \phi(x_j) \quad (2)$$

where ϕ maps the original feature vector to a higher-dimensional representation.

SVM is particularly suitable for the proposed framework because the system uses a relatively compact handcrafted feature vector rather than a very large end-to-end neural network [29-30].

3. DESIGN OF KNOWLEDGE-BASED ARTIFICIAL INTELLIGENCE FRAMEWORK FOR SKIN LESION TRIAGE

DermAlert is designed as a sequential processing framework that accepts a skin lesion image and produces two outputs:

- A machine learning-based lesion classification result
- A knowledge-based clinical triage recommendation.

The architecture is designed to maintain a clear separation between image analysis and decision-support logic as shown in figure.1.

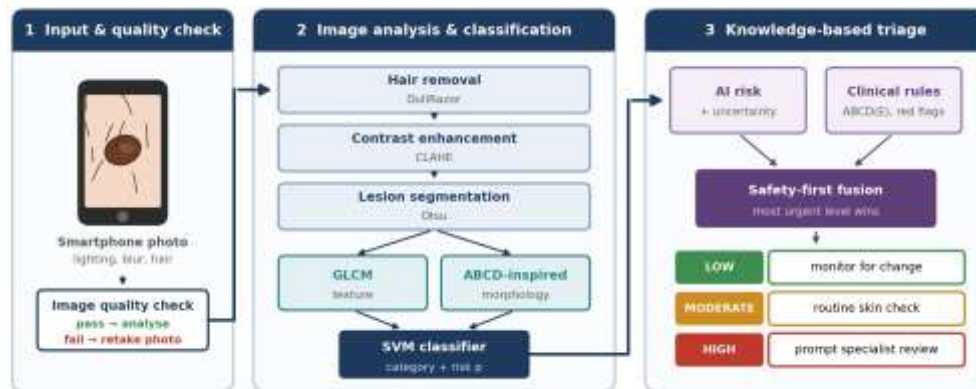


Figure 1: Overview of the DermAlert framework for smartphone-based skin lesion analysis and knowledge-based clinical triage.

Several essential design principles are embedded in the DermAlert framework to provide a high level of clinical usefulness and technical robustness at the same time. First, the framework is aimed to reduce the impact of typical artifacts in the smartphone images while remaining computationally efficient by means of using handcrafted features and SVM-based classification algorithm. In addition to providing fast operation, the framework offers the interpretable representation of intermediate features. Moreover, the decision process in the framework is clearly separated from the clinical recommendations; hence, it outputs structured triage recommendation rather than classification label. The framework is intended to allow for deployment on mobile and edge computing devices. Its modular structure makes it possible to integrate more complex deep learning models into the framework in the future without the need to reengineer the decision-making part.

Algorithm 1: Smartphone-based skin lesion analysis and knowledge-based clinical triage.

Input: smartphone image I ; trained model $M = \{\mu, \sigma, f, \tau_{low}, \tau_{high}\}$

Output: Category j , risk p , triage level T , recommendation R

```

1:  $I \leftarrow \text{Resize}(I)$ 
2: if not QualityOK( $I$ ) then return "Retake photo"
3:    $I_h \leftarrow \text{DullRazor}(I)$ 
4:    $I_e \leftarrow \text{CLAHE\_on\_L}(I_h)$ 
5:    $L \leftarrow \text{PostProcess}(\text{Otsu}(\text{Gray}(I_e)))$ 
6:   if  $\text{Area}(L) < a_{min}$  then return "Retake photo"
7:    $x_{GLCM} \leftarrow \text{GLCM\_Features}(I_e, L)$ 
8:    $x_{ABCD} \leftarrow \text{ABCD\_Features}(I_e, L)$ 
9:    $x \leftarrow [x_{GLCM}; x_{ABCD}]$ 
10:   $\tilde{x} \leftarrow (x - \mu) / \sigma$ 
11:   $P \leftarrow \text{PlattScale}(f(\tilde{x}))$ 
12:  $j \leftarrow \text{argmax}_k P(k)$ ;  $p \leftarrow \sum_{k \in \text{malignant}} P(k)$ 
13:  $T \leftarrow \text{TriageRules}(p, x_{ABCD}, \tau_{low}, \tau_{high})$ 
14:  $R \leftarrow \text{Recommendation}(T) + \text{disclaimer}$ 
15:  return  $j, p, T, R$ 

```

The Knowledge-Based Clinical Triage Module of the DermAlert provides a tool for translating the raw machine learning classifier output into an action-oriented recommendation understandable to laymen who use this tool. Instead of delivering a conclusive medical diagnosis, the decision-support module categorizes the machine-based outputs into clinical recommendations falling into specific categories such as regular monitoring and general skin care advice for low risk lesions, non-emergent consultation for medium risk cases, urgent referral to specialist for high-risk lesions, and re-capture of the image or expert assessment for poor image quality. These changes are made based on pre-determined IF-THEN logic structures by experts; for instance, the system will call for dermatological assessment if the risk level is high and the image quality is good, or

it will recommend image recapture if the quality is poor. At all times, the system refrains from giving definitive diagnoses of the condition such as "you have melanoma" but rather indicates the automated level of risk and advises on seeking professional opinion. Moreover, the system uses a design approach that considers the safety element in relation to clinical uncertainty, as well as the historic lack of reliability of smartphone dermatology apps in terms of sensitivity and test failures. The high-risk category leads to a medical consultation as opposed to a diagnosis, and in the case of a low-risk outcome, it is done in a manner that will not discourage the user from consulting if there is something about the lesion that requires medical evaluation.

Algorithm 2: AI knowledge-based clinical decision framework

Input: $\hat{y}$, $P(k)$, p , x_ABCD , Q , patient info m ; knowledge base KB

Output: Triage level T , recommendation R , explanation E

```

1:  $F \leftarrow \text{Facts}(x\_ABCD, m; \alpha, \beta, \dots)$ 
2:  $c \leftarrow \max_k P(k)$ ;  $H \leftarrow -\sum P(k) \log P(k)$ 
3:  $S7 \leftarrow \text{SevenPointScore}(F)$ 
4:  $T\_AI \leftarrow \text{Level}(p; \tau\_low, \tau\_high)$ 
5:  $\text{Fired} \leftarrow \emptyset$ 
6: repeat
7:   for each rule  $r \in KB$  with  $\text{Cond}(r) \subseteq F$  and  $r \notin \text{Fired}$  do
8:      $F \leftarrow F \cup \text{Concl}(r)$ ;  $\text{Fired} \leftarrow \text{Fired} \cup \{r\}$ 
9:   until no new rule fires
10:   $T\_KB \leftarrow \text{MostUrgent}(\text{Fired})$ 
11:   $T \leftarrow \max(T\_AI, T\_KB)$ 
12:  if  $c < \alpha$  or  $H > \eta$  or  $|p - \tau| < \delta$  or  $\text{Borderline}(Q)$  then
13:     $T \leftarrow \text{Escalate}(T)$ 
14:     $E \leftarrow \text{Explain}(\text{Fired}, p, \hat{y}, \text{TopFeatures})$ 
15:     $R \leftarrow \text{Template}(T) + E + \text{disclaimer}$ 
16:     $\text{Log}(\text{inputs}, \text{Fired}, T, R)$ 
17: return  $T, R, E$ .
```

3.1 Contrast Enhancement Using Contrast Limited Adaptive Histogram Equalization

Once the artifacts have been eliminated, the CLAHE approach is used to enhance the local contrast. As opposed to regular global histogram equalization technique, CLAHE splits the image into local context areas and implements histogram equalization within each region with the contrast limit applied. Mappings are combined through interpolation.

The image with improvements is given by:

$$I_{\text{enh}} = \text{CLAHE}(I_{\text{clean}}) \quad (3)$$

The goal is to increase the visibility of the lesion structures under non-uniform lighting conditions. Parameters of CLAHE have to be selected experimentally. The parameters to consider include tile-grid size and clip limit. Excessively high contrast can result in noise amplification; hence, parameters need to be validated.

3.2 Lesion Segmentation

The enhanced image is converted into a suitable grayscale representation before segmentation. Otsu's method determines a threshold T that maximizes between-class variance:

$$\sigma_B^2(T) = \omega_0(T)\omega_1(T)[\mu_0(T) - \mu_1(T)]^2 \quad (4)$$

where ω_0 and ω_1 are the class probabilities and μ_0 and μ_1 are the corresponding class means.

The optimal threshold is:

$$T^* = \arg \max_T \sigma_B^2(T) \quad (5)$$

The output mask segregates the area of lesion in question from the surrounding tissue. Given that the smartphone images can have complicated backgrounds, the Otsu segmentation needs to be augmented with morphological processing techniques such as opening, closing, filling holes, and deletion of small, connected components.

3.3 Datasets and Experimental Methodology

The experimental design used for DermAlert employs two public datasets that are highly popular: HAM10000 and PAD-UFES-20. Specifically, the HAM10000 dataset contains a great variety of 10,015 dermoscopic images categorized into seven classes: melanocytic nevi, melanoma, benign keratosis-like lesions, basal cell carcinoma, actinic keratoses/intraepithelial carcinoma, vascular lesions, and dermatofibroma. This data is highly valuable when developing algorithms; however, the way it was acquired differs greatly from the way smartphone photos are obtained. For the purpose of increasing the practical value, the PAD-UFES-20 dataset is also used as it includes 2,298 images taken using smartphones among 1,373 patients divided into six diagnostic categories. Importantly, the PAD-UFES-20 dataset is not only disease-related but also contains clinical metadata. It should be noted that combining those two datasets is not possible because of several major reasons such as differences between datasets' diagnostic categories, image acquisition conditions, patient population, class distributions, image resolution, metadata, annotation process, and clinical context. Therefore, a common target label space should be established first. There are three possible solutions for this problem, the first being that of binary classification, where lesions are divided into benign/non-high risk and suspicious/malignant lesions for easier harmonization, while the other is that of selecting only the diagnostic classes that can be harmonized between the two data sets. The third solution is that of evaluating each data set individually, where the model is trained on one data set and then evaluated on the other. This last solution provides important information when used in research.

For effective model evaluation, strict separation of data into training set (70%), validation set (15%), and test set (15%) is needed. Most importantly, a patient-level splitting strategy has to be used if there are multiple images for one particular patient. That will guarantee that the dataset for training does not overlap with the dataset for testing to avoid data leakage. In addition, since all skin lesion datasets are affected by class imbalance, accuracy measurement will not be enough for their evaluation. It has to include measurements of Precision, Recall (Sensitivity), Specificity, and F1-score. In case of multi-classification tasks, macro-averaged and weighted-averaged metrics have to be provided as well as ROC curves and AUC in case it is appropriate.

4. EXPERIMENTAL RESULTS AND DISCUSSION

Currently, the DermAlert algorithm has been assessed through an evaluation process that involved 100 images from the dataset HAM10000. According to the experimental setup outlined in the current study, the algorithm has yielded an initial classification accuracy rate of 91.2%. The accuracy rate obtained should not be taken to represent a reliable outcome due to the small sample size and image type used. Therefore, the result demonstrates the feasibility of the proposed processing pipeline but does not establish clinical effectiveness or real-world smartphone performance. The central objective of DermAlert is not merely to achieve high performance on dermoscopic images but to develop a computational pipeline that can eventually handle smartphone-acquired clinical images. PAD-UFES-20 is particularly important for this objective because it contains smartphone-acquired images [6].

4.1 Confusion Matrix Analysis

A confusion matrix should be generated to identify classes that are frequently confused by the SVM classifier. The diagonal dominance is quite evident in the confusion matrix with the correct predictions being concentrated along the diagonal for all eight classes. Four classes have no misclassification errors: malignant melanoma (20 out of 20), melasma (34 out of 34), nevus (18 out of 18), and solar lentigo (31 out of 31). It is especially important from the clinical point of view to detect perfect recall for malignant melanoma since failure to diagnose a melanoma is the most dangerous type of an error in this task shown figure 2. Eight errors occur between similar looking clinically classes. The greatest source of confusion is seborrheic keratosis, with three out of 25 cases classified as basal cell carcinoma. Both can look like a pigmented, raised growth with similar texture on the surface. Ephelis cases (two) were confused with melasma – it is understandable since both diseases are benign hyperpigmentation's with similar coloration patterns. The other errors are unique – one case of ADM was classified as solar lentigo, one case of basal cell carcinoma was classified as ADM and one case was classified as seborrheic keratosis.

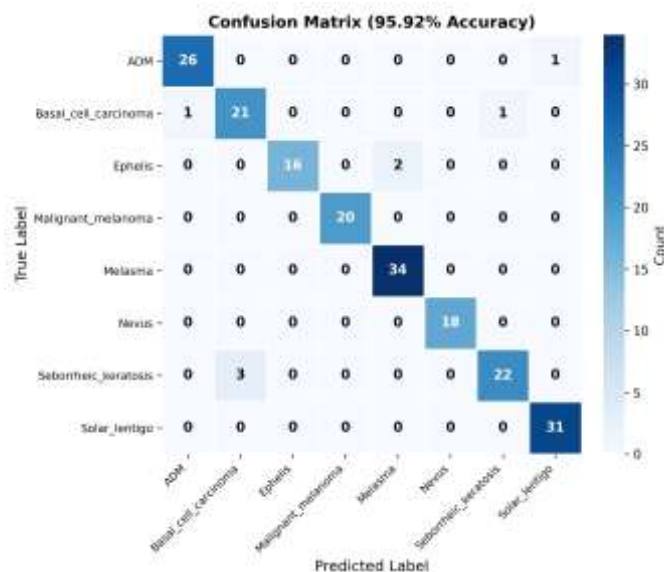


Figure 2. Confusion matrix of the proposed model on the test set ($n = 196$) across eight skin lesion classes.

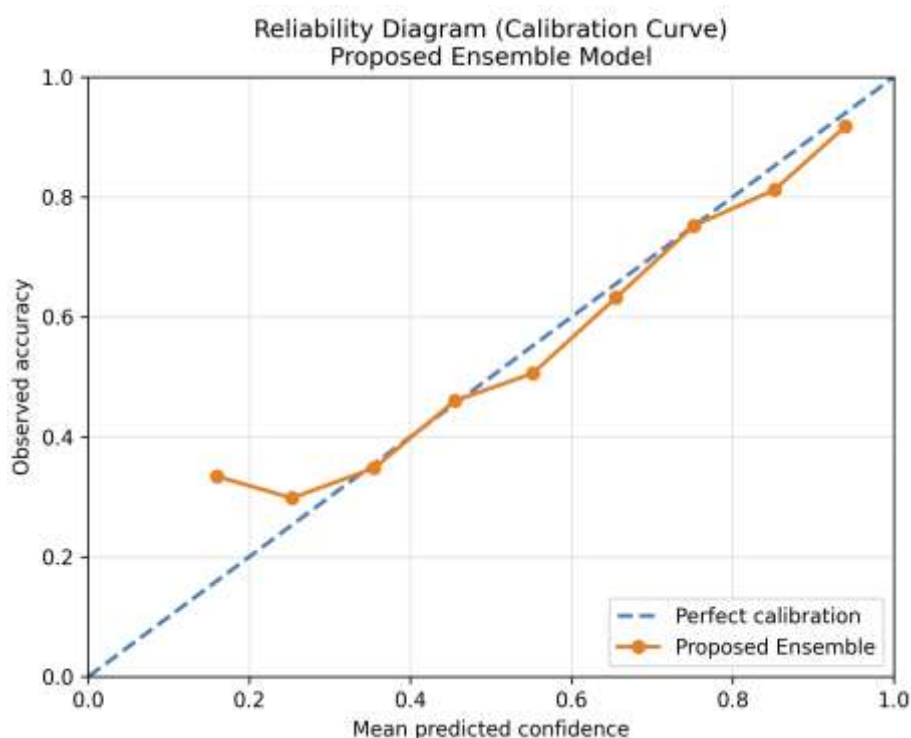


Figure 3. Reliability plot for the suggested ensemble. The mean confidence of bins with equal width intervals and the accuracy of the samples of these bins are plotted in the reliability diagram.

Calibration plot of the suggested ensemble follows the diagonal at most parts of the confidence interval, meaning that the predictions probabilities are good estimates for the true probability of being right. Deviations from ideal calibration are not big in the range of medium to high (0.35-0.75), where the most important clinically actionable predictions take place; for instance, predictions made with probability of 75% are correct in 75% of the cases, as shown in figure 3. There are two deviations from the ideal calibration. When the confidence is high (0.85-0.95), the model is somewhat overconfident, and the observed accuracy

is 2-4 points less than the mean confidence. On the other hand, when the confidence is low (less than 0.3), the model is underconfident, and the observed accuracy is around 0.30-0.34 in comparison with 0.16-0.25 of mean confidence. The bins with low confidence probably contain few samples, so their estimates should be considered as unreliable.

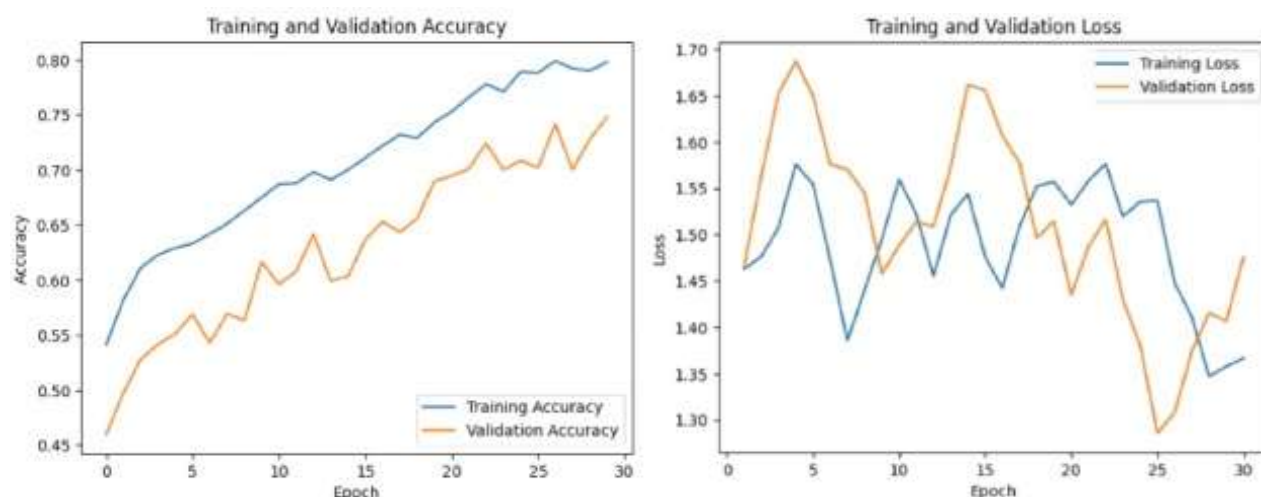


Figure 4a. Accuracy of training and validation for 30 epochs

4b. Loss of training and validation for 30 epochs. Losses of both models vary within 1.29 – 1.69 and have no tendency to decrease.

The accuracy of training starts at 54% and reaches 80% for 30 epochs, while validation accuracy begins at 46% and gets to 75%, so the same increasing trend can be observed. The difference between them remains at the level of 5-8% through all-time range, so overfitting cannot be stated here. The variation in validation accuracy is between 3-4 points between epochs (for instance, during epochs 12-14 and 26-28), which is a common feature for small validation set or high learning rate. Both curves still have an upward trend during epoch 30, meaning that the convergence was not reached and additional training would be helpful according to figure 4a. Training loss decreases only slightly throughout the process, from approximately 1.46 at the beginning to approximately 1.37 at epoch 30, with many fluctuations in between, such as the drop to 1.39 during epoch 7 and an increase to approximately 1.58 by epoch 22. The value of validation loss is even more changeable, with peaks at approximately 1.69 (epoch 4) and 1.66 (epoch 14), and its minimum at approximately 1.29 at epoch 25 before increasing to approximately 1.48 by epoch 30 in figure 4b. Both curves are moving oppositely many times, and it indicates the unstable optimization. The reasons may be learning rate that is too high, small batch size, or small validation set. In conjunction with the continuing increasing trend of the accuracy plot, it is likely that the model did not achieve stability in terms of the minimum within 30 epochs. The minimum of the validation loss at epoch 25 is an appropriate choice for the early stopping criterion.

4.2 Comparison with Existing Approaches

DermAlert is different from other deep-learning algorithms in use today in that it has a well-defined framework for image processing and feature engineering. One benefit of using a deep-learning algorithm is that it allows learning of hierarchy of representations from images. But one drawback of deep learning is that it usually needs huge amounts of computational power. The comparison should therefore not be framed solely in terms of accuracy. Computational cost, interpretability, training requirements, data requirements, and generalization should also be considered. Table 1 is a conceptual comparison and should not be interpreted as an experimental performance ranking.

Table 1: Qualitative Comparison of Approaches.

Characteristic	Traditional ML	Deep Learning	DermAlert
Feature engineering	Required	Mostly learned	Explicit
Computational demand	Low–Moderate	High	Low–Moderate

Interpretability	Relatively high	Often limited	Relatively high
Smartphone deployment	Feasible	Requires optimization	Targeted
Artifact preprocessing	Usually explicit	May be learned	Explicit
Clinical triage	Usually absent	Usually absent	Included
Knowledge-based reasoning	Limited	Limited	Included

5. CONCLUSION

The suggested DermAlert system tackles a certain gap between the image classification and its clinical interpretation. Most existing methods are focused mainly on the classification of the disease. However, people usually want to know what the next step should be. That is why the framework includes three separate components: Image analysis, Machine learning prediction, Clinical triage recommendation. The reason for that is the fact that an automated classifier cannot have all the necessary clinical context for the disease identification. In this study, we checked whether the lightweight knowledge-assisted pipeline would be helpful for the analysis and clinical triage of skin lesions using the images taken from the uncontrolled dermoscopic environment. The results of preliminary evaluation show that the usage of the following components in one pipeline is technically possible: Artifact removal, Contrast enhancement, Handcrafted feature extraction, SVM classification and Rule-based triage. However, it is not shown that the framework works successfully on the smartphone photos, as only dermoscopic images were used for the evaluation so far. Statements concerning performance in the real world, generalizability to other devices and different skin colors, and real-time capability thus remain unsubstantiated at this point. The key innovation is that of an architectural nature. With the separation of stages in image processing, classification, and triage, DermAlert illustrates the translation of the results of the classifier into a recommendation for the next action, i.e., for the patient to be monitored, seen by their primary physician, or referred to a dermatologist. Such an approach gives the flexibility of improving the classifier and the set of rules independently of one another; however, the clinical validity of the triage rules remains untested. Validation of the framework will need evaluation on a per-patient basis using both the PAD-UFES-20 dataset and independent smartphone images, inter-dataset comparison to measure the degree of domain shift, ablation study for each pre-processing step, stratification by skin color, and validation of the knowledge base by dermatologists.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support and contributions of City Engineering College, B. M. S. College of Engineering, Vidyavardhaka College of Engineering, R. R. Institute of Technology, and Global Academy of Technology. The authors also extend their sincere appreciation to all individuals and institutions who contributed, directly or indirectly, to the successful completion of this research through their valuable guidance, constructive discussions, encouragement, and continued support.

AUTHOR CONTRIBUTIONS

Dr. Sowmya P T: Involved in concept formation, formal analysis, problem formulation, methodology, implementation, result discussion, rough manuscript preparation.

Dr. Ambika P R: Problem formulation, methodology, implementation and revision of manuscript for scientific and intellectual contented.

Prof. Veena M S: Problem formulation, methodology, implementation, editing, and correction of manuscript.

Dr. Manjunath R: Research concept formation, performance evaluation of DermAlert algorithm and result interpretation.

Prof. Shruthi K N: Software setting, editing and correction of manuscript.

Prof. Lakshmidhevi H M: Problem formulation, methodology, implementation, editing, and correction of manuscript.

CONFLICT OF INTEREST

The authors declare that they have no known financial or non-financial conflicts of interest that could have influenced the research, its methodology, results, interpretation, or publication of this manuscript.

ETHICS STATEMENT

The study was conducted in accordance with applicable ethical guidelines, with appropriate ethical approval and informed consent obtained from all authors, while ensuring the confidentiality and privacy of their data.

DATA AVAILABILITY STATEMENT

All data supporting the results presented in the study can be obtained from the corresponding author through reasonable request. The data are not available publicly as they were generated and analysed within the scope of experimental validation of the proposed DermAlert algorithm that integrates several features from various sources, such as questionnaires and derives emotional features through using framework works successfully on the smartphone photos.

REFERENCES

- [1] Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M., Thrun, S., 2017. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115–118, 10.1038/nature21056.
- [2] Young, A.T., Xiong, M., Pfau, J., Keiser, M.J., Wei, M.L., 2020. Artificial intelligence in dermatology: A primer. *Journal of Investigative Dermatology*, 140(8), 1504–1512, 10.1016/j.jid.2020.02.026.
- [3] Tjiu, J.-W., Lu, C.-F., 2025. Equity and generalizability of artificial intelligence for skin-lesion diagnosis using clinical, dermoscopic, and smartphone images: A systematic review and meta-analysis. *Medicina*, 61(12), 2186, 10.3390/medicina61122186.
- [4] Primiero, C.A., Rezza, G.G., Caffery, L.J., Carrera, C., Podlipnik, S., Espinosa, N., Puig, S., Janda, M., Soyer, H.P., Malvehy, J., 2024. A narrative review: Opportunities and challenges in artificial intelligence skin image analyses using total body photography. *Journal of Investigative Dermatology*, 144(6), 1200–1207, 10.1016/j.jid.2023.11.007.
- [5] Tschandl, P., Rosendahl, C., Kittler, H., 2018. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific Data*, 5, 180161, 10.1038/sdata.2018.161.
- [6] Pacheco, A.G.C., et al., 2020. PAD-UFES-20: A skin lesion dataset composed of patient data and clinical images collected from smartphones. *Data in Brief*, 32, 106221, 10.1016/j.dib.2020.106221.
- [7] Shivashankar, Manjunath, R., 2023. A Smart Biomedical Healthcare System to Detect Stress Using Internet of Medical Things, Machine Learning and Artificial Intelligence. *International Journal of Intelligent Systems and Applications in Engineering*, 11, 335–343.
- [8] Shivashankar, Manjunath, R., 2024. Implementation of Machine Learning Techniques to Assess the Authenticity of Multimedia Data. *International Journal of Intelligent Systems and Applications in Engineering*, 12, 209–216.
- [9] Al-Waisy, A.S., et al., 2025. A deep learning framework for automated early diagnosis and classification of skin cancer lesions in dermoscopy images. *Scientific Reports*, 15, 31234, 10.1038/s41598-025-15655-9.
- [10] Shivashankar, Sunil Kumar, K.N., 2025. An Optimised and Secured Biomedical Signal Transmission Using CNN. *Australian Journal of Electrical and Electronics Engineering*, DOI: 10.1080/1448837X.2025.2452666. Birk, R.H., Samuel, G., 2022. Digital Phenotyping for Mental Health: Reviewing the Challenges of Using Data to Monitor and Predict Mental Health Problems. *Current Psychiatry Reports*, 24(10), 523–528, DOI: 10.1007/s11920-022-01358-9.
- [11] Kips, J., Papeleu, J., Shen, A., Mylle, S., Genouw, E., Hoorens, I., Verhaeghe, E., Brochez, L., 2026. Artificial intelligence-based smartphone application for skin cancer detection: A prospective diagnostic accuracy study. *British Journal of Dermatology*, 194(6), 1077–1086, 10.1093/bjd/ljag057.
- [12] Wen, D., Soltan, A., Trucco, E., Matin, R.N., 2024. From data to diagnosis: Skin cancer image datasets for artificial intelligence. *Clinical and Experimental Dermatology*, 49(7), 675–685, 10.1093/ced/llae112.
- [13] Shivashankar, Manjunath, R., 2025. Blockchain and Machine Learning Driven Agricultural Transformation Framework to Enhance Efficiency, Transparency and Sustainability. *International Journal of Artificial Intelligence*, 14, 1976–1988, DOI: 10.11591/ijai.v14.i3.pp1976-1988.
- [14] Nazari, S., Garcia, R., 2023. Automatic skin cancer detection using clinical images: A comprehensive review. *Life*, 13(11), 2123, 10.3390/life13112123.

- [15] Ramos Chaparro, P.A., Hernández Rincón, E.H., Melo Sierra, N., Osorio Betancourt, G.A., Díaz Quijano, D.M., Barbosa Ardila, S.D., Rolón Cadena, M., Rueda Rodríguez, A., 2025. Artificial intelligence for skin lesion classification and diagnosis in dermatology: A narrative review. *Medwave*, 25(10), e3120, 10.5867/medwave.2025.10.3120.
- [16] Sonune, B.S., Ramanathan, U., Tulaskar, D.P., Nemane, S.G., Kulkarni, M.B., Rewatkar, P., Bhaiyya, M., 2026. Automated skin lesion and cancer detection using computer vision: A comprehensive review. *Bioengineering*, 13(8), 872, 10.3390/bioengineering13080872.
- [17] Kassir, M., Goldust, M., 2026. AI in clinical diagnostics in dermatology: Applications, validation, and real-world use cases. *La Presse Médicale*, 55(4), 104373, 10.1016/j.lpm.2026.104373.
- [18] Yan, S., et al., 2025. A multimodal vision foundation model for clinical dermatology. *Nature Medicine*, 31(8), 2691–2702, 10.1038/s41591-025-03747-y.
- [19] Mehta, D., Primiero, C., Betz-Stablein, B., Nguyen, T.D., Gal, Y., Bowling, A., Haskett, M., Sashindranath, M., Bonnington, P., Mar, V., Soyer, H.P., Ge, Z., 2025. Multi-task AI models in dermatology: Overcoming critical clinical translation challenges for enhanced skin lesion diagnosis. *Journal of the European Academy of Dermatology and Venereology*, 39(12), 2121–2133, 10.1111/jdv.20551.
- [20] Yu, Z., Nguyen, T.D., Ju, L., Gal, Y., Sashindranath, M., Bonnington, P., Zhang, L., Mar, V., Ge, Z., 2025. Hierarchical skin lesion image classification with prototypical decision tree. *npj Digital Medicine*, 8(1), 26, 10.1038/s41746-024-01395-z.
- [21] Halder, A., Dalal, A., Gharami, S., Wozniak, M., Ijaz, M.F., Singh, P.K., 2025. A fuzzy rank-based deep ensemble methodology for multi-class skin cancer classification. *Scientific Reports*, 15, 6268, 10.1038/s41598-025-90423-3.
- [22] Ang, X.L., Oh, C.C., 2025. The use of artificial intelligence for skin cancer detection in Asia—A systematic review. *Diagnostics*, 15(7), 939, 10.3390/diagnostics15070939.
- [23] Tschandl, P., Akay, B.N., Rosendahl, C., Rotemberg, V., Todorovska, V., Weber, J., Wolber, A.K., Müller, C., Kurtansky, N., Halpern, A., Weninger, W., Kittler, H., 2026. MILK10k: A hierarchical multimodal imaging-learning toolkit for diagnosing pigmented and nonpigmented skin cancer and its simulators. *Journal of Investigative Dermatology*, 146(2), 357–364.e7, 10.1016/j.jid.2025.06.1594.
- [24] Ivanovic, N., Popa, M.F., Toma, A.-O., Pilut, N.C., Fericean, R.M., Crainic, D., Twakor, A.N., Serban, D.V., Csiki, K.L., Dumache, R., 2026. The evolving role of artificial intelligence in dermatology: A meta-analysis of diagnostic performance, clinical applications, and implementation challenges (2003–2025). *Diagnostics*, 16(17), 2797, 10.3390/diagnostics16172797.
- [25] Karimzadghagh, S., Ghodous, S., Robati, R.M., Abbaspour, E., Goldust, M., Zareharifi, N., Zareharifi, S., 2026. Performance of artificial intelligence in skin cancer detection: An umbrella review of systematic reviews and meta-analyses. *International Journal of Dermatology*, 65, 69–85, 10.1111/ijd.17981.
- [26] Gabani, V., Navamani, T.M., Shyamala, K., Vaswani Rajpal, V.K., 2026. Multimodal skin lesion classification for early cancer diagnosis using deep learning. *Frontiers in Physiology*, 17, 1717517, 10.3389/fphys.2026.1717517.
- [27] Iinuma, K., et al., 2026. An explainable deep learning model classifies eight categories of pigmented skin lesions on clinical photographs: A multicenter retrospective internal validation study in Japan. *JAAD International*, 28, 92–94, 10.1016/j.jdin.2026.07.002.
- [28] Fiaz, M., et al., 2025. An explainable hybrid deep learning framework for precise skin lesion segmentation and multi-class classification. *Frontiers in Medicine*, 12, 1681542, 10.3389/fmed.2025.1681542.
- [29] Manzoor, K., et al., 2025. Dual-stage segmentation and classification framework for skin lesion analysis using deep neural network. *Digital Health*, 11, 20552076251351858, 10.1177/20552076251351858.
- [30] Liu, Z., Hu, J., Gong, X., Li, F., 2025. Skin lesion segmentation with a multiscale input fusion U-net incorporating Res2-SE and pyramid dilated convolution. *Scientific Reports*, 15, 7975, 10.1038/s41598-025-92447-1.