Archives available at journals.mriindia.com

International Journal of Recent Advances in Engineering and Technology

ISSN: 2347 - 2812

Volume 15 Issue 01, 2026

Automated Skin Disease Diagnosis Using Deep Learning & Image Processing Techniques

[1]Bhupesh Dewangan, [2]Milind Sahu, [3]Jaynendra Kumar, [4]Mrs. Harsha Dubey

[1,2,3] Computer Science and Engineering, Shri Shankaracharya Institute Of Professional Management & Technology, Raipur, India

[4]Assistant Professor, Computer Science and Engineering, Shri Shankaracharya Institute Of Professional Management & Technology, Raipur, India

Email: [1]bhupeshdewangan160204@gmail.com[2]milindsahu.21@gmail.com, [3]jaynendrak18@gmail.com, [4]harsha.dubey@ssipmt.com

Peer Review Information

Submission: 02 Jan 2026

Revision: 23 Jan 2026

Acceptance: 15 Feb 2026

Keywords

Automated, Skin Disease Detection, Image Processing, HAM10000, RestNet

Abstract

Skin conditions must be treated promptly and appropriately because if left untreated they can develop into more serious conditions like melanoma, or skin cancer. They must be treated promptly and appropriately. Manual examinations by doctors take time, and the outcomes can vary from one to the next. In order to solve this problem, this study created an automated system that recognizes skin conditions using machine learning and image processing methods. First, skin images are cleaned—hair is removed using digital hair removal technology, and image quality and clarity are improved using Gaussian filtering. Next, the GrabCut algorithm is used to precisely isolate only the affected area (lesion). The structure, color, texture, and statistical features of the area are then extracted, which are used to identify the disease. These features are analyzed by machine learning models such as SVM, KNN, and Decision Tree to determine which skin disease is present. This system helps increase diagnostic accuracy, reduce human error, and speed up the identification process. In study conducted on datasets such as ISIC 2019 and HAM10000, after the test the SVM model demonstrated the highest accuracy. The results shows that this technique can be very helpful for doctors diagnose skin diseases earlier, and automatically.

Introduction

The human body is made up of several organs. Skin is the largest organ of the human body which covers the entire human body and act as shield against external factors such as infections, radiation, and chemical exposure [1]. However, it is also highly vulnerable to a wide range of diseases, including acne, eczema (atopic dermatitis), psoriasis, skin cancer, rosacea, alopecia areata (hair loss), vitiligo, and hives [2]. "These diseases are capable of causing skin symptoms, such as rashes, inflammation, itching, and other skin abnormalities. Some skin disorders may be genetic, while others may be associated with lifestyle choices. Skin disorders may be treated with medications, lotions or ointments, or changes in lifestyle." [3].

The global burden of skin diseases is significant. According to World Life Expectancy using WHO-data (2020), skin disease deaths in India reached 17,857, which accounts for about 0.21% of all deaths in India [4]. Beyond physical health, these conditions severely impact patients' personal, social, and professional lives. People with skin diseases try to avoid the community because they feel cursed.

These skin conditions can affect person's personal, social, and professional as well as the lives of their family members. Suicidal thoughts may result from the affected people's anxiety, stress, depression, and low self-esteem.

Any abnormal change in the skin's appearance or texture compared to normal, healthy skin is called Skin Lesion. In simple words, we can say that, it means a spot, bump, patch, or sore on the skin that looks or feels different.

There is main two kinds of skin lesions: primary and secondary.

The initial observable alterations in the skin are known as primary lesions. Typical instances consist of:

- Macules are flat, discolored patches on the skin.
- Papules are tiny, elevated lumps.
- Nodules: larger, solid lumps that develop beneath the epidermis.
- Mole: innocuous pigmented growths.

When primary lesions progress or the skin becomes even more irritated or damaged, secondary lesions develop. These consist of:

- Erosion: shallow spots in which the epidermis, the outermost layer of skin, is lost.
- Scars: fibrous tissue that forms when skin damage heals.
- Ulcers: deeper, open sores that penetrate the dermis, the skin's middle layer.

The severity and symptoms of lesions vary. Some may be painless or painful, and Some are permanent, and some are temporary. Melanoma is the most dangerous and fatal of these skin conditions. However, if a skin disease is detected early, about 95% of patients can recover.

There exists a significant gap between dermatologists and patients with skin diseases, largely because many individuals lack awareness of the various types, symptoms, and stages of these conditions [5].

Sometimes it takes a long time for visible signs to appear, which makes it even harder to detect them in time. Early and precise diagnosis is essential, but conventional techniques can frequently be difficult, expensive, and time-consuming, especially when identifying the precise kind and stage of an illness. Automatic computer-aided systems driven by machine learning have surfaced to overcome these obstacles, allowing for the quicker and more accurate diagnosis of various skin conditions.

Literature Review

Various researchers have suggested distinct alternative approaches for skin disease classification. These studies can generally be divided based on their used dataset, used feature extraction method, used feature selection method, and used classification model; most relevant contributions will be reviewed below to show used tools, used techniques methodologies and limitations of the past studies and find research gaps.

Computational Methods Evolution in Skin Disease Classification:

Despite their success, the first computational methods of skin disease detection employed classical image processing. For example, Jagdish et al. [6] used a fuzzily clustered image preprocessing optimized with K-Nearest Neighbor (KNN) and supported by a for squamous and basal cell carcinoma detection learned on the dataset of only 50 pictures. Although KNN showed 91.2% accuracy, the small size of the initial dataset limited the model's generalizability. Naeem et [7] al. developed a processing pipeline for cancer prediction that used the Gray Level Co-occurrence Matrix (GLCM) along with SVM for text-based feature extraction and picture augmentation.

Hybridization of Deep and Traditional Techniques:

Once deep learning became popular, the majority of researchers started to apply already trained models with Convolutional Neural Networks (CNN) for automatic feature extraction. Bandyopadhyay et al. [8] implemented a hybrid approach combining AlexNet, GoogLeNet, ResNet50, and VGG16 architectures with tradition classifiers such as SVM, Decision Tree, and Adaboost. They proved that combining the deep representations with classical properties might result in more considerable performance. Another study by AlDera et al. [9] combined the segmentations and the feature extraction methods. They applied Otsu's thresholding, Gabor Filters using classifiers including SVM, Random Forest, and KNN to detect a variety of diseases from the Atlas Dermatologico and DermNet NZ datasets, including psoriasis, melanoma, and acne, with 67.1% to 90.7% accuracy.

Advancements with Lightweight and Ensemble Models

There were even more additional advancements, First, these additional advancements are computationally lightweight architectures. Kshirsagar et al. [10] used MobileNetV2 with Long Short-Term Memory networks. As a result, its accuracy and efficiency were much higher than conventional CNN schemes, durably keeping an eye on the manner the tumor patterns grow. Second, they also tried ensemble models.

For example, Kalaivani et al. [11] adopted a dual propagation ensemble. It was a deep learning model's presence to the ISIC 2019 data set, showing remarkable performance on seven skin disease categories. These ensemble techniques highlighted the advantages of model fusion in enhancing classification stability and generalization.

Deep Architectures, Optimization Techniques, and Transfer Learning:
Following that, recent studies have explored architectural improvements in ensuring generalizability and data imbalance issues and optimized training strategies. For example, Hatem [12] developed a simple KNN model for binary lesion classification that demonstrated high accuracy but undue reduce diagnostic capabilities. In addition, to overcome data imbalance challenge, Yao et al. [13] designed multi-weighted loss function with DropOut layer, DropBlock layer, and Modified RandAugment, which was ultimately able to perform consistently with minimal added computational cost.
Kethana et al. [14] employed a total of 10,015 images from the ISIC 2019 dataset to construct CNN models that accomplished 92% accuracy in the determination of melanoma, nevus, and seborrheic keratosis. In the same way, Maduranga et al. utilized transfer learning with MobileNet based on the HAM10000 dataset to develop an AI mobile diagnostic device that achieves, on average, 85% accuracy in real-time detection. Extending these results, Jain et al. [15] presented a probabilistic approach to deep learning for post-hoc optimization of neural networks and called the Optimal Probability-Based Deep Neural Network OP-DNN), which resulted in an accuracy of 95% and excellent specificity 0.97 and sensitivity 0.91 values, thus proving the high efficiency of possible approaches to probabilistic optimization in the field of diagnostic deep learning. Padmavathi et al. [16] used deep learning models of pre-training and fine-tuning to classify skin lesion images autonomously. After evaluating the performance of the models showed by such parameters as sensitivity, precision, specificity, and accuracy, the researchers confirmed the effectiveness and applicability of the presented methods.

Identified Research Gaps and Study Direction:
These studies demonstrate that the majority of current methodologies can only detect a small number of skin diseases usually dedicated to a single disease. Furthermore, many methods implement small and/or imbalanced datasets, which adds to their limited generalizability. Many preprocessing methods are not comprehensive, and previous works primarily rely on basic methodologies including thresholding, edge detection, and clustering.
To tackle these issues, the current study employs the large-scale ISIC-2019 dataset which covers eight skin diseases. The study uses advanced preprocessing with digital hair removal and next employs a graph-based segmentation approach to improve lesion localization to enhance performance on multiclass classification.

Methodology

The proposed methodology for skin disease detection using image processing consists of several stages, starting from image acquisition and ending with classification. Each step plays a crucial role in ensuring accurate and efficient disease detection. The complete workflow is described below:

A. Dataset Gathering and Preparation

Utilize large and extensive dermoscopic image datasets, such as the ISIC 2019 Challenge and HAM10000, which are publicly accessible and cover a range of skin disease classes (e.g., melanoma, benign keratosis, vascular lesions, basal cell carcinoma, etc.).
Random Oversampling data balancing strategies are used for address dataset imbalances, which ensure that all classes are equal represented and enhance the model generalization.

B. Image Preprocessing

- Image Resizing: To Maintain consistency and to save computing costs, we set input images to a standard size (e.g., 224x224 or 512x512 pixels).
- Hair Removal: Used digital hair removal techniques for the elimination of hairs and other artifacts that impede lesion analysis, by combining morphological filters such as Black-Hat transformation with inpainting algorithms.
- Noise Removal: Apply Gaussian filtering to minimize noise in the picture while keep preserving significant texture of elements.

C. Image Segmentation

For accurate segment lesions from the surrounding healthy skins, used sophisticated approaches like the GrabCut algorithm, which is uses colour space and clustering, and provide efficient boundary detection.
The segmentation step isolates lesion regions to concentrate feature extraction and enhance classification accuracy.

D. *Feature Extraction using Deep Learning*

Use networks like deep convolutional neural networks (CNNs) for automatic feature extraction of segmented lesion images.

EfficientNet, a convolutional neural network (CNN) by google which achieved high accuracy and efficiency through compound scaling is applied for learning spatial hierarchies of features. It uses a single compound coefficient to uniformly measure the network's depth, width, and resolution rather than just one dimension. This is a suitable approach for both large-scale systems and mobile devices because while using fewer computational resources it also enhances performance.

E. *Model Training*

- To take advantage of the features learned on the comprehensive image collection, apply transfer learning by adapting pre-trained CNN models (such as EfficientNet, ResNet, and VGG) to skin disease datasets.
- For model training, use an optimizer with an appropriate learning rate, such as Adam.
- The clategorical cross-entropy is a suitable loss function for multi-category classification applications.
- A regularization method like Dropout can be helpful for detection of overfitting and prevention of it.
- Accuracy and loss calculation can be used to measure the performance of the model and if necessary, it can perform early termination in case of overfitting.

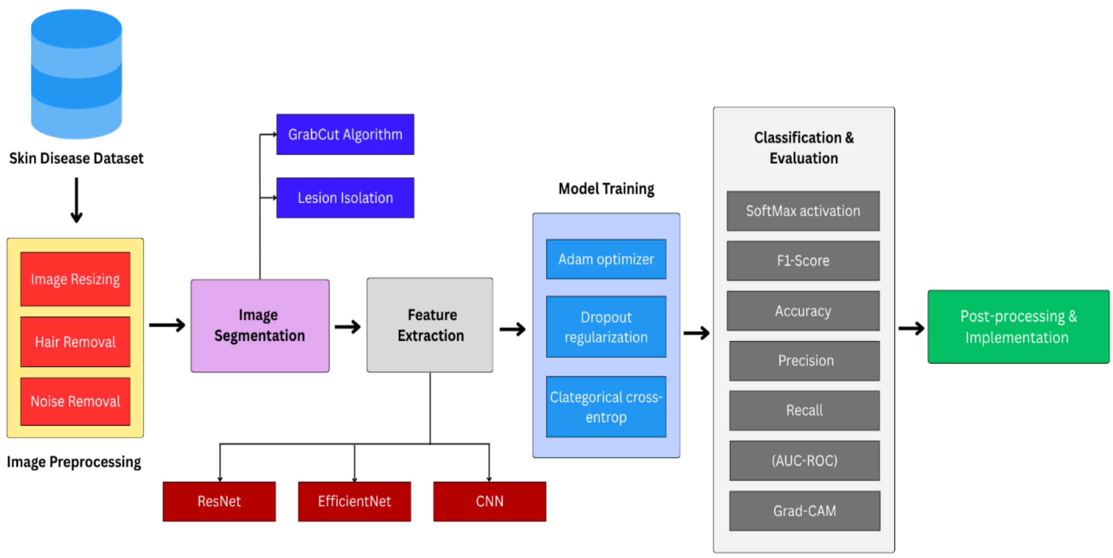
F. *Classification and Evaluation*

- To produce class probabilities for several skin disease categories. The Convolutional Neural Network (CNN) uses a SoftMax activation in the last layer.
- Parameter such as F1-Score, Accuracy, Precision, and Recall (Sensitivity) are used to measure performance of model.
- Across all possible boundaries, a single scalar value area under the receiver operating characteristic curve (AUC-ROC) that is used to encapsulate the performance of the binary classification model.
- Make a matrix of confusion to understand performance by class.
- To strengthen clinician confidence, use straightforward strategies such as Grad-CAM to pinpoint visual cues which impact assessments when observing model projections.

G. *Post-processing and Implementation*

By utilizing trained models in clinical decision support systems, dermatologists may identify skin diseases more rapidly, accurately, and automatically.

To maintain performance, regularly retrain the model and update it with fresh annotated data.



Result and Discussion

A. *Experiment Setup*

The ISIC 2019 Challenge dataset and the HAM10000 dataset were used as public benchmark datasets for testing the Python implementation of the proposed skin disease detection framework. There are eight classes of skin diseases in the ISIC dataset and seven classes in the HAM10000 dataset. The data was balanced by random oversampling as a part of preprocessing due to a natural imbalance between classes.

To greatly enhance the quality of the images for analysis, the images were resized to 512x512 pixels and went through various advanced preprocessing steps such as Gaussian noise filtering and digital hair removal (via Black-Hat morphological transformation and image inpainting).

B. *Segmentation and Feature Extraction*

In order to accurately identify lesion regions, we used an automatic GrabCut segmentation method along with k-means clustering in color space HSV. GrabCut consistently outperformed k-means segmentation alone by demonstrating precise extraction of the foreground, which was essential for classification in the subsequent tasks.
The lesion areas were segmented, and texture features were extracted from these lesion segments based upon the Gray Level Co-occurrence Matrix (GLCM) to capture critical amounts of contrast, homogeneity, energy, entropy, and correlation. To broaden the feature set, statistical color features were extracted from the RGB channels, which include mean, variance, standard deviation, and root mean square.

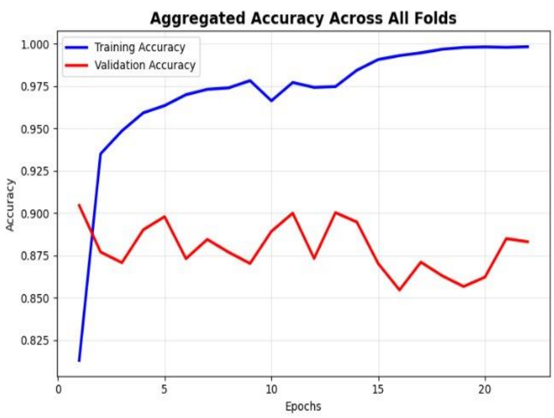
C. *Classification Performance*

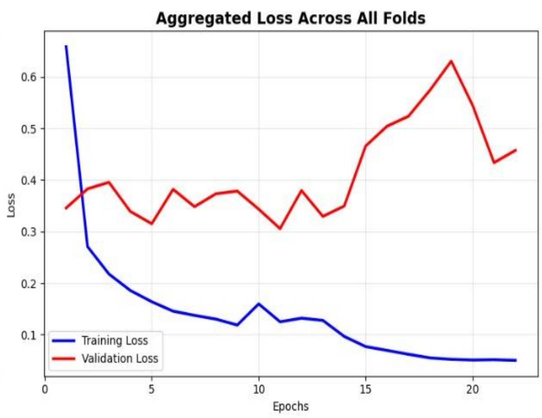
The three machine learning techniques – SVM, KNN and also Decision Tree were checked once again on balanced and unbalanced types of datasets for comparison. As expected, the classical classifiers gave very poor results. After oversampling, their performance increased and based on the dataset, they gave 93–97% accuracy.

However, using ResNet-50, the deep learning model showed
some improvement and on both the ISIC 2019 and HAM10000 datasets, it gave 95% accuracy. Similarly, Precision, Recall and F1-score also followed the same trend.
In which ResNet-50 outperformed and performed better than all classical methods. Instead of this, ResNet-50 performed
with less log-loss compared to SVM, KNN and DT models. Therefore, it became more reliable for the future.

Aggregated classification report:

Name	precision	recall	f1-score	support
Atopic Dermatitis	0.8609	0.8321	0.8463	1257
Normal Skin	0.9806	0.9919	0.9862	3315
Melanoma	0.9904	0.9904	0.9904	3140
Seborrheic Keratoses and other Benign Tumors	0.8951	0.9145	0.9047	1847
Nail Fungus	0.9763	1.0000	0.9880	2967
Chickenpox	0.9940	0.9614	0.9775	3293
Atopic Dermatitis	0.8609	0.8321	0.8463	1257





Individual Fold Results:

S.No.	Test Accuracy	Bal Accuracy	F1-score
Fold 1	0.9600	0.9358	0.9377
Fold 2	0.9685	0.9530	0.9540
Fold 3	0.9666	0.9563	0.9546

Final Report:

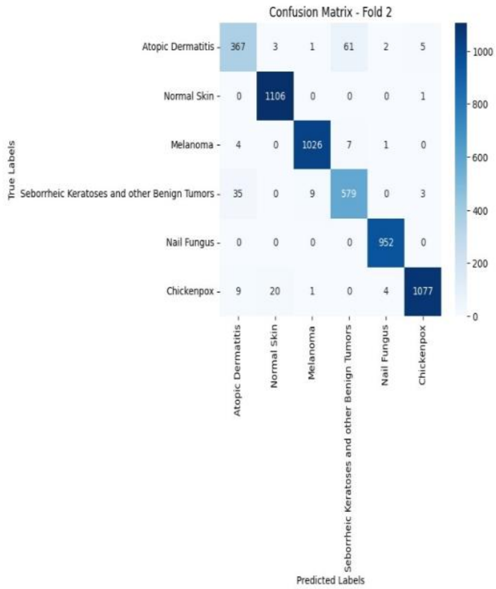
Metric	Mean	Std	Best Fold
Test Accuracy	0.9650	0.0037	0.9685
Balanced Acc	0.9484	0.0090	0.9563
F1-Score	0.9488	0.0078	0.9546
Accuracy Gap	0.0913	0.0469	0.1411

Performance Summary:
Test Accuracy: 0.9650 ± 0.0037
Balanced Accuracy: 0.9484 ± 0.0090
F1-Score (Macro): 0.9488 ± 0.0078

D. ROC and Confusion Matrix Analysis

Receiver Operating Characteristic (ROC) curves and Area Under the Curve (AUC) measurements also exhibited good classification performance with overall AUC values reaching 98% (ISIC) and 99% (HAM10000) through SVM.

The confusion matrix showed high per-class accuracy, especially for melanoma, chickenpox, and nail fungus, vascular lesions. All minor classes showed low identification accuracy. Balanced training and segmentation guided deep learning significantly reduced misclassifications.



E. Comparison with State-of-the-Art

The proposed ResNet-50 based framework has written the CNN variants and OP-DNNs' simple architecture correctly. The hybrid model proved to be better than the traditional one.

But deep-learning-based models are computationally intensive. GrabCut-based segmentation performs well, it reduces irrelevant pixel contribution, whose direct effect is seen in the model's output and its efficiency. Depending only on thresholding or manual segmentation, depending on traditional methods, causes problems in automatic segmentation in the RESNET pipeline and in localization; the system got improved flow. For this reason, compared to other modern deep learning models, the proposed approach is more accurate, reliable, and between computational complexity a computational balance has been achieved.

F. Limitations and Future Work

The limitation is due to GrabCut not always segmenting the lesions optimally and, consequently, affecting the reliability of the classification. Looking forward, future work will include working with ensemble and advanced deep learning techniques for automated and real-time skin disease detection, with continued efforts to optimize segmentation methods and improve the classification accuracy.

Conclusion

Because of Deep Learning concept many skin diseases identification becomes very successful. These methods can detect many skin diseases. Digital Hair Removal and noise reducing like technique, Image Processing, accurate Lesion

Segmentation (like GrabCut) together gives correct information of skin Lesion. From which our trust on these types of techniques increasing more. CNN is used in which EfficientNet like Architecture Features automatically extract and identifies many diseases. By doing experiment it is found that Deep Learning on big DataSet like ISIC2019 and HAM10000 is better compare to Machine Learning. Detects correctly, gets Precision Recall and F1 score, this Model helps correctly in starting which is quite effective for immediate treatment.
In future work for dealing with imbalance and different complex diseases also for easy explanation of Complex Model. Although technology and emerging techniques are reducing all these problems.
If we say in short, Deep Learning itself provides strong reliable help of science in detecting disease correctly Also for correcting human mistakes and reducing and also for timely correct treatment it is quite effective.

References

Anatomy of the human skin and it's Functional Layer, Stanford children's health, (2021) [Online].
Available:
https://www.stanfordchildrens.org/en/topic/default?id=anatomy-of-the-skin-85-P01336

Katherine Brind'Amour and Rachael Ajmera, MS, RD, "All About Common Skin Disorders", Healthline (2024) [Online]. Available: https://www.healthline.com/health/skin-disorders#list

Khasanova Saodat, Yuldoshaliyeva Dilshodakhon, Djuraeva Barno Gulamovna, "SKIN DISEASES ON MEDICINE", Zenodo, Jan. 2024.
doi:
https://www.doi.org/10.5281/zenodo.10472979
World Health Ranking, Leading Causes of death in india (2020) [Online] Available:
https://www.worldlifeexpectancy.com/india-skin-disease

Mohammad Shorif Uddin, Md. Al Mamun, Mostafiz Ahammed, "A machine learning approach for skin disease detection and classification using image segmentation", Healthcare Analytics, Vol. 2, p. 100122, ISSN 2772-4425. (2022)
doi:
https://doi.org/10.1016/j.health.2022.100122

María A. García López, Jhony A. de la Cruz-Vargas, Maleny Estefanía Roque Camacho, Mukta Jagdish, Sandra Paola Gualán Guamangate, "Advance study of skin diseases detection using image processing methods." 20220157042, English, Journal article, Turkey, 2148-9637, 9, (1), Eskisehir, Natural Volatiles & Essential Oils, (997–1007), Badebio Biotechnology Ltd., (2022)
doi:
https://www.cabidigitallibrary.org/doi/full/10.5555/20220157042

Z Bukhari, Z Naeem, G Zia, "A healthcare model to predict skin cancer using deep extreme machine." J. NCBAE, vol. 1, no. 2, pp. 23–30, (2022).
doi: http://refhub.elsevier.com/S2772-4425(22)00062-4/sb10

Payal Bose, Amiya Bhaumik, Sandeep Poddar, Samir Kumar Bandyopadhyay, "Machine Learning and Deep Learning Integration for Skin Diseases Prediction," International Journal of Engineering Trends and Technology, Volume-70 Issue-2, no. 3, pp. 13-21, (2022).
doi:
https://doi.org/10.14445/22315381/IJETT-V70I2P202

Mohamed Tahar Ben Othman, Shaden Abdulaziz, and AlDera, "A model for classification and diagnosis of skin disease using machine learning and image processing techniques." International Journal of Advanced Computer Science and Applications Vol. 13, Iss no. 5 (2022).
doi:
https://dx.doi.org/10.14569/IJACSA.2022.0130531

Praveen Kumar Balachandran, Abdulrhman M. Alshareef, ,Hariprasath Manoharan, Nabeel Albishry, S. Shitharth, Pravin R. Kshirsagar, "Deep learning approaches for prognosis of automated skin disease." Life, 12(3), 426; (2022)
doi: https://doi.org/10.3390/life12030426

S. Karpagavalli, A. Kalaivani, "Detection and classification of skin diseases with ensembles of deep learning networks in medical imaging", International Journal of Health Sciences, Vol. 6 No. S1, 13624–13637. (2022)
doi: https://doi.org/10.53730/ijhs.v6nS1.8402

Mustafa Qays Hatem, "Skin lesion classification system using a K-nearest neighbor algorithm", Visual Computing for Industry, Biomedicine, and Art, 5, Article No. 7 (2022).
doi: https://doi.org/10.1186/s42492-022-00103-6

Pengfei Shao, Jinyu Xing, Benjamin Kaffenberger, Fan Zhang, Ronald X. Xu, Shuwei Shen, Peng Liu, Peng Yao, Mengjuan Xu, "Single Model Deep Learning on Imbalanced Small Datasets for Skin Lesion Classification," in IEEE Transactions on Medical Imaging, Volume: 41, Issue: 5, May (2022)
doi:
https://doi.org/10.1109/TMI.2021.3136682

B. MS and A. KS. "Melanoma disease detection and classification using deep learning." Int. J. Res. Appl. Sci. Eng. Technol. 10, no. 7 (2022).
doi: http://refhub.elsevier.com/S2772-4425(22)00062-4/sb16

Chandra Sekhara Rao, Praphula Kumar Jain, Annavarapu, Ajith Abraham and Arushi Jain, "Multi-type skin diseases classification using OP-DNN based feature extraction approach", Multimedia Tools and Applications, Volume 81, pages 6451–6476, (2022).
doi: https://doi.org/10.1007/s11042-021-11823-x

Dilshan Nandasena and M.W.P. Maduranga, "Mobile-Based Skin Disease Diagnosis System Using Convolutional Neural Networks (CNN)", International Journal of Image, Graphics and Signal Processing (IJIGSP), Vol.14, No.3, pp. 47-57, (2022).
doi: https://doi.org/10.5815/ijigsp.2022.03.05