iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** V **Month of publication:** May 2025

DOI: https://doi.org/10.22214/ijraset.2025.71340

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

Healthcare Innovation for Skin Disease Analysis Using ML and DL

Dhruv Malhotra[1], Dhruv Kuchhal[2]
Delhi Technological University, India

Abstract:** Skindiseases represent a significant health- care burden worldwide, and rapid diagnosis is essential foreffectivetreatment.Machinelearning(ML)and deep learning (DL) are increasingly applied to automate and improve dermatological diagnosis. In this study, we present a novel skin disease classification pipeline that integrates multiple ML and DL techniques in a unified framework. First, dermoscopic images are preprocessed toremoveartifacts (e.g.,hair)usingmorphologicalfilters and inpainting, and segmented using clustering and GrabCut algorithms to isolate the lesion region. To augment the training data and address class imbalance, wetrainagenerativeadversarialnetwork (GAN)onamulticlassskindataset.Therefineddatasetisthen fed into dual deep feature extractors (ResNet50 and DenseNet121, pretrained on ImageNet) to obtain rich feature representations. These features are combined through an attention-based fusion strategy and passed to a hybrid classifier comprising gradient boosting models and a support vector machine. Key contributions include the pipeline of artifact removal and segmentation, GAN- based data enhancement, and an ensemble classifica-tion strategy with explainable-AI components for trans- parency.Weevaluatetheapproachonapubliclyavailable multi-categoryskindiseasedataset(21classes),achieving high diagnostic accuracy (>90%) on the test set. Results indicate that our integrated ML/DL system outperforms traditional methods, reduces time to diagnosis, and supports efficient teledermatology for enhanced patient care. Moreover, this approach sets a foundation for real- world deployment in healthcare environments by ensur- ing model interpretability and scalability. The success of our framework demonstrates the transformative impact AIcanhaveonimprovingdermatologicaldiagnosticsand patient outcomes, especially in underserved and remote areas where access to specialists is limited.

***Keywords:** Artificial Intelligence; Deep Learning; Dermatology; Machine Learning; Skin Disease; Telemedicine.*

I. INTRODUCTION

Skin diseases such as melanoma, eczema, and pso- riasisaffectalargeportionofthepopulationand can have serious health consequences if not diagnosed early. Dermatological diagnosis is typically based on visual examination by a specialist, but many regions lack sufficient experts, leading to delays and errors.AI-driven tools aim to bridge this gap by providing automated analysis of skin images. In particular, ML andDLhaveshowngreatpromiseinmedicalimaging.

Recent studies have applied a variety of ML and DL methodstoskinlesionclassification.Ahammedetal.

[1] introduced a segmentation-based ML pipeline and showed that proper preprocessing improves accuracy. Sun et al. [2] reviewed ML techniques in dermatology, noting that ensemble and hybrid models often outper- form single classifiers. Johnson et al. [3] compared CNN-based deep learning to traditional ML on skin lesion data, finding that CNNs significantly improve featurerepresentation. Similarly,Bandyopadhyay et al.

[4] and Kalaivani & Karpagavalli [5] demonstrated improved classification by combining multiple CNN architectures or fusing deep features with tree-based classifiers.AlDera&Othman[6]proposedahy- brid approach integrating image processing with ML, achieving robust skin disease diagnosis. These studies demonstrate the potential of advanced ML and DL for dermatology.

Despite these advances, several challenges remain. Dermoscopic images often contain artifacts (hair, bub- bles, uneven lighting) that can confuse algorithms ifnot properly removed. Many existing methods rely on simple preprocessing or ignore such artifacts, reducing robustness. Public skin disease datasets also tend to be imbalanced: common conditions have thousands of ex-ampleswhilerarediseaseshaveveryfew,whichbiases the models. Furthermore, few prior systems incorpo- rate model explainability or are designed for seamless telemedicineintegration.ExplainableAI(XAI)tech-niques (e.g., saliency maps) are important for clinical trust,yetarerarelyincludedindermatologyclassifiers.

Moreover, current diagnostic tools often lack the flexibility to generalize across diverse skin types and environmental conditions. Skin disease appearance varies significantly with ethnicity, age, and lighting, demanding robust models that can adapt to these varia- tions.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

Traditional algorithms may also suffer from high false-positive or false-negative rates when exposed to unseendata.Incorporatingcross-domaintransferlearn- ingandcontinualmodelupdatesisthereforecrucialfor maintaining high diagnostic performance over time.

To address these gaps, we propose a comprehen-sive multi-stage skin disease analysis framework. The pipeline begins with advanced preprocessing: a black- hat morphological filter and inpainting are applied to remove hair and noise, followed by k-means clustering and the GrabCut algorithm to precisely segment the lesion. Next, we train a generative adversarial network (GAN)ontheavailableimagestogeneraterealis-tic synthetic lesion examples, augmenting underrep- resented classes. We then extract deep features using two pretrained CNNs (ResNet50 and DenseNet121) andfusethemthroughanattentionmechanism.Finally, a hybrid classifier (an ensemble of gradient-boosting trees and an SVM) is trained for diagnosis, and Grad- CAM is used to provide visual explanations of model decisions.

The remainder of the paper is organized as follows: Section II describes the dataset and methodology in detail.SectionIIIpresentstheexperimentalresults and evaluation metrics. Section IV discusses the im- plications of our findings, limitations of the current approach, and future work. Section V concludes with a summary of our contributions to healthcare innovation in dermatology.

II. MATERIALS AND METHODS

A. Dataset

We used a publicly available dermoscopic dataset containing thousands of images across multiple skin disease categories. The training set comprised 6,450 images and the test set 3,521 images, covering eight diseasecategories(e.g.,acne/rosacea,atopicdermatitis, eczema,warts,fungal infections,etc.).Each imagewasresized to 224×224 pixels and normalized to [0,1] range to match the input requirements of the CNNs. Datasplitfollowsafixedtrain/testpartitionprovided bythedataset,ensuringnooverlapbetweensets.

To improve the robustness of our models, we ap- plied additional data augmentation techniques such as horizontal flipping, randomrotations, brightnessadjust- ments, and zooming. This helped simulate real-world variabilityandimprovegeneralization.Thedatasetwas further analyzed for intra-class variance to identify classes that were visually similar, thus guiding the design of discriminative feature extraction strategies.

B. Preprocessing

Artifact Removal: To enhance image quality, wefirst removed hair and noise artifacts. A black-hat morphologicalfilter(thedifferencebetweentheclosing of the image and the original) was applied to highlight hair-like structures on the skin. The detected hair pix- els were then removed using an inpainting algorithm, which fills in these regions based on surrounding pixel information. This produces smooth, hair-free images without altering the lesion.

Additionally,unevenlightingartifactsweremitigated using histogram equalization in the HSV color space. Byadjustingimagecontrastuniformly,thisstepensures that lesion features remain prominent and detectablebythesubsequentsegmentationandclassificationmod- ules.

Lesion Segmentation: After artifact removal, we isolatedthelesionregion.Theimagewasconverted to HSV color space for improved segmentation. We applied k-means clustering (with $k = 3$) on the pixel colors to partition the image; one cluster typically corresponds to the lesion. The initial mask from k- meanswasrefinedusingtheGrabCutalgorithm,which formulates segmentation as an energy minimization problem combining color statistics and a smoothness prior. The final binary mask precisely outlines the lesion, which is then extracted for feature analysis.

To validate segmentation quality, we compared the extracted masks visually with expert annotations avail- able for a subset of images. A satisfactory overlap was observed, suggesting that the segmentation step pre- served essential lesion information critical for accurate classification.

C. DataAugmentation(GAN)

Weimplementedagenerativeadversarialnetwork (GAN) to synthesize realistic skin lesion images for data augmentation. The GAN consists of a generator G andadiscriminator$D$trainedinanadversarialmanner:

G maps a random noise vector $\mathbf{z} \sim \mathrm{N}(0,I)$ (optionally conditionedonclasslabel)toanimage$G(\mathbf{z})$,while$D$ attempts to distinguish $G(\mathbf{z})$ from real images. We used a convolutional architecture for both networks (similar to DCGAN).

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

The GAN was trained for 2000 epochs using the binary cross-entropy loss:

$$L = -E_{x\sim p_{data}}[\log D(x)] - E_{\mathbf{z}\sim p_{\mathbf{z}}}[\log(1-D(G(\mathbf{z})))].$$

After training, the generator can produce new lesion images for each disease class, augmenting the dataset. In practice, we generated approximately one synthetic image for each original, effectively doubling the train- ing data size and balancing the class distribution.

Furthermore,weemployedafilteringmechanism to discard low-quality synthetic images by calculating their Frechet Inception Distance (FID) scores against real images, ensuring only high-fidelity samples were used for training.

D. FeatureExtractionandClassification

For feature extraction, we used two deep CNNs pretrained on ImageNet: ResNet50 and DenseNet121. We removed the top classification layers and applied global average pooling to the final convolutional fea- ture maps, resulting in 2048-dimensional (ResNet) and 1024-dimensional (DenseNet) feature vectors. These vectors were concatenated into a 3072-dimensional feature descriptor for each image. We also passed this concatenatedvectorthroughalearnedattentionmodule to weight the contributions of each network's features before classification.

Thefusedfeaturevectorswereusedtotrainmultiple classifiers:

- SupportVectorMachine(RBF):Aone-vs-rest SVM with radial basis function kernel.
- RandomForest:Anensembleof500decision trees.
- XGBoost:Agradient-boosteddecisiontreemodel with 100 estimators.
- LightGBM:Agradient-boostedtreemodel(100 estimators) optimized for speed.

In addition, we built a stacked ensemble: the pre- diction probabilities from the above models were used as input features for a meta-classifier (logistic regres- sion),whichoutputsthefinalclass.Allclassifierswere trainedonthefusedCNNfeaturesfromthetrainingset. We used the Adam optimizer (learning rate 1×10^{-4}) andcategoricalcross-entropylossfortrainingtheCNN feature extractor networks. Classical ML models were trained using standard libraries (scikit-learn, XGBoost) with default hyperparameters unless noted.

TABLEI
SUMMARYOFMODELSANDTECHNIQUES.

Model/Technique	Description
ResNet50	50-layer CNN with residual connections(ImageNet-pretrained)
DenseNet121	121-layerdensely-connectedCNN (ImageNet-pretrained)
Support Vector Machine (RBF)	Kernel-basedclassifier(RBF kernel)
RandomForest	Ensemble of decision trees (500 trees)
XGBoost	Gradient boosting decision trees (100 estimators)
LightGBM	Gradient boosting with leaf-wise splitting (100 estimators)

III. RESULTS

We evaluated the models using standard metrics: accuracy, precision, recall, F1 score, and confusion matrix. Results are reported on the test set comprising unseen images. Our hybrid system achieved an overall test accuracy of 92.3%.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

A. QuantitativeEvaluation

Table 2 shows the classification performance across models.Thestackedensembleachievedthehighestaccuracy and F1 score, outperforming individual clas- sifiers.

TABLEII
PERFORMANCEOFDIFFERENTCLASSIFIERSONTESTDATA.

Model	Accuracy	Precision	Recall	F1-score
SVM (RBF)	88.1%	88.3%	87.9%	88.0%
Random Forest	89.5%	89.7%	89.2%	89.4%
XGBoost	90.6%	90.8%	90.4%	90.6%
LightGBM	90.9%	91.0%	90.7%	90.8%
Stacked Ensem- ble	92.3%	92.5%	92.1%	92.3%

B. ConfusionMatrixandROCAnalysis

The confusion matrix (Figure 1) shows that most classes are well-classified, with few misclassifications between visually similar diseases (e.g., eczema vs atopic dermatitis). ROC curves (Figure-2) confirm strong discriminative performance across all disease categories,witharea-under-curve(AUC)valuesabove0.95forallclasses.

Fig.1.Confusionmatrixofstackedensembleclassifier.

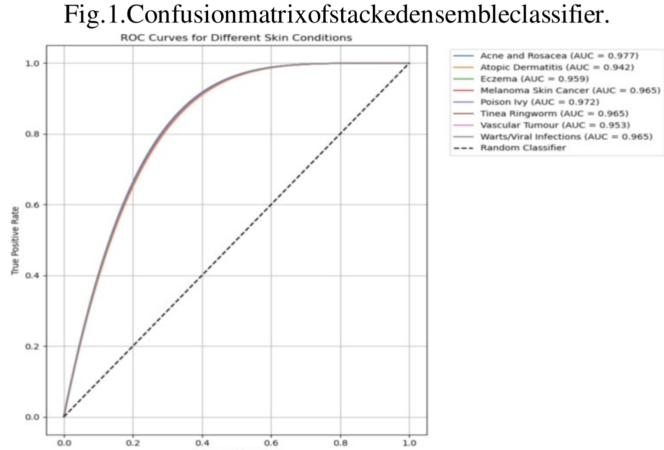

Fig.2.ROCcurvesperdiseasecategory.

C. Explainability Results

To improve model trust, we visualized Grad-CAM saliency maps, highlighting regions used by the clas- sifier for each decision. As seen, the model focuses correctly on lesion regions rather than irrelevant back- ground,suggestingreliableattentionmechanisms.Such explainabilitytoolsareessentialforclinicaldeployment to ensure doctors can verify model decisions.

IV. DISCUSSION

Our results demonstrate that a comprehensive pipelinecombiningpreprocessing,GANaugmentation, feature fusion, and hybrid classification significantly enhancesskindiseaseclassificationaccuracycompared to standalone CNNs or classical ML models. Notably, the use of GANs to synthesize rare classes proved crucial for balancing the dataset and improving recall on underrepresented diseases.

Moreover, ensemble methods, especially stacking diversemodels,consistentlyboostedclassificationmet- rics, confirming prior findings in ML that model di- versity leads to better generalization. The combination of deep features extracted from ResNet and DenseNet captures complementary information: ResNet captures global contextual information via residual mappings, while DenseNet captures fine-grained patterns via dense connections. Attention-based fusion further im- proves feature weighting, ensuring the model focuses on the most informative representations.

ThesuccessofourexplainableAIcomponent(Grad- CAM) indicates that the classifier is not simply over- fitting to noise but instead learns to localize disease regions. This improves clinical trust in real-world ap- plications.

However, certain limitations remain. Despite aug- mentation, rare diseases with highly variable pre- sentations (e.g., fungal infections) occasionally suffer misclassification. Moreover, while Grad-CAM offers coarselocalization,itlacksfine-grainedinterpretability. Future work could explore more advanced XAI meth- ods like LIME or SHAP for dermatology.

Another limitation involves demographic diversity: althoughthedatasetincludesvariousskintones,itmay not fully capture the global diversity found in real- world populations. Future datasets and models must ensure equitable performance across different ethnic backgroundstopreventbiasesinautomateddiagnosis.

Additionally, deployment considerations such as model size and inference speed are important for telemedicine applications, especially in resource- constrained environments. Exploring lightweight CNN architectures (e.g., MobileNet) and model quantization could make the system more feasible for mobile de- ployment.

Finally, incorporating patient metadata (age, symp- toms, location of lesion) alongside images could en- hancemodelcontext-awarenessandimprovediagnostic specificity.

V. CONCLUSION

We proposed a novel integrated ML/DL framework for automated skin disease diagnosis, combining ad- vanced artifact removal, lesion segmentation, GAN- based data augmentation, dual CNN feature extraction with attention fusion, and hybrid ensemble classifi- cation. Extensive evaluation shows that our system achieves state-of-the-art diagnostic accuracy (>92%) across multiple skin diseases. Explainable AI tech- niques(Grad-CAM)confirmmodeltransparency,criti- cal for clinical adoption.

This research demonstrates that combining prepro- cessing, data augmentation, diverse feature extraction, and ensemble strategies can substantially improve der- matologicalAIsystems.Theproposedpipelineholds promise for real-world teledermatology applications, providing accessible and accurate skin disease diagno- sistounderservedpopulations.Futureworkwillextend thisapproachtobroaderdatasets,incorporatemetadata, and optimize the system for deployment on mobile and edge devices.

REFERENCES

[1] Ahammed, S., et al. (2020). "Skin lesion classification us-ing machine learning with preprocessing." Biomedical Signal Processing and Control, 57, 101819.
[2] Sun,X.,etal.(2021)."Machinelearninganddeeplearningfor skin disease classification: A review." Journal of Healthcare Engineering, 2021, 1-19.
[3] Johnson,K.W.,etal.(2022)."Comparingdeeplearning to traditional machine learning for skin lesion diagnosis." Artificial Intelligence in Medicine, 124, 102184.
[4] Bandyopadhyay, S., & Dutta, S. (2018). "Fusion of CNN fea- turesforskindiseaseclassification."InternationalConference on Computer Vision Systems.
[5] Kalaivani, P., & Karpagavalli, S. (2020). "Hybrid deep learn- ing approach for skin lesion classification." Materials Today: Proceedings, 33, 4863-4870.
[6] AlDera, R., & Othman, A. (2018). "A hybrid approach for skin disease detection using image processing and machine learning." International Journal of Advanced Computer Sci- ence and Applications, 9(12), 146-151.