A Multimodal AI System integrating image and textual inputs for skin disease classification.

Tisha Jeswani
Computer Department
Vivekanand Education Society's Institute of Technology
Mumbai, India
2021.tisha.jeswani@ves.ac.in

Varsha Makhija
Computer Department
Vivekanand Education Society's Institute of Technology
Mumbai, India
2021.varsha.makhija@ves.ac.in

Dinky Khatri
Computer Department
Vivekanand Education Society's Institute of Technology
Mumbai, India
2021.dinky.khatri@ves.ac.in

Jiya Lund
Computer Department
Vivekanand Education Society's Institute of Technology
Mumbai, India
2021.jiya.lund@ves.ac.in

Mrs. Abha Tewari
Assistant Professor
Computer Department
Vivekanand Education Society's Institute of Technology
Mumbai-40074, India
abha.tewari@ves.ac.in

***Abstract*—*Skin disorders impact a significant portion of the global population, making early and precise diagnosis essential to prevent severe health issues. Traditional diagnostic techniques, often dependent on physical exams and lab tests, are not always accessible in remote or under-resourced areas. To overcome these limitations, this study proposes an AI-based solution that merges machine learning and deep learning approaches for efficient diagnosis. Users can input symptoms using four modalities: textual entries, voice input, live camera capture, or uploaded images, which are then used for segmentation and classification. The system leverages a Random Forest classifier for text-based predictions, CNNs for image classification, and U-Net for image segmentation. Infection severity is determined by calculating the proportion of infected pixels, classifying skin conditions into severe, moderate, or mild, and tagging unaffected cases as healthy. Through the integration of NLP, deep learning, and speech recognition, the model supports an inclusive, interactive, and accessible diagnostic system. Its multimodal data handling significantly boosts diagnostic accuracy and usability, particularly for populations with limited access to dermatology services.*

Keywords—Skin Disease Diagnosis, Machine Learning (ML), Deep Learning (DL), Random Forest Classifier, U-Net Segmentation,Skin Lesion Detection, Automated Diagnosis

I . Introduction

Millions suffer from skin conditions like acne, psoriasis, eczema, and melanoma, where early detection is vital to managing symptoms and avoiding complications [7]. However, in many rural or underserved regions, diagnostic tools such as dermoscopy, biopsies, and lab facilities are not readily available due to a lack of trained dermatologists and infrastructure [11]. In response, our solution, WellnessInsight, presents a comprehensive AI-powered platform that incorporates ML and DL techniques for skin disease classification [6]. It allows individuals to report symptoms via text, voice, live images, or uploaded images. A Random Forest model handles textual and voice-based diagnosis [12], while CNNs and U-Net are used for processing and analyzing images. Additionally , the system also calculates the infection percentage through analysing the number of infected and non infected pixels , which inturn also tells the severity levels or confirming healthy skin if no disease is detected. This multifaceted AI-driven model, combining NLP, voice recognition, and computer vision, offers an efficient, accessible diagnostic alternative, particularly tailored for healthcare delivery in isolated communities [19]. We mainly focused on three diseases which includes Eczema, Melanoma and Psoriasis

Ⅱ. Literature Review

A. Survey on Skin Lesion Detection and Classification using Machine Learning:

This paper basically focuses on how machine learning can give better review or tells about the disease instead of looking from the eyes [1].Sometimes the image might appear different due to various factors which leads to improper classification so this paper suggests also to use the medical history of a person. Also the transfer learning is used in this system for performing better and extracting important features [19].

B. Machine Learning Based Skin Disease Detection using Semi-Automatic Image Segmentation:

The paper discusses the weakness of detecting skin diseases by hands and also suggests the ways to improve it. It recommends using a ML based method to increase the accuracy of skin disease detection [2]. For classification, K-Nearest Neighbors (KNN) is used and for feature extraction, Gray-Level Co-occurrence Matrix (GLCM) is used which gives higher accuracy than the standard techniques. This is tested on multiple datasets and gives faster and more reliable skin condition analysis.

C. Image Segmentation based Approach for Skin Disease Detection and Classification using Machine Learning Algorithms:

This paper talks about how some medical machines are so costly so study shares are cost friendly to find skin problems using processing of images. It takes an image , changes the size of image accordingly and then uses CNN and SVM which helps to decide which skin disease it is. This method can find 4 types of skin disease.

D. Machine Learning and Cloud-based Mobile App for Real Time Skin Cancer Prediction:

This study looks at how to use machine learning to find skin cancer. It uses a method called DWT to pick important parts of skin then it tests three tools - LDA , Random forest , naive bayes to see which one works best. They also made a mobile app to help with easy checkup. Naive Bayes turns up to be best in finding skin cancer correctly. It tested a total of 2,357 skin pictures.

E. Lumpy Skin Disease Virus Detection on Animals Through Machine Learning Method:

This research explores the use of machine learning to detect Lumpy Skin Disease Virus (LSDV) in cattle. By working with a large dataset of 24,803 cases gathered over 16 years, the study integrates factors such as weather conditions and geographic information. Techniques like Support Vector Classifier (SVC), XGBoost, and Random Forest [5] are applied to improve detection accuracy, offering a more efficient alternative to traditional diagnostic methods.

F. Skin Disease Classification using Dermoscopy Images through Deep Feature Learning Models and Machine Learning Classifiers:

This study uses CNNs like VGG16, VGG19, and Inception V3 to extract features from dermoscopy images, followed by ML classifiers like SVM and Random Forest [6].

G.Skin Cancer Detection and Classification Using Machine Learning:

The study presents a machine learning-based approach for early detection and classification of skin cancer, including melanoma, using dermoscopic images [7]. The study employs image processing techniques like the Dull Razor method, Gaussian filter, and K-means clustering for preprocessing and segmentation. Feature extraction is performed using GLCM, while classification is carried out using Multi-class Support Vector Machine (MSVM).

H.Improving AI-Based Skin Disease Classification with StyleGAN3 for Minority Skin Tone Generation:

The study enhances skin disease classification by addressing the underrepresentation of minority skin tones in AI models [8]. The Fitzpatrick 17k dataset, focusing on Acne, Psoriasis, and Vitiligo, is utilized, with VGG16 applied for classification. The study compares models trained on original, traditionally augmented, and StyleGAN3-augmented datasets.

I.Deep Learning in Dermatology: Convolutional Neural Network-Based Classification of Skin Diseases and Cancer:

This study uses convolutional neural networks (CNNs) to classify 57 types of skin diseases and cancer [10]. A custom CNN model outperforms pre-trained models, achieving 96.64% accuracy. This research exhibits the accuracy of detecting skin disease using deep learning and also reduces the dependency to physically visit the dermatologists.

J.A Transfer Learning-based Pre-trained VGG16 Model for Skin Disease Classification:
The paper discusses the use of the VCG16 model to detect the skin diseases which gives the accuracy of 90.1%. The dataset is collected through Kaggle [8]. This shows the importance of deep learning and transfer learning to enhance the accuracy of the detection of skin diseases.

K.Deep Learning-Based Skin Diseases Classification using Smartphones:
This study uses smartphones to capture the images which is further used to detect the disease. Deep learning is used to classify skin diseases [14]. For the extraction of features and the classification of images, CNN is used.Datasets are used from diverse sources to increase the accuracy and generalization.

L.Intelligent Skin Disease Prediction System using Transfer Learning and Explainable Artificial Intelligence:
This study uses smart computer models like VGG16 and explainable AI to tell disease like chickenpox and monkeypox, it works really well, about 93 percent correct answers are identified and it helps doctor to understand easily how results were found.

M.Survival Outcomes of a Large Cohort of Acral Melanoma Patients Treated at a South:
This research talks about how the people in the south region who are being affected with rare acral cancer are surviving due to treatment at hospitals.

N.Application of Convolutional Neural Networks in Skin Disease Diagnosis:
This study explores CNNs like ResNet, VGG, and EfficientNet for classifying skin diseases using medical images [14]. It emphasizes the need for large datasets, data augmentation, and transfer learning. Explainable AI is suggested to enhance model transparency and trust.

O.Skin Disease Prediction using the MobileNet Approach:
This study uses MobileNet with transfer learning and data augmentation, achieving 94.4% accuracy in classifying eight skin categories [17]. It emphasizes the need for automated diagnosis in rural areas and proposes a mobile/web app for image-based classification.

P. Integration of Multimodal Data Sources for Enhanced Skin Disease Diagnosis:
This study combines images, text, and other data using ML and DL techniques like CNNs and NLP to improve diagnostic accuracy [20]. The multimodal approach enhances early detection and decision-making, especially in underserved areas, benefiting telemedicine and clinical use.

Table 1. Comparison of existing research for skin disease analysis

Project/App Name	Key Features	Model/Method Used	Input Type (Image, Text, Voice)	Limitations & Challenges	Solutions Provided
SkinLesNet	- Classifies 3 skin lesions (melanoma, nevus, seborrheic keratosis) - Trained on smartphone images (PAD-UFES-20, HAM10000, ISIC2017) - 96% accuracy (outperforms ResNet50 &	CNN (4-layer deep model) Compared with ResNet50 & VGG16	Image (Upload Only)	- No segmentation for infection percentage detection - No live camera or voice input - Not an app, only a research model on Google Colab - Needs clinical validation before real-world use	- Uses U-Net for segmentation to detect infection percentage - Supports live camera input & voice-based text entry - Developed as a full system (not just a research model) - Uses the ISIC dataset, ensuring real-world compatibility

	VGG16)				
Eczema Area and Severity Index (EASI) – DERMNET	- Manual scoring system for eczema severity - Body divided into 4 regions for area estimation - Uses clinician judgment	Clinical Scoring (Manual Assessment)	Visual Inspection by Dermatologist	- Subjective, depends on doctor's assessment - No AI-based automation - Only for eczema, not general skin diseases	- Automates infection area calculation using U-Net segmentation - Works for multiple skin diseases, not just eczema - Eliminates manual scoring bias with AI-based measurement
Skin Image Analysis for Quantitative Assessment – BMC MED INFORM DECIS MAK	- AI-based image processing for skin disease detection - Segmentation techniques to estimate affected area - Tracks disease progression over time	Deep Learning (CNN, U-Net, DeepLabv3+ Models)	Image Upload (AI-based analysis)	- Still under research, not clinically validated - Accuracy depends on dataset quality	- Enhances segmentation using U-Net for precise infection area calculation - Integrates multiple input methods (image upload, live camera, text, voice) - Uses a well-established dataset (ISIC) for higher accuracy - Provides a practical application rather than just a research model
AI Dermatologist: Skin Scanner	- AI-driven skin condition analysis- Identifies various skin issues including moles, angiomas, warts, and papillomas- Provides personalized guidance and information	CNN	Image Upload (via smartphone camera)	- Requires high-quality images for accurate analysis- Does not offer infection percentage calculation- Limited to predefined skin conditions	- Incorporates U-Net for precise infection area segmentation- Supports live camera input for real-time analysis- Calculates and displays infection percentage to users
Skinive AI	- AI-based skin health tracking- Risk assessment for multiple skin conditions- Compatible with web and mobile platforms	CNN	Image Upload (AI-based analysis)	- No segmentation for affected area analysis- Cannot calculate infection percentage- Lacks detailed medical diagnostics	- Enhances segmentation using U-Net for precise infection area calculation- Integrates multiple input methods (image upload, live camera, text, voice)- Uses a well-established dataset (ISIC) for higher accuracy

III. Materials

Disease-Specific Dataset (DermNet)

For the classification of skin diseases, images of eczema, psoriasis, and melanoma were sourced from the **DermNet** dataset, containing a total of 1114 images for eczema, 1175 for psoriasis, and 831 for melanoma. To enhance model performance, the melanoma dataset was expanded to 1996 images using data augmentation.

Eczema Melanoma Psoriasis

Figure 1. Skin disease samples from DermNet.

Normal Skin Dataset

The **Oily, Dry, and Normal Skin Types Dataset** from Kaggle was used to collect normal (non-infected) skin images. Only the Normal folder from this dataset was considered to ensure a clear distinction between healthy and diseased skin.

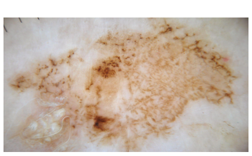
Normal Normal

Figure 2. Sample images of normal skin.

ISIC Dataset for Infection Percentage Calculation

The **ISIC 2016 dataset** was incorporated to calculate the infection percentage for skin diseases. It consists of two Folders: the Training Data folder, containing 900 images, and the Training Ground Truth folder, which includes 899 corresponding segmentation masks. These masks assist in identifying affected areas in eczema, psoriasis, and melanoma, aiding in infection percentage analysis.

Table 2. Example from the ISIC 2016 dataset showing a Training Data image and its corresponding ground truth mask.

Training Data Image	Ground Truth Mask
	

Text Dataset

Additionally, a **Text Dataset** from Kaggle was incorporated to enhance classification by leveraging textual descriptions of skin diseases. This dataset aids in mapping symptoms to diseases, improving classification when users input descriptions via text or voice

IV. Methodology

The proposed methodology is shown in Fig. 1

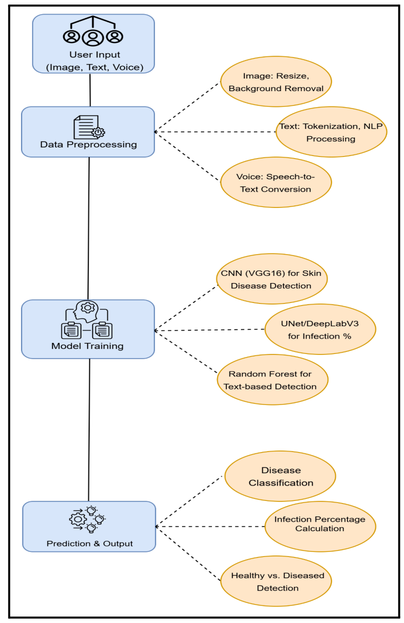


Figure 3. Workflow of the skin disease classification model.

The skin disease detection system utilizes multiple datasets, including ISIC for infection percentage detection, a healthy skin dataset for better classification, and a 2.8 GB text from Kaggle for symptom-based diagnosis.

Pre-processing techniques were applied to enhance image quality, including background removal, resizing, normalization, and noise reduction using Gaussian and Median Filtering. Text data underwent tokenization, normalization, and feature extraction, while voice data was converted to text and preprocessed using NLP techniques.

Lesion segmentation was performed using thresholding and contour detection to refine boundaries. Feature extraction focused on color, texture, and shape features to improve classification accuracy.

For image-based disease detection, a CNN with pretrained VGG16 was used, achieving 81% accuracy. Infection percentage estimation employed segmentation models like UNet and DeepLabV3, though version control issues arose with the latter. Classification was also implemented to distinguish healthy from diseased skin.

Symptom-based diagnosis used a Random Forest classifier for text input, while voice data was converted to text and processed before classification. The system predicts the disease category along with an estimated infection percentage.

V. Model Training Performance

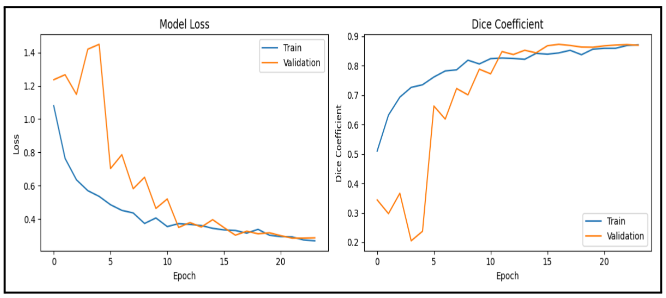


Figure 4. Model Training

Figure 4 presents how the model performed during training. As seen on the left, both training and validation loss decreased steadily, showing that the model was learning well without overfitting. On the right, the Dice Coefficient increased consistently, reaching close to **0.9**, which suggests the model was highly accurate in identifying and segmenting the affected skin areas.

VI. Future Scope

Future developments for WellnessInsight will use sophisticated deep learning models for illness classification, such as ResNet or EfficientNet, to improve the precision of skin disease detection. To assure fairness, the size of the dataset should be increased to have a greater variety in all the aspects such as skin tones and uncommon skin disorders. The improvements in various technologies like NLP and speech recognition can increase the text and voice-based symptom accuracy [19]. By expanding this to a mobile application can help the users to access the platform through their cell phones

V II. Conclusion

This WellnessInsight System provides a rare and different approach for predicting the skin diseases and also its percentage using various machine learning and deep learning techniques. The major advantage of this system includes the ways of input it is taking, as every possible input method is implemented in the system like real time image capturing,browsing images,input through text and also voice. For the text analysis and voice input we have used the random forest classifier , image classification includes CNN model and for calculating infection percentage we have used UNet model

One of the benefits the system incorporates is that it bridges the gap between healthcare centres and online healthcare applications. This system will be beneficial for the ones who are living in local areas where there is limited access to healthcare centres which leads to delay resulting in severe crises.

VIII. References

[1] N. Nirupama and Virupakshappa, "Survey on Classification of Skin Diseases Using Machine Learning Techniques," *2024 3rd Int. Conf. PARC*, Mathura, India, 2024, pp. 135-140.

[2] J. Sathya and A. Kalaivani, "Machine Learning Based Skin Disease Detection Using Semi-Automatic Image Segmentation," *2024 Int. Conf. ICSSEECC*, Coimbatore, India, 2024, pp. 215-219.

[3] N. H. Sany and P. C. Shill, "Image Segmentation based Approach for Skin Disease Detection and Classification using Machine Learning Algorithms," *2024 Int. Conf. ICICACS*, Raichur, India, 2024, pp. 1-5.

[4] R. K. Kaushal et al., "Machine Learning and Cloud-based Mobile App for Real Time Skin Cancer Prediction," *2023 7th Int. Conf. I-SMAC*, Kirtipur, Nepal, 2023, pp. 629-636.

[5] D. H. Patil et al., "Lumpy Skin Disease Prediction Using Machine Learning," *2023 4th IEEE GCAT*, Bangalore, India, 2023, pp. 1-5.

[6] S. Gupta, A. Panwar and K. Mishra, "Skin Disease Classification using Dermoscopy Images through Deep Feature Learning Models and Machine Learning Classifiers," *2021 IEEE EUROCON*, Lviv, Ukraine, 2021.

[7] A. Na, "Improving Convolutional Neural Networks Diagnostic Ability for Malignant Skin Diseases," *2024*

Int. Conf. ICECCE, Kuala Lumpur, Malaysia, 2024, pp. 1-6.

[8] G. Singh, K. Guleria and S. Sharma, "A Transfer Learning-based Pre-trained VGG16 Model for Skin Disease Classification," *2023 IEEE MysuruCon*, Hassan, India, 2023, pp. 1-6.

[9] P. K. G et al., "Enhanced Skin Disease Classification Using Modified ResNet50 Architecture," *2024 Int. Conf. ISENSE*, Kottayam, India, 2024, pp. 1-5.

[10] M. M. Ali et al., "Deep Learning in Dermatology: Convolutional Neural Network-Based Classification of Skin Diseases and Cancer," *2024 Int. Conf. MAC*, Dehradun, India, 2024, pp. 1-7.

[11] S. Murugan et al., "A Machine Learning Approach to Predict Skin Diseases and Treatment Recommendation System," *2023 5th Int. Conf. ICSSIT*, Tirunelveli, India, 2023, pp. 1157-1163.

[12] B. P T, P. G K and M. S, "Skin Disease Prediction Using Machine Learning Techniques," *2023 3rd Int. Conf. ICMNWC*, Tumkur, India, 2023, pp. 1-6.

[13] M. Pal and B. R. Roy, "Evaluating and Enhancing the Performance of Skin Disease Classification Based on Ensemble Methods," *2020 2nd Int. Conf. ICAICT*, Dhaka, Bangladesh, 2020, pp. 439-443.

[14] S. Verma, A. Kumar and M. Kumar, "Skin disease detection using Decision Tree and Support Vector Machine," *2024 1st Int. Conf. ACET*, Ghaziabad, India, 2024, pp. 1-6.

[15] K. S. Reddy et al., "Implementation of InceptionV3 Model for Skin Disease Detection and Classification," *2025 6th Int. Conf. ICMCSI*, Goathgaun, Nepal, 2025, pp. 1479-1485.

[16] M. K, S. S and T. E, "Streamlit-Powered Comprehensive Health Analysis and Disease Prediction System," *2023 Int. Conf. ICERCS*, Coimbatore, India, 2023, pp. 1-7.

[17] D. V. V. Rani, G. Vasavi and B. Maram, "Skin Disease Classification Using Machine Learning and Data Mining Algorithms," *2022 IEEE iSSSC*, Gunupur, India, 2022, pp. 1-6.

[18] P. Banditsingha et al., "A Decision Machine Learning Support System for Human Skin Disease Classifier," *2022 ECTI DAMT & NCON*, Chiang Rai, Thailand, 2022, pp. 200-204.

[19] J. Lin et al., "Automatic Classification of Clinical Skin Disease Images with Additional High-Level Position Information," *2019 Chinese Control Conf. (CCC)*, Guangzhou, China, 2019, pp. 8606-8610.

[20] T. Indupalli et al., "A Novel Deep-Learning Based Classification of Skin Diseases," *2024 5th IEEE GCAT*, Bangalore, India, 2024, pp. 1-6.