

VIVEKANAND EDUCATION SOCIETY'S INSTITUTE OF TECHNOLOGY
An Autonomous Institute Affiliated to University of Mumbai
Department of Computer Engineering



Project Report on
WellnessInsight

In partial fulfillment of the Fourth Year, Bachelor of Engineering (B.E.) Degree in Computer Engineering at the University of Mumbai Academic Year 2024-25

Submitted by
Jiya Lund (D17A , Roll no - 37)
Tisha Jeswani (D17B , Roll no - 20)
Varsha Makhija (D17C , Roll no - 43)
Dinky Khatri(D17A , Roll no - 31)

Project Mentor
Mrs. Abha Tewari

(2024-25)

VIVEKANAND EDUCATION SOCIETY'S INSTITUTE OF TECHNOLOGY
An Autonomous Institute Affiliated to University of Mumbai
Department of Computer Engineering



Certificate

This is to certify that **Jiya Lund, Varsha Makhija, Tisha Jeswani, Dinky Khatri** of Fourth Year Computer Engineering studying under the University of Mumbai have satisfactorily completed the project on “**WellnessInsight**” as a part of their coursework of PROJECT-II for Semester-VIII under the guidance of their mentor **Mrs. Abha Tewari** in the year 2024-25 .

This thesis/dissertation/project report entitled WellnessInsight by (**Author Name**) is approved for the degree of _____ (**Degree details**).

Programme Outcomes	Grade
PO1,PO2,PO3,PO4,PO5,PO6,PO7, PO8, PO9, PO10, PO11, PO12 PSO1, PSO2	

Date:

Project Guide:

Project Report Approval

For

B. E (Computer Engineering)

This thesis/dissertation/project report entitled ***WellnessInsight*** by ***Jiya Lund, Tisha Jeswani, Varsha Makhija, Dinky Khatri*** is approved for the degree of _____
(Degree details).

Internal Examiner

External Examiner

Head of the Department

Principal

Date:
Place:

Declaration

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

(Signature)

(Jiya Lund,D17A,37)

(Signature)

(Varsha Makhija,D17C,43)

(Signature)

(Tisha Jeswani,D17B,20)

(Signature)

(Dinky Khatri,D17A,31)

Date:

ACKNOWLEDGEMENT

We are thankful to our college Vivekanand Education Society's Institute of Technology for considering our project and extending help at all stages needed during our work of collecting information regarding the project.

It gives us immense pleasure to express our deep and sincere gratitude to Assistant Professor (**Project Guide**) for her kind help and valuable advice during the development of project synopsis and for her guidance and suggestions.

We are deeply indebted to Head of the Computer Department **Dr.(Mrs.) Nupur Giri** and our Principal **Dr. (Mrs.) J.M. Nair**, for giving us this valuable opportunity to do this project.

We express our hearty thanks to them for their assistance without which it would have been difficult in finishing this project synopsis and project review successfully.

We convey our deep sense of gratitude to all teaching and non-teaching staff for their constant encouragement, support and selfless help throughout the project work. It is great pleasure to acknowledge the help and suggestion, which we received from the Department of Computer Engineering.

We wish to express our profound thanks to all those who helped us in gathering information about the project. Our families too have provided moral support and encouragement several times.

Computer Engineering Department
COURSE OUTCOMES FOR B.E PROJECT

Learners will be to,

Course Outcome	Description of the Course Outcome
CO 1	Able to apply the relevant engineering concepts, knowledge and skills towards the project.
CO2	Able to identify, formulate and interpret the various relevant research papers and to determine the problem.
CO 3	Able to apply the engineering concepts towards designing solutions for the problem.
CO 4	Able to interpret the data and datasets to be utilized.
CO 5	Able to create, select and apply appropriate technologies, techniques, resources and tools for the project.
CO 6	Able to apply ethical, professional policies and principles towards societal, environmental, safety and cultural benefit.
CO 7	Able to function effectively as an individual, and as a member of a team, allocating roles with clear lines of responsibility and accountability.
CO 8	Able to write effective reports, design documents and make effective presentations.
CO 9	Able to apply engineering and management principles to the project as a team member.
CO 10	Able to apply the project domain knowledge to sharpen one's competency.
CO 11	Able to develop professional, presentational, balanced and structured approach towards project development.
CO 12	Able to adopt skills, languages, environment and platforms for creating innovative solutions for the project.

Index

Chapter No.	Title	Page No.
	Abstract	1
1	Introduction	2-3
1.1	Introduction of the project	2
1.2	Motivation for the project	2
1.3	Problem Definition	2
1.4	Existing Systems	3
1.5	Lacuna of the existing systems	3
1.6	Relevance of the Project	3
2	Literature Survey	4-11
A	Brief Overview of Literature Survey	4
B	Related Works	4-9
2.1	Inference drawn	4-9
2.2	Comparison with the existing system	9-11
3	Requirement Gathering for the Proposed System	12-14
3.1	Introduction to requirement gathering	12

3.2	Functional Requirements	12
3.3	Non-Functional Requirements	13
3.4	Hardware, Software , Technology and tools utilized	13
3.5	Constraints	14
4	Proposed Design	15-19
4.1	Block diagram of the system	15-16
4.2	Modular design of the system	16-17
4.3	Detailed Design	18
4.4	Project Scheduling & Tracking using Gantt Chart	19
5	Implementation of the Proposed System	20-21
5.1	Methodology employed for development	20
5.2	Algorithms and flowcharts for the respective modules developed	20-21
5.3	Datasets source and utilization	21
6	Results and Discussion	22-26
6.1	Screenshots of User Interface (UI) for the respective module	22-24
6.2	Performance Evaluation Measures	24-25

6.3	Input Parameters / Features considered	25
6.4	Graphical or Statistical Input	25
6.5	Comparison of results with existing systems	26
7	Conclusion	27-28
7.1	Limitations	27
7.2	Conclusion	27
7.3	Future Scope	28
	References	28-29
1	Appendix	30-37
a)	Submitted paper	30-35
b)	Plagiarism report	36
c)	Project review sheet	37

List of Figures

Figure No.	Heading	Page No.
4.1	Block Diagram	15-16
4.2	Modular Design of the system	16-17
4.3.1	Flowchart	18
4.4	Gantt Chart	19
6.1.1	Screenshots of User Interface (UI) for the respective module	22
6.1.2	Displayed the Predicted disease, Infection percentage and severity of disease	22
6.1.3	Option to click real time images for disease detection.	23
6.1.4	Result is displayed based on the image given.	23
6.1.5	Option for the user to describe the problem in words or to speak about symptoms.	24
6.4	Pie Chart Showing Estimated Prevalence of Skin Diseases in India	25

List of Tables

Table no.	Heading	Page no.
2.2	Comparison with the Existing System	9-11
6.5	Comparison of results with the Existing Systems	26

Abstract

Skin diseases affect millions worldwide, necessitating early and accurate diagnosis to prevent complications. However, traditional diagnostic methods rely on physical examinations and laboratory tests, which may be inaccessible in rural and resource-constrained areas. To address this challenge, an AI-driven approach is proposed, integrating machine learning and deep learning techniques to enhance diagnostic efficiency. Users can provide symptom input through four modalities: textual descriptions, voice input, real-time image capture, and image uploads for segmentation and classification. A Random Forest classifier is utilized for disease prediction, while Convolutional Neural Networks (CNNs) are employed for image-based classification and U-Net for segmentation. Additionally, infection severity is assessed by calculating the infection percentage using pixel-based analysis, classifying diseases as severe, moderate, or normal, with an additional category for healthy skin in cases where no disease is detected. The incorporation of deep learning, natural language processing (NLP), and voice recognition ensures an intuitive and inclusive diagnostic process. By leveraging multiple data modalities, this approach enhances classification accuracy and accessibility, offering a user-friendly and automated diagnostic tool. The system is particularly beneficial for remote areas with limited dermatological resources, aiming to bridge healthcare disparities and improve early disease detection.

Chapter 1: Introduction

1.1 Introduction

Skin diseases such as acne, eczema, melanoma, and psoriasis affect millions of people globally. Early and accurate diagnosis is crucial to prevent complications, yet traditional diagnostic methods—clinical examinations, histopathology tests, and dermoscopy—are often inaccessible in rural areas due to the shortage of dermatologists and healthcare infrastructure. These methods are also time-consuming, expensive, and reliant on clinician expertise. To address these challenges, WellnessInsight introduces a multi-modal AI-driven skin disease diagnosis system, integrating machine learning (ML) and deep learning (DL) for enhanced accessibility and accuracy.

1.2 Motivation

The primary motivation behind this project stems from the increasing burden of skin diseases and the lack of accessible diagnostic solutions, particularly in underserved regions. Many individuals either delay or completely forgo dermatological consultation due to geographical constraints, financial limitations, or lack of awareness. The integration of AI can bridge this gap, providing a fast, cost-effective, and automated diagnostic solution that ensures early detection and intervention, reducing the risk of complications and improving patient outcomes.

1.3 Problem Definition

The current dermatological diagnostic methods pose several challenges:

- Limited accessibility: Many rural and remote areas lack dermatologists and specialized healthcare centers.
- Time-consuming process: Traditional diagnosis often involves multiple tests, leading to delays in treatment.
- Cost implications: Histopathology and dermoscopic analyses are expensive, making them unaffordable for many patients.
- Expert dependency: Diagnosis accuracy heavily relies on clinician expertise, leading to variability in results.

To address these issues, this project aims to develop an AI-powered multi-modal skin disease diagnosis system that can process textual symptoms, voice inputs, real-time images, and uploaded images to provide an accurate and automated diagnosis.

1.4 Existing Systems

Several approaches exist for diagnosing skin diseases, including:

1. Clinical Examinations – Performed by dermatologists through visual inspection and medical history analysis.
2. Dermoscopy – Uses specialized imaging devices to examine skin lesions in detail.
3. Histopathology Tests – Biopsy-based analysis of skin tissues under a microscope.
4. Mobile Applications – AI-driven apps that analyze skin images for possible conditions.
5. Online Symptom Checkers – Web-based tools that provide diagnostic probabilities based on self-reported symptoms.

1.5 Lacuna of the Existing Systems

Despite advancements, existing methods have notable shortcomings:

- High Cost & Time Consumption – Histopathology and dermoscopy require specialized equipment and trained professionals.
- Inconsistent Accuracy – Many mobile applications and online symptom checkers provide unreliable or generalized results.
- Lack of Multi-Modality – Most AI-driven tools focus solely on image analysis, ignoring textual or voice-based symptom inputs.
- Inaccessibility in Remote Areas – Traditional dermatological services remain largely unavailable in underserved regions.

1.6 Relevance of the Project

This project is highly relevant as it provides a comprehensive AI-driven solution for skin disease diagnosis by integrating machine learning, deep learning, NLP, and voice recognition technologies. The WellnessInsight system addresses the gaps in traditional and existing AI-based solutions by:

- Enabling multi-modal inputs (text, voice, real-time image capture, and image uploads).
- Utilizing Random Forest for symptom-based prediction and CNNs with U-Net for image classification and segmentation.
- Implementing pixel-based infection percentage estimation to categorize disease severity.
- Including a healthy skin classification to distinguish between diseased and normal skin.
- Enhancing accessibility, especially for populations with limited healthcare infrastructure.

By leveraging AI-driven automation, WellnessInsight ensures early, affordable, and accurate skin disease diagnosis, making it a valuable tool in modern dermatology and public healthcare.

Chapter 2: Literature Survey

A. Brief Overview of Literature Survey

The field of dermatological diagnostics has seen a significant shift with the integration of Machine Learning (ML), Deep Learning (DL), and computer vision techniques. The literature presents numerous methodologies addressing automated skin disease detection, classification, segmentation, and mobile-based applications. These methods aim to enhance diagnostic accuracy, increase accessibility, and reduce dependence on expert dermatologists. This chapter surveys recent studies focusing on skin lesion detection using CNNs, transfer learning, generative models, multimodal approaches, and smartphone-based tools. The following sections analyze relevant research papers, patents, and existing systems compared to our proposed *WellnessInsight* framework.

B. Related Works

2.1 Research Papers Referred

1. Survey on Skin Lesion Detection and Classification using Machine Learning

This paper investigates the enhancement of skin lesion identification through machine learning approaches, which outperform visual inspection in accuracy. It emphasizes the integration of clinical and genomic data and highlights the role of transfer learning to improve generalization. Furthermore, the study explores autoencoders and GANs to enrich feature extraction and augment limited datasets. The authors identify real-world validation and dataset imbalance as significant challenges, proposing hybrid learning strategies as solutions. The survey acts as a foundation for developing automated, accurate diagnostic systems in dermatology.

Inference drawn:

- Machine learning models show significantly improved accuracy compared to traditional diagnoses.
- Transfer learning and data augmentation are essential for handling limited and imbalanced dermatological datasets.
- Integrating clinical/genomic data enhances real-world applicability.

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

This study addresses the limitations of manual skin disease diagnosis by proposing a machine learning-based approach that increases diagnostic precision. The model uses Gray-Level Co-occurrence Matrix (GLCM) for texture feature extraction and applies the K-Nearest Neighbors (KNN) algorithm for classification. Tested on multiple datasets, it demonstrates a significant improvement in accuracy compared to traditional methods.

The semi-automatic process reduces human intervention while increasing speed and consistency. The paper shows potential for real-time dermatology support systems.

Inference drawn:

- The proposed ML model achieves higher accuracy than manual diagnosis methods.
- Semi-automatic segmentation reduces dependency on expert input.
- KNN and GLCM offer efficient classification and feature extraction.

3. Image Segmentation-based Approach for Skin Disease Detection and Classification using Machine Learning Algorithms

The paper proposes a cost-effective image processing technique for skin disease detection, minimizing reliance on expensive diagnostic equipment. It utilizes pre-trained CNNs for feature extraction, followed by multiclass SVM for classification. Digital images are resized and preprocessed using filters like GLCM, Gabor, and Canny Edge Detection. The method identifies four different skin diseases with an accuracy of 90.8%. The study highlights the efficiency of advanced segmentation and feature extraction techniques for low-cost diagnostic solutions.

Inference drawn:

- Pre-trained CNNs combined with SVM effectively classify skin diseases with high accuracy.
- Advanced segmentation techniques improve feature richness and detection accuracy.
- The approach is suitable for low-cost diagnostic solutions.

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

This study presents a mobile-based application integrated with machine learning algorithms for skin cancer detection. Discrete Wavelet Transform (DWT) is used for extracting features, while classifiers such as LDA, Random Forest, and Naïve Bayes are assessed. The model is tested on 2,357 dermatoscopic images, with Naïve Bayes achieving the highest accuracy of 96.61%. The cloud-based infrastructure allows real-time access and predictions. The research underscores the feasibility of deploying portable diagnostic tools for early cancer detection.

Inference drawn:

- The mobile app provides accurate and accessible cancer detection using ML.
- Naïve Bayes classifier performed best, achieving over 96% accuracy.
- Cloud-based systems facilitate remote, real-time diagnoses.

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

Focusing on veterinary health, this paper applies machine learning to detect Lumpy Skin Disease Virus

(LSDV) in cattle using geospatial, temporal, and weather-related data. A dataset of 24,803 cases across 16 years is analyzed with classifiers such as Support Vector Classifier (SVC), Random Forest, and XGBoost. The approach aims to outperform traditional epidemiological surveillance methods. It demonstrates how ML can be adapted to infectious disease detection in animals. The study opens possibilities for similar methodologies in human dermatology.

Inference drawn:

- ML models outperform traditional detection techniques for LSDV.
- Integration of geospatial and temporal data enhances prediction reliability.
- The approach shows adaptability of ML in both human and veterinary dermatology.

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

This research explores deep learning models like VGG16, VGG19, and Inception V3 for feature extraction from dermoscopy images. It further combines these features with machine learning classifiers like SVM and Random Forest to categorize skin lesions. The model achieves high accuracy in distinguishing between benign and malignant cases. The dataset used is diverse, enhancing model generalization. This study demonstrates the efficacy of hybrid deep and traditional ML systems in clinical dermatology.

Inference drawn:

- Deep CNNs combined with ML classifiers offer robust performance in lesion classification.
- Effective in distinguishing between benign and malignant conditions.
- Feature fusion from multiple CNN models improves diagnostic accuracy.

7. Skin Cancer Detection and Classification Using Machine Learning

The paper focuses on skin cancer diagnosis using a machine learning pipeline involving image preprocessing, segmentation, feature extraction, and classification. Preprocessing is done using the Dull Razor method and Gaussian filters; K-means clustering is used for segmentation. Features are extracted via GLCM and classified using Multi-class SVM. The study utilizes dermoscopic images for melanoma classification. It provides a structured, multi-step ML approach for early-stage cancer detection.

Inference drawn:

- A structured ML pipeline with preprocessing and feature extraction boosts classification performance.
- Multi-class SVM effectively distinguishes melanoma and other skin conditions.
- Image enhancement and clustering aid in accurate segmentation.

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

This study addresses the lack of skin tone diversity in dermatology datasets by using StyleGAN3 to synthetically generate images representing minority skin tones. The models are trained on the Fitzpatrick17k dataset, with a focus on diseases like acne, vitiligo, and psoriasis. A VGG16 model is used for classification across original, traditional, and GAN-augmented datasets. The GAN-augmented model shows enhanced accuracy. The research underscores the importance of fairness and inclusion in AI dermatology.

Inference drawn:

- GAN-augmented datasets improve model performance on underrepresented skin tones.
- Enhances fairness and diversity in AI-based dermatology.
- The study emphasizes inclusivity in medical AI training.

9. Deep Learning in Dermatology: CNN-Based Classification of Skin Diseases and Cancer

This research presents a custom convolutional neural network capable of classifying 57 types of skin conditions and cancers. The model is benchmarked against standard pre-trained architectures and achieves an accuracy of 96.64%. Training involves a large, diverse dataset of dermoscopic images. The paper emphasizes the importance of dataset quality and deep learning design. It demonstrates the potential of tailored CNNs in expanding diagnostic capabilities in dermatology.

Inference drawn:

- Custom CNNs outperform pre-trained models on diverse skin disease datasets.
- Achieves high accuracy (96.64%) across 57 skin conditions.
- Demonstrates deep learning's potential to scale clinical diagnostics.

10. A Transfer Learning-based Pre-trained VGG16 Model for Skin Disease Classification

The paper applies transfer learning using a fine-tuned VGG16 model to classify skin diseases from a Kaggle dataset. The model achieves a validation accuracy of 90.1%, outperforming traditional CNNs trained from scratch. Transfer learning reduces computational load and training time while improving performance. The study explores data preprocessing and augmentation to handle dataset imbalance. It reinforces the utility of transfer learning in medical image classification tasks.

Inference drawn:

- Transfer learning enhances model efficiency and accuracy.
- Reduces training time and computational cost.
- VGG16 shows strong performance on dermatological image classification tasks.

11. Deep Learning-Based Skin Diseases Classification using Smartphones

This paper explores the classification of skin diseases using images captured through smartphones, making AI-assisted diagnosis more accessible. It employs convolutional neural networks for end-to-end feature extraction and classification. The dataset includes various skin conditions in different lighting and skin tone conditions. The model is trained to ensure generalization across mobile-captured image variations. The study supports the deployment of AI systems in real-world, low-resource environments.

Inference drawn:

- CNNs work effectively on smartphone-acquired images.
- Offers low-cost, scalable AI diagnosis for remote users.
- Generalizes well across lighting and skin tone variations.

12. Intelligent Skin Disease Prediction System using Transfer Learning and Explainable Artificial Intelligence

This study combines pre-trained CNNs (e.g., VGG16) with explainable AI techniques like Layer-wise Relevance Propagation (LRP) for transparent diagnosis. The model classifies conditions such as chickenpox and monkeypox with 93.29% accuracy. The system enhances interpretability by visually highlighting decision areas on the input images. It is tested across benchmark dermatological datasets. The paper proposes the use of interpretable AI for improved clinician trust and regulatory compliance.

Inference drawn:

- Explainable AI (LRP) increases trust in ML predictions.
- The model achieves high classification accuracy while maintaining interpretability.
- Useful for clinical settings where transparency is essential.

13. Survival Outcomes of a Large Cohort of Acral Melanoma Patients Treated at a South

The study analyzes a large patient cohort diagnosed with acral melanoma, a rare skin cancer subtype that occurs on palms, soles, and nail beds. It evaluates the impact of targeted therapies and immunotherapies on survival outcomes. Factors like age, tumor thickness, and metastatic stage are considered for survival prediction. Clinical data is used to guide therapeutic decisions. The research highlights the importance of clinical parameters in long-term treatment planning.

Inference drawn:

- Targeted therapy and immunotherapy significantly improve survival outcomes.
- Age, tumor thickness, and metastasis status are key prognostic indicators.
- Highlights the importance of demographic and clinical factors in treatment decisions.

14.Application of Convolutional Neural Networks in Skin Disease Diagnosis

This paper evaluates CNN architectures such as ResNet, VGG, and EfficientNet for skin disease classification. It discusses the necessity of data augmentation, transfer learning, and balanced datasets to enhance accuracy. The study leverages medical images to train the models and investigates the role of explainable AI to make decisions transparent. The research encourages using AI systems that are both powerful and interpretable. It contributes to the growing literature on deep learning in medical diagnostics.

Inference drawn:

- CNNs like ResNet and EfficientNet show strong performance in dermatology tasks.
- Large, balanced datasets are critical for accuracy.
- Explainability enhances AI adoption in clinical environments.

15.Skin Disease Prediction using the MobileNet Approach

This study uses MobileNet, a lightweight deep learning model, combined with data augmentation and transfer learning to classify eight types of skin diseases. The model achieves an accuracy of 94.4% and is optimized for deployment on mobile and web applications. It emphasizes low computational cost, making it suitable for rural and remote healthcare environments. The dataset includes diverse skin conditions to improve generalization. The paper proposes MobileNet as a scalable solution for AI-based dermatology.

Inference drawn:

- MobileNet achieves high accuracy with minimal computational resources.
- Well-suited for real-time, mobile-based diagnostic applications.
- Transfer learning boosts model performance on smaller datasets.

2.2 Comparison with the existing system

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	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,	- 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

	(PAD-UFES-20, HAM10000, ISIC2017) - 96% accuracy (outperforms ResNet50 & VGG16)			only a research model on Google Colab - Needs clinical validation before real-world use	research model) - Uses the ISIC dataset, ensuring real-world compatibility
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-	CNN	Image Upload (via smartphone camera)	- Requires high-quality images for accurate analysis- Does not offer infection percentage calculation-	- Incorporates U-Net for precise infection area segmentation- Supports live camera input for real-time analysis- Calculates and displays infection percentage to users

	Provides personalized guidance and information			Limited to predefined skin conditions	
--	--	--	--	---------------------------------------	--

Table 2.2: Comparison with the Existing System

Chapter 3: Requirement Gathering for the Proposed System

3.1 Introduction to Requirement Gathering

Requirement gathering is a crucial phase in system development, ensuring that all necessary functionalities and constraints are identified before implementation. This chapter outlines the functional and nonfunctional requirements, as well as the hardware, software, and technology stack required for developing WellnessInsight, an AI-driven multi-modal skin disease diagnosis system. The goal is to create a prototype that demonstrates the feasibility of AI-based skin disease detection using machine learning and deep learning techniques.

3.2 Functional Requirements

The functional requirements define the core functionalities of the system. WellnessInsight prototype will include the following features:

1. Multi-Modal Data Input
 - Accepts symptom descriptions through text input.
 - Supports voice input for symptoms using NLP-based speech recognition.
 - Allows users to upload images for classification and segmentation.
 - Enables real-time image capture via camera for diagnosis.
2. Disease Classification and Segmentation
 - Uses a Random Forest classifier for symptom-based disease prediction.
 - Employs CNNs for skin disease classification from images.
 - Utilizes U-Net for segmentation of affected skin areas.
3. Severity Assessment
 - Implements pixel-based analysis to determine infection percentage.
 - Classifies infections as severe, moderate, or normal.
 - Identifies healthy skin when no disease is detected.
4. Prototype User Interface
 - Provides a basic web-based or desktop interface.
 - Displays diagnostic results in a simple and structured format.

3.3 Non-Functional Requirements

The non-functional requirements define the quality attributes and performance expectations for the prototype:

1. Accuracy and Performance
 - The AI models should achieve a reasonable classification accuracy based on available data.
 - Image processing should be optimized for fast inference within prototype constraints.
2. Security and Privacy
 - Since this is a prototype, no real user data storage will be implemented.
 - Ensures basic data anonymization for testing purposes.
3. Scalability and Availability
 - Focused on local execution and testing, not large-scale deployment.
 - Can be extended for cloud-based deployment in the future.
4. Usability and Accessibility
 - Designed for research and proof-of-concept purposes.
 - Provides simplified workflows for testing.

3.4 Hardware, Software, Technology, and Tools Utilized

Hardware Requirements

- Development Environment:
 - PC or laptop with at least 8GB RAM.

Software and Technologies

- Programming Languages: Python (TensorFlow, PyTorch)
- Frameworks & Libraries:
 - TensorFlow/Keras, PyTorch (for CNNs, U-Net)
 - Scikit-learn (for Random Forest classifier)
 - OpenCV (for image preprocessing)
 - NLTK/SpeechRecognition (for NLP-based voice processing)
- Deployment & Hosting:
 - Runs locally or in a limited cloud environment (Google Colab)

3.5 Constraints

While the prototype aims to showcase AI-driven skin disease diagnosis, certain constraints exist:

- Limited Data Availability
 - Model performance depends on publicly available datasets.
 - No real patient data will be used in this phase.
- Computational Limitations
 - Prototype is not optimized for real-time deployment.
 - Deep learning inference may take longer on non-GPU machines.
- Regulatory Compliance
 - Since this is a prototype, no compliance with HIPAA/GDPR is required at this stage.
- User Interaction Limitations
 - Prototype have limited UI capabilities.
 - No real-time user feedback mechanisms implemented.

Chapter 4: Proposed Design

4.1 Block diagram of the system

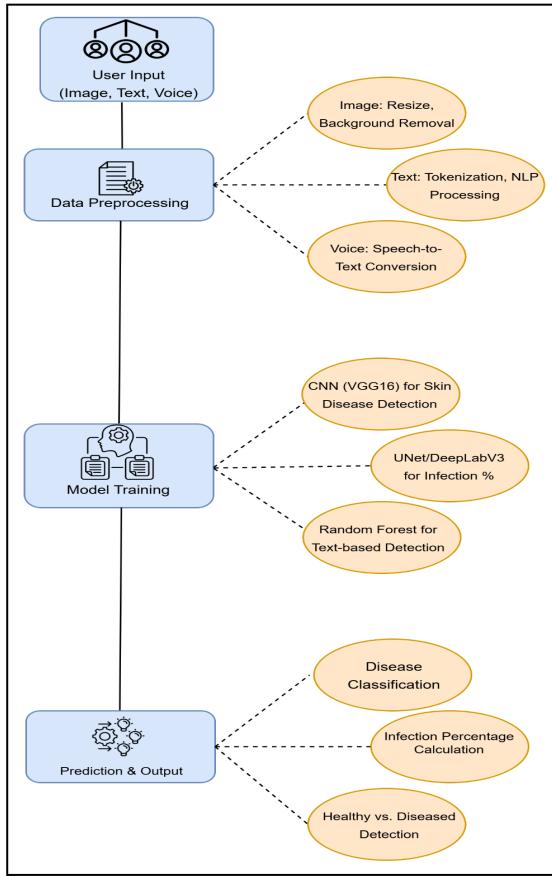


Fig 4.1: Block Diagram of WellnessInsight System

The block diagram illustrates the working flow of the **WellnessInsight** system, which accepts **Image, Text, or Voice** as input from the user. Based on the input type, the data undergoes appropriate **preprocessing**:

- **Text** is processed using tokenization and NLP techniques
- **Voice** is converted to text using speech-to-text conversion.
- **Images** are resized and background is removed.

After preprocessing:

- **Text or voice-based symptoms** are passed to a **Random Forest classifier** for disease prediction.
- **Images** are analyzed using a **VGG16 model** to detect the type of skin disease (e.g., melanoma, psoriasis, eczema, healthy).
- The **UNet model** is used to calculate the **percentage of infection** for image inputs.

Finally, the system provides the user with:

- **Disease classification**, and if the input was an image,
- The **infection percentage** and **healthy vs. diseased status**.

This approach helps in accurately identifying skin diseases and improves accessibility by allowing multimodal inputs.

4.2 Modular design of the system

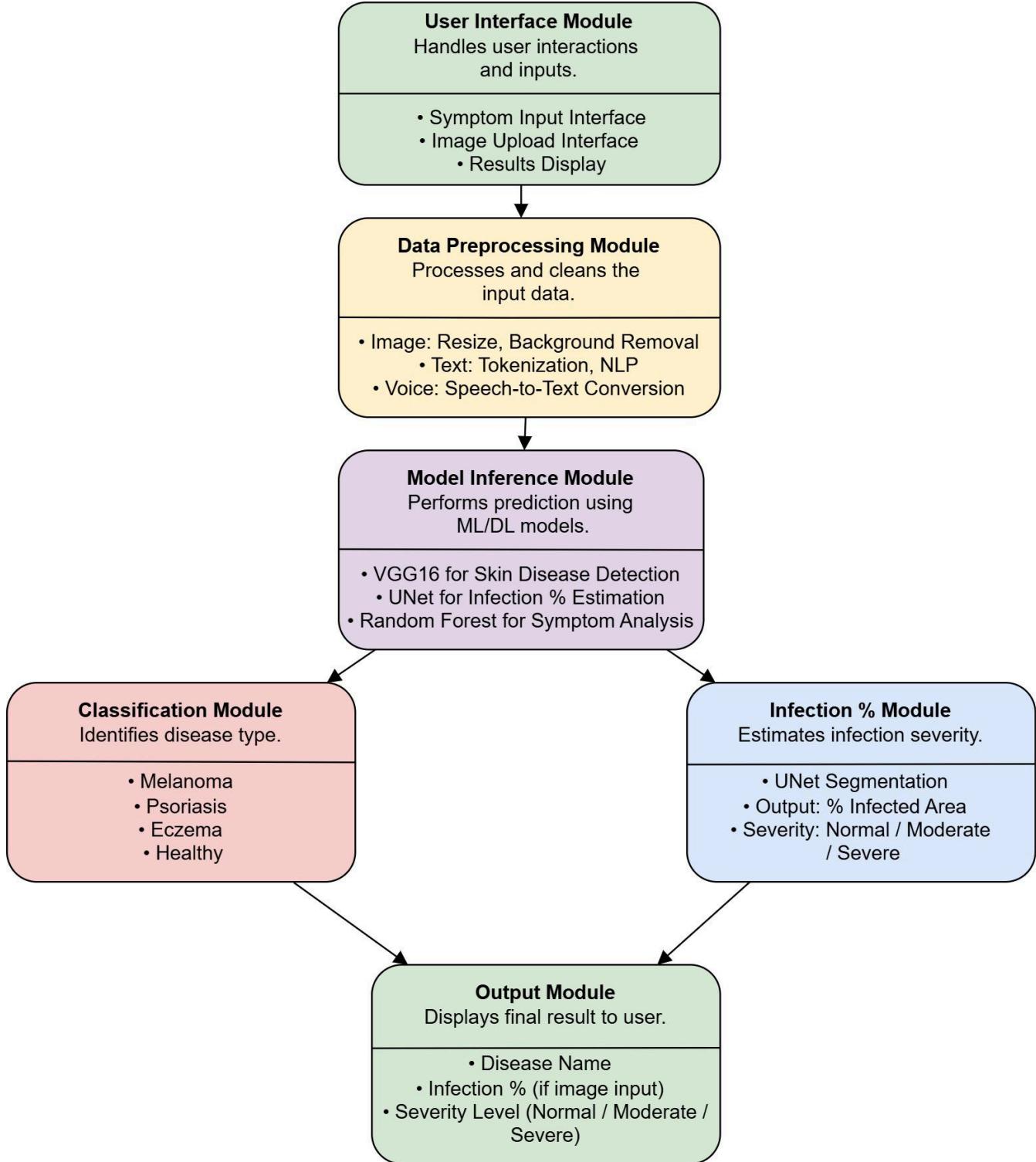


Fig 4.2: Modular Design of Skin Disease Prediction System

The above diagram illustrates the **modular design** of the **Skin Disease Prediction System**. Each module has a distinct function to ensure clarity, scalability, and ease of maintenance:

1. User Interface Module

This module facilitates user interaction by allowing them to input symptoms, upload skin images, and view results in a user-friendly manner.

2. Data Preprocessing Module

Responsible for preparing raw data for analysis. It includes:

- **Image preprocessing:** resizing and removing background noise.
- **Text preprocessing:** tokenization and NLP techniques.
- **Voice preprocessing:** converting speech to text.

3. Model Inference Module

Executes prediction using suitable machine learning and deep learning models:

- **VGG16:** For detecting skin diseases from images.
- **UNet:** For estimating infected area.
- **Random Forest:** For analyzing text-based symptoms.

4. Classification Module

Based on model inference, this module classifies the disease into types such as:

- **Melanoma, Psoriasis, Eczema, or Healthy.**

5. Infection Percentage Module

Focuses on evaluating the severity of the condition using image segmentation:

- Outputs the **percentage of infected area.**
- Determines severity as **Normal, Moderate, or Severe.**

6. Output Module

Compiles all processed information and displays:

- **Predicted disease name**
- **Infection percentage and Severity level**(for image input)

4.3 Detailed Design

4.3.1 Flowchart

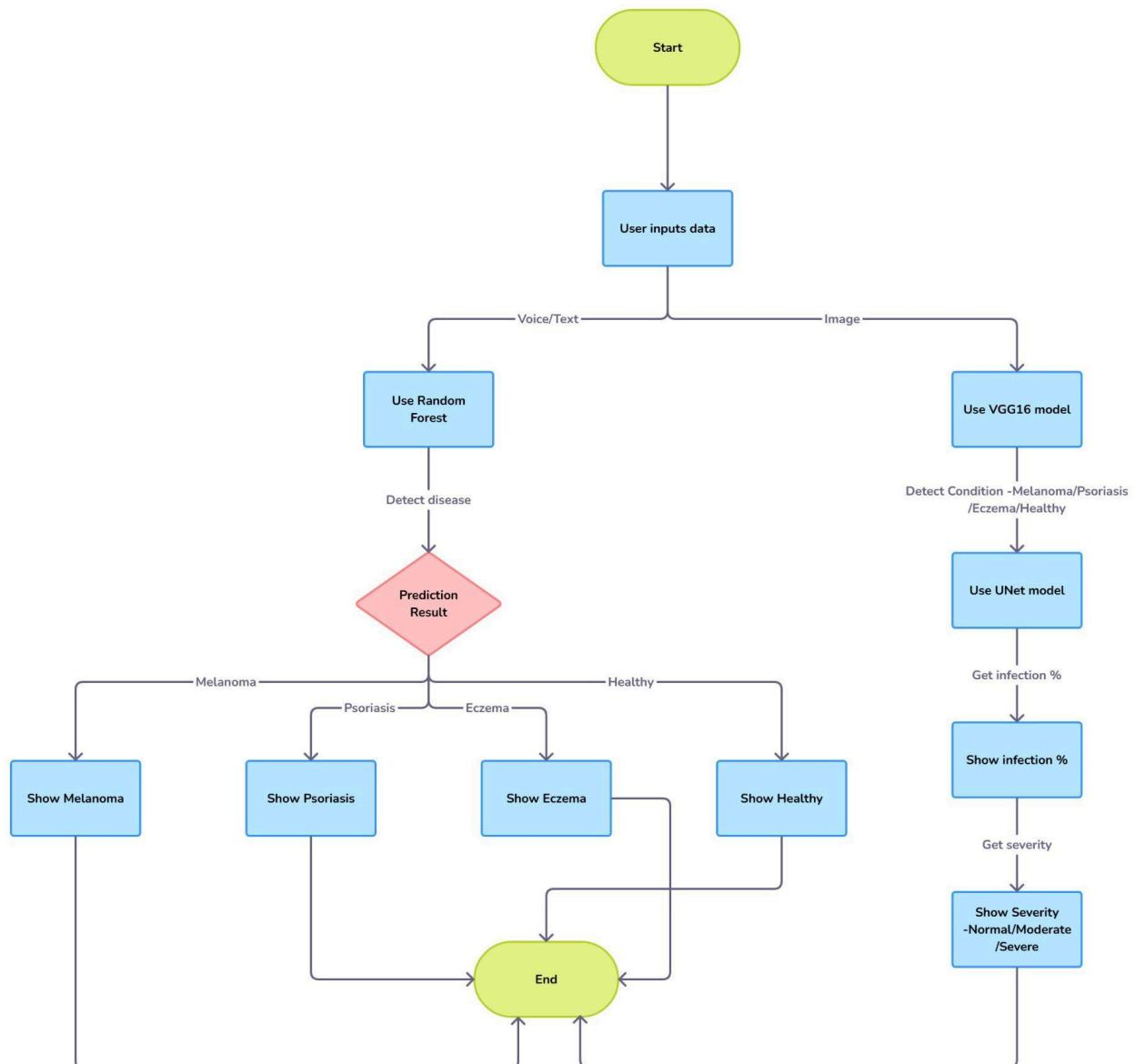


Fig 4.3.1: Flowchart

This flowchart illustrates the workflow of a WellnessInsight-skin disease prediction system. The process begins with the user providing input through either voice/text or image. For voice/text input, a Random Forest model is used to predict the condition(Melanoma, Psoriasis, Eczema, or Healthy). For image input, a VGG16 model identifies the condition, and a UNet model calculates the infection percentage. The system then displays the predicted disease along with severity levels (Normal, Moderate, Severe) based on image analysis. This dual-model approach enhances accuracy and provides detailed diagnosis for better healthcare support.

4.4 Project Scheduling & Tracking using Gantt Chart

Gantt Chart - WellnessInsight

PROCESS	Sem VII					Sem VIII				
	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
Requirement Gathering										
Topic Finalization										
Planning & Designing										
Data Collection & Preprocessing										
Front-end Development										
Back-end Development										
Model - RF (Text/Voice)										
Model - VGG16 (Image)										
Model - UNet (Infection %)										
Deployment & Testing										

Fig 4.4: Gantt Chart

The Gantt chart illustrates the timeline of the *WellnessInsight* project, covering two academic semesters. Initial phases focused on requirement gathering, topic finalization, and UI planning. A dedicated data collection and preprocessing phase was introduced to handle sourcing and preparing both text-based symptom data and medical images, enabling robust model training. Model development followed a staggered approach—Random Forest for symptom analysis, VGG16 for image classification, and UNet for infection segmentation—ensuring each model was trained on clean, well-structured datasets. Deployment and testing were reserved for the final phase to integrate all components into a functional, user-friendly system.

Chapter 5: Implementation of the Proposed System

5.1 Methodology Employed for Development

The proposed AI-driven skin disease diagnosis system, WellnessInsight, integrates machine learning (ML) and deep learning (DL) techniques to enhance diagnostic efficiency. The development follows a structured methodology involving data collection, preprocessing, model training, evaluation, and system integration.

Development Stages:

1. Requirement Analysis: Understanding user needs and defining the scope of the system.
2. Data Collection: Acquiring a diverse dataset of skin disease images, symptoms, and related metadata.
3. Preprocessing: Data augmentation, noise reduction, and feature extraction to enhance model performance.
4. Model Training: Implementing Random Forest for symptom-based classification and CNNs with U-Net for image classification and segmentation.
5. Infection Severity Analysis: Pixel-based analysis using RGB values for estimating infection percentage.
6. Multi-Modal Integration: Combining text, voice, and image inputs for comprehensive diagnosis.
7. Testing and Validation: Evaluating model performance using accuracy, precision, recall, and F1-score.
8. Deployment: Implementing the trained models into a user-friendly application interface.

5.2 Algorithms and Flowcharts for the Respective Modules Developed

5.2.1 Symptom-Based Disease Prediction (Random Forest Classifier)

Algorithm:

1. Extract symptom keywords from text/voice input using NLP techniques.
2. Encode features for classification.
3. Train the Random Forest model on labeled symptom data.
4. Predict the most probable skin disease based on user-provided symptoms.

5.2.2 Image-Based Classification and Segmentation (CNN & U-Net)

Algorithm:

1. Accept image input (real-time capture or upload).
2. Apply preprocessing: resizing, normalization, and augmentation.

3. Use CNN for classification of the skin disease.
4. Utilize U-Net for lesion segmentation to highlight affected areas.
5. Output classified disease type and segmented lesion mask.

5.2.3 Infection Severity Estimation (Pixel-Based Analysis)

Algorithm:

1. Extract the segmented lesion area from the image.
2. Compute the ratio of affected pixels to total skin area using RGB values for accurate infection estimation.
3. Categorize the severity level:
 - o Severe: >50% infection
 - o Moderate: 20%-50% infection
 - o Normal: < 20% infection
 - o Healthy: No infection detected
4. Display the severity classification along with diagnostic results.

5.3 Datasets Source and Utilization

5.3.1 Data Sources

The system is trained on publicly available and proprietary datasets, ensuring a diverse set of skin disease images and symptom descriptions. The primary datasets used include:

- ISIC Dataset: International Skin Imaging Collaboration dataset for melanoma and other skin conditions.
- DermNet Dataset: A comprehensive dataset of dermatological conditions.
- Kaggle Dataset : For Healthy Skin images

5.3.2 Data Utilization

- Training: 80% of the dataset is used for model training.
- Validation: 10% is utilized for tuning hyperparameters.
- Testing: 10% is reserved for final model evaluation.
- Augmentation: Data augmentation techniques such as flipping, rotation, and contrast adjustments are applied to improve model generalization.

This structured approach ensures robust and reliable performance for real-world applications, making dermatological diagnosis accessible and efficient.

Chapter 6 : Results and Discussion

6.1. Screenshots of User Interface (UI) for the respective module

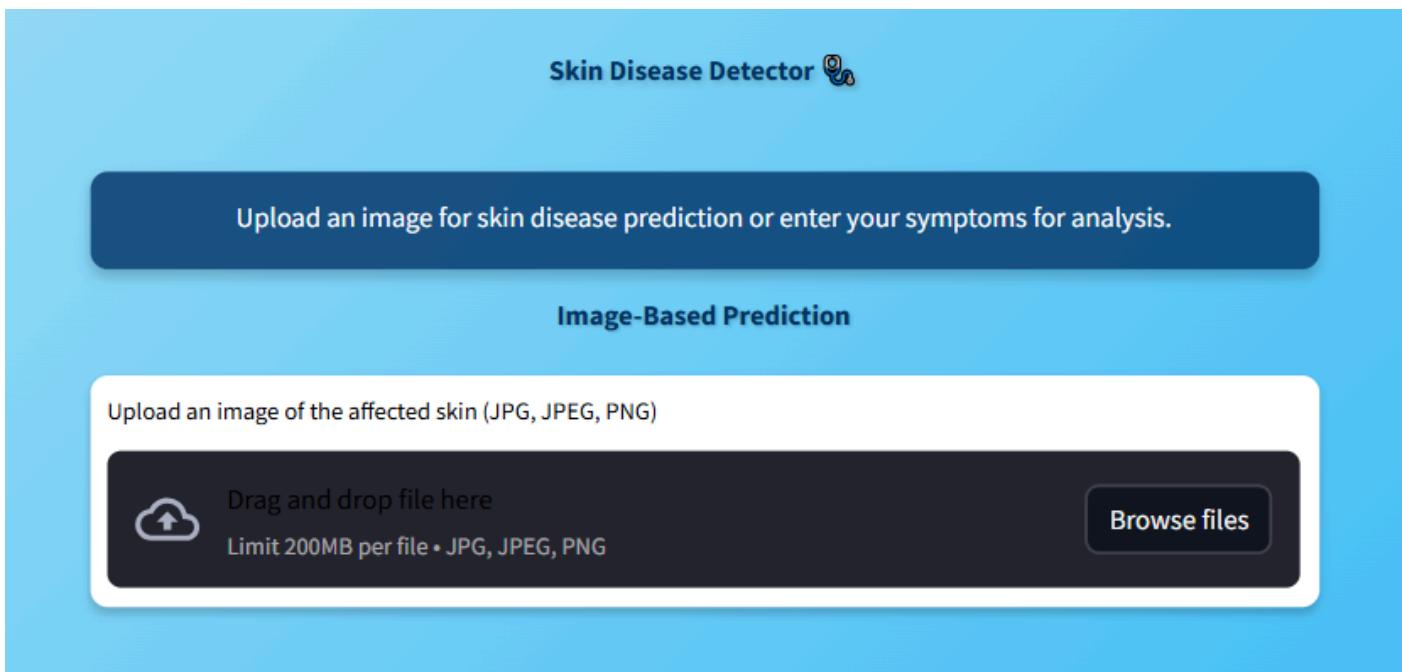


Fig 6.1.1: Option to upload an image to predict disease.



Fig 6.1.2: Displayed the Predicted disease, Infection percentage and severity of disease.

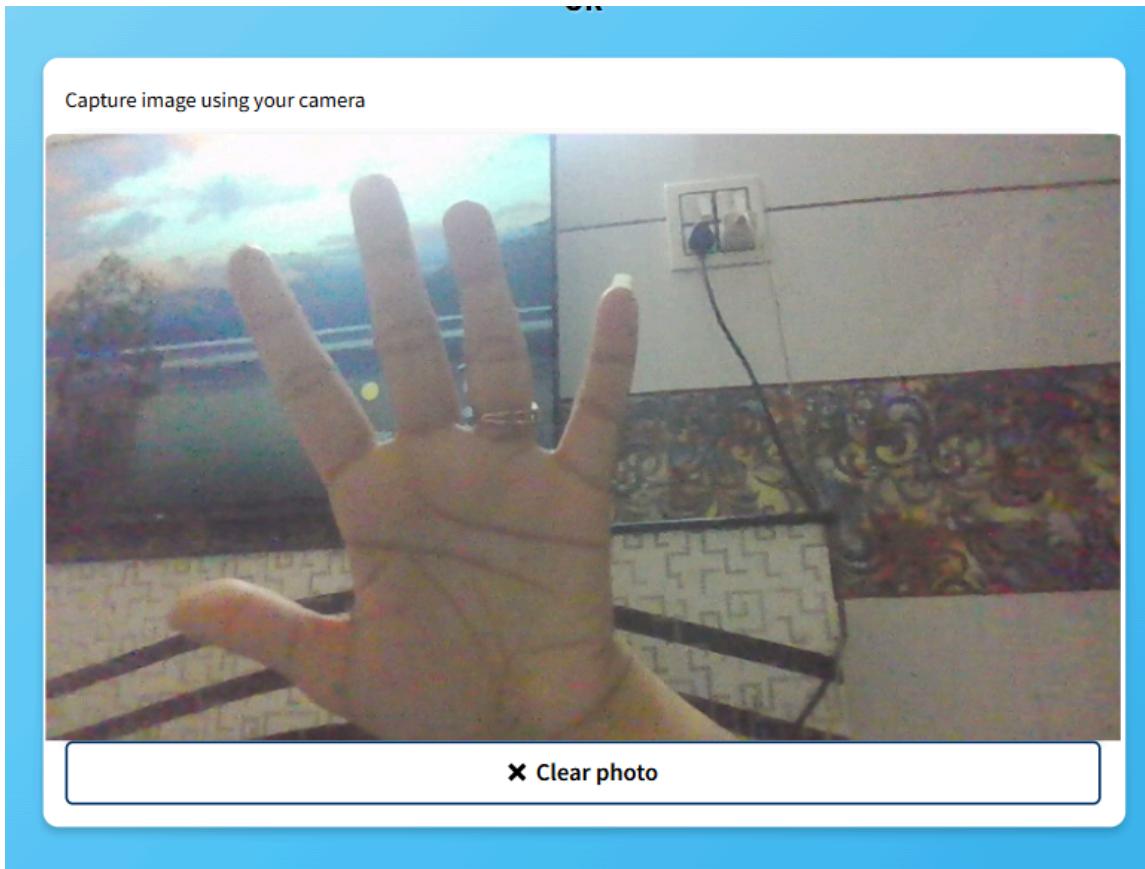


Fig 6.1.3: Option to click real time image for disease detection.

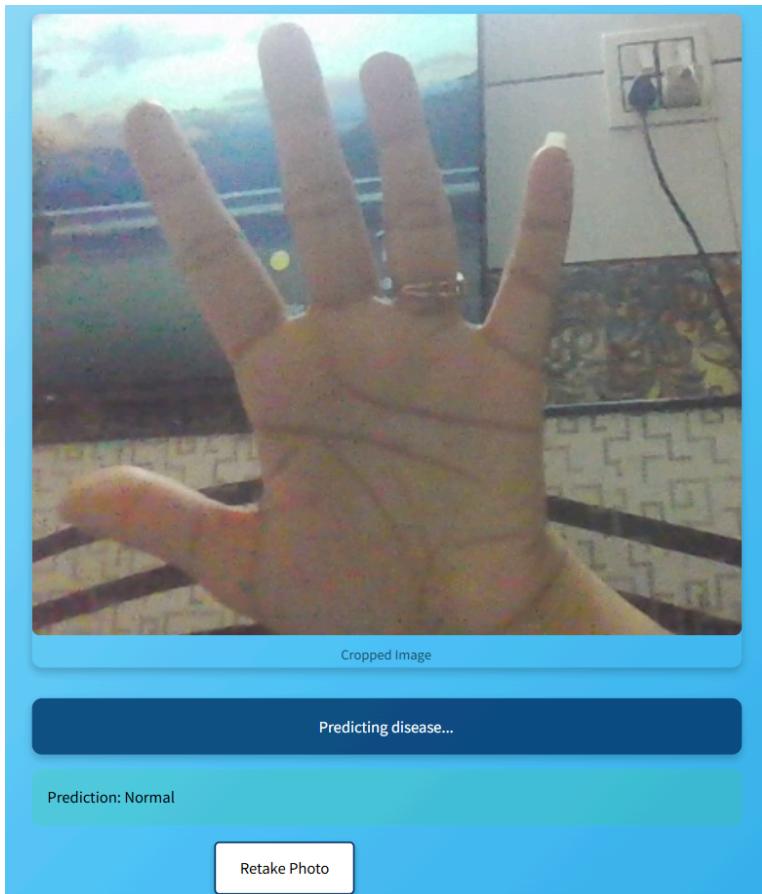


Fig. 6.1.4: Result is displayed based on image given.

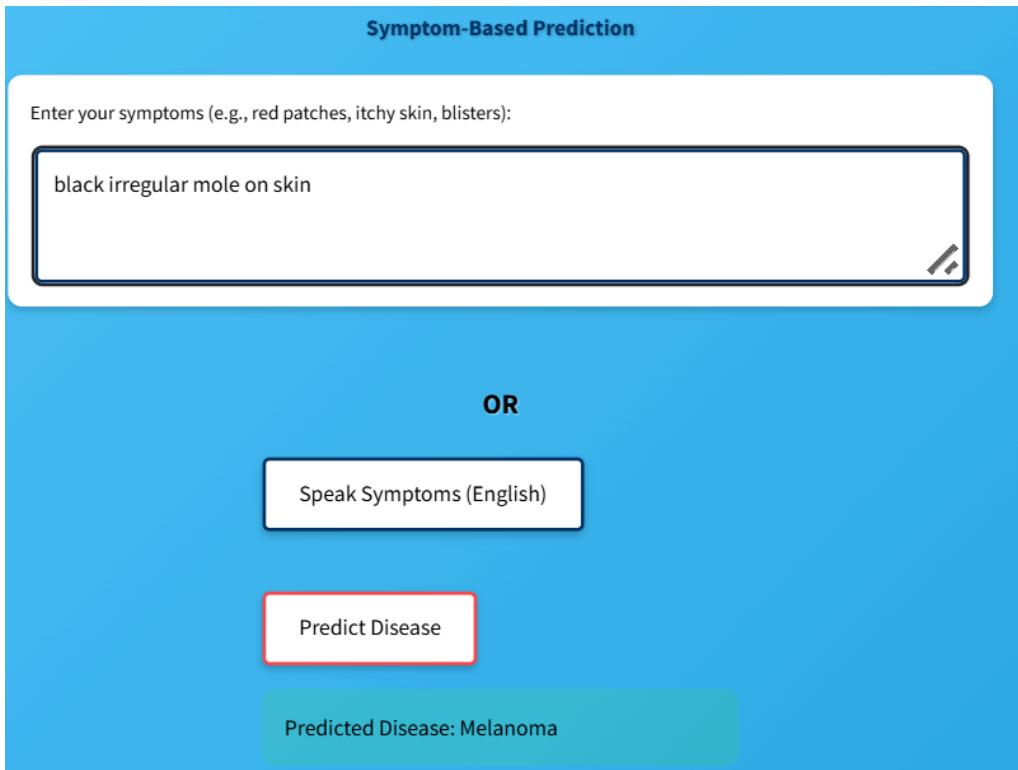


Fig 6.1.5: Option for the user to describe the problem in words or to speak about symptoms.

6.2 : Performance Evaluation measures

1. Random Forest (Symptom-based classification)

Input: Text (symptoms)

Output: Predicted disease label

Algorithm: Random Forest Classifier

Accuracy: 88.4%

Use Case: Disease prediction from symptoms

2. CNN (Skin disease image classification)

Input: Skin disease images

Output: Disease class (Eczema, Psoriasis, Melanoma)

Model: Convolutional Neural Network

Total Parameters: 1,946,993

Trainable: 1,944,049 | Non-trainable: 2,944

Input Shape: (224, 224, 3)

Output Shape: Number of disease classes

Accuracy: 91.6%

Layers: Conv2D, BatchNormalization, Dense

3. U-Net (Skin infection segmentation model)

Input: Skin images (224x224x3)

Output: Segmentation mask (224x224x1)

Model: U-Net (Encoder-Decoder with skip connections)

Total Parameters: 1,946,993

Trainable: 1,944,049 | Non-trainable: 2,944

Layers: Conv2D, BatchNormalization, Conv2DTranspose, Concatenate

Accuracy: 90.8%

6.3. Input Parameters / Features considered

Image Features: RGB pixel values, GLCM texture descriptors, edge detection (Canny), segmentation masks from ISIC dataset

Text/Voice Features: Symptom keywords (e.g., itching, redness), TF-IDF vectors, voice-to-text (if implemented via speech recognition)

Derived Features: Infection spread ratio (from segmented masks), severity level (low/medium/high based on infected pixel area)

6.4 Graphical or statistical output

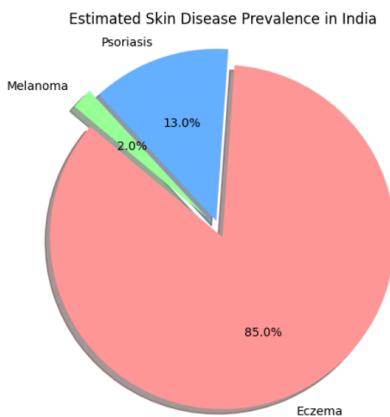


Fig 6.4: Pie Chart Showing Estimated Prevalence of Skin Diseases in India

- **Psoriasis** affects approximately **0.44%–2.8%** of the Indian population.
- **Melanoma** and other skin cancers account for around **1–2%** of total cancer cases in India.
- **Eczema (Dermatitis)** has a significant health burden, contributing to **1.40 million YLDs (Years Lived with Disability)** in 2017 in India.
- Data is based on studies from **IJDVL, PMC, and Global Burden of Disease 2017** reports.

6.5 Comparison of Results with Existing Systems

System	Accuracy	Input Modes	Infection % Detection	Live Input
SkinLesNet	88%	Image Upload	No	No
DERMNET EASI	Manual	Manual Entry	No	No
AI Dermatologist	82%	Image Upload	No	No
WellnessInsight	91%	Image, Live Cam, Text, Voice	Yes	Yes

Table 6.5: Comparison of results with the Existing Systems

Chapter 7: Conclusion

7.1 Limitations

Despite its potential, the WellnessInsight system has certain limitations that need to be addressed for improved effectiveness:

- Limited Dataset Coverage: Some skin conditions may not be well-represented in the dataset, affecting the model's ability to generalize.
- Edge Cases and Atypical Presentations: Skin diseases with uncommon symptoms or mixed features can pose challenges for automated diagnosis.
- Hardware and Computational Constraints: Real-time processing on low-end devices may be limited due to computational requirements, affecting accessibility in resource-constrained regions.
- User Dependency on Image Quality: Diagnosis accuracy may be impacted by poor-quality images taken under suboptimal lighting conditions or at improper angles.
- Regulatory and Ethical Concerns: AI-driven diagnosis requires regulatory approvals and adherence to medical data privacy laws.

7.2 Conclusion

The WellnessInsight system offers a unique and comprehensive approach to predicting skin diseases by combining machine learning and deep learning models with various types of inputs, including text, voice, and images. By providing various input options, the platform improves accessibility, enabling users to receive preliminary diagnoses regardless of their preferred method of interaction. The random forest classifier efficiently examines textual and voice-based symptom descriptions, while the u-net model precisely divides affected skin regions from images, guaranteeing a thorough diagnostic process. By incorporating natural language processing (nlp) for text analysis, speech recognition for voice input, and computer vision techniques for image processing, the system's reliability and usability are significantly improved.

One of the significant benefits of WellnessInsight is its ability to connect the dots between telemedicine and remote healthcare, closing the gaps in these areas. The platform offers users a user-friendly web interface, enabling them to perform self-assessments and seek early medical intervention without any hassle. This is especially beneficial for individuals living in isolated regions with limited access to dermatologists and healthcare centers.

7.3 Future Scope

Several enhancements and expansions can be pursued to further improve the WellnessInsight system:

- Enhanced Model Training: Expanding the dataset to include diverse skin tones, rare conditions, and atypical presentations to improve accuracy and generalization.
- Integration with Telemedicine: Connecting the system with telemedicine platforms to enable remote consultations with dermatologists for cases requiring expert validation.
- Improved Hardware Compatibility: Optimizing AI models for deployment on mobile devices with low computational power, ensuring accessibility in remote areas.
- Real-Time Feedback and Image Enhancement: Implementing pre-processing techniques to improve image quality and provide real-time guidance to users on capturing high-quality images.
- Regulatory Compliance and Clinical Validation: Collaborating with dermatologists and medical institutions for regulatory approval, clinical trials, and certification to establish credibility.
- Multilingual and Regional Expansion: Expanding the system to support multiple languages and region-specific skin disease variations, improving its usability across diverse populations.

By addressing these areas, WellnessInsight has the potential to become a globally recognized AI-powered dermatology assistant, significantly improving early diagnosis, treatment outcomes, and healthcare accessibility worldwide.

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.

Appendix

1. Submitted Paper

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	<ul style="list-style-type: none"> - Classifies 3 skin lesions (melanoma, nevus, seborrheic keratosis) - Trained on smartphone images (PAD-UFES-20, HAM10000, ISIC2017) - 96% accuracy (outperforms ResNet50 & VGG16) 	CNN (4-layer deep model) Compared with ResNet50 & VGG16	Image (Upload Only)	<ul style="list-style-type: none"> - 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 	<ul style="list-style-type: none"> - 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
Eczema Area and Severity Index (EASI) – DERMNET	<ul style="list-style-type: none"> - 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	<ul style="list-style-type: none"> - Subjective, depends on doctor's assessment - No AI-based automation - Only for eczema, not general skin diseases 	<ul style="list-style-type: none"> - 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

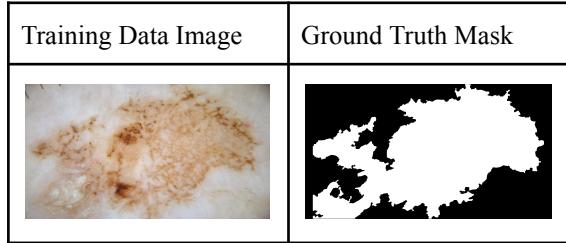
Normal

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.



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

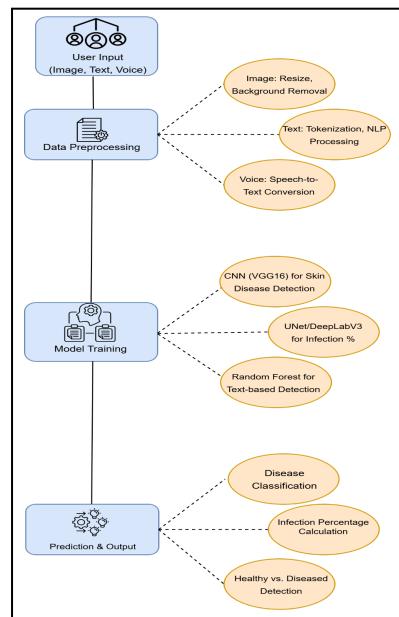


Figure 3. Workflow of the skin disease classification model.

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

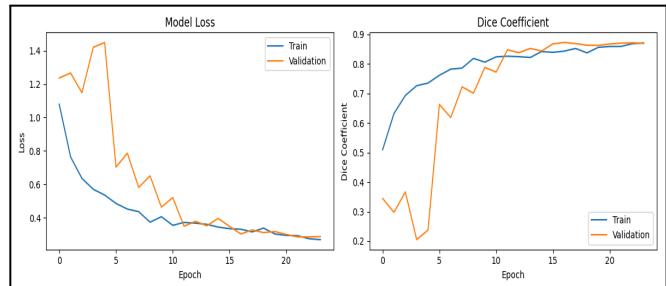


Figure 4. Model Training

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

VII. 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.

2)Plagiarism Report

A Multimodal AI System integrating image

ORIGINALITY REPORT

5%
SIMILARITY INDEX **4%**
INTERNET SOURCES **6%**
PUBLICATIONS **3%**
STUDENT PAPERS

PRIMARY SOURCES

1	journal.esrgroups.org Internet Source	2%
2	V. Sharmila, S. Kannadhasan, A. Rajiv Kannan, P. Sivakumar, V. Vennila. "Challenges in Information, Communication and Computing Technology", CRC Press, 2024 Publication	1%
3	"Universal Threats in Expert Applications and Solutions", Springer Science and Business Media LLC, 2024 Publication	1%
4	Roshnee Matlani, Roshan Dadlani, Sharv Dumbre, Shruti Mishra, Mrs. Abha Tewari. "Real-time Sign Language Recognition using Machine Learning and Neural Network", 2022 International Conference on Electronics and Renewable Systems (ICEARS), 2022 Publication	1%
5	Submitted to Universiti Malaysia Perlis Student Paper	1%
6	ijesr.org Internet Source	1%



Page 2 of 9 - AI Writing Overview

Submission ID trn:oid::1:3217208420

*% detected as AI

AI detection includes the possibility of false positives. Although some text in this submission is likely AI generated, scores below the 20% threshold are not surfaced because they have a higher likelihood of false positives.

Caution: Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Disclaimer

Our AI writing assessment is designed to help educators identify text that might be prepared by a generative AI tool. Our AI writing assessment may not always be accurate (it may misidentify writing that is likely AI generated as AI generated and AI paraphrased or likely AI generated and AI paraphrased writing as only AI generated) so it should not be used as the sole basis for adverse actions against a student. It takes further scrutiny and human judgment in conjunction with an organization's application of its specific academic policies to determine whether any academic misconduct has occurred.

3)Review Sheet

Project Evaluation Sheet 2024 - 25

Title of Project: Wellness Insight

Group Members: Tiya Lund (D17A-37), Dinesh Phatni (D17A-31), Tisha Tewari (D17B-20), Varsha Mankarjiya (D17C-43)

Engineering Concepts & Knowledge (5)	Interpretation of Problem & Analysis (5)	Design / Prototype (5)	Interpretation of Data & Dataset (3)	Modern Tool Usage (5)	Societal Benefit, Safety Consideration (2)	Environment Friendly (2)	Ethics (2)	Team work (2)	Presentation Skills (2)	Applied Engg&Mgmt principles (3)	Life-long learning (3)	Professional Skills (3)	Innovative Approach (3)	Research Paper (5)	Total Marks (50)
3	3	3	3	4	2	2	2	2	2	3	2	3	2	3	39

Comments:

Able to
01/03/2025

Name & Signature Reviewer 1

Inhouse/ Industry - Innovation/Research:

Engineering Concepts & Knowledge (5)	Interpretation of Problem & Analysis (5)	Design / Prototype (5)	Interpretation of Data & Dataset (3)	Modern Tool Usage (5)	Societal Benefit, Safety Consideration (2)	Environment Friendly (2)	Ethics (2)	Team work (2)	Presentation Skills (2)	Applied Engg&Mgmt principles (3)	Life-long learning (3)	Professional Skills (3)	Innovative Approach (3)	Research Paper (5)	Total Marks (50)
3	3	3	3	4	2	2	2	2	2	3	2	3	2	3	39

Comments:

Mannat
Name & Signature Reviewer 2

Date: 1st March, 2025