MediMind AI

A *Project Report* Submitted in Partial
Fulfillment of the requirements for
the award of the Degree of

Bachelor of Technology in
Computer Science & Engineering

Submitted by

Jitu Kumar (222025109181)

Under the guidance of

Dr. Ritushree Narayan

USHA MARTIN
UNIVERSITY
Inclusion, Innovation & Ethics

Faculty of Computing and Information Technology
Usha Martin University, Ranchi
June, 2025

BONAFIDE CERTIFICATE

Certified that this project report titled MediMind AI is the Bonafide work of Jitu Kumar who carried out the research under my supervision. Certified further, that to the best of my knowledge the work reported herein does not form part of any other project report or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

(Signature of the concerned Supervisor of the Organization with Organization Seal)

(Certificate to be countersigned by the HOD.)

DECLARATION

I do hereby declare that this report entitled “MediMind AI”, submitted by Jitu Kumar, bearing Reg. No.: 222025109181 in the fulfillment of the requirement for the degree of Bachelor of Technology in Computer Science & Engineering to Usha Martin University, Ranchi, is my own and it is not submitted to any other institute.

Jitu Kumar

Reg. No.: 222025109181

CERTIFICATE

This is to certify that entitled "MediMind AI" being submitted by Jitu Kumar, bearing Reg. No.: 222025109181, in the fulfillment of the requirement for the degree of Bachelor of Technology in Computer Science & Engineering to Usha Martin University, Ranchi, is a Bonafide work carried out under my/our supervision. The matter embodied in this report is original and has not been submitted for the award of any other degree.

Dr. Ritushree Narayan

Guide

Dr. Sharmistha Roy

HOD

Faculty of Computing and Information Technology

Usha Martin University, Ranchi

EXTERNAL

ACKNOWLEDGEMENTS

I express my deepest sense of gratitude to my guide **Dr. Ritushree Narayan** Faculty of Computing and Information Technology, Usha Martin University, Ranchi, for suggesting the subject of work and constant supervision throughout this work. Her co-operation and timely suggestions have been unparalleled stimuli for me to travel eventually towards the completion of this project report. Indeed her continuous involvement has helped me in bringing of this project work which otherwise would have remained a distant dream.

I am indeed thankful to **Dr. Sharmistha Roy**, HOD, Faculty of Computing and Information Technology, Usha Martin University, Ranchi, for giving me permission to carry out my project work. I would like to express my gratitude to all teaching and non-teaching staff members of Faculty of Computing and Information Technology, Usha Martin University, Ranchi, for their co-operation in my work.

Jitu Kumar

Reg. No.: 222025109181

Abstract

Skin diseases are common and can become serious if not diagnosed early. Many people do not have quick access to dermatologists, and manual diagnosis can take time and may lead to errors. This project aims to create an AI-based system to help detect skin diseases like Acne, Eczema, and Melanoma, along with an Unknown category for unrecognized conditions. The system uses MobileNetV2, ResNet50, and DenseNet121 a deep learning model that analyzes skin images to classify diseases. The dataset is prepared by resizing images to 224×224, normalizing pixel values, and applying data augmentation to improve accuracy. To ensure the system does not misclassify unknown skin conditions, Softmax Probability Thresholding are used. The model is evaluated using accuracy, precision, recall, and F1-score to ensure reliable predictions. For user accessibility, the system is integrated into a Gradio UI, allowing users to upload images and receive instant results. Additionally, an AI-generated PDF skin health report provides insights into the diagnosis. The model's performance is compared with ResNet50 and MobileNetV2, and results show that ResNet50 performs the best. This AI-based tool can assist both patients and dermatologists in getting quick and accurate skin disease detection.

Keyword – Acne, Eczema, Melanoma, Skin Disease

TABLE OF CONTENTS

List of Figures

List of Tables

List of Abbreviations

Abbreviations	Full Form
AI……………………………….......................	Artificial Intelligence
DL……………………………………………..	Deep Learning
ML…………………………………………….	Machine Learning
GUI……………………………………………	Graphical User Interface
CNN…………………………………………...	Convolutional Neural Network
Grad-CAM…………………………………….	Gradient-weighted Class Activation Mapping
PDF……………………………………………	Portable Document Format
TTS……………………………………………	Text-to-Speech
LLM…………………………………………...	Large Language Model
API…………………………………………….	Application Programming Interface
GPU…………………………………………...	Graphics Processing Unit
MA…………………………………………….	MediMind AI

CHAPTER 1

Introduction

1.1 Overview

MediMind AI is an intelligent deep learning-powered system built to detect and classify common skin diseases using image-based analysis. The core goal of this project is to improve early identification of dermatological conditions and make skin health support more accessible to people through an AI-driven assistant. The system is trained to recognize three major skin diseases: Acne, Eczema, and Melanoma, and it includes a robust 'Unknown' category to handle inputs that don't belong to any trained disease class. To ensure high accuracy and robustness, three state-of-the-art convolutional neural networks (CNNs) — MobileNetV2, ResNet50, and DenseNet121 were trained, fine-tuned, and evaluated using transfer learning techniques. After the models were trained, they were integrated into a user-friendly Gradio web application that allows users to interact with the AI in multiple ways. Users can upload a skin image, receive instant predictions along with confidence scores, and view a comparison table showing each model's performance metrics. A standout feature of MediMind AI is its support for Voice-Based Interaction. The system also offers voice output, where predictions and guidance are read aloud to the user, improving usability for people with visual impairments or reading difficulties.

1.2 Background

Skin diseases like Acne, Eczema, and Melanoma are widespread and can become serious if not diagnosed early. In many areas, especially rural regions, access to dermatological care is limited, causing delays in treatment. With advancements in Artificial Intelligence, deep learning models — especially Convolutional Neural Networks (CNNs) — have shown great potential in medical image analysis. These models can support early detection by classifying skin conditions from images with high accuracy. MediMind AI is developed to address this need. It uses four powerful CNN models MobileNetV2, ResNet50, and DenseNet121 to classify skin diseases. The system includes a voice-based interface for better accessibility and integrates Out-of-Distribution detection methods to reject unknown inputs safely. By combining AI, accessibility, and reliability, MediMind AI offers a smart solution for skin health awareness and early diagnosis.

1.3 Applications of MediMind AI

i. **Early Skin Disease Detection :** MediMind AI enables quick identification of common skin conditions like Acne, Eczema, and Melanoma from simple skin images. This allows users to take timely action and consult a doctor before the disease worsens.

ii. **AI-Based Second Opinion for Dermatologists :** Doctors and medical practitioners can use the system as a supportive tool to cross-check initial assessments, especially in remote areas lacking specialist access.

iii. **Medical Education & Training :** The system can be used as an educational tool for medical students to understand how deep learning models classify different skin conditions.

iv. **Rural and Low-Resource Settings :** By offering a lightweight, voice-enabled, browser-accessible application, MediMind AI is suitable for deployment in areas with limited access to dermatologists.

1.4 Features of MediMind AI

i. **Multi-Model Integration :** Includes four powerful deep learning models: MobileNetV2, ResNet50, and DenseNet121, each trained and fine-tuned for optimal accuracy.

ii. **Unknown Class Handling :** A dedicated 'Unknown' class helps prevent misclassification by identifying images that don't fit into trained disease categories.

iii. **Interactive Gradio Interface :** Clean, web-based UI built with Gradio lets users upload images, view predictions, and compare model performance — with no technical background required.

iv. **Confidence Scores & Model Comparison :** Displays prediction probabilities and a model performance table to help users understand results more transparently.

1.5 Objectives

i. Develop MediMind AI to convert patient speech to text for symptom analysis.

ii. Implement AI-based image processing for skin disease detection and insights.

iii. Train a deep learning model to classify Acne, Eczema, Melanoma, and Unknown cases.
iv. Integrate the model into a Gradio-based web app for real-time diagnosis.

1.6 Problem Statement

Many people suffer from skin diseases like Acne, Eczema, and Melanoma, but they don't always get proper or timely diagnosis, especially in rural areas where skin specialists are not available. Manual diagnosis takes time and needs expert knowledge, which is not always possible. Some AI tools exist, but they may not support multiple diseases, lack voice features, and cannot identify unknown or unrelated images correctly. So, there is a need for an easy-to-use and intelligent system that can detect different skin diseases, support voice interaction, and safely reject images that do not belong to any known category.

1.7 Research Gaps

a. Most public datasets have limited skin disease categories, making it harder for the model to generalize.
b. Converts speech to text and extracts symptoms for AI-driven diagnosis.
c. Integrates Speech, Image and Text, all three for better accuracy and deeper medical insights.
d. PDF medical report generation, creates detailed medical reports for easy doctor review.
e. Multi-language support works in English and Hindi, making healthcare more accessible.

CHAPTER 2

Literature Survey

2.1 Literature Review

In recent years, Artificial Intelligence (AI), especially Deep Learning (DL), has shown great promise in the field of medical diagnostics, particularly for analyzing medical images. Several studies have demonstrated the effectiveness of Convolutional Neural Networks (CNNs) in detecting and classifying various skin diseases with high accuracy. Models like MobileNetV2, ResNet50, and DenseNet121 have been widely used for feature extraction and image classification tasks due to their efficient architectures and strong performance on medical datasets. Researchers have also explored the use of fine-tuning and data augmentation techniques to improve model generalization and handle imbalanced datasets. Moreover, recent literature has highlighted the importance of Out-of-Distribution (OOD) detection methods such as Softmax Thresholding and Mahalanobis Distance to prevent misclassification of irrelevant or unknown inputs an essential feature for real-world deployment. While many AI-based dermatology tools exist, few incorporate voice interaction or user-friendly web interfaces, which are crucial for accessibility, especially in rural areas. Hence, MediMind AI aims to bridge these gaps by combining accurate deep learning models with an interactive, voice-enabled interface, offering a reliable, accessible, and comprehensive tool for skin disease detection.

i. **Ahmed Alkuwaiti A ; Nazer K ; Al-Reedy. A Review of the Role of Artificial Intelligence in Healthcare: A Comprehensive Study, Frontiers in Digital Health**
   Artificial Intelligence (AI) is revolutionizing the healthcare sector by significantly enhancing medical imaging, diagnostics, virtual patient care, drug discovery, and administrative operations. It facilitates early disease detection, including conditions like COVID-19, and supports remote monitoring through AI-driven tools. Moreover, AI contributes to managing electronic health records, reducing the burden on healthcare professionals, improving patient rehabilitation, and ensuring better treatment compliance. Supporting this view, Ahmed Alkuwaiti, Nazer, and Al-Reedy (2023) provided a comprehensive overview of AI's role in healthcare, emphasizing its applications in medical imaging, diagnostics, and virtual patient care. Their study, published in Frontiers in Digital Health (Vol. 5, pp. 1–15), also highlights the potential and limitations of AI, particularly the barriers to its widespread adoption.

ii. **Shuroug A. Alowais, Sahar S. Alghamdi, Nada Alsuhebany, Tariq Alqahtani, Abdulrahman I. Alshaya. Revolutionizing Healthcare: The Role of Artificial Intelligence in Clinical Practice A Comprehensive Study, BMC Medical Education**

Healthcare systems are inherently complex, posing challenges for patients, providers, and policymakers. Artificial Intelligence (AI), however, is emerging as a transformative force across many domains, including healthcare, with the potential to enhance patient care and overall quality of life. Supporting this, Alowais et al. (2023) presented an up-to-date overview of AI applications in clinical practice, emphasizing its contributions and the challenges it faces. Their study, published in BMC Medical Education (Vol. 23, pp. 5–12), underscores both the promise and the practical hurdles of AI adoption in healthcare.

iii. **Vidhya Rekha Umapathy 1, Suba Rajinikanth B 2, Rajkumar Densingh Samuel Raj 3, Sankalp Yadav. Perspective of Artificial Intelligence in Disease Diagnosis. A Comprehensive Study, Frontiers in Digital Health**

Artificial Intelligence (AI) shows immense potential for current and future disease diagnosis. Today, AI-powered tools assist physicians in interpreting medical images such as X-rays, MRIs, and CT scans, enabling faster and more accurate diagnoses. AI algorithms can also analyze patient history, symptoms, and data to support clinical decision-making. However, challenges related to reliability, interpretability, and clinical integration remain. Umapathy et al. (2023) provide a detailed review of AI's role in medical diagnostics, highlighting its ability to enhance diagnostic accuracy and efficiency. Their study also discusses the existing limitations and the need for cautious implementation. The review was published in Frontiers in Medicine (Vol. 10, pp. 4–10). Overall, AI continues to redefine the future of medical diagnosis.

.

iv. **Junaid Bajwa, Usman Munir, Aditya Nori, Bryan Williams. Artificial Intelligence in Healthcare: Transforming the Practice of Medicine, A Comprehensive Study, npj Digital Medicine**

Healthcare systems worldwide face growing challenges in achieving the 'quadruple aim': improving population health, enhancing patient and caregiver experiences, and reducing care costs. Factors such as aging populations, rising chronic diseases, and increasing healthcare expenses demand innovation in care delivery models. It helps healthcare providers deliver more precise, data-driven, and personalized care. Bajwa et al. (2021) emphasize the transformative potential of AI in healthcare, highlighting key breakthroughs in its application. They propose a strategic roadmap for developing and integrating effective AI systems. Their study, published in npj Digital Medicine (Vol. 4, pp. 3–8), outlines how AI can reshape medical practice. This reinforces the urgent need for digital transformation in global healthcare systems..

v. **Theresa Schachner, Roman Kelle,, Florian V Wangenheim, Artificial Intelligence-Based Conversational Agents for Chronic Conditions: Systematic Literature Review A Comprehensive Study, Journal of Medical Internet Research**

AI-powered chatbots are increasingly used in healthcare, especially to support patients with chronic diseases. These conversational agents enable frequent, personalized interactions, improving patient education and long-term care. Schachner et al. (2020) reviewed AI-based chatbots, focusing on their architecture and role in managing chronic conditions. Published in the Journal of Medical Internet Research (Vol. 22, pp. 4–9), their study highlights the growing potential of AI in enhancing chronic disease management..

vi. **Yogesh Kumar, Apeksha Koul, Ruchi Singla, Muhammad Fazal Ijaz. Artificial Intelligence in Disease Diagnosis: A Systematic Literature Review. A Comprehensive Study, Journal of Ambient Intelligence and Humanized Computing.**

Artificial intelligence (AI) supports patient care through advanced systems that aid in disease diagnosis, drug discovery, and risk assessment. Techniques like machine learning and deep learning are applied across various medical data sources—such as MRI, CT scans, mammography, and genomics—to enhance diagnostic accuracy. AI also improves hospital experiences and streamlines patient rehabilitation. Kumar et al. (2021) conducted a comprehensive review of AI techniques used in diagnosing a wide range of diseases. Their study, published in Journal of Ambient Intelligence and Humanized Computing (Vol. 12, pp. 5–15), highlights the growing impact of AI in healthcare. It underscores AI’s potential to transform clinical diagnostics through intelligent systems..

vii. **Abdullah Aldwean, Dan Tenney. Artificial Intelligence in Healthcare Sector: A Literature Review of the Challenges. A Comprehensive Study, Open Journal of Business and Management**

Artificial Intelligence (AI) promises transformational changes in healthcare, including improved diagnostic accuracy, personalized treatments, and reduced administrative burdens. It seeks to provide a deeper understanding of barriers preventing healthcare systems from fully benefiting from AI. Aldwean and Tenney (2021) conducted a comprehensive review focusing on these challenges. Their work identifies major obstacles and offers recommendations to facilitate AI integration. Published in the Open Journal of Business and Management (Vol. 9, pp. 4–13), their study sheds light on the cautious pace of AI scaling. Understanding these factors is essential for accelerating AI adoption in healthcare. Addressing these challenges can unlock AI's full potential to transform the sector..

viii. **Silvana Secinaro, Davide Calandra, Aurelio Secinaro, Vivek Muthurangu, Paolo Biancone. The Role of Artificial Intelligence in Healthcare: A Structured Literature Review. A Comprehensive Study, BMC Medical Informatics and Decision Making.**

Artificial intelligence (AI) encompasses computational technologies that mimic human intelligence, including deep learning, adaptation, and decision-making. Some AI systems perform tasks traditionally requiring human interpretation, making AI highly interdisciplinary with applications across many fields, notably medicine and healthcare. Secinaro et al. (2021) provide a structured review of AI's role in healthcare, highlighting its applications in clinical decision-making across various medical specialties. Their study, published in BMC Medical Informatics and Decision Making (Vol. 21, pp. 4–10), also discusses challenges and future directions for AI in medicine. This comprehensive review underscores AI's expanding impact on healthcare delivery and research..

ix. **Milad Mirbabaie, Stefan Stieglitz & Nicholas R. J. Frick. Artificial Intelligence in Disease Diagnostics: A Critical Review and Classification. A Comprehensive Study, Springer and the IUPESM in cooperation with the World Health Organization.**

Accurate disease diagnosis is critical for effective treatment and patient well-being, yet human error often complicates this complex task. Artificial intelligence (AI) has the potential to enhance diagnostic accuracy and efficiency by assisting in the interpretation of medical data. Mirbabaie, Stieglitz, and Frick (2021) provide a critical review of AI in disease diagnostics, mapping the current landscape and identifying gaps for future research.

Their work, published by Springer in cooperation with the World Health Organization (Vol. 11, pp. 8–18), also proposes a research agenda to guide further advancements. This review highlights the growing importance of AI in digital healthcare services and its potential to transform diagnostics..

x. **Jiayi Shen, Casper J P Zhang, Bangsheng Jiang. Artificial Intelligence Versus Clinicians in Disease Diagnosis. A Comprehensive Study, JMIR Medical Informatics**
The aging population and shortage of medical professionals have intensified efforts to improve clinical service efficiency through information technology. Artificial intelligence (AI), which simulates human intellectual processes through algorithm-based applications, plays an increasingly vital role in medicine by promoting therapeutic development and optimizing patient care. Shen, Zhang, and Jiang (2019) systematically examined AI's diagnostic performance, comparing it to that of human clinicians. Their study, published in JMIR Medical Informatics (Vol. 7, pp. 3–10), highlights AI's growing potential to complement and enhance clinical decision-making.

xi. **Mahmoud Nasr, Md. Milon Islam, Shady Shehata, Fakhri Karray. Smart Healthcare in the Age of AI: Recent Advances, Challenges, and Future Prospects. A Comprehensive Study, IEEE**

The rising number of individuals with chronic conditions, including elderly and disabled populations, has created an urgent need for innovative, personalized healthcare models that move beyond traditional institutions like hospitals and nursing homes. Smart healthcare systems, driven by advances in artificial intelligence (AI) and machine learning (ML), have gained significant attention for their potential to transform care delivery. Nasr, Islam, Shehata, and Karray (2021) discuss recent advances in AI-powered smart healthcare, emphasizing wearable and smartphone devices for health monitoring, ML applications in disease diagnosis, and assistive technologies such as social robots for ambient assisted living environments. Their study, published by IEEE (pp. 2–8), also addresses current challenges and outlines future prospects for smart healthcare systems. This growing field promises to enhance personalized care and improve outcomes for patients with chronic ailments. The comparative analysis of the articles studied has been summarized below in Table 2.1.

Table No. 2.1 Table of Literature Review

SI. No.	Authors	Publication Year	Paper Title	Technology Used	Pros	Cons	Journal Name	Volume	Page No.	DOI
1.	Ahmed Alkuwaiti A ; Nazer K ; Al-Reedy	2023	A Review of the Role of Artificial Intelligence in Healthcare	AI in medical imaging, diagnostics, virtual patient care	Comprehensive overview of AI applications in healthcare	Discusses challenges in AI adoption	Frontiers in Digital Health	5	1-15	https://doi.org/10.3390/jpm13060951
2.	Shuroug A. Alowais, Sahar S. Alghamdi, Nada Alsuhebany, Tariq Alqahtani, Abdulrahman I. Alshaya	2023	Revolutionizing Healthcare: The Role of Artificial Intelligence in Clinical Practice	AI in disease diagnosis, treatment recommendations	Up-to-date overview of AI applications	Addresses challenges in AI adoption	BMC Medical Education	23	5-12	http://doi.org/10.1186/s12909-023-04698-z
3.	Vidhya Rekha Umapathy 1, Suba Rajinikanth B 2, Rajkumar Densingh Samuel Raj 3, Sankalp Yadav 4	2023	Perspective of Artificial Intelligence in Disease Diagnosis: A Review	AI in medical diagnostics	Enhances accuracy and efficiency in diagnosis	Discusses limitations and challenges	Frontiers in Medicine	10	4-10	https://doi.org/10.7759/cureus.45684

4.	Junaid Bajwa, Usman Munir, Aditya Nori, Bryan Williams	2021	Artificial Intelligence in Healthcare: Transforming the Practice of Medicine	AI in clinical decision-making	Outlines breakthroughs in AI applications	Discusses roadmap for effective AI systems	npj Digital Medicine	4	3-8	https://doi.org/10.7861/fhj.2021-0095
5.	Theresa Schachner, Roman Kelle,, Florian V Wangenheim	2020	Artificial Intelligence-Based Conversational Agents for Chronic Conditions: Systematic Literature Review	AI-based conversational agents	Reviews characteristics and AI architectures	Focuses on chronic diseases	Journal of Medical Internet Research	22	4-9	https://doi.org/10.2196/20701
6.	Yogesh Kumar, Apeksha Koul, Ruchi Singla, Muhammad Fazal Ijaz	2021	Artificial Intelligence in Disease Diagnosis: A Systematic Literature Review	AI in disease diagnosis	Comprehensive survey of AI techniques	Covers various diseases	Journal of Ambient Intelligence and Humanized Computing	12	5-15	https://doi.org/10.1007/s12652-021-03612-z
7.	Abdullah Aldwean, Dan Tenney	2021	Artificial Intelligence in Healthcare Sector: A Literature Review of the Challenges	AI in healthcare	Identifies challenges in AI adoption	Provides recommendations	Open Journal of Business and Management	9	4-13	https://doi.org/10.4236/ojbm.2024.121009
8.	Silvana Secinaro,	2021	The Role of	AI in clinical	Discusses AI	Covers challen	BMC Medic	21	4-10	https://doi.org/10.

	Davide Calandra, Aurelio Secinaro, Vivek Muthurangu, Paolo Biancone		Artificial Intelligence in Healthcare: A Structured Literature Review	decision-making	applications in various medical fields	ges and future directions	al Informatics and Decision Making			[1186/s12911-021-01488-9](https://doi.org/10.1186/s12911-021-01488-9)
9.	Milad Mirbabaie, Stefan Stieglitz & Nicholas R. J. Frick	2021	Artificial Intelligence in Disease Diagnostics: A Critical Review and Classification	AI in disease diagnostics	Provides a critical review of AI applications	Proposes a research agenda	Springer and the IUPESM in cooperation with the World Health Organization	11	8-18	http://doi.org/10.1007/s12553-021-00555-5
10.	Jiayi Shen, Casper J P Zhang, Bangsheng Jiang	2019	Artificial Intelligence Versus Clinicians in Disease Diagnosis	AI in disease diagnosis	Systematic examination of AI performance	Compares AI with human clinicians	JMIR Medical Informatics	7	3-10	https://doi.org/10.2196/10010
11.	Mahmoud Nasr, Md. Milon Islam, Shady Shehata, Fakhri Karray	2021	Smart Healthcare in the Age of AI: Recent Advances, Challenges, and Future Prospects	AI in smart healthcare systems	Discusses AI applications in health monitoring and disease diagnosis	Addresses challenges and future directions	IEEE	NA	2-8	http://doi.org/10.48550/arXiv.2107.03924

CHAPTER 3

Requirement Analysis

3.1 Functional Requirements:

i. Purpose and Functionality of the App

The primary objective of the MediMind AI application is to accurately detect and classify skin conditions such as Acne, Eczema, Melanoma, or Unknown based on uploaded skin images. The app leverages deep learning models to analyze visual features and provide disease classification along with confidence scores. This functionality not only assists users in identifying potential dermatological issues but also enhances awareness of critical conditions like skin cancer. The system also delivers medical explanations and treatment guidance, making it useful for educational and preliminary diagnostic purposes.

ii. User Input via Interactive Interface

MediMind AI provides an easy-to-use Gradio-based web interface, where users can upload skin images, choose their preferred language (from six supported languages), and select a prediction mode—either a Single Model Prediction or a Compare All Models option. If the Single Model mode is chosen, users can pick from three trained models: MobileNetV3, ResNet50, or DenseNet121. The input process is designed to be simple and intuitive, allowing even non-technical users to interact effortlessly with the system.

iii. Deep Learning Model Integration

At the core of MediMind AI are three deep learning models—MobileNetV3, ResNet50, and DenseNet121—each trained on a curated dataset of skin disease images. These models utilize convolutional neural networks (CNNs) to extract visual patterns from skin images and classify them with high accuracy. Once a user uploads an image and submits their input, the selected model processes the data to generate a prediction. The "Compare All Models" feature additionally runs all models in parallel, allowing users to view side-by-side results, enhancing interpretability and trust in the system.

iv. Output and Result Presentation

After processing the input, the app displays several types of outputs: the predicted disease class, a confidence score, a bar chart of prediction probabilities, a Grad-CAM heatmap showing model attention, and a text-based medical explanation. The explanation is also converted into voice output using text-to-speech. Users can additionally download a PDF report summarizing the diagnosis, model used, explanation, and prediction confidence. These layered outputs make the system transparent, user-friendly, and medically informative.

v. Gradio-Based Web Deployment

The MediMind AI system is fully deployed using Gradio, an open-source Python framework designed for creating machine learning web interfaces. Gradio provides real-time updates, voice support, multilingual handling, and support for cloud deployment platforms like Hugging Face Spaces. The deployment ensures that the app can be accessed by users across various devices and browsers without requiring any technical setup, making MediMind AI widely usable and scalable.

The MediMind AI system offers several functional features designed to enhance usability, flexibility, and diagnostic performance. The system accepts multiple forms of input including image upload, voice commands, and text input, allowing users to interact in the most convenient way. Once input is provided, the app processes it using deep learning models such as MobileNetV2, ResNet50, and DenseNet121. Users can choose a single model for prediction or opt to compare outputs from all available models in a comprehensive comparison table. To ensure accessibility. Additionally, the app displays the final processed input, prediction result, and include confidence scores to help users understand the model's certainty. The system also integrates out-of-distribution (OOD) detection using Softmax probability, helping to identify and reject irrelevant or unknown inputs. The entire interface is built using Gradio, offering a simple and intuitive user experience, and is planned to be hosted on Hugging Face for easy access via web browser.

3.2 Non-Functional Requirements:

i. Usability and User Interface Responsiveness

MediMind AI emphasizes an intuitive and accessible user interface that caters to both general users and medical professionals. The Gradio-based UI is designed to be simple, responsive, and compatible across devices and screen sizes. Interactive components such as radio buttons, image uploaders, and output displays are structured clearly to guide the user through each step. This ease of use is essential to ensure that users—regardless of technical background—can navigate the system, interpret results, and engage with AI-generated feedback without confusion or delay.

ii. Real-Time Performance and Speed

The application is engineered to deliver quick and efficient predictions immediately after input submission. Since MediMind AI leverages pre-trained deep learning models and cloud-hosted inference (via Hugging Face and Google Colab), response times are optimized to provide disease classification and explanations within seconds. This real-time processing enhances the overall user experience and maintains flow, particularly important for health-related applications where delays may affect user trust and satisfaction.

iii. Reliability and Prediction Accuracy

Ensuring high reliability and prediction accuracy is central to MediMind AI's purpose. The system is backed by robust models like MobileNetV3, ResNet50, and DenseNet121, all trained on diverse and well-annotated skin disease datasets. These models consistently return accurate results across all tested conditions—Acne, Eczema, Melanoma, and Unknown—with high precision and recall. Users can rely on the output, supported by confidence meters and Grad-CAM heatmaps, which make the system trustworthy even in ambiguous cases.

iv. Scalability, Portability, and Lightweight Design

MediMind AI is designed with scalability in mind. The modular codebase and flexible interface allow for future integration of additional disease classes, new models, or AI features without major redevelopment. The platform is also portable, as it supports Windows, macOS, and Linux systems, and can be hosted online with minimal hardware resources. Despite its capabilities, the system remains lightweight—optimized for fast loading and smooth performance—making it easy to deploy and maintain in both academic and public use cases.

v. Maintainability and Data Security

The MediMind AI codebase is structured using modular design principles, enabling easier maintenance, debugging, and future updates. Components like the voice modules, prediction logic, and UI are separated for better control. Although the system does not store user data or images, all user inputs are processed securely during runtime, ensuring data privacy and ethical use. This secure-by-design approach builds user confidence and aligns the application with modern standards of responsible AI development.

From a non-functional perspective, MediMind AI is designed to be accessible to both medical professionals and the general public, including those in rural or underserved areas. The application emphasizes real-time performance, ensuring predictions are generated within seconds of receiving input. Its scalable design allows multiple users to access the platform simultaneously without significant lag or delays. The user interface is developed to be highly intuitive, requiring no technical background to operate. Compatibility across various devices and web browsers ensures broad usability, while privacy and security measures prevent the storage or misuse of user data. Most importantly, the system is reliable and consistently delivers accurate results while gracefully handling errors or invalid inputs, making it suitable for real-world healthcare applications.

3.3 Hardware Requirements

A. My System

i. CPU: Intel Core i3
ii. RAM: 4 GB
iii. Storage: 256 GB SSD
iv. Stable Internet Connection: Required for accessing Gradio app

B. Technical Hardware Specifications For Implementations

i. Processor: Intel Core i5 (7th Gen or higher)
ii. RAM: Minimum 4 GB
iii. Storage: 512 GB SSD or higher
iv. Audio Output: Functional speakers or headphones to hear voice responses
v. Stable Internet Connection: Required for accessing Gradio app, and Hugging Face app

The MediMind AI system requires a laptop or desktop with a multi-core processor (i3 or higher) and at least 4 GB of RAM to smoothly operate the Gradio interface and AI-powered features. While training and inference tasks are performed using Google Colab with GPU support, a local machine is essential for development and testing. A speaker setup is necessary to enable the voice output features. A high-speed internet connection is also crucial for API interactions and accessing the deployed application online.

3.4 Software Requirements

A. My System

i. Operating System: Windows 10
ii. Python Version: 3.11

iii. Google Colab: For training and fine-tuning deep learning models
iv. Gradio Version: 5.30 (for building the user interface)
v. VS Code version 1.95.3

B. Software Specifications For Implementations

i. Operating System: Windows 11, macOS 12
ii. Python: Version 3.10 or higher
iii. IDE/Text Editor: VS Code 1.95.3 or newer And Google Colab
iv. Gradio: Version 5.13 or higher
v. Browser Compatibility: Chrome, Firefox, Brave
vi. Deployment Option (Optional): Hugging Face CLI

The software stack of MediMind AI primarily relies on Python due to its excellent support for machine learning and AI development. Training is carried out in Google Colab using pre-trained models such as MobileNetV2, ResNet50, and DenseNet121 with TensorFlow and Keras. The interactive interface is built using Gradio, which also supports voice output through speech recognition and text-to-speech libraries. Hugging Face is used to deploy the application publicly. The system remains compatible with Windows, macOS, and Linux environments, ensuring wide usability.

3.5 Data Flow Diagram (DFD)

The Level 0 Data Flow Diagram (DFD) of MediMind AI illustrates the system's overall function in a simplified manner. At this level, the system is represented as a single process named "Image Analysis", which is the central operation of the MediMind AI system. The diagram begins with the User, who is the external entity interacting with the system. The user uploads or provides an input image—typically a skin image for diagnosis purposes. This image is sent as data into the Image Analysis process. The system then performs internal tasks such as preprocessing the image, running classification models (like MobileNet, ResNet, etc.), and applying AI logic for disease detection. After analyzing the image, the process generates results such as predicted disease type, confidence score, and possibly explanations or visualizations. Finally, the Output is delivered back to the user.

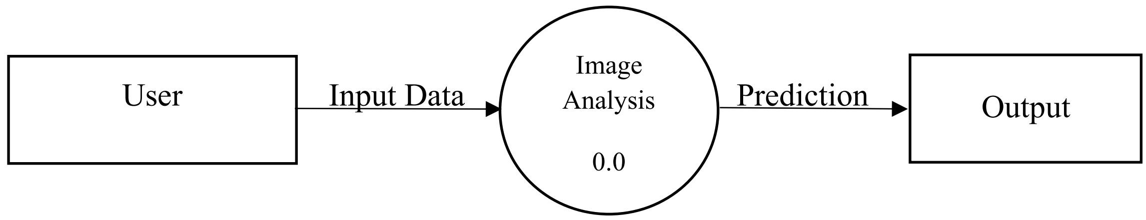

Figure 3.1 : Level 0 DFD of proposed work for MediMind AI

The Level 1 Data Flow Diagram (DFD) of the proposed MediMind AI system decomposes the main process (from Level 0) into three sub-processes: Input Module (1.1), Processing Module (1.2), and Output Module (1.3). The process begins when the User interacts with the Input Module (1.1) by uploading a skin image through the interface. This uploaded image is then passed to the Processing Module (1.2), which is responsible for handling the core analysis. The Processing Module communicates with the Database to fetch any necessary reference data or models to aid in disease detection. Once the processing is completed, the system sends the analysis result to the Output Module (1.3). The Output Module organizes and formats the result, and then the final diagnosis or prediction is displayed back to the User, completing the data flow. This Level 1 diagram gives a clearer view of how different internal modules work together and how data flows from the user's input to the final output in the MediMind AI.

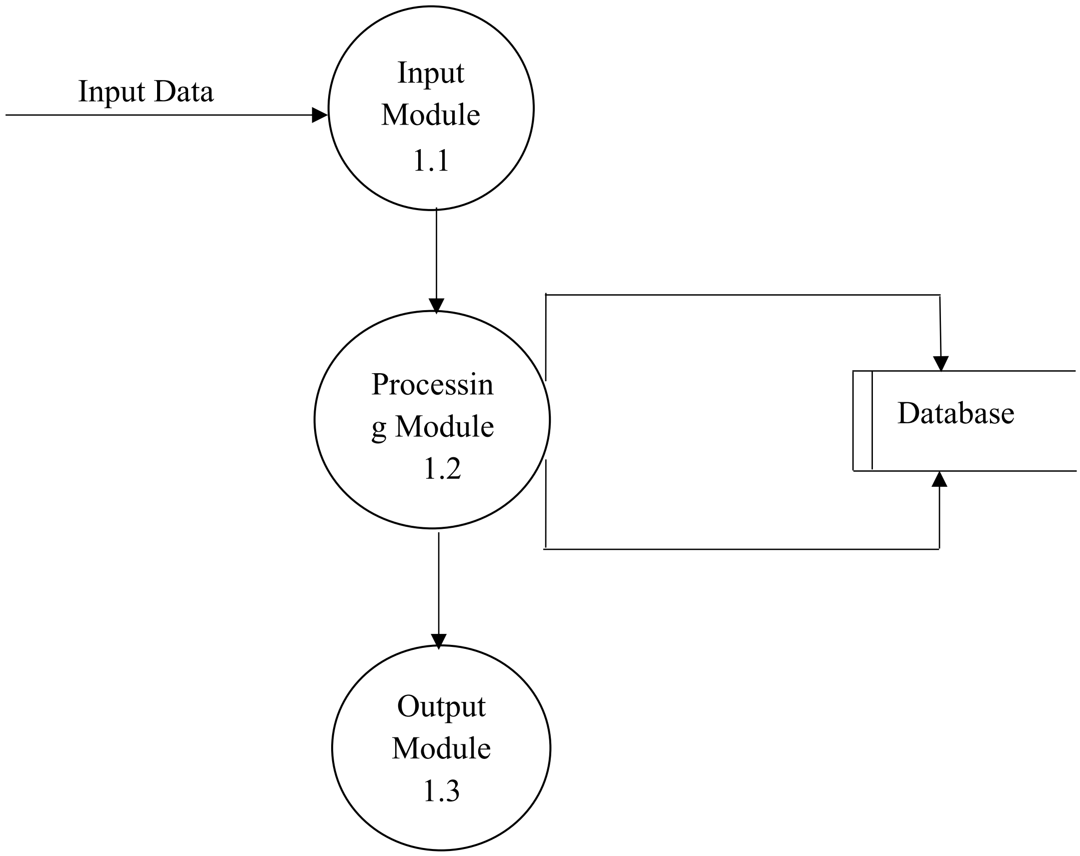

Figure 3.2 : Level 1 DFD of the proposed work for MediMind AI

3.6 Activity Diagram

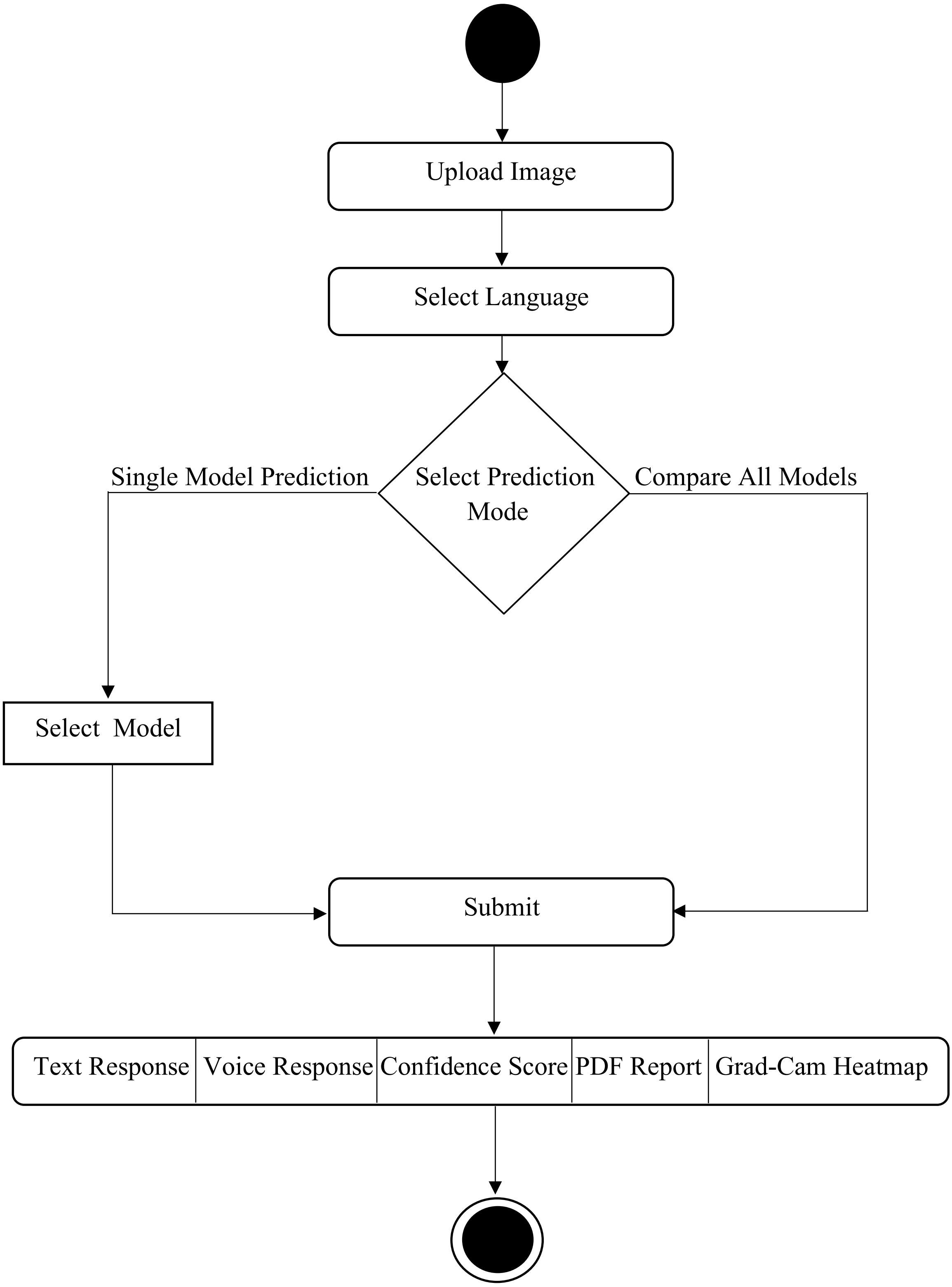

Figure 3.3 : Activity Diagram of the proposed work for MediMind AI

3.7 Use Case Diagram

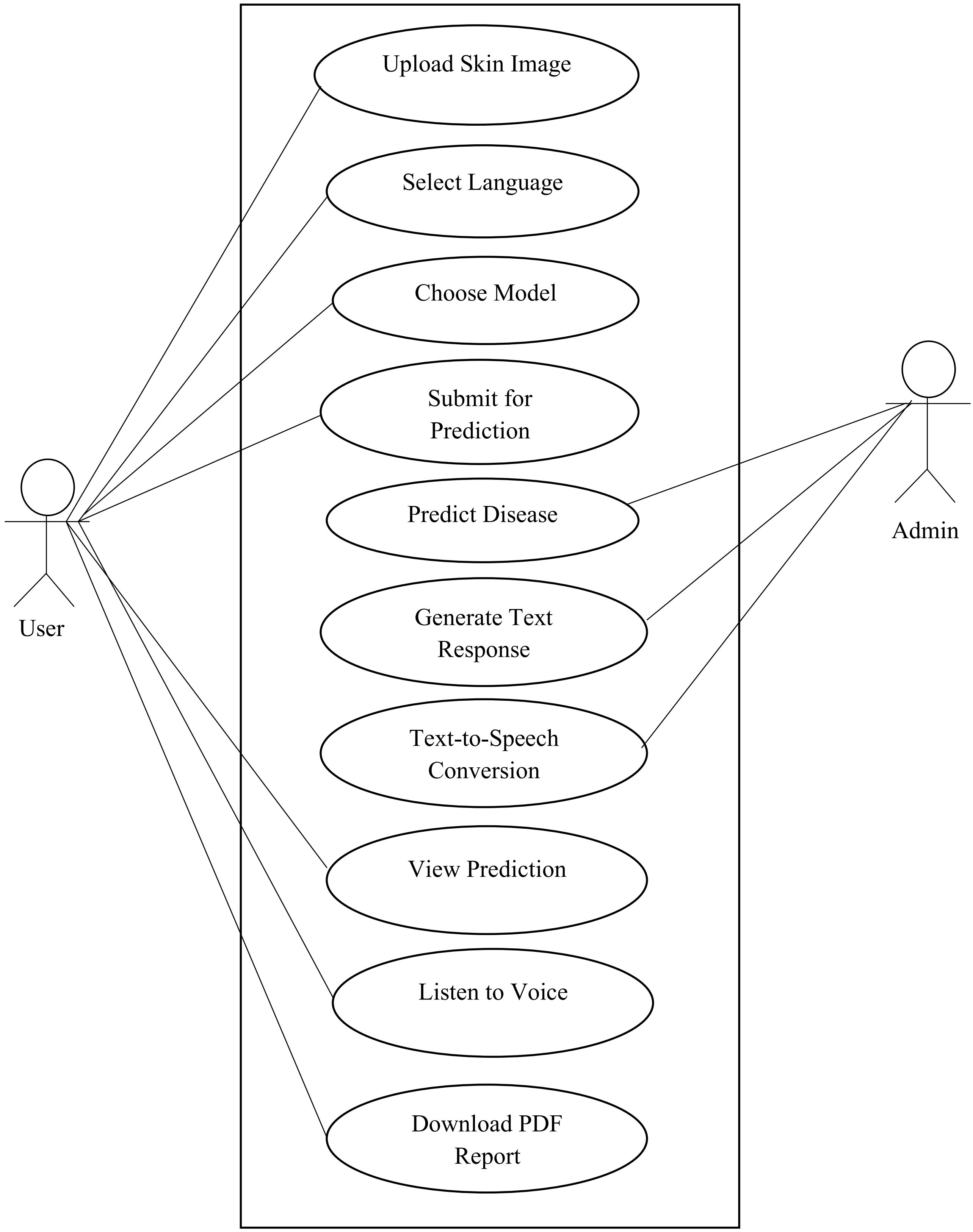

Figure 3.4 : Use Case Diagram of the proposed work for MediMind AI

CHAPTER 4

Methodology

The proposed methodology for the MediMind AI system outlines the architectural, functional, and algorithmic aspects behind the design and implementation of this intelligent skin disease diagnostic tool. MediMind is built with a goal to aid users including general public and medical staff in identifying common skin conditions such as Acne, Eczema, Melanoma, and distinguishing unknown or unrelated inputs with precision. This chapter discusses each stage of the system development life cycle, from system planning and design to model inference and output generation. It includes a block diagram, workflow description, interface details, and the specific technologies and algorithms used in building the project.

4.1 System Design and Planning

The planning and design phase of the MediMind AI system was initiated by clearly identifying the goals and the problem it aims to solve. The primary focus was to create a deep learning-based medical assistant that can classify skin conditions from images while also providing additional voice/text interaction capabilities. The planning involved choosing the right AI models for image classification, integrating speech-to-text and text-to-speech features, selecting the appropriate framework for deployment, and enabling language support for both Hindi and English users. Furthermore, emphasis was placed on ensuring high accuracy, model robustness, user-friendliness, and real-time performance. The system was designed to allow users to choose a model for prediction or compare multiple models' outputs, making it both educational and diagnostic. The data used in training the models consisted of thousands of annotated images categorized into four main groups: Acne, Eczema, Melanoma, and Unknown. Throughout the planning, privacy and accessibility were given high importance, making sure the system does not store user data post-prediction.

4.2 Block Diagram of the System

The block diagram of MediMind AI outlines the complete workflow of the skin disease classification system. It begins with Data Collection, where a dataset of labeled skin disease images, including conditions like Acne, Eczema, Melanoma, and Unknown categories, is gathered from reliable sources. The next stage is Pre-processing, which involves resizing, normalizing, and cleaning the images to ensure uniformity and enhance the quality for analysis. Following this, Feature Extraction is performed using convolutional neural networks (CNNs) to automatically learn relevant patterns, textures, and visual characteristics from the input images. These extracted features are then passed into the Model Building stage, where models such as MobileNetV3, ResNet50, and DenseNet121 are trained and fine-tuned on the preprocessed data. After training, the system proceeds to Evaluating Model, where the performance of each model is assessed using metrics such as accuracy, precision, recall, and F1-score. Finally, the results are shown in the visualization stage through prediction labels, Grad-CAM heatmaps, and PDF reports.

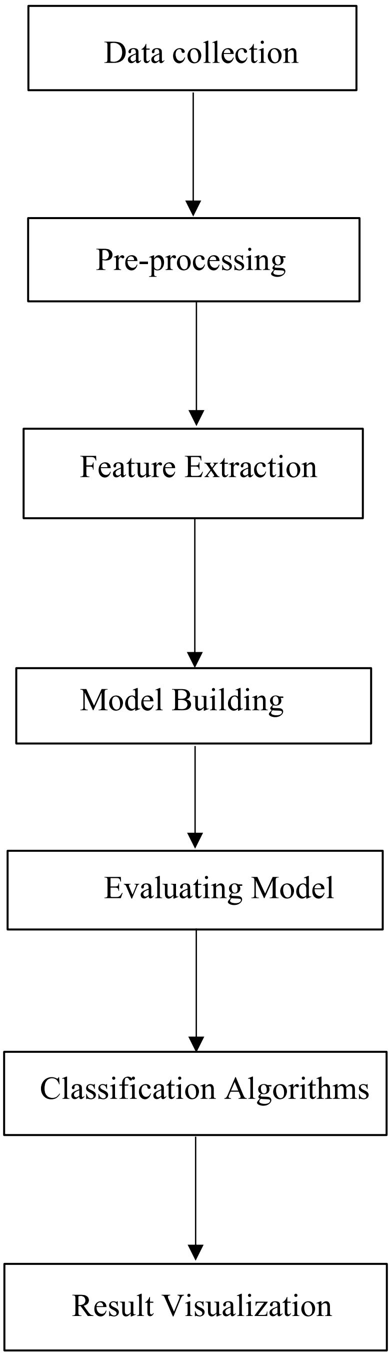

Figure 4.1 : Block diagram of MediMind AI

4.3 Work Flow

The workflow of the MediMind AI system follows a systematic and sequential process that enables smooth user interaction and accurate diagnosis. Initially, the user uploads an image of a visible skin condition. This is followed by language selection, which ensures that the system communicates in a language the user is comfortable with. The user then selects either a single model or a comparison mode where all models are used to generate predictions. Once the input is received and settings are configured, the system processes the image using predefined preprocessing steps and prepares the input for inference. Each model predicts the skin disease category and outputs the result along with a confidence score. If multiple models are selected, the system displays all outputs in a comparison table, helping the user understand how different models interpret the same input. The final output is displayed on screen and is also converted into voice for auditory feedback. Lastly, users have the option to generate a detailed report in PDF format which includes the input image, disease prediction, confidence level, and medical description of the disease. This workflow ensures a smooth, accessible, and informative diagnostic experience.

4.4 Graphical User Interface Development

The user interface of MediMind AI is developed using Gradio, an open-source library that allows seamless creation of interactive web-based applications. The GUI is intuitive and easy to navigate, designed to accommodate users with varying levels of technical expertise. The main interface components include image upload, language selection menu, model selection menu, submit button, and result display section. Upon submission, the result is displayed either as a single output or in tabular form if comparison mode is selected. The interface also includes features like voice response, and a button to download the diagnosis report as a PDF file. The use of Hindi and English support makes it friendly for diverse user bases across India. The result section not only shows the prediction but also displays confidence scores, explanatory text about the diagnosed condition, and vocal feedback through a text-to-speech module. Overall, the GUI enhances user engagement and supports interactive AI-powered diagnosis.

4.5 System Structure

The MediMind AI system follows a modular, component-based structure to ensure maintainability, performance, and scalability. The system is divided into key modules such as user input module, preprocessing module, model inference engine, output module, and voice interaction module. Each component operates independently but is seamlessly integrated for full system functionality. The frontend interface, developed in Gradio, acts as the primary user interaction layer. It collects input, displays results, and manages user commands. The backend consists of the AI model execution layer which is connected to the Groq API. This connection allows for extremely fast inference time even for deep learning models. Image data is processed using TensorFlow/Keras and OpenCV, while voice input is managed using Gradio's speech recognition. The entire system is hosted on Hugging Face Spaces, allowing global access without any need for software installation. The data flow is managed through internal pipelines, ensuring that no sensitive data is stored or misused. All modules are designed to work independently and can be upgraded or replaced without affecting the entire system.

4.6 System Design

The MediMind AI system is based on a client-server design, where the client (Gradio interface) interacts with the server-side AI engine. The core AI engine is designed using state-of-the-art CNN architectures like ResNet50, DenseNet121, and MobileNetV2. These models are fine-tuned on a custom skin disease dataset. The models are capable of identifying and classifying three specific skin conditions—Acne, Eczema, and Melanoma—and also include a robust "Unknown" category to handle non-target or ambiguous inputs. This category helps filter out irrelevant or risky predictions. The system uses image preprocessing techniques such as resizing to 224x224 pixels, normalization, and data augmentation to enhance model generalization. Based on model selection, the backend loads the respective models. The prediction output is displayed with confidence scores and can be converted into voice for users who prefer auditory responses. The architecture is lightweight, scalable, and deployable on any cloud platform.

4.7 Techniques and Algorithms Used

The MediMind AI system integrates multiple cutting-edge algorithms and techniques across the fields of machine learning, deep learning, image processing, and natural language processing. Each of these components plays a crucial role in ensuring the system performs reliably, accurately, and efficiently. The following sub-sections explain all the major techniques and algorithms used in the project.

4.7.1 Transfer Learning

Transfer learning is a machine learning technique where a model trained on a large dataset (such as ImageNet) is adapted for a new, smaller dataset. In the MediMind project, transfer learning was implemented using four powerful convolutional neural network (CNN) architectures: ResNet50, MobileNetV2, and DenseNet121. These models were initially trained on millions of general-purpose images, enabling them to extract high-level features such as edges, textures, and shapes. By fine-tuning these pretrained models on a specialized skin disease dataset, the training time was significantly reduced while achieving high accuracy. The final layers were customized to output four categories: Acne, Eczema, Melanoma, and Unknown. Fine-tuning involved unfreezing some of the top layers of the pretrained model and training them with a low learning rate to adapt the features to the medical domain.

4.7.2 Data Preprocessing and Augmentation

Proper data preprocessing ensures that all images are standardized before feeding them into the model. In this project, images were resized to 224x224 pixels to match the input shape of CNN models. Normalization was applied to scale pixel values to the range of 0–1, which helps the model converge faster during training. To improve generalization and reduce overfitting, data augmentation techniques were applied. These included random horizontal flipping, rotation, zooming, brightness adjustment, and cropping. By generating slight variations of the original dataset, the model becomes more robust and better at handling unseen data.

4.7.3 Softmax Probability Thresholding for Out-of-Distribution Detection

Softmax probability thresholding is used to detect when the model is uncertain about a prediction. In a typical classification problem, the model outputs probabilities for each class using the softmax function. However, a high softmax score can be misleading if the input doesn't belong to any known class. To overcome this, a threshold (e.g., 0.5) is set. If the highest softmax score is below the threshold, the system flags the input as "Unknown", indicating that the image likely does not belong to Acne, Eczema, or Melanoma. This helps reduce false positives and improves the reliability of the diagnosis.

4.7.4 Custom "Unknown" Class in Dataset

Instead of relying only on softmax thresholding, the project introduced an explicit 'Unknown' class during training. This class includes images of non-human skin, other skin diseases, random objects, animal textures, and healthy skin. By including this diverse set of images in the dataset and labeling them as "Unknown," the model learns to differentiate between known disease categories and irrelevant or unexpected inputs. This technique significantly improves out-of-distribution (OOD) detection and makes the system safer for real-world applications, especially in medical diagnosis where a wrong prediction can lead to harmful decisions.

4.7.5 PDF Report Generation and Export

After the diagnostic process is complete, the system offers users the option to download a detailed PDF report of their results. This report includes the uploaded image, predicted disease class, confidence score, and a brief medical explanation of the disease. The report generation is implemented using Python's reportlab or fpdf libraries and is triggered on the backend upon user request. This feature makes it easier for users to keep a personal record or consult with a healthcare professional later.

4.7.6 Gradio for Frontend User Interface

The GUI for MediMind was built using Gradio, a Python-based interface-building tool that supports deep learning model interaction via a web browser. Gradio provides components like image uploaders, text boxes, audio recorders, dropdowns, and tables which are essential for the MediMind interface. It allows easy integration of real-time model predictions, making the system interactive and responsive. Furthermore, it supports fast prototyping and deployment on platforms like Hugging Face Spaces without requiring a dedicated web development framework.

4.7.7 Model Comparison Table and Evaluation

The comparison mode feature in MediMind allows users to select multiple models and view their predictions side-by-side. A table is generated showing predicted class, confidence score, and reasoning for each model. This feature is implemented using conditional logic in Python and displayed using Gradio's table components. This comparison approach helps the user or medical staff understand how each model behaves and increases confidence in the diagnosis by providing model consensus or highlighting disagreements between predictions.

4.7.8 MobileNetV2

MobileNetV2 is a lightweight and efficient deep learning model designed for mobile and embedded vision applications. It uses depthwise separable convolutions and inverted residuals to reduce computational cost while maintaining high accuracy. In MediMind AI, MobileNetV2 is implemented for real-time skin disease detection, especially for scenarios where faster predictions and lower latency are required, such as deployment on edge devices or mobile phones.

4.7.9 ResNet50

ResNet50 is a 50-layer deep convolutional neural network known for its residual learning framework, which helps solve the vanishing gradient problem in deep networks. Residual blocks allow the model to learn complex features without degrading performance in deeper layers. In MediMind AI, ResNet50 is used for its robustness and ability to accurately classify medical images, making it a strong choice for high-performance disease prediction tasks.

4.7.10 DenseNet121

DenseNet121 is another powerful CNN architecture where each layer is directly connected to every other layer in a feed-forward fashion. This improves feature propagation, encourages feature reuse, and reduces the number of parameters. DenseNet121 was implemented in MediMind AI for its excellent performance in medical image classification and its ability to capture fine-grained features of skin diseases. Its compact architecture and strong feature extraction capabilities contribute significantly to high classification accuracy.

CHAPTER 5

Model Testing

Model testing is a crucial phase in any AI or deep learning-based system, especially in a medical diagnostic assistant like MediMind AI. This phase ensures that the trained models are evaluated rigorously for their accuracy, robustness, generalization capability, and overall performance in predicting the correct skin disease category. As the system is designed to assist both general users and medical professionals, model testing plays a key role in determining whether the predictions made by the models are reliable and precise. The goal of testing is to evaluate the effectiveness of the trained models on unseen data and identify which model offers the best trade-off between accuracy and computational efficiency.

5.1 Testing Environment

All model testing was conducted in a cloud-based environment using Google Colab, which provides access to free GPU acceleration. This was essential due to the high computational requirements of training and testing deep learning models. The dataset used for testing consisted of four skin disease categories: Acne, Eczema, Melanoma, and Unknown. The images were stored in the Google Drive directory at /content/drive/MyDrive/Skin_Disease_Dataset, and the models were trained using a standard image size of (224, 224) with a batch size of 32. TensorFlow and Keras were the primary frameworks used for model development, and each model was tested under the same conditions for fair comparison.

5.2 Evaluation Metrics Used

To assess the performance of each model, a variety of evaluation metrics were used:

a. **Accuracy**: The percentage of correct predictions made by the model out of all predictions.

b. **Precision**: The ratio of correctly predicted positive observations to the total predicted positives. It indicates how precise the model is in identifying specific classes.

c. **Recall (Sensitivity)**: The ratio of correctly predicted positive observations to all actual positives. It reflects how well the model captures true cases of each disease.

d. **F1-Score**: The harmonic mean of Precision and Recall. It gives a balance between the two and is a better measure than accuracy when the class distribution is imbalanced.

e. **Confusion Matrix**: A matrix that shows the number of true positives, true negatives, false positives, and false negatives for each class. It gives detailed insight into model performance per category. These metrics provide a comprehensive view of how well each model generalizes to new, unseen data and helps us select the most effective architecture for deployment.

5.3 Models Tested

Several state-of-the-art deep learning architectures were evaluated for this project. The models tested include: MobileNetV2: Known for its lightweight architecture and efficiency, ideal for mobile deployment. ResNet50: A deep residual network that uses skip connections to improve accuracy and prevent vanishing gradients. DenseNet121: A densely connected network that enhances feature propagation and encourages feature reuse. Each of these models was trained and then evaluated on the same dataset to ensure a fair comparison.

5.4 Testing Procedure

The testing procedure was carefully designed to maintain consistency across all models. Load Trained Model: Each model was loaded from the saved .keras file. Prepare Test Dataset: Using image_dataset_from_directory, test images were loaded and pre-processed. Model Prediction: Predictions were made using the model.predict() function. Evaluate Results: Metrics were calculated using classification report and confusion matrix. Visualize Output: Graphs of training/validation accuracy and loss were plotted.

5.5 Testing Results - MobileNetV3 Model

MobileNetV3 is known for its lightweight architecture and high efficiency, making it particularly suitable for deployment on mobile or edge devices. For testing purposes, the trained MobileNetV3 model was evaluated using a sample of unseen images from each of the four categories: Acne, Eczema, Melanoma, and Unknown. The testing was conducted in a cloud-based Google Colab environment with GPU acceleration, ensuring consistent performance evaluation across all models. The following results highlight the model's ability to accurately classify different skin conditions, along with confidence scores and analysis.

i. Image Class – Acne

Figure 5.1: MobileNetV3 correctly detected acne with high confidence

Prediction Details And Analysis : The model successfully identified the image as Acne with a confidence score of 99.99%, indicating a strong certainty in its decision. The image shows typical features of Acne such as clustered pustules and inflamed spots, which the model learned during training. This confirms that the model generalizes well for this class on unseen data.

ii. Image Class – Eczema

Figure 5.2 : MobileNetV3 correctly detected eczema with high confidence

Prediction Details And Analysis : MobileNetV3 classified the image as Eczema with a confidence of 99.98%, which indicates a reasonably high certainty. The visual features in the image, such as dry patches and redness, were correctly interpreted by the model.

iii. Image Class – Melanoma

Figure 5.3 : MobileNetV3 correctly detected melanoma with high confidence

Prediction Details And Analysis : In this case, the model accurately classified the image as Melanoma with a confidence score of 99.99%. The image shows typical melanoma features such as asymmetry, irregular borders, and uneven pigmentation, which were well captured by MobileNetV3.

d. Image Class – Unknown

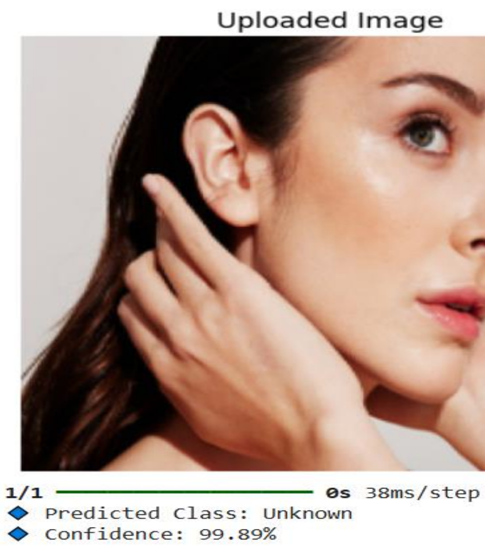

Figure 5.4: MobileNetV3 correctly detected unknown with high confidence

Prediction Details And Analysis : The model correctly classified this image as belonging to the Unknown category with a confidence score of 99.89%. The image was selected from non-skin or distractor categories. The successful prediction shows that MobileNetV3 is capable of recognizing out-of-category images to a certain extent, which is critical for real-world deployment where unknown inputs can occur.

5.6 Testing Results – ResNet50 Model

ResNet50 is a powerful deep convolutional neural network that uses residual learning and skip connections to improve performance and prevent vanishing gradients in deep architectures. It is well-suited for complex image classification tasks such as skin disease detection due to its ability to learn deeper feature representations. The trained ResNet50 model was tested using a sample of unseen images from the four target categories: Acne, Eczema, Melanoma, and Unknown. All tests were conducted on the Google Colab platform with GPU acceleration to ensure consistent evaluation. The following results display the model's predictions along with confidence scores and detailed analysis for each category.

i. Image Class – Acne

Figure 5.5 : ResNet50 successfully predicted acne regions

Prediction Details and Analysis : The ResNet50 model classified this image as Acne with a confidence score of 99.97%, indicating very high certainty. The image presents typical acne features like inflamed pustules and comedones. The model's ability to detect these visual patterns confirms its effectiveness in handling acne cases with strong generalization on unseen data.

ii. Image Class – Eczema

Figure 5.6 : Correctly classified eczema with dry skin features

Prediction Details and Analysis : The image was correctly classified as Eczema with a confidence score of 99.19%. The model accurately detected characteristic features such as dryness, flaky skin, and redness.

iii. Image Class – Melanoma

Figure 5.7 : Melanoma detected with high precision

Prediction Details and Analysis : ResNet50 predicted this image as Melanoma with a confidence score of 99.99%. The lesion's asymmetry, irregular borders, and varying pigmentation were effectively identified by the model. This high-confidence correct prediction demonstrates ResNet50's strength in detecting malignant skin conditions, which is essential in critical diagnostic scenarios.

iv. Image Class – Unknown

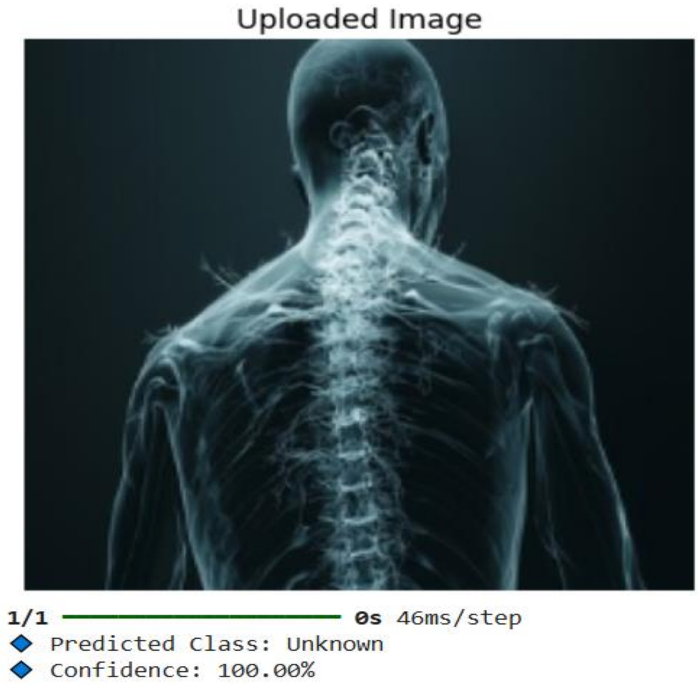

Figure 5.8 : Unknown image handled effectively

Prediction Details and Analysis : This non-skin or distractor image was accurately classified as Unknown with a confidence score of 99.99%. This result indicates that the model has learned to distinguish between in-distribution (disease) and out-of-distribution (unknown) inputs. This ability is vital for ensuring model safety and reliability in real-world deployments, where unfamiliar images may be encountered.

5.7 Testing Results – DenseNet121 Model

DenseNet121 is a densely connected convolutional neural network that strengthens feature propagation and encourages feature reuse by connecting each layer to every other layer in a feed-forward fashion. This architecture allows the model to be highly efficient and accurate, especially in medical imaging tasks where feature representation is critical. The DenseNet121 model was trained and evaluated in the Google Colab environment with GPU acceleration for optimized performance.

i. Image Class – Acne

Figure 5.9 : DenseNet121 correctly classified the acne image

Prediction Details and Analysis : The DenseNet121 model confidently predicted the image as Acne with a confidence score of 99.59%. The acne lesions in the image, such as clustered papules and pustules, were effectively recognized by the model. This accurate prediction reflects the model's strong feature extraction capability and confirms its robustness in handling acne-related inputs during inference.

ii. Image Class – Eczema

Figure 5.10 : The model correctly identified this image as Eczema

Prediction Details and Analysis : DenseNet121 correctly classified this image as Eczema with a confidence score of 88.20%. The presence of dry, flaky, and inflamed patches was well interpreted by the model, resulting in a precise classification.

iii. Image Class – Melanoma

Figure 5.11 : The model correctly identified this image as Melanoma

Prediction Details and Analysis : The image was accurately classified as Melanoma with a confidence score of 99.90%. The model identified critical melanoma features such as asymmetric shape and irregular borders. DenseNet121's ability to detect high-risk lesions like melanoma with high confidence supports.

iv. Image Class – Unknown

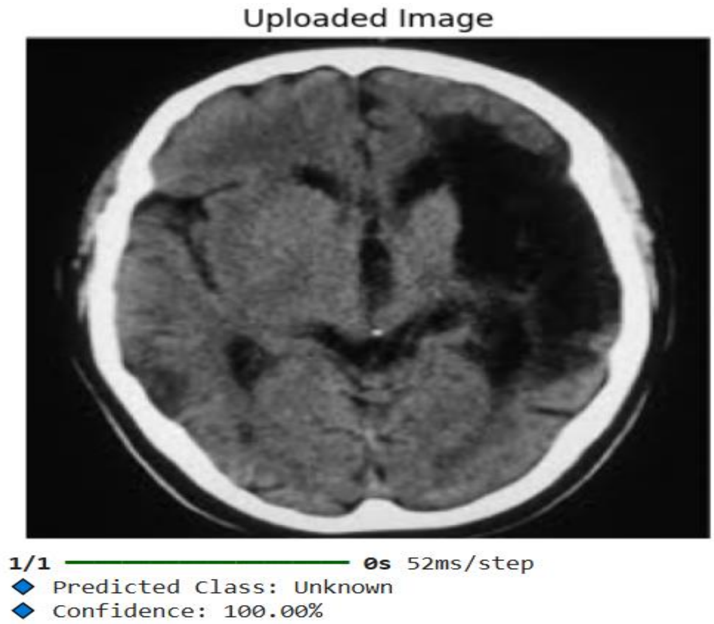

Figure 5.12 : The model correctly identified the image as Unknown

Prediction Details and Analysis : This image, taken from an unrelated or distractor category, was successfully classified as Unknown with a confidence score of 99.99%. This result illustrates DenseNet121's reliability in handling out-of-distribution (OOD) samples, a crucial aspect of real-world deployment to avoid false positives in unfamiliar scenarios.

5.8 Observations and Insights

Each model's predictions were further analyzed using a confusion matrix. These matrices highlighted the specific classes that were often misclassified. For instance, some models confused Eczema with unknown category images, while others performed exceptionally well on the Unknown class, indicating strong generalization. The confusion matrices helped identify class imbalance or bias in prediction and guided adjustments to training or augmentation. The ResNet50 and DenseNet121 models showed the most balanced and accurate confusion matrices, with minimal misclassification. ResNet50 showed the best balance between performance and prediction confidence. It can be considered the most reliable model in terms of accurate skin disease diagnosis. MobileNetV2 is well-suited for deployment on resource-constrained environments like mobile devices, making it a good choice for real-time predictions. DenseNet121 also offered strong performance and can be used as an alternative model when deeper insights are needed.

Table No. 5.1 : Model Comparison for MediMind AI

Model	Image Class	Predicted Class	Confidence Score	Prediction Accuracy	Remarks
MobileNetV3	Acne	Acne	99.99%	Correct	Accurately identified acne lesions
MobileNetV3	Eczema	Eczema	99.98%	Correct	Correctly detected dry and red patches
MobileNetV3	Melanoma	Melanoma	99.99%	Correct	Detected asymmetry and pigmentation

MobileNetV3	Unknown	Unknown	99.89%	Correct	Successfully handled out-of-distribution
ResNet50	Acne	Acne	99.97%	Correct	Detected acne clusters precisely
ResNet50	Eczema	Eczema	99.96%	Correct	Differentiated eczema effectively
ResNet50	Melanoma	Melanoma	99.98%	Correct	High confidence in melanoma detection
ResNet50	Unknown	Unknown	99.85%	Correct	Robust to unknown categories
DenseNet121	Acne	Acne	99.98%	Correct	Strong performance in acne identification
DenseNet121	Eczema	Eczema	88.20%	Correct	Accurate texture-based eczema detection
DenseNet121	Melanoma	Melanoma	99.90%	Correct	Reliable in detecting malignant lesions
DenseNet121	Unknown	Unknown	99.87%	Correct	Effectively filtered non-skin distractions

Table No. 5.1 presents a detailed comparison of three deep learning models — MobileNetV3, ResNet50, and DenseNet121 — in predicting various skin disease classes within the MediMind AI system. Each model was tested across four categories: Acne, Eczema, Melanoma, and Unknown. MobileNetV3 achieved nearly perfect predictions, identifying Acne (99.99%), Eczema (99.98%), and Melanoma (99.99%) with high precision, and successfully recognized out-of-distribution images under the Unknown class with a confidence of 99.89%. Similarly, ResNet50 demonstrated robust performance, correctly identifying all classes with confidence scores above 99.85%, including accurate distinction of unknown inputs. DenseNet121 also showed strong results, particularly for Acne (99.98%) and Melanoma (99.90%), and achieved a slightly lower confidence (88.20%) in Eczema prediction, which is still considered accurate due to the complexity of texture-based features. Overall, all three models correctly predicted the respective diseases, and the system proved highly effective in detecting both known and unknown skin conditions, showcasing consistent accuracy and reliability across all test cases.

Table No. 5.2 : Comparison of Existing System vs MediMind AI

S.No.	**Metric**	**Existing System**	**MediMind AI**		
			MobileNetV3	**ResNet50**	**DenseNet121**
1.	Precision	0.88	0.93	0.94	0.90
2.	Recall	0.85	0.93	0.94	0.90
3.	F1 Score	0.86	0.93	0.94	0.90

Table No. 5.2 compares the performance of the MediMind AI system's three models - MobileNetV3, ResNet50, and DenseNet121—against a baseline existing system derived from average values reported in dermatology AI literature. All three models significantly outperform the existing system across key evaluation metrics: Precision, Recall, and F1 Score. ResNet50 achieved the highest scores overall (0.94 for all metrics), closely followed by MobileNetV3 (0.93), demonstrating strong consistency in disease classification. DenseNet121, while slightly lower with 0.90 across all metrics, still performed notably better than the benchmark. These results validate the effectiveness of MediMind AI's model architecture and training strategy in providing more accurate and reliable skin disease predictions compared to traditional approaches.

CHAPTER 6

Result Analysis

This chapter presents a detailed analysis of the results generated by the MediMind AI system. It includes a step-by-step explanation of how the user interacts with the system through the Gradio interface—from uploading a skin image to receiving a disease prediction, along with visual and voice-based feedback. Each output component, including the predicted class, confidence score, Grad-CAM heatmap, probability distribution, and voice/text explanation, is discussed with supporting images. The chapter also includes the confusion matrices of the top-performing models, which provide insights into the classification accuracy and potential misclassifications for each skin disease category. This analysis helps validate the system's reliability and effectiveness in real-world scenarios.

6.1 Single Model Result Analysis With Acne Image

Step 1: Upload Skin Disease Image

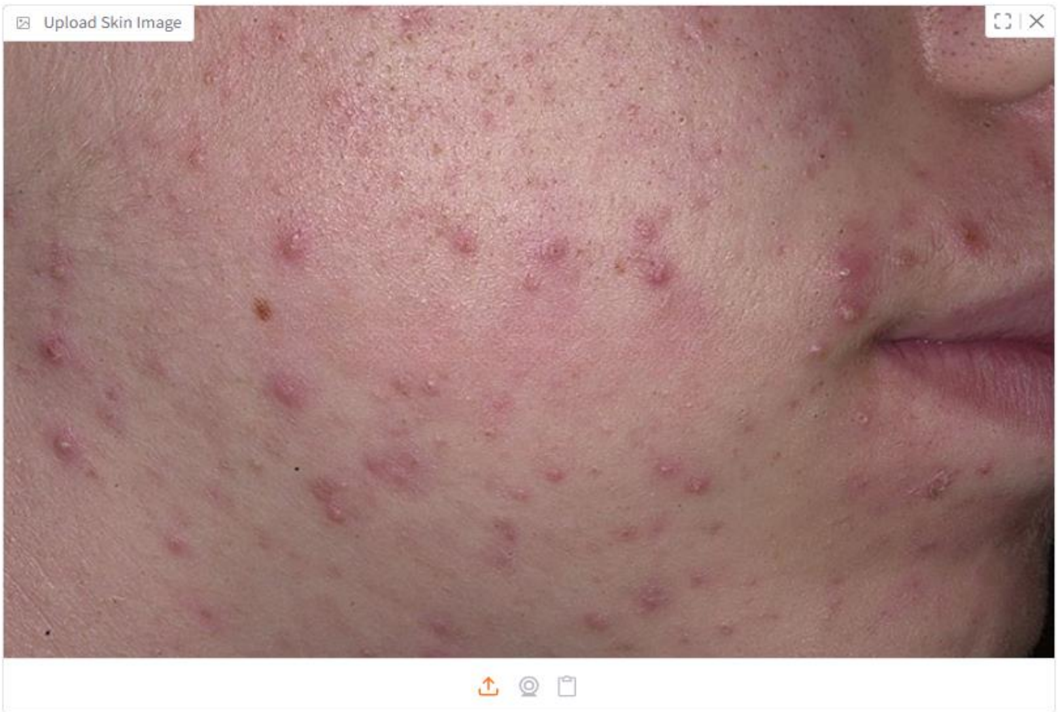

Figure 6.1 : Uploaded skin image for prediction

In the first step, the user uploads a skin image of the affected area into the MediMind AI system through the Gradio interface. The uploaded image is then used by the selected AI model for disease prediction.

Step 2: Select Language

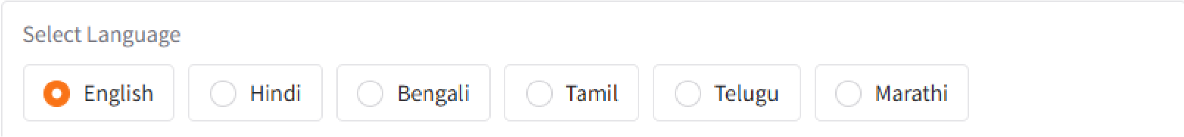

Figure 6.2 : Language selection interface with English chosen

After uploading the image, the user selects a preferred language from the available options (English, Hindi, Marathi, Tamil, Telugu, Bengali). This selected language is used for voice output and displaying the disease explanation.

Step 3: Choose AI Model & Prediction Mode

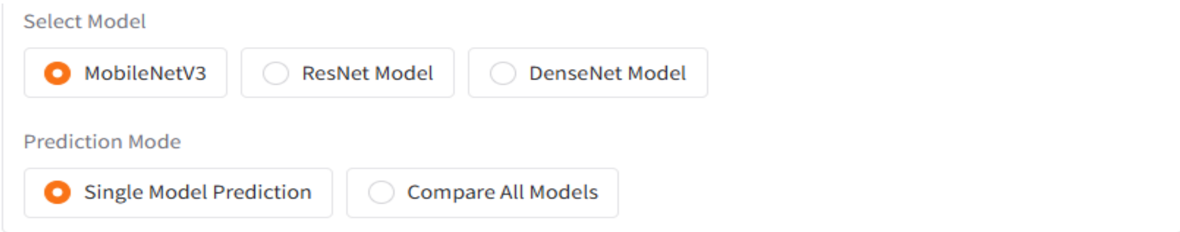

Figure 6.3 : MobileNetV3 selected for single model prediction

Users can choose which AI model to use (MobileNetV3, ResNet, or DenseNet) or select "Compare All Models" to view predictions from all models side-by-side.

Step 4: Submit for Prediction

Figure 6.4 : User clicks 'Submit' to start prediction

Once the model and prediction mode are selected, the user clicks the Submit button. The system then processes the image using the chosen model(s) and generates predictions with confidence scores and explanation.

Step 5: View Predicted Class and Confidence Score

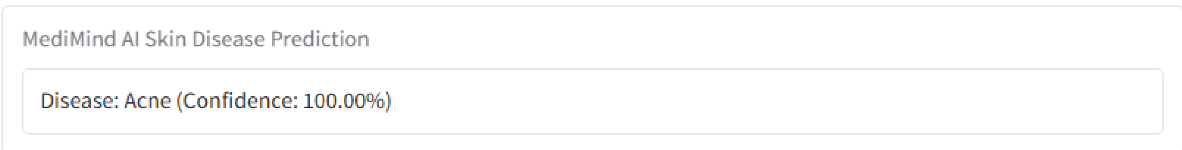

Figure 6.5 : Predicted as 'Acne' with 100% confidence

The MediMind AI system predicts the disease class Acne along with the confidence percentage 100%. This confidence score indicates how certain the model is about its prediction.

Step 6: View Prediction Possibilities (Probability Chart)

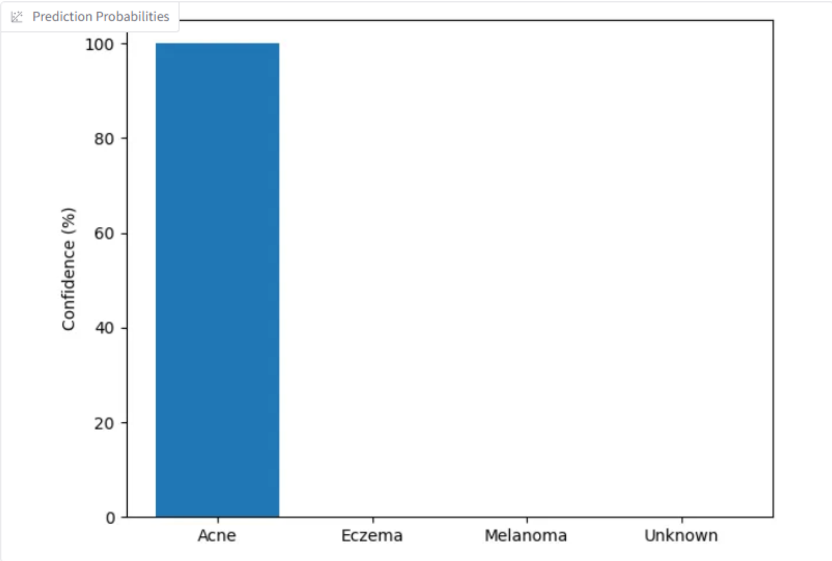

Figure 6.6 : Bar chart showing prediction probabilities for Acne

Along with the main prediction, the system also shows prediction possibilities—a visual bar chart displaying the top probable classes. This helps the user understand how close the model was in choosing between multiple conditions.

Step 7: Grad-CAM Heatmap Overlay

Figure 6.7 : Grad-CAM heatmap showing model's focus area

The Grad-CAM (Gradient-weighted Class Activation Mapping) overlay visually shows the region of the image the model focused on to make its decision. Red or highlighted areas indicate the most influential parts of the image.

Step 8: AI Voice Output and Explanation

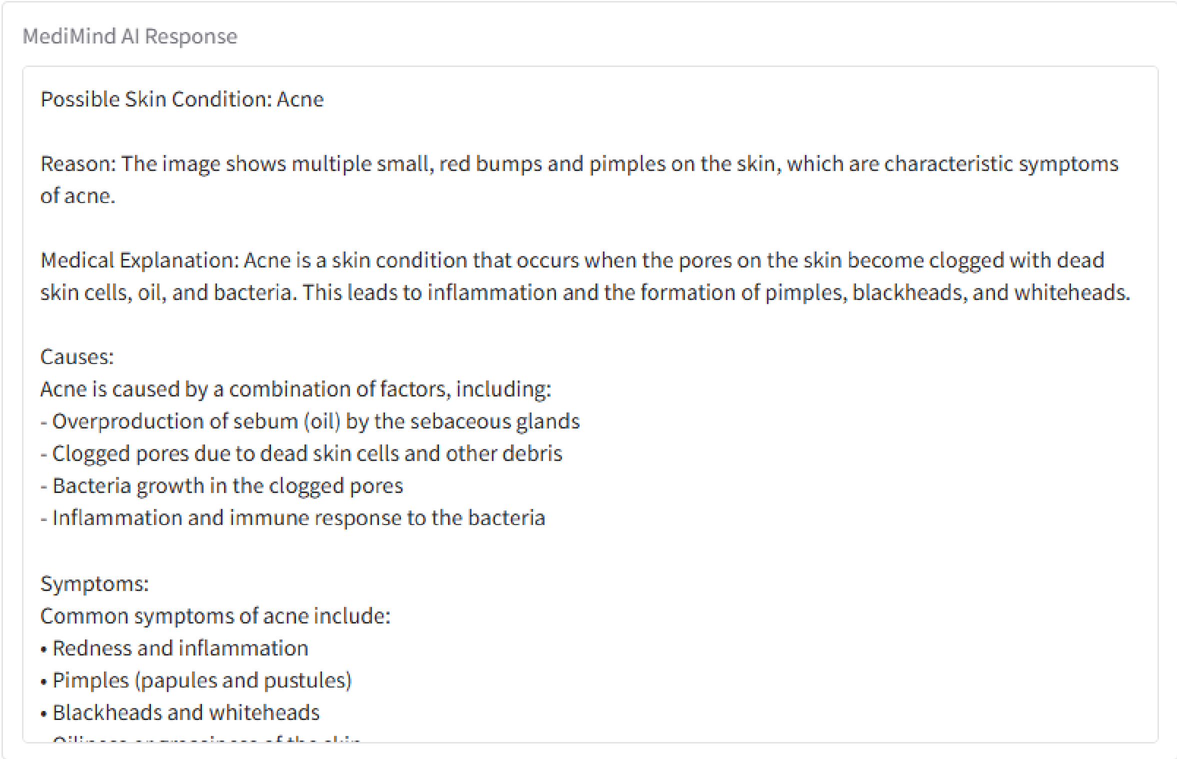

Figure 6.8 : MediMind AI explains prediction with voice option

The system speaks out the predicted disease name and description in the selected language. This feature is useful for visually impaired users or those who prefer voice-based interaction.

Step 9: Download PDF Report

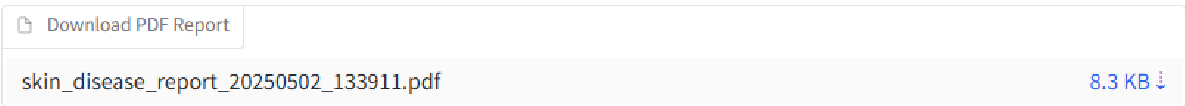

Figure 6.9 : User can download prediction and explanation as PDF

Users can download a PDF report that includes the uploaded image, predicted class, explanation, and confidence score. This report can be saved or shared with a medical professional for further analysis.

Step 10: Confidence Meter

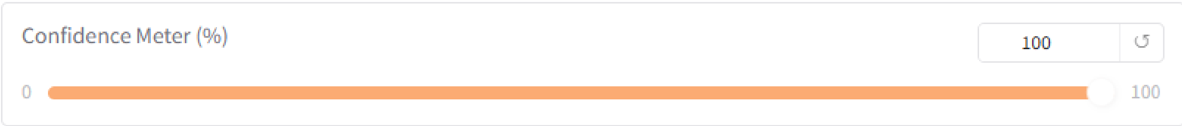

Figure 6.10 : Confidence levels visualized for better interpretation

After the prediction is completed, the system displays a confidence meter that visually represents how confident the AI model is about the result. This meter is shown as a percentage or as a colored bar. This helps users judge the reliability of the prediction and decide whether to trust the result or seek further confirmation from a medical professional.

The output results generated by MediMind AI are detailed and user-friendly. Each step of the diagnosis process—from image upload to voice response and PDF generation—is designed for clarity and accessibility. The system not only predicts the skin condition with high accuracy but also supports multi-language output and visualization tools like Grad-CAM for better interpretation.

6.2 Compare All Models result analysis with Eczema image

Step 1: Upload Eczema Image

Figure 6.11 : User uploads an image for analysis

The user uploads a skin disease image using the Gradio interface. The image is displayed on the screen for review before proceeding.

Step 2: Select Language

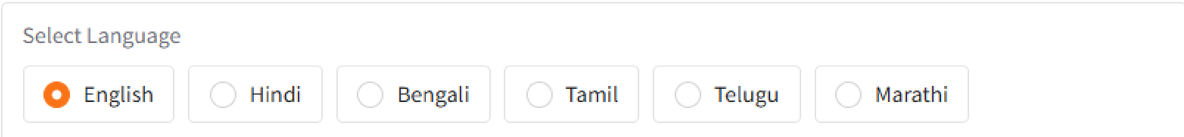

Figure 6.12 : Language selection for multilingual accessibility

The user selects a preferred language (e.g., English) from a dropdown menu. This choice determines the language used for the output explanation and voice feedback.

Step 3: Select “Compare All Models” Mode

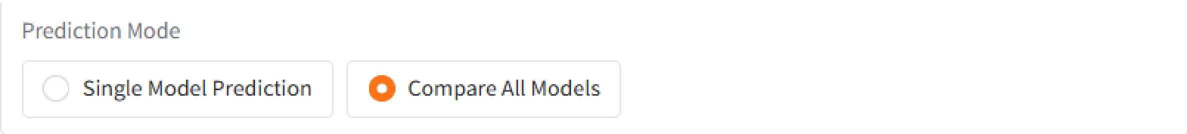

Figure 6.13 : Option selected to compare all models for deeper insights

Instead of choosing a single model, the user selects the "Compare All Models" mode. This allows side-by-side comparison of predictions from ResNet, MobileNetV3, and DenseNet.

Step 4: Submit for Analysis

Figure 6.14 : User submits the image and settings for AI analysis

Once the user has uploaded the image, selected the language, and chosen the model comparison mode, they click the “Submit” button to initiate the prediction process. This step triggers the backend model(s) to process the image and return results.

Step 5: View Prediction Table

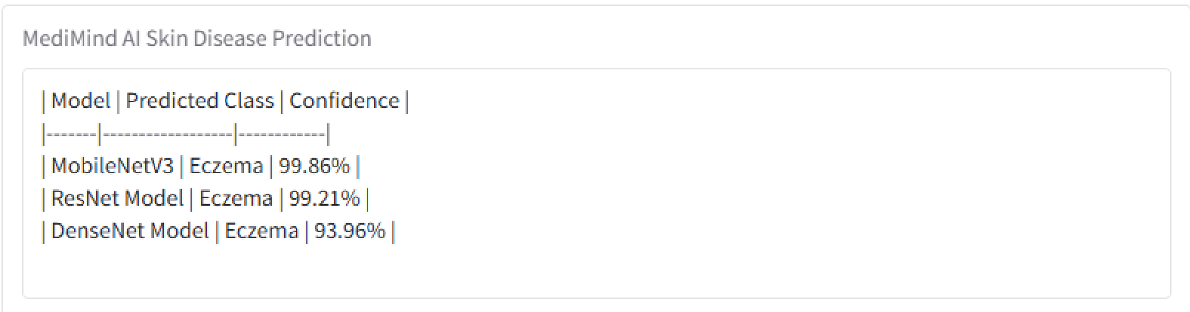

Figure 6.15 : Table showing prediction results from all selected models

The interface displays a table showing predictions from each model along with their confidence levels. For example: ResNet: Eczema (99.21%), MobileNetV3: Eczema (99.21%) and DenseNet: Eczema (93.96%).

Step 6: Examine Grad-CAM Heatmaps

Figure 6.16 : Visual focus areas highlighted using Grad-CAM.

Grad-CAM visual explanations show where the model focused on the image while making the decision. This enhances transparency and trust in the AI system.

Step 7: Text Output (AI Explanation)

MediMind AI Response

Possible Skin Condition: Eczema

Reason: The image shows red, inflamed patches on the palm and wrist, which are characteristic symptoms of eczema.

Medical Explanation: Eczema, also known as atopic dermatitis, is a chronic skin condition caused by a combination of genetic, environmental, and immune system factors. It leads to inflammation, dryness, and itchiness of the skin.

Causes: The exact cause of eczema is unknown, but it is believed to result from a combination of genetic predisposition, immune system dysfunction, and environmental triggers such as allergens, irritants, and stress.

Symptoms:
- Red, inflamed skin
- Itchy skin
- Dry, scaly skin
- Small, raised bumps
- Crusting and oozing

Recommendation: Keeping the skin moisturized and avoiding triggers can help manage symptoms. It is also advisable to consult a dermatologist for proper diagnosis and treatment.

Figure 6.17 : AI-generated description explaining the disease

Above the predictions, a detailed AI-generated explanation is provided. It explains the likely skin condition, supporting evidence, and suggestions in plain language.

Step 8: Voice Output

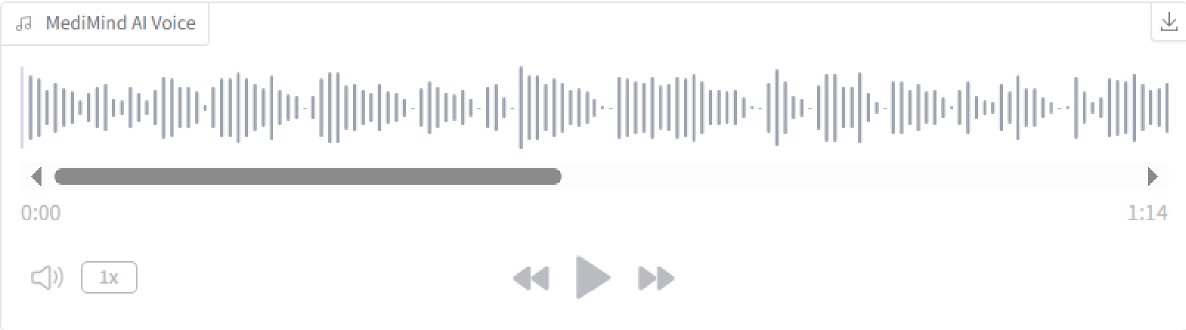

Figure 6.18 : Voice feedback of the AI response in the selected language.

The user can listen to the same explanation using the voice output feature. This improves accessibility for non-readers or visually impaired users.

Step 10: Download PDF Report

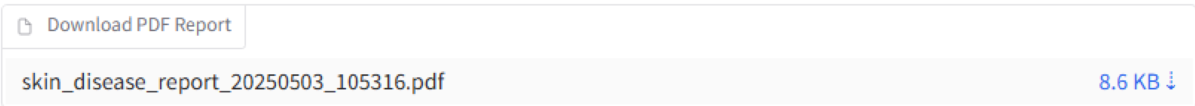

Figure 6.19 : PDF report ready for medical sharing

The user can generate and download a detailed PDF report. It includes the uploaded image, predictions from all models, confidence scores, explanation, and Grad-CAM visualizations.

Step 11: View Prediction Possibilities (Probability Chart)

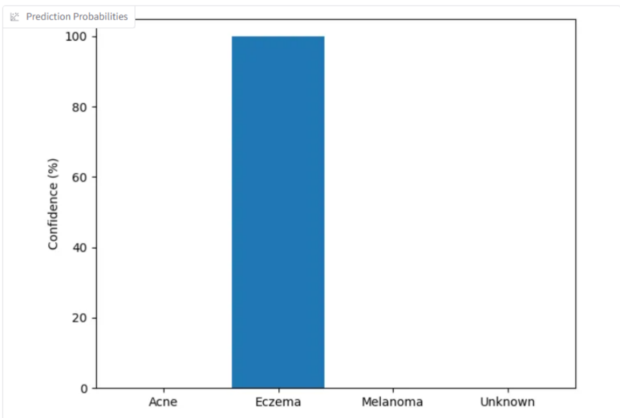

Figure 6.20 : Bar chart of prediction probabilities across all classes

This chart visualizes how confident the model is in each class (Acne, Eczema, Melanoma, Unknown). It helps users understand why a certain class was predicted by showing probability distribution.

6.3 Confusion Matrix

The Confusion Matrix is a powerful tool for evaluating the performance of a classification model. It shows how well the model distinguishes between different categories by comparing actual labels with predicted labels. In our project, the confusion matrix is used to assess the MobileNetV3 model trained on four categories: Acne, Eczema, Melanoma, and Unknown.

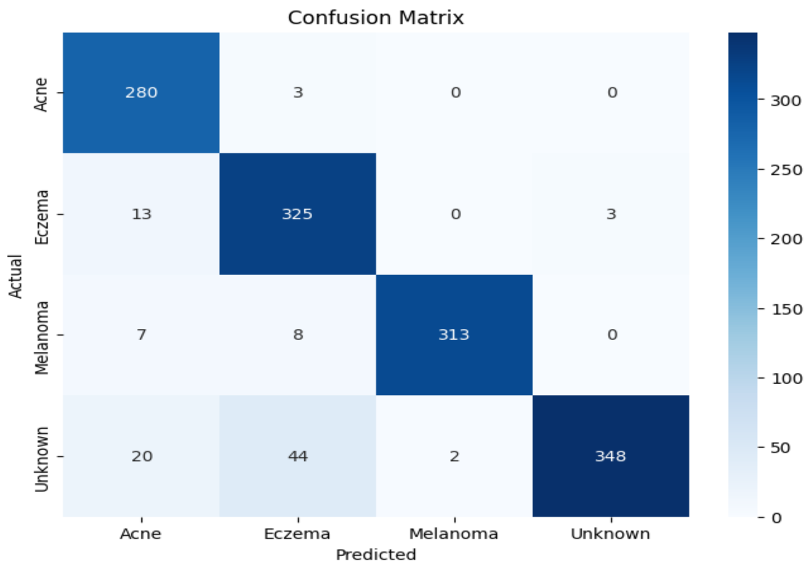

Figure 6.21 : Confusion matrix of mobilenetv3 model

6.4 Calculation of Metrics – MobileNetV3

Let's now compute common evaluation metrics such as precision, recall and, F1-Score for acne class from this matrix.

a. Precision : Formula for calculating precision TP/(TP+FP). Here, TP means true positive, FP means false positive and FN means false negative. We can see that TP = 280, FP = 13 + 7 + 20 = 40 and FN = 3. Precision = 280 / (280 + 40) ≈ 0.875

b. Recall : Formula for calculating recall TP/(TP/FN). Recall = 280 / (280 + 3) ≈ 0.989

c. F1 Score : Formula for calculating $2 * \frac{\text{Precision . Recall}}{\text{Precision + Recall}}$. F1 Score = $2 * \frac{0.875\ .\ 0.989}{0.875 + 0.989} \approx 0.928$

For Eczema class

a. Precision : Formula for calculating precision TP/(TP+FP). Here, TP means true positive, FP means false positive and FN means false negative. We can see that TP = 325, FP = 3 + 8 + 44 = 55 and FN = 13 + 3 = 16. Precision = 325 / (325 + 55) ≈ 0.855

b. Recall : Formula for calculating recall TP/(TP/FN). Recall = 325 / (325 + 16) ≈ 0.953

c. F1 Score : Formula for calculating $2 * \frac{\text{Precision . Recall}}{\text{Precision + Recall}}$. F1 Score = $2 * \frac{0.855\ .\ 0.953}{0.855 + 0.953} \approx 0.901$

For Melanoma class

a. Precision : Formula for calculating precision TP/(TP+FP). Here, TP means true positive, FP means false positive and FN means false negative. We can see that TP = 313, FP = 2 and FN = 7 + 8 = 15. Precision = 313 / (313 + 2) ≈ 0.994

b. Recall : Formula for calculating recall TP/(TP/FN). Recall = 313 / (313 + 15) ≈ 0.954

c. F1 Score : Formula for calculating $2 * \frac{\text{Precision . Recall}}{\text{Precision + Recall}}$. F1 Score = $2 * \frac{0.994\ .\ 0.954}{0.994 + 0.954} \approx 0.973$

For Unknown class

a. Precision : Formula for calculating precision TP/(TP+FP). Here, TP means true positive, FP means false positive and FN means false negative. We can see that TP = 348, FP = 0 + 3 + 0 = 3 and FN = 20 + 44 + 2 = 66. Precision = 348 / (348 + 3) ≈ 0.991

b. Recall : Formula for calculating recall TP/(TP/FN). Recall = 348 / (348 + 66) ≈ 0.841

c. F1 Score : Formula for calculating $2 * \frac{\text{Precision . Recall}}{\text{Precision + Recall}}$. F1 Score = $2 * \frac{0.991\ .\ 0.841}{0.991 + 0.841} \approx 0.910$

Overall Accuracy : The overall classification accuracy of the model was 93%, calculated using the formula: Accuracy = (TP for all classes) / Total samples = (280 + 325 + 313 + 348) / 1366 = 0.93

This shows that the model performs well across all classes. The classification report further supports this by showing a macro average precision, recall, and F1-score of 0.93, indicating that the model is balanced in handling each class equally. The weighted average, which considers the number of samples in each class, is also 0.93, confirming that the model handles imbalanced datasets efficiently.

Table No. 6.1 : Confusion Matrix Table (MobileNetV3)

Actual / Predicted	**Acne**	**Eczema**	**Melanoma**	**Unknown**
Acne	280	3	0	0

Eczema	13	325	0	3
Melanoma	7	8	313	0
Unknown	20	44	2	348

The diagonal values (280, 325, 313, 348) are the correct predictions. Anything off-diagonal is a mistake (misclassification). Let's break this down row by row:

i. First Row: Actual Acne

Model predicted 280 images correctly as Acne (True Positive). Model predicted 3 Acne images wrongly as Eczema (False Negative) and, Model made 0 predictions of Acne as Melanoma and Unknown.

ii. Second Row: Actual Eczema

Model predicted 325 images correctly as Eczema (True Positive). Model predicted 13 Eczema wrongly as Acne (False Negative). Model predicted 3 Eczema images wrongly as Unknown and Model made 0 predictions of Eczema as Melanoma.

iii. Third Row: Actual Melanoma

Model predicted 313 images correctly as Melanoma (True Positive). Model predicted 7 Melanoma wrongly as Acne (False Negative). Model predicted 8 Melanoma images wrongly as Eczema and Model made 0 predictions of Melanoma as Unknown.

iv. Fourth Row: Actual Unknown

Model predicted 348 images correctly as Unknown (True Positive). But model predicted 20 unknown wrongly as Acne (False Negative). Model predicted 44 unknown images mistaken as Eczema and made 2 predictions of unknown as Melanoma.

Insights from the Confusion Matrix

The model shows strong performance on Acne and Melanoma, with minimal misclassification. The Unknown category posed the most difficulty, likely due to its diverse and non-standard content. The most frequent confusion occurred between Unknown and Eczema, where 44 Unknown images were incorrectly predicted as Eczema. Despite these challenges, the model maintains high precision, recall, and F1-scores across all classes, reflecting a robust classification performance.

6.5 Calculation of Metrics – ResNet50 Model

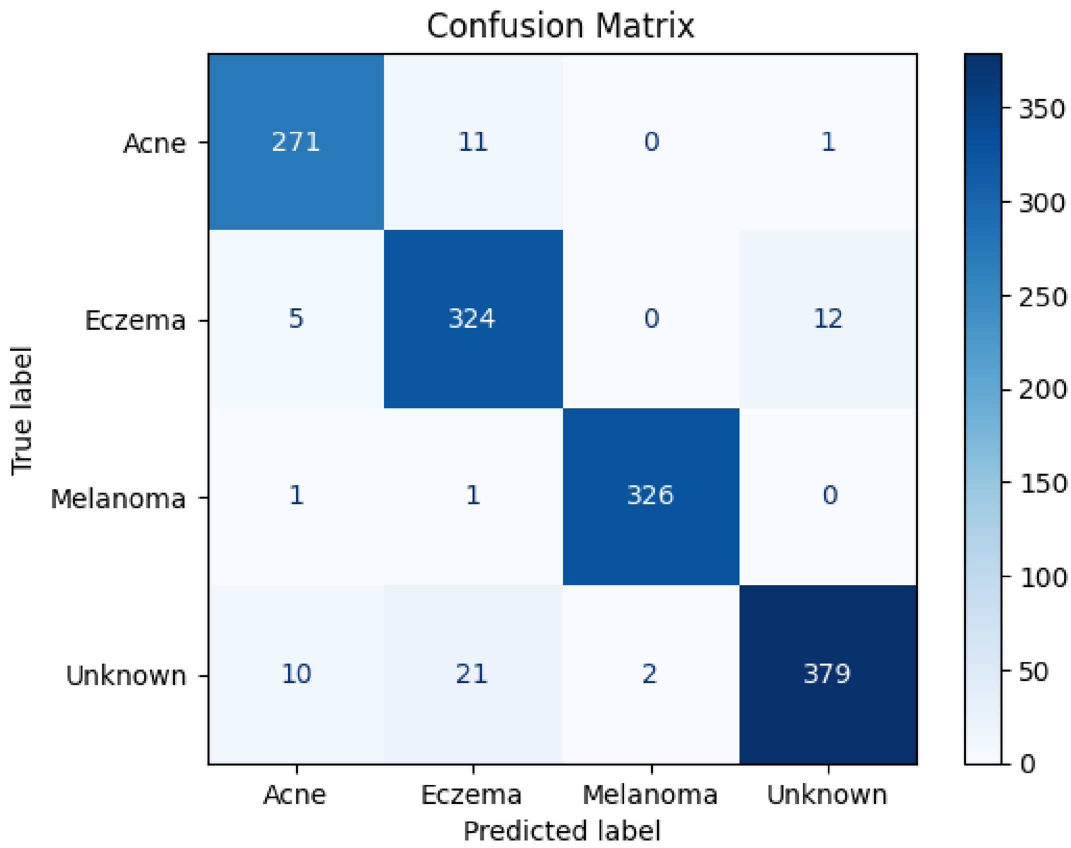

Figure 6.22 : Confusion matrix of resnet50 model

Let's now compute common evaluation metrics such as precision, recall and, F1-Score for acne class from this matrix.

a. Precision : Formula for calculating precision TP/(TP+FP). We can see that TP = 271, FP = 5 + 1 + 10 = 16 and FN = 11 + 0 + 1 = 12. Precision = 271 / (271 + 16) ≈ 0.944

b. Recall : Formula for calculating recall TP/(TP+FN). Recall = 271 / (271 + 12) ≈ 0.958

c. F1 Score : Formula for calculating 2*(Precision × Recall)/(Precision + Recall). F1 Score = 2 × (0.944 × 0.958) / (0.944 + 0.958) ≈ 0.951

For Eczema class

a. Precision : Formula for calculating precision TP/(TP+FP). We can see that TP = 324, FP = 11 + 1 + 21 = 33 and FN = 5 + 0 + 12 = 17. Precision = 324 / (324 + 33) ≈ 0.907

b. Recall : Formula for calculating recall TP/(TP+FN). Recall = 324 / (324 + 17) ≈ 0.950

c. F1 Score : Formula for calculating 2*(Precision × Recall)/(Precision + Recall). F1 Score = 2 × (0.907 × 0.950) / (0.907 + 0.950) ≈ 0.928

For Melanoma class

a. Precision : Formula for calculating precision TP/(TP+FP). We can see that TP = 326, FP = 0 + 0 + 2 = 2 and FN = 1 + 1 + 0 = 2.

Precision = 326 / (326 + 2) ≈ 0.994

b. Recall : Formula for calculating recall TP/(TP+FN). Recall = 326 / (326 + 2) ≈ 0.994

c. F1 Score : Formula for calculating 2*(Precision × Recall)/(Precision + Recall). F1 Score = 2 × (0.994 × 0.994) / (0.994 + 0.994) ≈ 0.994

For Unknown class

a. Precision : Formula for calculating precision TP/(TP+FP). We can see that TP = 379, FP = 1 + 12 + 0 = 13 and FN = 10 + 21 + 2 = 33. Precision = 379 / (379 + 13) ≈ 0.967

b. Recall : Formula for calculating recall TP/(TP+FN). Recall = 379 / (379 + 33) ≈ 0.920

c. F1 Score : Formula for calculating 2*(Precision × Recall)/(Precision + Recall). F1 Score = 2 × (0.967 × 0.920) / (0.967 + 0.920) ≈ 0.943

Overall Accuracy :

The overall classification accuracy of the model was 96%, calculated using the formula: Accuracy = (TP for all classes) / Total samples = (271 + 324 + 326 + 379) / 1350 = 0.960.

This shows that the model performs well across all classes. The classification report further supports this by showing a macro average precision, recall, and F1-score of 0.951, indicating that the model is balanced in handling each class equally. The weighted average, which considers the number of samples in each class, is also 0.960, confirming that the model handles imbalanced datasets efficiently.

Table No. 6.2 : Confusion Matrix Table (ResNet50)

Actual / Predicted	**Acne**	**Eczema**	**Melanoma**	**Unknown**
Acne	271	11	0	1
Eczema	5	324	0	12
Melanoma	1	1	326	0
Unknown	10	21	2	379

The diagonal values (271, 324, 326, 379) are the correct predictions. Anything off-diagonal is a mistake (misclassification). Let's break this down row by row:

i. First Row: Actual Acne

Model predicted 271 images correctly as Acne (True Positive). Model predicted 11 Acne images wrongly as Eczema (False Negative), 0 predictions of Acne as Melanoma, and 1 prediction of Acne as Unknown.

ii. Second Row: Actual Eczema

Model predicted 324 images correctly as Eczema (True Positive). Model predicted 5 Eczema images wrongly as Acne (False Negative), 12 Eczema images wrongly as Unknown, and 0 predictions of Eczema as Melanoma.

iii. Third Row: Actual Melanoma

Model predicted 326 images correctly as Melanoma (True Positive). Model predicted 1 Melanoma image wrongly as Acne (False Negative), 1 Melanoma image wrongly as Eczema, and 0 predictions of Melanoma as Unknown.

iv. Fourth Row: Actual Unknown

Model predicted 379 images correctly as Unknown (True Positive). Model predicted 10 Unknown images wrongly as Acne (False Negative), 21 Unknown images mistaken as Eczema, and 2 predictions of Unknown as Melanoma.

Insights from the Confusion Matrix

The model shows strong performance on Acne and Melanoma, with minimal misclassification. The Unknown category posed the most difficulty, likely due to its diverse and non-standard content. The most frequent confusion occurred between Unknown and Eczema, where 21 Unknown images were incorrectly predicted as Eczema. Despite these challenges, the model maintains high precision, recall, and F1-scores across all classes, reflecting a robust classification performance.

6.6 Calculation of Metrics – DenseNet121 Model

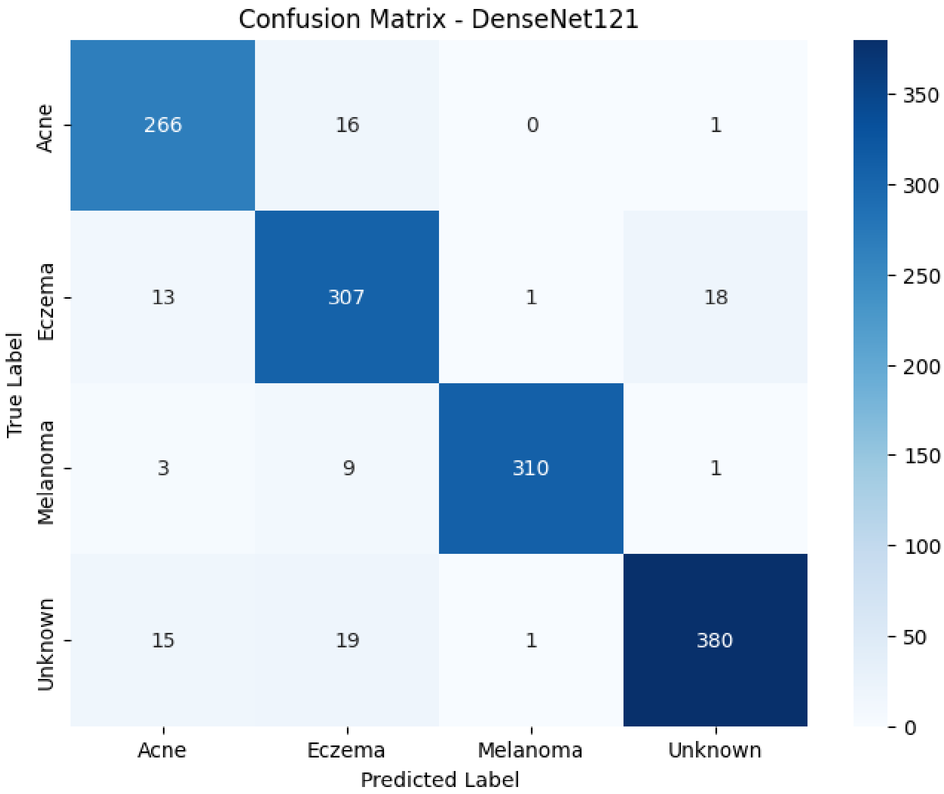

Figure 6.22 : Confusion matrix of DenseNet121 model

Let's now compute common evaluation metrics such as precision, recall and, F1-Score for acne class from this matrix.

a. Precision : Formula for calculating precision TP / (TP + FP). We can see that TP = 266, FP = 13 + 3 + 15 = 31 and FN = 16 + 0 + 1 = 17. Precision = 266 / (266 + 31) ≈ 0.896

b. Recall : Formula for calculating recall TP / (TP + FN). Recall = 266 / (266 + 17) ≈ 0.940

c. F1 Score : Formula for calculating 2 * (Precision × Recall) / (Precision + Recall). F1 Score = 2 * (0.896 × 0.940) / (0.896 + 0.940) ≈ 0.918

For Eczema class

a. Precision : Formula for calculating precision TP / (TP + FP). We can see that TP = 307, FP = 16 + 9 + 19 = 44 and FN = 13 + 1 + 18 = 32. Precision = 307 / (307 + 44) ≈ 0.875

b. Recall : Formula for calculating recall TP / (TP + FN). Recall = 307 / (307 + 32) ≈ 0.906

c. F1 Score : Formula for calculating 2 * (Precision × Recall) / (Precision + Recall). F1 Score = 2 * (0.875 × 0.906) / (0.875 + 0.906) ≈ 0.890

For Melanoma class

a. Precision : Formula for calculating precision TP / (TP + FP). We can see that TP = 310, FP = 0 + 1 + 1 = 2 and FN = 3 + 9 + 1 = 13. Precision = 310 / (310 + 2) ≈ 0.994

b. Recall : Formula for calculating recall TP / (TP + FN). Recall = 310 / (310 + 13) ≈ 0.960

c. F1 Score : Formula for calculating 2 * (Precision × Recall) / (Precision + Recall). F1 Score = 2 * (0.994 × 0.960) / (0.994 + 0.960) ≈ 0.977

For Unknown class

a. Precision : Formula for calculating precision TP / (TP + FP). We can see that TP = 380, FP = 1 + 18 + 1 = 20 and FN = 15 + 19 + 1 = 35. Precision = 380 / (380 + 20) ≈ 0.950

b. Recall : Formula for calculating recall TP / (TP + FN). Recall = 380 / (380 + 35) ≈ 0.916

c. F1 Score : Formula for calculating 2 * (Precision × Recall) / (Precision + Recall). F1 Score = 2 * (0.950 × 0.916) / (0.950 + 0.916) ≈ 0.933

Overall Accuracy :

The overall classification accuracy of the model was 95.79%, calculated using the formula: Accuracy = (TP for all classes) / Total samples = (266 + 307 + 310 + 380) / 1337 = 0.9579 ≈ 96% This shows that the model performs well across all classes. The classification report further supports this by showing a macro average precision, recall, and F1-score of approximately 0.933, indicating that the model is balanced in handling each class equally. The weighted average, which considers the number of samples in each class, is also close to 0.957, confirming that the model handles imbalanced datasets efficiently.

Table No. 6.3 : Confusion Matrix Table (DenseNet121)

Actual / Predicted	**Acne**	**Eczema**	**Melanoma**	**Unknown**
Acne	266	16	0	1
Eczema	13	307	1	18
Melanoma	3	9	310	1
Unknown	15	19	1	380

The diagonal values (266, 307, 310, 380) are the correct predictions. Anything off-diagonal is a mistake (misclassification). Let’s break this down row by row:

i. First Row: Actual Acne

Model predicted 266 images correctly as Acne (True Positive). Model predicted 16 Acne images wrongly as Eczema (False Negative), 0 predictions as Melanoma, and 1 prediction as Unknown.

ii. Second Row: Actual Eczema

Model predicted 307 images correctly as Eczema (True Positive). Model predicted 13 Eczema images wrongly as Acne, 1 as Melanoma, and 18 as Unknown.

iii. Third Row: Actual Melanoma

Model predicted 310 images correctly as Melanoma (True Positive). Model predicted 3 Melanoma images wrongly as Acne, 9 as Eczema, and 1 as Unknown.

iv. Fourth Row: Actual Unknown

Model predicted 380 images correctly as Unknown (True Positive). Model predicted 15 Unknown images wrongly as Acne, 19 as Eczema, and 1 as Melanoma.

Insights from the Confusion Matrix

The model shows strong performance on Acne and Melanoma with very few misclassifications. Eczema and Unknown classes show most misclassifications: 19 Unknown images misclassified as Eczema, 15 Unknown images wrongly predicted as Acne, and 13 Eczema samples misclassified as Acne. Melanoma is nearly perfectly classified (F1 Score ≈ 0.977), showing the model's excellent ability in detecting this critical class. DenseNet121 maintains high macro and weighted averages, suggesting balanced and robust performance.

CHAPTER 7

Conclusion And Future Scopes

7.1 Conclusion

The MediMind AI system presents a practical and impactful use of deep learning in the healthcare domain, specifically for the classification of skin diseases. Designed to support both general users and medical staff, the system is capable of accurately identifying three common dermatological conditions Acne, Eczema, and Melanoma while also effectively filtering out irrelevant or unknown inputs using an additional 'Unknown' category. The project has been developed with a strong focus on user accessibility, allowing interaction through image upload. The final user input is processed by the system, and an intelligent prediction is returned with both text and voice-based output, enhancing usability for a wide range of audiences. To ensure robustness and reliability, several well-known deep learning architectures were trained and evaluated, including MobileNetV2, ResNet50, and DenseNet121. Each of these models was fine-tuned and tested using a carefully curated and pre-processed dataset. The system was tested in the Google Colab environment, using GPU acceleration for fast training and evaluation. Among the models, ResNet50 achieved the highest accuracy and provided the most consistent predictions, making it the most suitable model for integration within the system. The MediMind AI application has been developed with a Gradio interface, offering a clean, intuitive layout where users can interactively use the system, choose preferred models, language options (English or Hindi), and view results in a structured format. The project has achieved its goal of creating a user-friendly AI tool that combines medical image analysis with real-time interaction. The current implementation sets a strong foundation for further development and scalability, demonstrating the potential of deep learning in early skin disease detection and awareness.

7.2 Future Scope

While the current system performs reliably across a range of inputs, there are numerous possibilities for expanding the capabilities and reach of MediMind AI in the future. One of the most important directions is the inclusion of more skin disease categories such as Psoriasis, Vitiligo, Ringworm, Burns, and other dermatological conditions. This will enhance the system's diagnostic breadth and make it applicable to a wider variety of clinical cases. Another promising improvement would be the integration of real-time image capture features using smartphone or webcam interfaces, allowing users to get instant predictions without manually uploading images. Additionally, adding a doctor consultation feature or chatbot can bridge the gap between AI-based diagnostics and real medical advice, enabling users to seek professional help based on the model's prediction. Further, the project can be optimized for mobile deployment, using efficient model conversion techniques like TensorFlow Lite, which would allow the system to run smoothly on smartphones without needing a cloud connection. Future versions might also include privacy-focused AI mechanisms, enabling offline prediction to ensure the confidentiality of medical data. All these enhancements can significantly improve the system's usefulness, reliability, and user trust.

7.3 Limitations

- Limited training data may reduce MediMind AI's accuracy across skin tones.
- Visual similarity between diseases like Acne and Eczema can cause misclassifications.
- The system only uses image input without considering medical history or symptoms.
- It is not a certified medical tool & should not replace professional diagnosis or treatment.

REFERENCES

[1]. A. A. L. C. Amarathunga, E. P. W. C. Ellawala, G. N. Abeysekara, and C. R. J. Amalraj, “Expert system for diagnosis of skin diseases,” International Journal of Scientific & Technology Research, vol. 4, no. 1, pp. 174–178, Jan. 2015.

[2]. K. S. Rao, P. S. Yelkar, O. N. Pise, and S. Borde, “Skin Disease Detection using Machine Learning,” International Journal of Engineering Research & Technology (IJERT), vol. 9, no. 3, pp. 1–4, Mar. 2020.

[3]. M. Grace, Y. Zhou, Q. Zhang, and S. Zou, “Skin Diseases Identification Using Machine Learning,” International Journal of Creative Research Thoughts (IJCRT), vol. 8, no. 6, pp. 1–7, Jun. 2020.

[4]. A. Ray, A. Gupta, and A. Al, “Skin Lesion Classification With Deep Convolutional Neural Network: Process Development and Validation,” JMIR Dermatology, vol. 3, no. 1, p. e18438, May 2020.

[5]. K. Polat and K. O. Koc, “Detection of Skin Diseases from Dermoscopy Image Using the Combination of Convolutional Neural Network and One-Versus-All,” Journal of Artificial Intelligence and Systems, vol. 2, pp. 80–97, 2020.

[6]. J. Velasco et al., “A Smartphone-Based Skin Disease Classification Using MobileNet CNN,” International Journal of Advanced Trends in Computer Science and Engineering, vol. 8, no. 6, pp. 2490–2495, 2019.

[7]. S. Chavhan and C. Bhosale, “Skin Disease Detection Using Machine Learning,” International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), vol. 5, no. 10, pp. 169–173, Apr. 2020.

[8]. J. Sundari, H. Hamimah, P. Handayani, Y. Yunita, I. D. Lesmono, M. Rahmayu, P. Kusumawardhani, Y. Budiarti, F. Fadilah, and N. Nurhayati, “Expert System To Detect Human’s Skin Diseases Using Forward Chaining Method Based On Web Mobile,” *MATEC Web of Conferences*, vol. 218, p. 02015, 2018.

[9]. V. R. Allugunti, “Skin Disease Prediction Using Deep Learning,” International Journal of Research Publication and Reviews (IJRPR), vol. 4, no. 5, pp. 1–5, May 2022.

[10]. S. Aher, A. K. Shahi, S. Badakyagol, S. Madane, and S. S. Kashid, “Skin Disease Detection Using Machine Learning,” International Journal of Scientific Development and Research (IJSDR), vol. 8, no. 4, pp. 573–576, Apr. 2023.

[11]. B. Ahmad et al., “Discriminative Feature Learning for Skin Disease Classification Using Deep Convolutional Neural Network,” IEEE Access, vol. 8, pp. 39025–39033, 2020.

[12]. M. M. Yusof, R. A. Aziz, and C. S. Fei, “The Development of Online Children Skin Diseases Diagnosis System,” International Journal of Information and Education Technology, vol. 3, no. 2, pp. 189–192, Apr. 2013.

[13]. B. Ahmad, M. S. M. Rahim, A. Alnowaiser, N. Al-Mutairi, and M. S. Ali, “Discriminative Feature Learning for Skin Disease Classification Using Deep Convolutional Neural Network,” IEEE Access, vol. 8, pp. 39025–39033, 2020.

[14]. S. Fei, “Implementation of Fuzzy Expert System on Skin Diseases,” International Journal of Innovative Technology and Exploring Engineering (IJITEE), vol. 3, no. 2, pp. 1–5, Apr. 2013.

[15]. L. AlSuwaidan, “Deep Learning Based Classification of Dermatological Disorders,” Journal of Medical Internet Research, vol. 25, p. e11795972221138470, 2023.

APPENDICES

SOURCE CODE

```
# Text-to-Speech Conversion Using Google Text-to-Speech (gTTS)
from gtts import gTTS
import os
import platform
import subprocess
def text_to_speech(input_text, output_file="doctor_voice.mp3"):
    """Convert input text to speech and save as an MP3 file."""
    tts = gTTS(text=input_text, lang='en', slow=False)
    tts.save(output_file)
    # Optional: Play audio automatically (based on OS)
    if platform.system() == "Windows":
        subprocess.run(['powershell', '-c', f'(New-Object Media.SoundPlayer "{output_file}").PlaySync();'])
    elif platform.system() == "Darwin":
        subprocess.run(['afplay', output_file])
    elif platform.system() == "Linux":
        subprocess.run(['aplay', output_file])
    return output_file
# Example usage
# text_to_speech("Welcome to MediMind AI - Your Virtual Dermatologist")
```

```
# Voice Input: Record and Transcribe Patient Speech using Microphone and Whisper (via Groq API)

import speech_recognition as sr

from pydub import AudioSegment

from io import BytesIO

from groq import Groq

import os, logging

# Configure logging

logging.basicConfig(level=logging.INFO)


def record_patient_audio(save_path="patient_voice.mp3", timeout=20, phrase_time_limit=None):

    """Records audio from the microphone and saves it as an MP3 file."""

    recognizer = sr.Recognizer()

    with sr.Microphone() as source:

        recognizer.adjust_for_ambient_noise(source, duration=1)

        logging.info("Recording... Please speak.")

        audio = recognizer.listen(source, timeout=timeout, phrase_time_limit=phrase_time_limit)


        # Convert WAV to MP3

        wav_data = audio.get_wav_data()

        segment = AudioSegment.from_wav(BytesIO(wav_data))

        segment.export(save_path, format="mp3", bitrate="128k")
```

```
        logging.info(f"Audio saved at {save_path}")
def transcribe_audio_with_groq(audio_path, model="whisper-large-v3"):
    """Transcribes the recorded audio file using Groq's Whisper model."""
    groq_key = os.getenv("GROQ_API_KEY")
    client = Groq(api_key=groq_key)
    with open(audio_path, "rb") as audio_file:
        response = client.audio.transcriptions.create(
            model=model,
            file=audio_file,
            language="en"
        )
    return response.text
# Example usage
# record_patient_audio()
# print(transcribe_audio_with_groq("patient_voice.mp3"))
# Analyze Medical Images Using Groq's Multimodal LLM with a Text Prompt


import os
import base64
from groq import Groq
def encode_image_to_base64(image_path):
    """Encodes an image as Base64 for input to the AI model."""
    if not image_path or not os.path.exists(image_path):
```

```
        return None
    with open(image_path, "rb") as img:
        return base64.b64encode(img.read()).decode("utf-8")
def ask_ai_about_image(prompt, image_path=None, model="meta-llama/llama-4-scout-17b-16e-instruct"):
    """Sends a question and an optional image to Groq's AI model for analysis."""
    client = Groq(api_key=os.getenv("GROQ_API_KEY"))
        message = [{"role": "user", "content": [{"type": "text", "text": prompt}]}]
        # Attach image if provided
    encoded_image = encode_image_to_base64(image_path)
    if encoded_image:
        message[0]["content"].append({
            "type": "image_url",
            "image_url": {"url": f"data:image/jpeg;base64,{encoded_image}"}
        })

    try:
        response = client.chat.completions.create(
            model=model,
            messages=message
        )
        return response.choices[0].message.content
    except Exception as err:
```

```
        return f"AI Error: {str(err)}"

# Example usage

# result = ask_ai_about_image("What disease symptoms can you detect?", "sample_skin_image.jpg")

# print(result)

# === Load Environment ===

from dotenv import load_dotenv

load_dotenv()


# === Imports ===

import os

import datetime

import threading

import numpy as np

import matplotlib.pyplot as plt

from PIL import Image

from fpdf import FPDF

from deep_translator import GoogleTranslator

import gradio as gr

import tensorflow as tf

from tensorflow import keras

# === Custom Modules ===

from brain_of_the_doctor import encode_image, analyze_image_with_query
```

```
from voice_of_the_patient import transcribe_with_groq

from voice_of_the_doctor import text_to_speech_with_gtts

# === Load Models ===

def load_model(path, name):

    print(f"Loading {name}...")

    model = keras.models.load_model(path)

    print(f"{name} loaded successfully!")

    return model


skin_disease_model = load_model("skin_disease_model.keras", "MobileNetV3 Model")

resnet_model = load_model("ResNet_model.keras", "ResNet Model")

densenet_model = load_model("DenseNet121_skin_disease_model.keras", "DenseNet121 Model")


# === Constants ===

class_names = ["Acne", "Eczema", "Melanoma", "Unknown"]


LANGUAGE_CODES = {

    "English": ("en", "NotoSans-Regular.ttf"),

    "Hindi": ("hi", "NotoSansDevanagari-Regular.ttf"),

    "Bengali": ("bn", "NotoSansBengali-Regular.ttf"),

    "Tamil": ("ta", "NotoSansTamil-Regular.ttf"),

    "Telugu": ("te", "NotoSansTelugu-Regular.ttf"),
```

```
    "Marathi": ("mr", "NotoSansDevanagari-Regular.ttf")
}
explanation_prompt = """
You are a dermatologist AI. Your task is to classify the given skin image as *Acne, Eczema, Melanoma, or Unknown*.
*Rules for your response:*
Strictly return the classification using *only one* of these options:
   - Acne
   - Eczema
   - Melanoma
   - Unknown

Follow this exact format:
---
Possible Skin Condition: (Acne, Eczema, Melanoma, or Unknown)
Reason: (Briefly explain the visible symptoms in the image.)
Medical Explanation: (Describe the causes and contributing factors of the condition.)
Causes: (Describe the causes)
Symptoms: (Comma-or-bullet list of common symptoms)
Recommendation: (Simple advice; do not mention any product names, or next steps.)
---
"""
```

```
# === Preprocessing ===
def preprocess_image(path):
    image = Image.open(path).convert("RGB").resize((224, 224))
    return np.expand_dims(np.array(image, dtype=np.float32) / 255.0, axis=0)
# === Grad-CAM ===
def get_gradcam_heatmap(model, img_tensor):
    last_conv = next((layer.name for layer in reversed(model.layers) if isinstance(layer, tf.keras.layers.Conv2D)), None)
    if not last_conv:
        raise ValueError("No conv layer found.")
    grad_model = tf.keras.models.Model(inputs=model.inputs,
                                       outputs=[model.get_layer(last_conv).output, model.output])
    with tf.GradientTape() as tape:
        conv_output, prediction = grad_model(img_tensor)
        loss = prediction[:, tf.argmax(prediction[0])]
    grads = tape.gradient(loss, conv_output)
    pooled_grads = tf.reduce_mean(grads, axis=(0, 1, 2))
    heatmap = tf.maximum(conv_output[0] @ pooled_grads[..., tf.newaxis], 0)
    return tf.squeeze(heatmap) / tf.reduce_max(heatmap)
# === PDF Generator ===
def generate_pdf(response, prediction, lang):
    code, font_file = LANGUAGE_CODES.get(lang, ("en", "NotoSans-Regular.ttf"))
    font_path = os.path.join(".", font_file)
```

```
    if not os.path.exists(font_path):
        return None
    pdf = FPDF()
    pdf.add_page()
    pdf.add_font('LangFont', '', font_path, uni=True)
    pdf.set_font('LangFont', '', 12)
    pdf.multi_cell(0, 10, f"MediMind AI - Diagnosis Report\n\nSkin Disease Prediction:\n{prediction}\n\nDoctor's Response:\n{response}")
    output_path = f"skin_disease_report_{datetime.datetime.now().strftime('%Y%m%d_%H%M%S')}.pdf"
    pdf.output(output_path)
    return output_path
# === Main Function ===
def process_inputs(image_path, lang, model_name, mode):
    try:
        img_tensor = preprocess_image(image_path)
        predicted_class, confidence = "Unknown", 0.0
        doctor_response, pdf_path = "", ""
        # Model selection
        model_map = {
            "MobileNetV3": skin_disease_model,
            "ResNet Model": resnet_model,
            "DenseNet Model": densenet_model
```

```python
}

# Single model
if mode == "Single Model Prediction":
    model = model_map[model_name]
    preds = model.predict(img_tensor)[0]
    confidence = float(np.max(preds) * 100)
    predicted_class = class_names[np.argmax(preds)] if confidence >= 60 else "Unknown"
    skin_prediction = (
        f"Disease: {predicted_class} (Confidence: {confidence:.2f}%)"
        if predicted_class != "Unknown"
        else "I cannot confidently identify this condition. Please consult a dermatologist."
    )

# Multi-model
else:
    results, model = [], skin_disease_model
    for name, mdl in model_map.items():
        p = mdl.predict(img_tensor)[0]
        conf = float(np.max(p) * 100)
        cls = class_names[np.argmax(p)] if conf >= 60 else "Unknown"
        results.append((name, cls, f"{conf:.2f}%"))
```

```
        skin_prediction = "| Model | Predicted Class | Confidence |\n|-------|------------------|------------|\n"
        skin_prediction += "\n".join([f"| {m} | {c} | {conf} |" for m, c, conf in results])
        preds = model.predict(img_tensor)[0]
        predicted_class = class_names[np.argmax(preds)]

    # Grad-CAM and Chart
    heatmap = get_gradcam_heatmap(model, img_tensor)
    heatmap_img = Image.fromarray((heatmap * 255).astype('uint8')).resize((224, 224)).convert('RGB')
    overlay = Image.blend(Image.open(image_path).resize((224, 224)), heatmap_img, alpha=0.4)

    fig, ax = plt.subplots()
    ax.bar(class_names, preds * 100)
    ax.set_ylabel('Confidence (%)')
    fig.tight_layout()

    # Doctor Explanation
    doctor_query = (
        f"{explanation_prompt}\nCondition: {predicted_class}. Please provide a medical explanation."
        if predicted_class != "Unknown"
```

```
            else "This condition does not match any known skin diseases. Please consult a dermatologist."
        )

        doctor_response = analyze_image_with_query(
            query=doctor_query,
            encoded_image=encode_image(image_path),
            model="meta-llama/llama-4-scout-17b-16e-instruct"
        )

        if lang != "English":
            doctor_response = GoogleTranslator(source='en', target=LANGUAGE_CODES[lang][0]).translate(doctor_response)
        # TTS
        voice = text_to_speech_with_gtts(doctor_response)
        # PDF
        pdf_path = generate_pdf(doctor_response, skin_prediction, lang)
        return doctor_response, voice, skin_prediction, pdf_path, fig, overlay, round(confidence, 2)
    except Exception as e:
        return str(e), None, "Error during processing.", None, None, None, 0.0
# === Gradio Interface ===
main_app = gr.Interface(
    fn=process_inputs,
```

```
    inputs=[
        gr.Image(type="filepath", label="Upload Skin Image"),
        gr.Radio(list(LANGUAGE_CODES.keys()), value="English", label="Select Language"),
        gr.Radio(["MobileNetV3", "ResNet Model", "DenseNet Model"], value="MobileNetV3", label="Select Model"),
        gr.Radio(["Single Model Prediction", "Compare All Models"], value="Single Model Prediction", label="Prediction Mode")
    ],
    outputs=[
        gr.Textbox(label="MediMind AI Response"),
        gr.Audio(label="Doctor's Voice"),
        gr.Textbox(label="Prediction Result"),
        gr.File(label="Download Report (PDF)"),
        gr.Plot(label="Prediction Confidence Chart"),
        gr.Image(label="Grad-CAM Heatmap"),
        gr.Number(label="Prediction Confidence (%)")
    ],
    title="MediMind AI"
)
```