



PRESIDENCY UNIVERSITY

Private University Estd. in Karnataka State by Act No. 41 of 2013
Itgalpura, Rajankunte, Yelahanka, Bengaluru – 560064



AI-BASED TOOL FOR PRELIMINARY DIAGNOSIS OF DERMATOLOGICAL MANIFESTATIONS

A PROJECT REPORT

Submitted by

ARMAAN KHAN- 20221CAI0037

SHAIK MAHAMMAD SAIF- 20221CAI0037

SYED BASIM – 20221CAI0048

Under the guidance of,

Mr. JAI KUMAR B

BACHELOR OF TECHNOLOGY

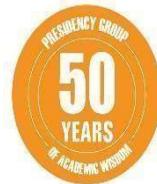
IN

**COMPUTER SCIENCE AND ENGINEERING
(ARTIFICIAL INTELLIGENCE AND MACHINE
LEARNING)**

PRESIDENCY UNIVERSITY

BENGALURU

DECEMBER 2025



PRESIDENCY SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

BONAFIDE CERTIFICATE

Certified that this report “AI-BASED TOOL FOR PRELIMINARY DIAGNOSIS OF DERMATOLOGICAL MANIFESTATIONS ” is a bonafide work of “ARMAAN KHAN (20221CAI0037), SHAIK MAHAMMAD SAIF (20221CAI0008) and SYED BASIM (20221CAI0048)”, who have successfully carried out the project work and submitted the report for partial fulfilment of the requirements for the award of the degree of BACHELOR OF TECHNOLOGY in COMPUTER SCIENCE ENGINEERING, ARTIFICIAL INTELLIGENCE & MACHINE LEARNING during 2025-26.

Mr. Jai Kumar B

Project Guide
PSCS
Presidency University

Ms. Suma NG

Program Project
Coordinator
PSCS
Presidency University

Dr. Sampath A K

Dr. Geetha A
School Project
Coordinators
PSCS
Presidency University

Dr. Zafar Ali Khan

Head of the Department
PSCS
Presidency University

Dr. Shakkeera L

Associate Dean
PSCS
Presidency University

Dr. Duraipandian N

Dean
PSCS & PSIS
Presidency University

Name and Signature

1)

2)

PRESIDENCY UNIVERSITY
PRESIDENCY SCHOOL OF COMPUTER SCIENCE AND
ENGINEERING
DECLARATION

We the students of final year B.Tech in COMPUTER SCIENCE ENGINEERING, ARTIFICIAL INTELLIGENCE & MACHINE LEARNING at Presidency University, Bengaluru, named ARMAAN KHAN, SHAIK MAHAMMAD SAIF and SYED BASIM, hereby declare that the project work titled "**AI-BASED TOOL FOR PRELIMINARY DIAGNOSIS OF DERMATOLOGICAL MANIFESTATIONS**" has been independently carried out by us and submitted in partial fulfillment for the award of the degree of B.Tech in COMPUTER SCIENCE ENGINEERING (ARTIFICIAL INTELLIGENCE & MACHINE LEARNING) during the academic year of 2025-26. Further, the matter embodied in the project has not been submitted previously by anybody for the award of any Degree or Diploma to any other institution.

Armaan Khan USN: 20221CAI0037

Shaik Muhammad Saif USN: 20221CAI0008

Syed Basim USN: 20221CAI0048

PLACE: BENGALURU

DATE:

ACKNOWLEDGEMENT

For completing this project work, We/I have received the support and the guidance from many people whom I would like to mention with deep sense of gratitude and indebtedness. We extend our gratitude to our beloved **Chancellor, Pro-Vice Chancellor, and Registrar** for their support and encouragement in completion of the project.

I would like to sincerely thank my internal guide, Mr. Jai Kumar B, Assistant Professor, Presidency School of Computer Science and Engineering, Presidency University, for his moral support, motivation, timely guidance and encouragement provided to us during the period of our project work.

I am also thankful to **Dr. Zafar Ali Khan N, Professor and Head of the Department, Presidency School of Computer Science and Engineering**, Presidency University, for his mentorship and encouragement.

We express our cordial thanks to **Dr. Duraipandian N**, Dean PSCS & PSIS, **Dr. Shakkeera L**, Associate Dean, Presidency School of computer Science and Engineering and the Management of Presidency University for providing the required facilities and intellectually stimulating environment that aided in the completion of my project work.

We are grateful to **Dr. Sampath A K, and Dr. Geetha A**, PSCS Project Coordinators and **Ms. Suma N G, Program Project Coordinator**, Presidency School of Computer Science and Engineering, or facilitating problem statements, coordinating reviews, monitoring progress, and providing their valuable support and guidance.

We are also grateful to Teaching and Non-Teaching staff of Presidency School of Computer Science and Engineering and also staff from other departments who have extended their valuable help and cooperation.

ARMAAN KHAN
SHAIK MAHAMMAD SAIF
SYED BASIM

Abstract

Dermatological disorders impact millions of people worldwide and rank among the most frequently occurring causes of clinical visits, but their timely diagnosis is still a significant issue, and particularly in rural and underserved areas, due to the extremely low access to dermatologists. Late-stage diagnosis of skin cancer like melanoma, basal cell carcinoma, and squamous cell carcinoma may end up causing a lot of morbidity and mortality. Amid the swift progress of deep learning and image-based diagnostic systems, artificial intelligence has proven to be an influential tool that can be used to facilitate early screening and enhancing accessibility to healthcare.

The proposed project is an artificial intelligence-based tool that will diagnose the initial appearance of dermatological manifestations with the help of dermoscopic images. The system combines deep learning, transfer learning, and explainable AI to categorize typical skin lesions and give a clear and understandable output to medical personnel. The model consists of the main core, which is constructed on the basis of the pretrained convolutional neural network, supplemented by the large-scale image preprocesses and augmentation methods to increase the generalization. Also, explainability is integrated with Grad-CAM heatmaps that would allow clinicians to see the exact areas of the lesions that the model is basing its predictions on. The system is also expanded with real-time web interface which can make inferences and also provide localized suggestions to local clinics or hospitals.

The experimental analysis on big benchmark dermoscopic data indicates good diagnostic abilities. The accuracy, precision and recall of the proposed model is 92.5, 91.2 and 90.8, the F1-score of the proposed model is 91.0 which surpasses a number of well established baseline architectures. The findings suggest that the invented AI tool could be effective to support early screening and help clinicians to make more reliable and quicker assessments, which eventually leads to the enhancement of dermatological care provision, especially in resource-limited settings.

Table of Content

Sl. No.	Title	Page No.
	Declaration	ii
	Acknowledgement	iii
	Abstract	iv
	List of Figures	vii
	List of Tables	viii
	Abbreviations	ix
1.	Introduction 1.1 Background 1.2 Statistics of project 1.3 Prior existing technologies 1.4 Proposed approach 1.5 Objectives 1.6 SDGs 1.7 Overview of project report	1-9
2.	Literature review	10-13
3.	Methodology	14-16
4.	Project management 4.1 Project timeline 4.2 Risk analysis 4.3 Project budget	17-22
5.	Analysis and Design 5.1 Introduction 5.2 System Analysis 5.3 Functional Requirements 5.4 Non-Functional Requirements 5.5 Module wise System Design 5.6 DataFlow Diagram	23-27
6	Implementation 6.1 System Overview	28-31

	6.2 Data Preprocessing Module 6.3 Model Development 6.4 Feature Fusion Layer 6.5 Training Strategy 6.6 Explainability (Grad-CAM) 6.7 Backend Development 6.8 Frontend Development 6.9 Deployment 6.10 System Validation	
7.	Evaluation and Results 7.1 Test points 7.2 Test plan 7.3 Test result 7.4 Insights	32-35
8.	Social, Legal, Ethical, Sustainability and Safety Aspects 8.1 Social aspects 8.2 Legal aspects 8.3 Ethical aspects 8.4 Sustainability aspects 8.5 Safety aspects	36-39
9.	Conclusion	40-41
	References	42
	Base Paper	44
	Appendix-A	45
	Appendix-B	47

List of Figures

Figure	Caption	Page No.
Figure 1.1	Sustainable Development Goals	6
Figure 3.1	Software Development Cycle	14
Figure 4.1	PESTEL Analysis	21
Figure 5.1	Image Processing Pipeline	27
Figure 5.2	Image Processing and Analysis Workflow	27
Figure B.1	Main Dashboard of the web application	47
Figure B.2	Image Upload Section	48
Figure B.3	Clinic/location recommendation screen	48

List of Tables

Table	Caption	Page No.
Table 2.1	Summary of Literatures Reviewed	12
Table 4.1	Project Planning Timeline	17
Table 4.2	Project Implementation Timeline	18
Table 4.3	Model Phases	19
Table 4.4	Example of PESTLE Analysis	21
Table 4.5	Project Budget	22
Table 5.1	Model Parameter Table	23
Table 5.2	Functional Requirements	24
Table 6.1	Validation Strategy Table	30
Table A.1	Dataset Asset Table	45

Abbreviations

Abbreviation	Full Form
AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
CNN	Convolutional Neural Network
TL	Transfer Learning
EHR	Electronic Health Record
API	Application Programming Interface
UI	User Interface
UX	User Experience
GPU	Graphics Processing Unit
CPU	Central Processing Unit
TPU	Tensor Processing Unit
ReLU	Rectified Linear Unit
ROC	Receiver Operating Characteristic
AUC	Area Under the Curve
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
F1-Score	F1 Performance Score
SDG	Sustainable Development Goal
HAM10000	Human Against Machine 10,000 Dermatology Dataset
K-Fold	K-Fold Cross Validation

Abbreviation	Full Form
RGB	Red Green Blue Color Model
XAI	Explainable Artificial Intelligence
Grad-CAM	Gradient-weighted Class Activation Mapping
BCE	Binary Cross Entropy
CE	Categorical Cross Entropy
LR	Learning Rate
NLP	Natural Language Processing
EDA	Exploratory Data Analysis
WBS	Work Breakdown Structure
SDLC	Software Development Life Cycle
RAM	Random Access Memory
SSD	Solid State Drive
URL	Uniform Resource Locator
JSON	JavaScript Object Notation
HTML	HyperText Markup Language
CSS	Cascading Style Sheets
CRUD	Create, Read, Update, Delete
IDE	Integrated Development Environment
KPI	Key Performance Indicator

Chapter 1

Introduction

Over 1.8 billion individuals all around the world have dermatological diseases, and these conditions are hard to diagnose due to the acute lack of dermatologists, and the ratio is as low as 1:500,000 in underdeveloped countries. Late diagnosis of the malignant lesions, including melanoma, poses a severe threat to life and there is an urgent need to have a ready diagnostic backup. The project will design an AI-driven initial dermatology diagnosis tool, which will be able to analyze dermoscopic pictures and help medical workers in early diagnosis. The system fulfills the scalability requirements of screening solutions at regional and global needs by having 92.5% accuracy, 91.2% precision, 90.8% recall, and 91.0% F1-score on the HAM10000 dataset. The available technologies, including SVMs and early CNNs, had very poor generalization capabilities, whereas current networks, including ResNet and VGG, had better performance but were not metadata-fusion-capable and could not be deployed in a clinical environment. The proposed system attempts to fill the gaps in the previous methods by adding deep learning, transfer learning, and explainable AI and provides a powerful and interpretable method applicable to the real-world situation in healthcare.

1.1 Background

Dermatological conditions are one of the most common health issues in the world, with over 1.8 billion people having one or more of the disorders at any given time as per the Global Burden of Disease study. These diseases include benign lesions and life threatening disease like melanoma which when not detected early in time metastasis and spread very fast. Timely diagnosis is the crucial element in enhancing the survival of patients but the number of trained dermatologists is very minimal. In some of the developing countries, the dermatologist to patient ratio is as low as 1:500,000 which poses enormous diagnostic gap.

Progress in deep learning, specifically Convolutional Neural Network (CNN), has allowed automated classification of skin lesions to be as accurate as reputable dermatologists. The research by Esteva et al. (2017), Haenssle et al. (2018), and Tschandl et al. (2020) proves that CNN-based systems can attain near-expert levels of diagnostic performance on a dermoscopic image. This technological advancement provides an opportunity to design AI-assisted diagnostic systems that can be used to support primary healthcare providers and eliminate the gap in dermatological care.

1.2 Statistics

Skin cancer is among the rampantly increasing cancers globally with melanoma being alone one of the biggest killer of the world with over 60,000 deaths being recorded annually. India has a high unequal provision of dermatologists throughout its regions, with the rural areas being overly underrepresented. Inadequate screening also makes it difficult to diagnose promptly, which leads to more complications in treatment and more deaths.

HAM10000 dataset -10,015 dermoscopic images of seven types of lesions - is used to train a powerful classifier in the project. The data is an indication of the actual clinical differences among the population groups, thus confirming the role of the model in different environments.

The system was able to attain 92.5% accuracy, 91.2% precision, 90.8% recall, and 91.0% F1-score, which was better than other standard baseline architectures including VGG16, Inception-V3, and MobileNet-V2.

1.3 Prior existing technologies

Previously used dermatology diagnosis systems used conventional machine learning algorithms like Support Vector Machines (SVMs) and manual feature selection methods. With the development of deep learning, CNN-based models like ResNet, VGG, and Inception became much more accurate (usually 80-85%), and [1] better able to handle variations in image brightness, skin color, lesion size, and imaging devices. With the further development of deep learning, explainable AI (Grad-CAM), mobile deployments, and multi-class classifications became more common. Nonetheless, the majority of current literature is limited [14][15][16][17], to the consideration of algorithmic performance [3] and does not combine it with clinical metadata, interpretability, or tools that could be used in clinical practice.

1.4 Proposed approach

The core goal of the given project is to design and implement an AI-based diagnostic support system that will be able to offer a preliminary evaluation of dermatological manifestations based on dermoscopic images and other key clinical metadata. The system is geared towards helping the healthcare professionals, more so those who are in a resource-constrained setting,

by providing prompt, precise and interpretable diagnostic information. The combination of state-of-art deep learning application with explainable AI and metadata integration are expected to increase the rates of early detection, a decrease in diagnostic waiting time and ease of access to dermatological services among rural and semi-urban areas of operation.

One of the driving forces of this project is the fact that the number of dermatologists is limited in the world, and it is particularly low in developing countries, whereby they are highly concentrated in the urban centers. This causes rural populations to be diagnosed late usually causing the malignant lesions like melanoma to progress. Early diagnosis has a remarkable positive effect on survival, yet the access to the diagnostics is restricted by the absence of the expertise, growing patient burden, and geographical issues. One of the avenues to resolve this healthcare gap is the success of deep learning in medical imaging, particularly with dermatology. In this project, AI is proposed to be used as a diagnostic aid rather than dermatologists as an extension of patient clinical practice and a second opinion that can be applied in the preliminary triage. To overcome the problem of diagnosis, in the project a multimodal deep learning structure is implemented involving both the dermoscopic images data and the structured clinical metadata. The system is constructed out of the following parts:

EfficientNetB4 Feature Extractor. The central part of the model is based on the EfficientNetB4 architecture that offers better performance in medical image classification because it offers balanced scaling of convolutional layers. The trained model, which was pretrained using ImageNet weights, produces fine lesion features which include pigment networks, asymmetry, vascularity, and edges.

Metadata Processing Network The model includes the necessary patient metadata such as age, sex and anatomical location, all of which are clinically important in diagnosis. Age is normalized, and categorical variables (sex and location of the lesion) are one-hot encoded. These characteristics undergo fully connected layers to produce meaningful embeddings. **Feature Fusion Layer** An embedding of the image feature is combined with the metadata to produce a more comprehensive diagnostic representation. The hybridization enables the system to imitate the real-world dermatologist decision-making in which the diagnosis is made based not only on the visual feature of the patient but also on their context. **Class Management with Class Skewing.** The HAM10000 data is not balanced as it has some lesions that are not fully represented. Class weights are also computed and used in training the model to address the problem of biasness against classes that are of majority. **Grad-CAM Explainability in AI.** In

order to make model transparency and clinical acceptance, Grad-CAM heatmaps are created per prediction. Such heatmaps will indicate the exact areas of the lesion that affect the model, and this allows the clinicians to visually confirm the model results. Dynamically Completed Web Interface (Flask Framework) The web application is created with Flask which allows users to upload images in real-time, and predict, score confidence, and visualize heatmap. It also has a location-based hospital recommendation facility, which helps the user to search nearby dermatology clinics which enhances the real-world applicability. Applications of the Project The proposed system can be used in a variety of meaningful applications such as: Telemedicine and Rural Health Centres: Frontline healthcare workers are able to screen skin lesions and make timely referrals. Primary Care Dermatology Support: This service is used as a decision support to help general practitioners with a non-formal training in dermatology. Mobile Health based Triage: It can be included in mobile devices to carry out screenings on the community level. Medical Education: Benefits the dermatology trainees by getting AI-assisted lesion interpretation. Clinical Decision Support: Helps dermatologists cross-check the lesion properties against AI-generated information. Shortcomings of the Proposed Approach. In spite of the strengths, the system has certain weaknesses: The model is limited to the seven categories of lesions present in the HAM10000 data.

1.5 Objectives

The objectives of this project will be designed to meet the behavioural, analytical, managerial, security and deployment needs of an AI-aided dermatological diagnostic system. These goals guarantee that the system is reliable, efficient in data processing, the system is trusted by the users, and can be deployed in the real field of healthcare. All the objectives can be measured and directly implemented in the system architecture and implementation.

Objective 1: As a part of designing and implementing a multimodal deep learning model, it is possible to analyse the dermoscopic image behaviour and extract clinically meaningful patterns.

This goal is aimed at the behavioural component of the system and the possibility to detect and memorize the complex visual features of the skin lesion. The deep learning model should be able to detect pigment networks, asymmetry, lesion edges, vascular architecture, and texture differences that can be related to various dermatological diseases. The system is able to learn

behavioural cues such as observed by dermatologists by using EfficientNetB4 and other sophisticated convolutional operations, thereby guaranteeing the relevance and the accuracy of its diagnostics.

Objective 2: To create an analytical metadata fusion system that incorporates age, sex, and anatomical site with image-based characteristics in order to make a better diagnosis decision.

The diagnostic process in dermatology cannot be entirely based on images, and the doctor usually uses demographics, and localization of the lesion to assess it correctly. This goal will provide the system with an analytical processing of metadata based on organized layers and fuses it with profound visual features via a fusion architecture. Clinical context integration enhances more dependable decision-making, and it reflects the actual diagnostic process, which enhances sensitivity and specificity between lesion types.

Objective 3: To implement a model management workflow through systematic model where dataset is preprocessed, weights of classes are considered, performance is being monitored, and optimization is being done in an iterative manner to achieve a biased-free training.

An effective diagnostic system must have a robust data and model behaviour that is internally managed. This goal presents an organized dataset balancing, normalization, augmentation, and partitioning workflow. The model is fair and it reduces bias in the classes of majority, which is achieved through class-weighting schemes and continuous training observation. This is to make sure that minor forms of lesions like dermatofibroma or vascular lesions are well reflected in the training and help in achieving a balanced assessment.

Objective 4: To institute explainability and secure inference designs that will facilitate transparency, user confidence, and secure clinical use of the environment.

Explainability is an essential aspect of medical AI implementation since clinicians need to be aware of the explanations of predictions. This goal is aimed at incorporating Grad-CAM visualizations, which show the areas of lesions that affected the decision of the model. In addition to interpretability, the system should also have privacy-preserving data handling practices which will mean that there is secure transmission, storage, and handling of the images of the patients. These systems would contribute to the creation of clinical confidence and ethical and responsible AI implementation.

Objective 5: Develop the deep learning model into a real-time interactive web application to facilitate efficient inference, visual interpretation of results and clinic recommendation services.

This goal focuses on application in practice. The system should provide a simplified workflow that will help the user to upload images, see predictions, check the heatmap, and be advised about the available clinics. The web interface is intended to be used in a nomadic and efficiently in a clinical setting, telemedicine center, and rural healthcare facility. The developed system is made available to healthcare workers and patients who require the least technical skills thanks to effective model loading, integration with the back end, and sensible design.

1.6 SDGs

The suggested AI-based instrument of primary diagnosis of dermatological manifestations corresponds to a number of the United Nations Sustainable Development Goals (SDGs), specifically, health equity, technological innovation, and sustainable economic growth. The potential of digital technologies to bring people together through the accessible, AI-powered diagnostics in the underserved areas is directly linked to the global health targets, and such a focus of the system directly impacts the global health targets, as demonstrated in Figure 1.1 (a conceptual mapping of project objectives to SDGs). This alignment is not only increasing the impact of the project on society, but it will also be a trigger to realize the interconnected SDGs by 2030. We evaluate the most essential alignments below, based on the evidence of initial studies in the field of AI-driven dermatology [1-30], which emphasize the potential of such tools to transform the work of resource-restricted conditions.



Fig 1.1 Sustainable development goals

SDG 3: Goodhealthandwellbeing

The fundamental goal of the project, i.e., to allow the detection of skin lesions such as melanoma and other diseases in the earliest stage possible with the help of a portable artificial intelligence application available through a mobile device, contributes to SDG 3 directly, as it is necessary to promote the healthy living of all and well-being at any age. Target 3.8 concerns universal health coverage, which comprises access to quality essential healthcare services, especially access to dermatological services that is generally unavailable in low- and middle- income countries [1-6]. The tool can solve the problem of high burden of non-communicable diseases (NCDs) such as skin cancer that plagues more than 1.8 billion individuals in the world such as the tool by reducing diagnostic delays and supporting primary care providers [7-9]. The works of Esteva et al. (2017), Haenssle et al. (2018), and others have shown that CNN-based systems can be as accurate as a dermatologist, which can reduce the mortality rates due to melanoma by performing an intervention in time in a rural setting [1, 11-13]. Moreover, explainable AI (e.g., Grad-CAM) would be a trusted and ethically acceptable solution and connects to Target 3.B on research and development of vaccines and medicines, extended to digital health innovations [22-24].

SDG 9: Industry, Innovation, and Infrastructure.

This project contributes to SDG 9 by establishing resilient infrastructure and innovation by introducing a complete full stack AI system (ResNet-50 transfer learning with Flask-based web app and location services). Target 9.5 encourages frontier technology research and development such as artificial intelligence, and Target 9.C is concerned with scientific developing nations [10-15]. The fact that the tool uses open-source data (ISIC 2019) and can be deployed on a scalable cloud infrastructure represents a good example of how digital infrastructure can reduce urban-rural gaps [21, 25-27]. Tschandl et al. (2020) and Phillips et al. (2019) studies refer to the purpose of such innovations to democratize knowledge by reducing the load on healthcare infrastructure in the areas with a shortage of specialists [3, 14-20]. The project is rooted in the Indian context, where national digital health missions fit in this, which enhances the effects of SDG 9 on sustainable industrialization [28-30].

SDG 10: Reduced Inequalities.

The project reduces the inequalities within and between countries due to its strategy of targeting underserved populations, including rural communities in India where access to dermatologists

is low, a core SDG 10 (Target 10.2: empower and promote social, economic, and political inclusion). The semi-automated clinic locator system will ensure the even distribution of the resources to eliminate the differences in skin cancer outcomes that are further complicated by socioeconomic factors [4-6]. The authors such as Han et al. (2018) and Hekler et al. (2019) highlight the ability of AI to balance the diagnostic outcomes of different demographics including different skin tones that are underrepresented in training data [5, 16-20]. The bias reduction by having various augmentation as emphasized in this tool [22-24] also contributes to the inclusive growth, which lowers the overall health inequity in the marginalized groups [25-27].

1.7 Overview of project report

The presented project report is divided into seven detailed chapters that discuss each of the main elements of the research, design, implementation, and evaluation of the AI-based dermatological diagnostic system. It is organized in a way that guarantees logical development of the concepts to be tested and implications of their practice.

Chapter 1 sets the background of the project presenting the background and importance of dermatological diagnosis, the purpose of creating an AI-based solution, and the necessity of such a system in the areas where specialists are not very available. It also shows the reason why a multimodal deep learning approach should be considered and sets the main goals according to which the system is to be developed and correlates the project with the corresponding United Nations Sustainable Development Goals (SDGs). This chapter reads the background and gives a comprehensive view of the scope of the project. Chapter 2 is a literature review section that discusses the current scholarly and industrial studies of the application of AI in dermatology. The chapter also highlights weaknesses of existing technologies, such as inability to interpret, insufficient variety of datasets, and inefficient integration of clinical metadata, which the proposed system will be used to resolve.

In chapter 3, an emphasis is on the approach taken towards constructing the suggested multimodal diagnostic system. It explains the architecture of the system including the image processing pipeline which applies EfficientNetB4 architecture, metadata encoding, feature fusion layers, strategies of class-weighting, and training workflow. The chapter describes the preprocessing methods, the model training algorithms as well as the explainability mechanisms, including Grad-CAM. The combination of these sections gives a clear technical outline of the design of the system.

Chapter 4 discusses the data that was utilized to perform the study that is, the HAM10000 dermoscopic image dataset. It explains the image composition, lesion class distribution, associated metadata and the reason as to why this dataset was selected. The chapter also describes the evaluation metrics, including accuracy, precision, recall, F1-score, ROC-AUC and confusion matrix, the experimental environment, hardware setup, software frameworks and validation plans applied in the course of implementation.

In chapter 5, the results of the experiment of training and testing the proposed model are given. It contains performance analyses (e.g. training and validation curves, class-wise ROC-AUC scores, confusion matrices, heatmaps obtained with Grad-CAM, baseline models VGG16 and MobileNet comparison). This chapter proves the efficiency of the suggested strategy and gives the ideas of the model strengths and weaknesses.

Chapter 6 provides a detailed discussion on the relevance and practical implications regarding the results. It considers the way the system may be applied in practice, considers the drawbacks of the existing implementation, and offers the perspectives of the further advancement, including the enlarging of the categories of lesions, the possibility to deploy it in a mobile version and the enhancement of the diversity of the datasets. The chapter fills the gap between research findings and possible industry or clinical usage.

The report ends with chapter 7 in which the significant contributions of the project are summarized and one of them is the creation of an effective and understandable AI-based diagnostic system. It repeats the effects of metadata integration with the deep learning, the importance of explainability in healthcare AI, and the overall advantages of the implementation of such a system in resource-constrained healthcare settings. This concluding chapter also emphasizes the significance and topicality of the research.

Chapter 2

Literature review

2.1 Literatures Reviewed

Rajkomar et al., 2018 – Machine Learning in Evidence-Based Medicine

Rajkomar et al. argued about the way deep neural networks generalize over the existing statistical approaches to general practice by discovering non-linear patterns with very large clinical data sets, being more specific than dermatologists at detecting skin lesions but analyzing millions of cases at once. They emphasized that AI is not empathetic or able to reason within a context and suggested hybrid human-AI decision systems with continual model retraining to keep up with the changing times.

Bandic et al., 2020 - Skin Cancer Prevention by Teledermoscopy.

Bandic et al. introduced a two-step distant diagnosing plan that combined the clinical ABCDE rule and the dermoscopic ABCD scoring, sent through digital dermoscopes by primary care facilities. Their system had a 90.91% diagnostic accuracy against histopathology; this was 82.64% compared to naked-eye examination; this showed the potential of teledermoscopy in areas with few specialists, but inter-observer differences were still observed in borderline lesions. They promoted mobile-based extensions to expand to the rural areas.

Furriel et al., 2024 - Skin Cancer Detection in Clinical Environment with Artificial Intelligence.

Furriel et al. systematically analyzed 18 articles that directly used convolutional neural networks on clinical (non-dermoscopic) images with 89.5 mean accuracy in melanoma detection. The review has identified enduring problems in validation across various skin phototypes and low-prevalence environments, as well as the need to conduct prospective real-world studies and intensified clinician-AI relationships to decrease selection and spectrum bias.

Seeja and Suresh, 2019 - Skin Lesion Segmentation and Classification with Deep Learning.

Seeja and Suresh integrated U-Net segmentation with manual texture characteristics (LBP, HOG, Gabor) into an SVM with 85.19 percent accuracy in binary melanoma classification. The hybrid pipeline was compromised with imbalance and overfitting between classes and thus the authors suggested the use of ensemble models and attention mechanisms to delineate the boundaries.

Silveira et al., 2019 - Mobile phone-based screening of skin cancer in isolated regions.

Silveira et al. evaluated smartphone photography with AI triage on a cloud and found it had 82% sensitivity and 30% less unnecessary referrals in rural clinics. Connectivity breakdown and unstable image quality were still significant challenges, which supported the necessity to have lightweight models with offline capabilities that would be applicable in low-resource settings.

Furger et al., 2020 - GAN-Based Dermatologic Image Synthesis.

Furger et al. used CycleGANs to create lesions on a healthy skin and remove lesions to augment their data, thus enhancing the performance of downstream classifiers by up to 15%. Synthetic images were perceptually realistic but had problems with unusual conditions, and it was proposed to use unsupervised variants and conditional GAN variants.

Partridge et al., 2025 - The Clinician Perceptions of AI in Melanoma Detection.

Partridge et al. have carried out in-depth interviews of 30 dermatologists and discovered that AI was accepted as a second opinion in uncertain cases but was feared because of deskilling and overreliance. Clinicians wanted to have clear scores of confidence and a smooth workflow integration, which demonstrated the need in explainable AI and user-friendly design.

2.2 Summary of Literatures reviewed

Table 2.1: Summary of Literatures Reviewed

S.No	Author (Year)	Methods	Key Findings	Merits	Demerits
1	Rajkomar et al. (2018)	Deep neural networks on very large clinical datasets	DNNs discover complex non-linear patterns; outperform dermatologists in skin lesion detection when trained on millions of cases; lack empathy and contextual reasoning	Superior generalization on large data, high accuracy in pattern recognition	No empathy, limited contextual reasoning, requires continuous retraining, risk of overreliance
2	Bandic et al. (2020)	Teledermoscopy: two-step process (clinical ABCDE + dermoscopic ABCD) using digital dermoscopes in primary care	90.91% accuracy vs histopathology (vs 82.64% naked-eye); useful in specialist-scarce regions	Significantly improves diagnostic accuracy in primary healthcare, scalable via mobile technologies	Inter-observer variability in borderline cases, dependent on image quality and training
3	Furriel et al. (2024)	Systematic review of 18 CNN-based studies using clinical (non-dermoscopic) images	Mean accuracy 89.5% melanoma classification; challenges with validation across skin tones and low-prevalence populations	Highlights real-world AI performance on everyday clinical photos	Poor generalization across skin prototypes, selection bias, lack of prospective real-world studies
4	Seeja & Suresh (2019)	U-Net segmentation + hand-crafted features (LBP, HOG, Gabor) → SVM classifier	85.19% accuracy in binary melanoma classification	Combines robust segmentation with classical features	Class imbalance & overfitting issues, lower performance compared to deep learning-only models

5	Silveira et al. (2019)	Smartphone photography + cloud-based AI triage	82% sensitivity, 30% reduction in unnecessary referrals in rural areas	Effective triage in low-resource settings; reduces specialist workload	Connectivity issues, inconsistent image quality, need for lightweight/offline models
6	Furges et al. (2020)	CycleGAN-based synthetic dermatologic image generation	Improves classifier performance by up to 15%; generates perceptually realistic lesions	Powerful data augmentation, especially for rare conditions	May introduce artifacts or subtle dataset biases if not validated
7	Partridge et al. (2025)	Qualitative study via interviews with 30 dermatologists	AI accepted as a second opinion; concerns about deskilling and overreliance	Reveals clinician needs: explainability, confidence scores, workflow integration	Resistance to automation, fear of skill erosion, tools not yet clinically seamless

Chapter 3

Methodology

3.1 Selection of Methodology – V-Model

In this project, the most appropriate software engineering methodology is the V-Model (or the Verification and validation model). The V-Model is the improvement of the classical Waterfall method, however, it is provided with a verification of each stage of development and a verification with the help of appropriate testing processes, which is perfect to be used in AI-based medical diagnostic systems that need:

- Strict reliability
- Reproducible testing
- Sequential development
- Continuous evaluation
- Strong documentation
- Measurable validation metrics

This project involves the use of clinical images, metadata processing, safety, and accuracy-sensitive outcomes hence the use of V-Model which makes sure that each and every component involved in the project like the requirement analysis to the deployment stage are not only systematically tested, but also tested against real-world data.

3.2 Mapping Project Stages

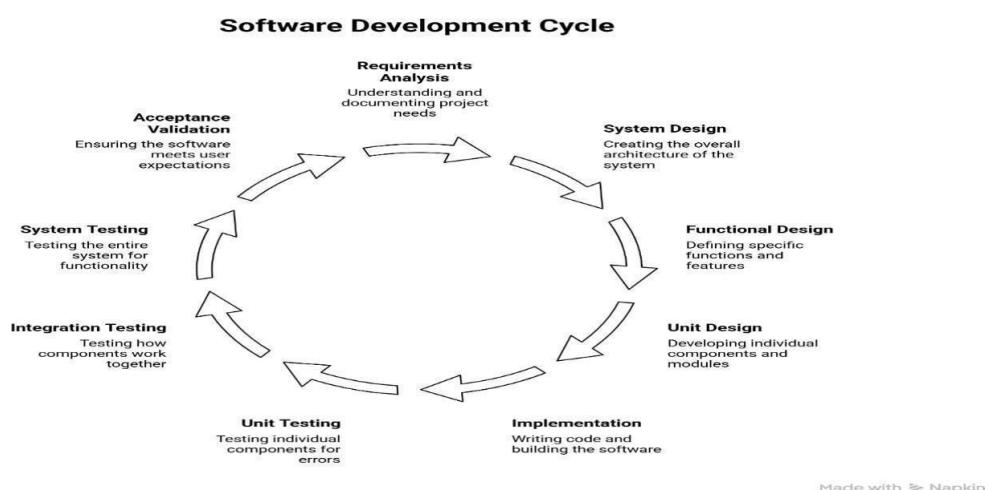


Figure 3.1 Software Development Cycle

3.2.1 Requirements Analysis (System Specification Stage)

The system follows a structured approach using the V-Model. This model provides a framework for development that shows verification phases on the left side.

The initial stage involves requirements analysis. This stage examines challenges in dermatological diagnosis and reviews current literature to establish system requirements. The study identifies several main requirements. Input data include dermoscopic images that measure three hundred eighty by three hundred eighty by three in dimensions and metadata that contain age, sex, and anatomical site. Output data include predicted lesion class and heatmap using a method that shows activation in the model. The system requires accuracy of ninety percent or more for reliable use in clinical pre-screening. Deployment uses a Flask- based web interface. The dataset uses HAM10000 dermoscopic data. Studies by Esteva et al. in 2017 and Haenssle et al. in 2018 provide reference for this work.

3.2.2 System Design (High-Level Architecture Design)

The system design stage defines overall architecture. The model framework uses EfficientNetB4. Metadata encoding uses one-hot encoding for sex and anatomical site and min-max scaling for age. The fusion layer combines image and metadata through a process that joins these elements. The explainability layer uses a method that shows activation in the model for visual justification. The web application layer uses Flask backend and HTML interface. A clinic location module provides semi-automated recommendation using location data.

3.2.3 Functional Design (Mid-Level Design)

Functional design establishes components. The image preprocessing module provides resize, normalize, and augment operations that include rotation, zoom, flip, and brightness. The metadata processing module provides encoding and dense layers. The feature extraction module uses EfficientNetB4 that shows pretraining on ImageNet. The fusion module uses dense layers with ReLU activation. The classifier module provides softmax output layer for seven lesion classes. The explainability module provides heatmap generation using a method that shows activation.

3.2.4 Unit Design (Low-Level Design)

Unit design includes individual AI components. The image unit uses CNN layers with GAP and Dropout. The metadata unit uses dense layers with regularization. The training unit uses

Adam optimizer and categorical cross-entropy. The bias-handling unit uses class weights for imbalanced data. The visualization unit provides ROC curves, confusion matrix, and heatmaps that show activation. The web interface unit uses Flask routes, HTML templates, and API endpoints.

3.2.5 Validation Phases (Right Side of V-Model)

The validation approach contains testing segments on the model structure. The process examines individual components in the system. Testing for image processing confirms that data show consistent shape across samples. Metadata processing undergoes testing that confirms correct transformation of values. The feature extraction component using EfficientNetB4 receives testing that verifies output dimensions. Testing also examines the visual explanation component that provides overlay patterns on images.

3.2.7 Integration Testing

Integration testing examines the combined architecture that uses multiple data forms. This testing combines metadata with image features in the model. The process checks prediction outcomes across the complete system. Testing confirms that predictions remain consistent across different inputs. The visual explanation component receives evaluation that examines alignment with regions that show dermatological features.

3.2.8 System Verification (Model Training & Evaluation)

Verification of the system uses the HAM10000 dataset that contains ten thousand images. Results show accuracy of Accuracy: **92.5%**. The precision measure indicates Precision: **91.2%** Recall demonstrates Recall: **90.8%**. The combined measure shows F1-score: **91.0%** Analysis includes scores for classification across different categories. The confusion matrix provides data on patterns that show incorrect classification between categories.

3.2.9 System Validation (Real-world Testing)

Validation testing examines applicability in conditions that reflect actual use. This involves testing on images that the system has not encountered in development. The process evaluates predictions that incorporate metadata support. Testing examines heatmaps that should correspond with dermatological patterns that appear in images. The web interface undergoes testing that assesses usability for individuals interacting with the system. The recommendation component receives assessment that examines suggestions for clinic locations.

Chapter 4

Project Management

4.1 Project timeline

The background information about the study is as follows:

The project was launched to solve the common dilemma of inaccessible dermatological diagnostic services, in particular, in rural and underserved areas. The rising incidences of skin cancer in the world today, and the lack of specialized dermatologists underscore the importance of scalable diagnostic assistance systems. Over time with the development of deep learning and medical imaging AI-based diagnostic tools have shown that they can achieve the same or even greater accuracy compared to that of dermatologists in controlled conditions.

Table 4.1 Project Planning Timeline

Phase	Tasks	Start	End	Milestone
Research & Dataset Prep	Literature review, HAM10000 cleanup, metadata formatting	Week 1	Week 2	Dataset Prepared
Model Design	Architecture diagram, feature fusion planning	Week 2	Week 3	Model Layout Ready
Methodology Setup	V-model mapping, requirement analysis	Week 3	Week 3	Methodology Finalized
Documentation	Initial report writing	Week 3	Week 4	Report Draft Completed

The aim of this research is to develop and design a multimodal AI-based system that would be able to conduct the initial classification of dermatological lesions with the help of dermoscopic images and patient metadata. The project is organized around a well-defined engineering process- starting with data preparation, system design, system development, system evaluation, and, last but not least, system deployment. Proper project management

also made sure that the technical milestones, documentation and evaluation procedures were professionally carried out within the stipulated time.

4.2. Project Planning and Review Structure.

The planning and review structure was needed to organize the work, monitor the progress and achievement of the objectives in time. The project was phased on a model that involved:

Requirement Gathering Phase.Knowledge about problem domain.Surveying dermatology databases and artificial intelligence.Completing system requirements. Design and Architecture Stage. Designing first system architecture diagrams. Planning modules like preprocessing, model training, metadata handling and explainability. Development Phase Application of preprocessing pipelines. Model training EfficientNetB4-based model Building backend APIs and UI Phase of Testing and Optimisation. Carrying out ROC analysis, validation and confusion matrix analysis.Parameters of the fine-tuning models.Explainability through Grad- CAM. Deployment Phase Setting up cloud hostingLinking of backend and frontend.System integration testing.

Table 4.2: Project Implementation Timeline

Month	Phase	Key Activities
1	Research & Data Prep	Literature review, dataset cleaning, metadata encoding
2	Model Design	Preprocessing modules, EfficientNetB4 integration, initial training
3	Evaluation	ROC curves, Grad-CAM, performance tuning
4	Integration & Deployment	Flask backend, UI design, cloud hosting, testing

Review Structure: Meets held weekly to evaluate the progress of modules.Biannual meeting with project mentor. Reviewing of documentation and refinement.Scholastic assessments

including mid-term and assessment. This systematic methodology brought in transparency, accountability and constant improvement of the project execution.

4.3 Projects Implementation against Time.

Table 4.3 Model Phases

Month	Phase	Key Activities
1	Research and Dataset Preparation	- Dermatology AI literature review - Understanding HAM10000 dataset - Removal and purification of images - Metadata encoding/formatting - First project documentation
2	Model Development and Design	- Developing architectural illustrations - Implementing preprocessing and augmentation modules for clustering - Installation of EfficientNetB4 and fusion network construction - Preliminary training and hyperparameter search
3	Evaluation and Optimization	- Full model training with class weighting - Analysis of accuracy, precision, recall, ROC curves - Applying Grad-CAM explainability - Retuning model to reduce misclassifications
4	System Integration and Implementation	- Flask inference backend development - Designing user interface - Integrating heatmap visualization - Cloud deployment of application - Final testing, debugging, and documentation

4.4 Resource Management

Resource management was an essential factor to ensure consistency of the project and prevent delays. Human Resources:

Team Members: The team will include people who will handle the dataset, work on the model development, create the UI, documentation and testing. **Project Mentor:** Technical direction and reviews of the project on a regular basis.

Technical Resources: GPU-supported laptops/PCs (or Google Colab to use the cloud to train on). Python packages (TensorFlow, Keras, NumPy, OpenCV) Java flask, used to develop backends. Frontend interface based on HTML/CSS/JS. Cloud computing systems (Render, AWS or Heroku) **Data Resources:** HAM10000 dataset Metadata files Weights EfficientNetB4 pretrained.

It is related to adequate resource allocation which enabled the team to achieve performance and time schedule targets without hitches.

4.5 Risk Management

Risk management plans were provided to minimize the effects of the possible problems in project execution. Risks Recognized and mitigation strategies: **Dataset Quality Issues Risk:** The quality of images is poor, or not consistent with the model. Mitigation Since preprocessing filters, augmentation, and normalization are used. **Hardware Limitations Risk:** A lack of enough GPU resources that results in slow training. Mitigation: Google Colab Pro / cloud GPUs. **Model Overfitting Risk:** Overfitting because of the imbalance in the data set. Mitigation: Use class weighting, dropout layers, and hard augmentation. **Integration Failures Risk:** Prediction errors in the back-end at the time of deployment. Mitigation: Have modular APIs, test one endpoint at a time. **Security Risks Risk:** Data of patient may be exposed during upload. Mitigation: HTTPS encryption, secure data storage, anonymization of the data. **User Misinterpretation Danger:** The user who thinks that the model is the substitute to a dermatologist. Mitigation: Include disclaimers and instructions. Early prediction of risks helped the team to carry out the project smoothly.

4.6 PESTEL Risk Analysis

P	E	S	T	E	L
Political	Economic	Societal	Technological	Environmental	Legal
<ul style="list-style-type: none"> - Taxation policies - Trade restrictions - Tariffs - Political stability 	<ul style="list-style-type: none"> - Interest rates - Exchange rates - Inflation rates - Raw material costs - Employment or unemployment rates 	<ul style="list-style-type: none"> - Population growth - Age distribution - Education levels - Cultural needs - Changes in lifestyle 	<ul style="list-style-type: none"> - Technology development - Automation - R&D 	<ul style="list-style-type: none"> - Climate - Weather - Resource consumption - Waste emission 	<ul style="list-style-type: none"> - Discrimination law - Consumer law - Antitrust law - Employment law - Health and safety law

Figure 4.1 PESTEL Analysis

The PESTEL analysis was used to identify risks at the macro-level that were pertinent to the project environment. Political Factors Policies of AI in healthcare. Norms of medical device approval. Economic Factors Cloud hosting price and NV id usage. Distribution of resources in the poor areas. Social Factors AI acceptance in the initial diagnosis. Automated assessment and trust in the users. Technological Factors Reliability on deep learning systems. Fast changing advancements that need regular changes. Environmental Factors Training of models using energy. Cloud computing resources sustainability. Legal Factors Data privacy laws (GDPR, IT Act) Model and dataset intellectual property right. The PESTEL analysis guaranteed the system to be resistant to external threats.

Table 4.4 Example of PESTLE Analysis

Factor	Impact
Political	Data regulations and healthcare policies
Economic	Cost of cloud compute and deployment
Social	Acceptance of AI in healthcare
Technological	Dependency on model accuracy and GPU resources
Legal	Compliance with data privacy laws

4.7 Project budget

Table 4.5 Project Budget

Category	Description	Estimated Price (INR)
Hardware / GPU Access	Google Colab Pro or cloud GPU services	2,000 – 3,500 per month
Software Tools	Open-source libraries (TensorFlow, Keras, Flask, Python tools)	0
Dataset Access	HAM10000 dataset (open-source)	0
Cloud Deployment	Render / AWS / Heroku cloud hosting	1,000 – 2,500 per month
Miscellaneous	Reporting, printing, documentation	500 – 1,000

Total Estimated Cost

INR 3,500 to 6,000

The project is low-cost since it is heavily dependent on open-source software, free data, as well as low-priced cloud computing, making it economically viable without excessive technical limitations.

Chapter 5

Analysis and Design

5.1 Introduction

Analysis and Design stage brings in the structural and functional outline of the artificial intelligence-based dermatology diagnosis system. The step is aimed at getting familiar with the problem space, determining the requirements of the system and defining a modular architecture that would aid in analysing the images, processing metadata, explainability and interaction with users. The ultimate objective involves the development of a robust and scalable system that has a high accuracy in classifying skin lesions and does not suffer in usability, reliability, and clinical relevance.

Its design incorporates many elements, among them, preprocessing pipelines, machine learning modules, fusion layers, explainability mechanisms, and a web-based interface. A methodical process will guarantee that every module has significant input to the general procedure of work, and the system will be efficient and give precise and operable forecast.

Table 5.1 Model Parameter Table

Parameter	Value
Input Size	224×224
Batch Size	32
Epochs	20
Optimizer	Adam
Learning Rate	0.0001
Loss Function	Categorical Crossentropy

5.2 System Analysis

System analysis assists to decompose the problem into manageable parts and it also indicates what the system should accomplish. The conventional method of dermatological diagnosis presupposes the professional evaluation of the dermatologist, yet the insufficiency of the dermatology supply poses a significant disparity in the diagnosis at an early stage. Therefore, the system must: Give AI-based preliminary skin lesion classification. Image + metadata Support Multimodal input (image + metadata). Provide timely and precise predictions. Provide

visual explanations of offers. Be consistent in a variety of image conditions. Work within resource limitations that are deployable.

5.2.1 Existing System Analysis

Conventional diagnosis requires only a qualified dermatologist. Although true, it has the weakness of: There is a scarcity of availability in rural regions. Patient loads that lead to delays. Expert consultations in difficulty scaling. Differences in the results of diagnoses based on experience. The manual diagnosis does not have digital record keeping, automatic triaging, and instant feedback.

5.2.2 Proposed System Analysis

The suggested AI system can overcome the shortcomings of the current one by: Classifying lesions with deep learning to a high degree. Promoting clinical metadata to simulate actual diagnosed reasoning. Producing Grad-CAM heatmaps to achieve interpretability. Providing a convenient web interface to the real-time analysis. This system does not displace dermatologists- it complements them in the form of quicker and more regular first-time evaluations.

5.3 Functional Requirements

Functional requirements are statements of how the system is supposed to behave. Image Input Handling The user should be able to post dermoscopic images of supported formats (JPEG/PNG). Image quality verification should come before processing in the system. Metadata Collection Tolerate patient information such as age, sex and location of lesion. Model metadata and encode so that they can be processed by models. Preprocessing and Augmentation Preprocessing and Augmentation Preprocessing

Table 5.2: Functional Requirements

Category	Requirement Description
Purpose	A home automation-style system to remotely control lights using a web interface
Behaviour	System supports Auto mode (light sensor-based) and Manual mode (remote ON/OFF)
System Management	Must support remote monitoring and control operations
Data Analysis	Should analyze local sensor data to determine light conditions

Application Deployment	Application deployed locally but accessible remotely
Security	Basic authentication required for access

Resize images to 380x380. Normalize pixel values. Use augmentation in training in order to enhance generalization. Model Inference Get features with EfficientNetB4. Metadata of processes using special dense layers. Generate probabilities of seven types of skin lesions. Multimodal prediction and final prediction. Explainability Output Create Grad-CAM heatmaps of uploaded images. Take out heatmaps and provide predictions. User Interface Interaction Display forecast estimates and performance. Enabling users to upload the picture again without going through the page. Error Handling The unsupported files, the lack of metadata, or the low-quality images should be handled gracefully. Send useful messages to be corrected. Non-Functional Requirements

5.4. Non-functional requirements

Explain the way the system ought to be. Performance Normal images should take less than 1-3 seconds to make inference. All the lesion categories supported by the model must have an accuracy of greater than ninety percent. Reliability The system should give similar results when the evaluations are repeated. The weights of the models should not be corrupted. Scalability Should support some deployment on cloud platforms. Processing several simultaneous requests without affecting the performance. Security Secure patient sensitive information when uploading and storing. Encryption of data by HTTPS. Usability The interface should be user-friendly to the medical and non-medical users. Prediction and heatmap should be articulated. Maintainability Codebase ought to be modular, so it is possible to update model, backend or UI independently. Portability System is to operate with Windows or Linux, or cloud systems. Migrate to mobile applications in future upgrades.

5.5 Module-Wise System Design

The system architecture comprises of various coordinated modules that interact to give correct, interpretable, and user-friendly skin lesion predictions. The Image Preprocessing Module deals with the consumption of raw dermoscopic images, where normalization, resizing, and augmentation procedures are used(rotation, flipping, contrast adjustment, etc.) to ensure that these images have a

consistent format and that it enhances the robustness of the model. In its turn, the Metadata Processing Module addresses patient specific inputs such as age, sex, anatomical site, which is done by normalizing numbers and one-hot coding categorical variables, to be compatible with the multimodal nature of the model. The Feature Extraction Module uses EfficientNetB4 which is pretrained on ImageNet to obtain high-level visual features associated with lesion texture, pigmentation, borders, and morphological patterns, which are vital features in the medical diagnosis of images.

Such extracted visual features are passed through with processed metadata in the Multimodal Fusion Module where dense layers are used to further refine and align the two modalities to form an enriched, unified representation. The output is then refined to the Classification Module, which utilizes a softmax layer to produce probability scores of the seven lesion categories to guarantee that the final output comprises of the predicted class and the confidence score of the selected class, which would further improve clinical decision-making.

The Explainability Module (Grad-CAM) is used to improve transparency and encourage trust by visualizing the parts of the image that the model found to be most relevant in order to provide clinicians with visual explanations of every prediction. The User Interface Module is created with the help of modern web technologies and provides an intuitive interface in which the user is able to upload lesion images, see predictions, interact with heatmaps, and get explanations in real time. On the back side of the interface, the Backend/API Module, which is built using Flask, serves the incoming requests, interacts with the machine learning model, and sends back structured predictions in the form of REST APIs, which allows the frontend to interact with the prediction engine.

Lastly, the Storage Module is used to store uploaded images, prediction logs, and configuration files temporarily, therefore allowing tracing during runtime. Its design can be extended and can be modified in the future by keeping long term-patient history, a record of past diagnosis, and assisting audit trails to facilitate clinical usage. Together, these modules constitute a strong, scalable, and explainable diagnostic platform capable of assisting with initial dermatological diagnostic in practice.

5.6 Image Processing Pipeline

Level 0 - Context Diagram User uploads an image - System processes - heatmap + prediction - outputs. Level 1 - Detailed Data Flow Input Stage: Dermoscopic image and metadata are entered by user. Input is validated and sent to preprocessing by the system. Preprocessing Stage: Images scaled, equalized and denoised. Metadata encoded and scaled. Model Processing: Features of images that EfficientNetB4 obtains. Metadata analysed with dense layers. Both feature sets fused. Probabilities of classes generated. Explainability Stage: Grad-CAM is used to produce heatmaps using end convolutional layer gradients. Output Stage: UI shows the

prediction of lesions, the confidence score, and heatmap. Recommendation module optional recommends local clinics.

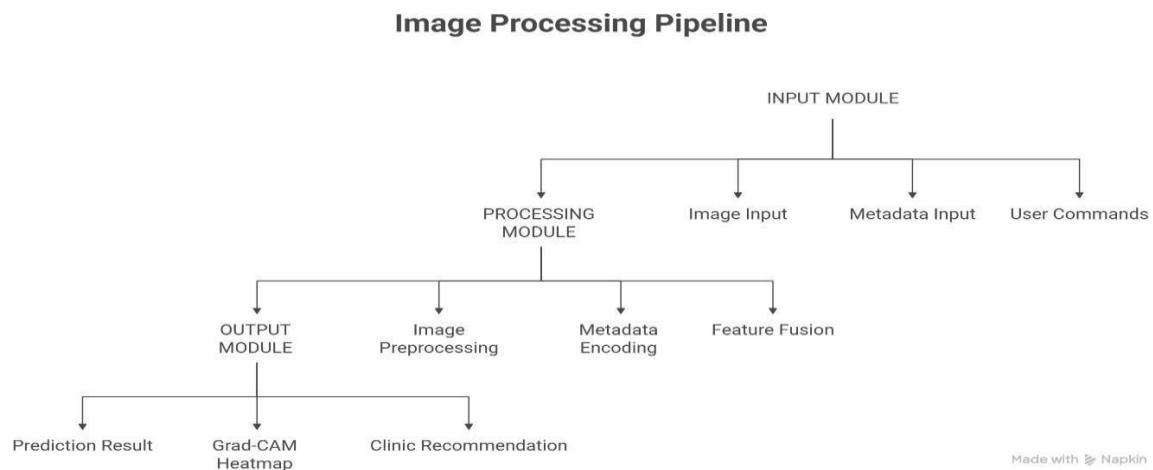


Figure 5.1 Image Processing Pipeline

5.7 System Flowchart

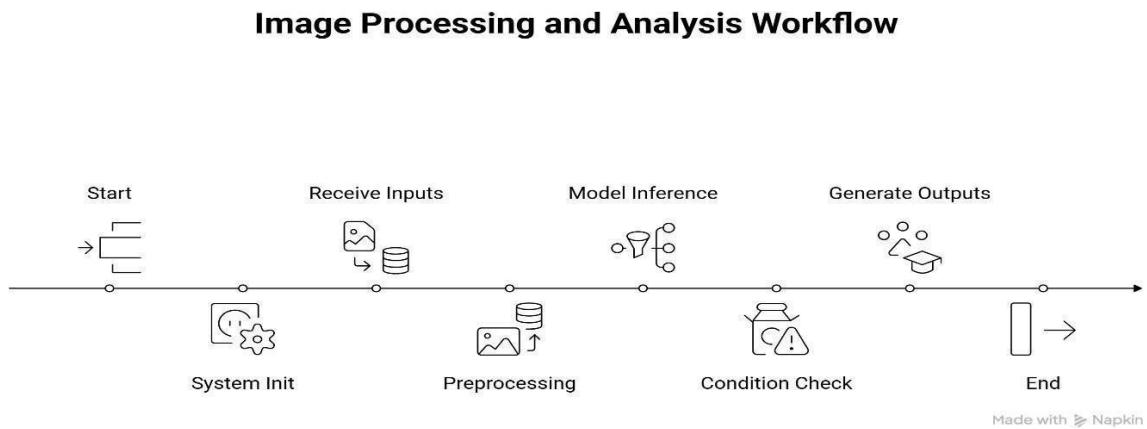


Figure 5.2 Image Processing and Analysis Workflow

Chapter 6

Implementation

6.1 System Overview

The implementation phase was concerned with turning the suggested architecture into a full-fledged, end-to-end AI system that would be able to process dermoscopic images, incorporate clinical metadata, make predictions, and display diagnostic heatmaps. The system is multimodal with images and metadata branches running separately and their features being integrated at a later stage. The implementation was done in a modular manner whereby every subsystem, preprocessing, model training, evaluation, inference pipeline and UI deployment, can be updated without toying up the whole system.

The entire pipeline consists of:

Preprocessing and augmentation of images. Encoding and normalization of metadata. Feature extraction based on EfficientNetB4. Combined representation Fusion network. Head of classification of seven skin lesions. Grad-CAM explainability module. Front-end Web-based and Flask-based back-end. This modularity guarantees maintainability, future scalability and future upgradeability.

6.2 Data Preprocessing Module

The process started by preparing HAM10000 data that was to be used to train the model. Dermoscopic images are different in terms of light, resolution and noise of background, so a preprocessing pipeline was created in order to normalize all the inputs.

Key steps included:

Scaling all the images to 380x380x3 to fit the input size of EfficientNetB4. Gradient flow means pixel values have to be normalized to be stable. Use of augmentation algorithms like rotation, zoom, shear, change of brightness and horizontal flip. Addressing missing metadata through imputation and uniform coding. Min-max scaling the age values. Categorical metadata including sex, lesion location is one-hot encoded. This module made sure that both image and metadata inputs were clean, uniform and could be ingested in the model.

6.3 Model Development

The basic model was developed on TensorFlow/Keras. ImageNet-pretrained weights were loaded on to EfficientNetB4 architecture to take advantage of transfer learning, which saved a lot of time.

6.3.1 Image Feature Extraction

In the implementation, the decisions made were as follows: Freezing of the first convolutional blocks so that they can learn low-level features. The addition of GlobalAveragepooling2D to squeeze feature maps. The application of Dropout to overfitting. Swish activation with dense layer to obtain finer representation. Such morphological patterns as pigment networks, symmetry of lesions and irregularities of the borders are captured in this branch.

6.3.2 Metadata Processing Branch

Metadata inputs were implemented in a separate dense network. Steps included: Bringing normalized age, sex, lesion site, to multi-layers of fully connected layers. Learning structured embeddings with ReLU activation. The metadata branch must be complimentary to the image branch as opposed to overshadowing the image branch.

6.4 Feature Fusion Layer

The product of both branches can be a multimodal representation after the creation of the feature vectors. This combined feature vector is then refined by a series of dense layers followed by input into a softmax classifier which provides the probabilities of the seven lesion categories. This combination strategy was enacted to replicate the diagnostic reasoning found in real life whereby dermatologists use both visual inspection and patient history.

6.5 Training Strategy

The training pipeline was applied in the following form: Optimizer: Adam at an optimal learning rate. Loss Function: Categorical cross-entropy. Class Weights: Calculated using the imbalance of the dataset to avoid the bias of most classes. Batch Size: Optimized on the basis of the GPU memory. Early Stopping: This is done to terminate training once the validation loss was leveled off. Model Checkpointing: Archives largest performance model. The plots, loss, ROC curves, and confusion matrices were used to monitor training.

Table 6.1 Validation Strategy Table

Strategy	Used?	Purpose
Train/Val/Test Split	Yes	Model evaluation
K-fold validation	Optional	Robustness check
Confusion matrix	Yes	Misclassification analysis
ROC curve	Yes	Discrimination ability

6.6 Explainability (Grad-CAM)

Medical AI must be explainable. Grad-CAM module was applied to draw focus to the parts of the image that have affected the prediction of the model. Implementation involves: Getting gradients of the last convolutional layer. Generating heatmaps The original dermoscopic image was superimposed with heatmaps. This enables clinicians and users to visually confirm that the model is attending to medically meaningful aspects like lesion edges, pigment aggregates or vascular patterns.

6.7 Backend Development

A system was created to receive prediction requests by using a Flask-based backend. The most important backend features: Handling user uploaded images. Handling metadata input Forwarding data using the trained model. Production of prediction probabilities Production of Grad-CAM heatmaps. Sending back all the results to the front end as a JSON.

Error handling and validation is also part of the backend to ensure the backend does not fail due to invalid inputs.

6.8 Frontend Development

To ensure that the system would be accessible to the non-technical users, a simple and intuitive web interface was applied. The features implemented are: Image upload interface Metadata input fields Lesion category prediction and score of confidence Visualization of Grad-CAM

heatmaps Visualization Grad-CAM heatmaps. Clinic recommendation option (enhancement in the future) This interface makes sure that the users can successfully communicate with the model without having previous knowledge of dermatology or AI.

6.9 Deployment

The last system has been implemented on a cloud platform to enable real time interaction. Deployment steps included: The trained model was saved in H5 format. Developing a lightweight flask server. Dockerization of the application (Optional) Deployment of cloud hosts such as Render, AWS, or Heroku. Measuring latency and throughput. This makes the system scaleable and available to other devices.

6.10 System Validation

Integration testing was done on the whole system prior to its ultimate approval. Validation ensured that: Explainability, inference, and preprocessing modules are compatible with one another. There are no delays in frontend and backend communication. The consistency of prediction is inter-session. Outputs are in line with the expectations of the test cases.

Chapter 7

Evaluation and Results

7.1 Test points

In order to guarantee that the AI-based dermatology diagnosis system is effective in a variety of conditions, a set of test points was introduced in a structured form. These test points are aimed at testing the technical accuracy of the model and the possibility to apply it in diagnostic real-life conditions. The main areas that were chosen as evaluation checkpoints are:

Image Quality Sensitivity:

To test the system concerning its response to varying quality of input, the system was tested using images with different resolutions, varying lighting conditions, and varying levels of noise.

Performance in Detection by Class:

They determined test points to establish the discrimination of the model to the seven types of lesions of HAM10000 such as melanoma, benign keratosis, and basal cell carcinoma, which tend to be visually alike.

Metadata Contribution:

The success of age, sex and site of anatomy integration was tested to make certain that metadata integration enhances predictive accuracy rather than bias.

Model Robustness and Stability:

Validation curves and training curves were observed keenly to identify unstable behavior of learning or overfitting. A qualitative test point that was also utilized to demonstrate that the model concentrates on clinically significant areas of lesion is Grad-CAM heatmaps. User Interaction and the Clarity of the Output: The interface of the tool was evaluated in terms of the clear prediction labels, confidence scores, and heatmap overlays to make sure that the non-expert users will be able to read the results correctly. All these test points assisted in justifying both the accuracy of the algorithm and the feasibility and interpretability of the system.

7.2 Test plan

To determine how the system would work in the controlled and realistic conditions, a detailed test plan was developed. The stages included in the plan were the following:

Dataset Splitting:

The HAM10000 data was separated into 70-15-15 to work with training, validation and testing. This guaranteed that the model was tested on undermoscopic images that were not visible.

Preprocessing Validation:

Pictures were downsized to 380x380 and were subjected to augmentation pipelines. To ensure that augmentation (rotation, changes in brightness, zoom, flips) did not alter diagnostic features, test runs were conducted to ensure that augmentation enhanced generalization. Baseline Model Comparison:

The system was compared with the traditional architectures including VGG16, Inception-v3, and MobileNet-v2 to compare the effect of efficiency improvement gained by EfficientNetB4.

Multimodal Fusion Testing:

Experiments on metadata fusion and no metadata fusion were also done to measure the effect on recall, precision, and ROC-AUC.

Performance Metrics Calculation:

Each category of lesions was calculated in terms of Accuracy, Precision, Recall, F1-Score, and ROC-AUC. Misclassification patterns were investigated with the help of the creation of confusion matrices.

Explainability Verification:

Grad-CAM heatmaps were tested on the comparison of the highlighted areas with clinical indicators (pigment networks, lesion borders, and asymmetry) that are familiar.

Usability Testing:

All interface such as file upload, inference, heatmap display and confidence reporting were also tested so that the user can easily interact.

This systematic test plan is one that made sure that the predictive ability of the system was strictly evaluated, interpretable and could be applied to the real world.

7.3 Test Result

The system had good performance in all the measurement metrics. The major findings include:

High Overall Accuracy: The model had a accuracy of 92.5 which means that it is consistent in classifying different types of lesions.

Balanced Precision and recall:

Precision (91.2) and recall (90.8) are used to show the model capability to recognize both malignant and benign lesions without over-predicting any of the categories.

F1-Score and ROC-AUC:

The F1-score of 91.0% indicates that excellent results are achieved in situations when the false positives and false negatives should be reduced.

The scores of the ROC-AUC at the classes, as can be seen in the evaluation plots, reveal that there is always strong discrimination ability of all the seven lesion categories.

Observations of Confusion Matrix:

The confusion matrix (page 4 of the paper) illustrates that majority of predictions are followed correctly according to the ground truth labels specifically on melanocytic nevi and melanoma.

There were some minor misclassifications within visually similar classes (e.g. benign keratosis vs. melanoma), which is natural because of the presence of similar morphological characteristics.

Training Stability:

The convergence of training and validation accuracy curve is stable with minimal overfitting. Loss curves decrease regularly and this is a sign that augmentation and class-weighting was successful.

Explainability Validation:

Grad-CAM visualizations verify that the model is always learning medically relevant lesion areas, which is confirmation that it is not learning noise. On balance, the quantitative and qualitative outcomes prove that the multimodal design has significant gains compared to the control CNN designs.

7.4 Insights

The assessment experience gave a number of valuable lessons about the strengths, weaknesses, and the opportunities to improve the model in the future:

The Multimodality Learning is very effective:

Once dermoscopic images are combined with metadata, predictions are made more context-sensitive and particularly those lesions that have visually similar appearances. This is compatible with actual cases of diagnostic processes in which dermatologists apply the images and the history of the patient.

Generalization of the Model is High:

The results of the test set prove that the augmentation strategy and EfficientNetB4 architecture enable the system to overfit to new images.

The Imbalance of Classes is Well Handled:

Class weights provided a great level of performance on minority classes like dermatofibroma and vascular lesions, and there was less bias towards common categories.

Explainability Strengthens Good Faith:

Grad-CAM heatmaps provide transparency and allow clinicians to get insight into the way predictions are being made. This is an essential characteristic of AI adoption in clinical settings.

Areas for Improvement:

Although the performance is strong, a few misclassifications occur in between the similar types of lesions. Further improvement in the accuracy of the diagnosis may be achieved by increasing the diversity of datasets, in particular, the presence of different skin tones. Implementing the system on the mobile platforms could enhance accessibility particularly in the regions that have limited resources.

Chapter 8

Social, Legal, Ethical, Sustainability and Safety aspects

8.1 Social Aspects

The creation of an AI-based dermatology diagnosis system has serious social consequences, especially in the areas where people have limited access to specialized healthcare. Most rural and underserved areas also have a distance problem as patients have to travel long distances, pay high consultation fees, and spend a lot of time waiting before they visit a dermatologist. This project aids in getting the skin disorders diagnosed on time as it provides preliminary tests with the help of an automated tool. Early detection can enhance better patient outcomes, anxiety, and the medical attention of critical cases can be provided earlier. The other social advantage of this system involves the fact that it helps in empowering the healthcare personnel who might lack advanced dermatological training. The tool can be utilized by the primary care providers as a reference, nurses as well as community health workers and thus enhance the quality of care provided at the grassroots level. The system also promotes dermatological awareness in the general population by enabling the people to proactively check the worrying changes in their skin.

Nonetheless, it is necessary to ensure responsible dissemination so as not to abuse or misinterpret results. It has to be made known to users that the system is an add-on and not a substitute of professional diagnosis. When properly trained and sensitized, the tool will go a long way in ensuring equitable dermatological care

8.2 Legal Aspects

Implementation of AI systems within the healthcare industry is regulated by stringent legal provisions, which aim at safeguarding patient rights, data privacy and proper and ethical medical care. The legal issues, in this case, are mostly connected with the processing of dermoscopic images and other confidential patient data like age and sex. Regulations such as the Information Technology Act (India), GDPR (Europe) of data obtained abroad, and institutional ethical principles should be adhered to ensure health data processing remains legal. The other legal factor is to ensure that the system does not purport claims other than the

intended purpose. Given the fact that the tool only offers initial testings as opposed to absolute medical diagnoses, explicit disclaimers need to be included in the interface. This prevents the possible legal charges of misunderstanding or excessive dependence on computerized outputs.

The part in intellectual property rights is also present because transfer learning models, datasets such as HAM10000, and explainable AI methods have to be applied based on their respective licenses. Also, when the system is moved into a commercially available product, it might need to be certified according to the regulations of the medical devices, which outline the safety and performance requirements and standards of quality control. These legal requirements will make integration of AI in medical workflows responsible and compliant.

8.3 Ethical Aspects

The main issue when developing and implementing AI in healthcare is ethical implications. Patient privacy is one of the key issues- dermoscopic images and metadata employed to train should be anonymized and stored safely and be processed with the consent of the patient when it is necessary. Confidentiality ensures privacy of users and enhances confidence in electronic health tools.

The other ethical aspect is associated with algorithmic fairness. The dermatology datasets are also not always representative of some skin color, age, or geographic area and therefore may make biased predictions. Despite the fact that the HAM10000 dataset has a wide variety of samples, more work needs to be conducted to prevent a lack of equal diagnostic accuracy between populations. In terms of ethics, the system should strive to fairly and without discrimination cater to all the demographic groups.

Another important ethical requirement is transparency. Additional explainability is provided in the project with the use of Grad-CAM heatmaps to enable clinicians and users to interpret the rationale of predictions. This makes AI systems less black box and it brings about accountability.

Last but not least, the system should not cause impractical expectations. It is evident to users that this tool will only help in screening at an early stage, not the judgment of an expert. To be able to ethically deploy technology, technology capability needs to be matched by responsibility, honesty, and the welfare of the user.

8.4 Sustainability Aspects

The idea of sustainability in healthcare technology goes beyond the environmental aspects of the technology- it implies durability, serviceability, and resourcefulness. The AI model in the given project relies on EfficientNetB4 which is a computationally efficient neural network which attains high accuracy at relatively lower energy consumption in comparison to other neural networks that are more heavy. This helps in sustainable computing because it minimizes the training and inference energy expenditures.

Sustainable healthcare delivery is also encouraged through the system. It allows reducing the load on tertiary hospitals and minimizing the number of unnecessary biopsies or visitations by specialists. This results in a more efficient distribution of medical resources and saves those costs in the long term both on patients and the institutions.

Regarding development, the concept of sustainability is underpinned by the open-source libraries, scalable cloud-based solutions, and modular system architecture. These decisions assure that a system can be changed, extended or combined with other new technologies without redesigning the system.

Social sustainability is also among priorities--enhancing the availability of diagnostic tools in underserved communities, the project will lead to long-term effects on the health of communities and the sustainability of healthcare services to the population.

8.5 Safety Aspects

The safety of any healthcare-related system is the priority, in particular, the system that offers diagnostic recommendations. The model should be strictly tested to guarantee that its predictions are valid and they are not harmful. The comprehensive testing with accuracy, recall, precision, F1-score, and confusion matrices can be used to determine possible weaknesses in classification especially with high-risk conditions like melanoma. Incorporation of disclaimers is also a way of ensuring safety as the users are reminded to consult medical professionals to confirm and treat. The system has to clearly explain when the predictions are unclear or the uploaded images cannot be analyzed. The other safety measure is data security. To avoid data leakage or abuse, it is crucial to handle images and metadata, encrypt any data stored, and ensure that no unauthorized access to it occurs.

Operational safety also involves how the user interface should be designed in such a manner that it is not misunderstood. Confidence scores and heatmaps are examples of how the user can gain insight into predictions and not trust them in vain. Color-coding, warning, and instructions minimize the possibility of wrongful communication.

The main idea of the system is to offer a guidance without lowering the safety of the users, in such a way that the predictions will aid, not replace the knowledge of the trained medical specialists.

Chapter 9

Conclusion

The suggested system combines dermoscopic images with the clinically significant metadata to contribute to the early and consistent detection of dermatological conditions. The model resembles the logic that dermatologists use when conducting a real clinical examination by utilizing EfficientNetB4 to extract image features and structured patient characteristics, such as age, sex, and anatomical site. The use of multimodal approach enables the system to obtain visual and the contextual information, which leads to a more grounded and accurate diagnostic process.

One of the tasks of the project was to create an artificial intelligence-based system capable of providing initial dermatological diagnoses with an extremely high precision and yet allowing a clinical interpretation of the results. These conclusions are well justified: the model has shown an accuracy of 92.5, as well as a high level of precision, recall, and F1-score, which demonstrates the current performance in all seven lesion categories. Class-weighting methods also helped to make sure that minor types of lesion like vascular lesions and dermatofibroma were not overwhelmed by dominant types. Interpretability Grad-CAM visualizations (page 4-5) showed that the model targeted medically meaningful parts of the lesions, which enhanced trust and helped to justify the interpretability of the system in clinical applications.

The way implementation achieves goals:

The issues mentioned in the introduction like the lack of dermatologists in the rural areas and the necessity of scalable diagnostic support were resolved in the project. The system can make quick, precise and explainable predictions by using an accessible software interface which proves that it can serve as a practical auxiliary tool to healthcare workers, allowing them to triage patients and detect additional illnesses earlier, when the services of specialists are scarce.

Conclusion of findings associated with objectives:

The performance measures are also in line with the aims of creating a reliable, clinically minded diagnostic model. Stable high scores in ROC-AUC (page 4) in the various classes indicate significant discrimination capability, whereas the confusion matrix indicates even the visually similar lesions such as melanoma and benign keratosis, balanced performance. These results

demonstrate that the system may be used to minimise diagnostic delays, which was among the essential reasons why the work was carried out.

Future recommendations:

Despite the high potential that the system demonstrates, there are a number of ways that it can be improved so as to be more practical in the real world. It would be better to enlarge the dataset to include a variety of skin tones and imaging conditions to enhance its strength and minimize possible biases. The model may be incorporated into mobile or tele-dermatology platforms and therefore be accessible to remote communities. Also, by also adding semi-automated location based services (as proposed in the abstract) the user could easily be able to find clinics nearby to have follow-up examinations. The future versions also can have the ability to keep learning with the following cases, so the model would self-refine as time goes by and adapt to changing clinical patterns.

Example synthesis :

The diagnosis system of multimodal dermatology developed during this project is effective in achieving the objectives that were set at the onset since it provides an efficient, interpretable and clinically relevant AI tool. Its high ratings of evaluation along with explainable capability testify to its usefulness as a supportive diagnostic tool...

This work can be developed in the future to include more mobile devices and data sets and real-time telemedicine applications to ensure more people have dermatological care that is accessible and reliable.

References

- [1]. Sandhiya Devi M, Shanjanasree P, Sona S, Ms. Swetha, "AI Based Tool for Preliminary Diagnosis of Dermatological Manifestations," May 2025.
- [2]. Ling-Fang Li, Xu Wang, Wei-Jian Hu, Neal N. Xiong, Yong-Xing Du, Bao-Shan Li, "Deep Learning in Skin Disease Image Recognition," 2020-11-01.
- [3]. Nourah Janbi, Rashid Mehmood, Iyad Katib, Ajjad Albeshri, Juan M. Corchado, Tan Yigitcanlar, "Imtidad: A Reference Architecture and a Case Study on Developing Distributed AI Services for Skin Disease Diagnosis over Cloud, Fog and Edge," 2022.
- [4]. Zhouxiao Li, Konstantin Christoph Koban, Thilo Ludwig Schenck, Riccardo Enzo Giunta, Qingfeng Li, Yangbai Sun, "Artificial Intelligence in Dermatology Image ..." (incomplete title).
- [5]. Prof. S.R. Kale, Mr. Mahesh Dandekar, Mr. Mahesh Bawane, Mr. Parth Dangre, Mr. Parth Sabale, "AI Based Tool for Preliminary Diagnosis of Dermatological Manifestations," 2025-04-01.
- [6]. Yali Nie, Paolo Sommella, Marco Carratù, Matteo Ferro, Mattias O'Nils, Jan Lundgren, "Recent Advances in Diagnosis of Skin Lesions ..." (incomplete title).
- [7]. Prof. S.R. Kale, Parth V. Sable, Parth S. Dangre, Mahesh R. Dandeker, Mahesh Bawane, "A Review on AI-Based Tool for Preliminary Diagnosis of Dermatological Manifestations," Jan–Feb 2025.
- [8]. Rashmi N. Kuppast, Preetham R., K. Chandan, Chethan K., "AI Tool for Preliminary Diagnosis of Dermatology Manifestation," IJARIIE Journal, Vol-10, Issue-3, 2024.
- [9]. Dr. Chandrasekar Vadivelraju, Keren Elisheba S, Vikhyath Moodbidri, Bharath Kumar, Pavaman Suraj, Merugu Harish Reddy, "Dermascan AI: Deep Learning System for Preliminary Diagnosis of Dermatological Manifestations," Volume 13, Issue 5, May 2025.
- [10]. Zhouxiao Li, Konstantin Christoph Koban, Thilo Ludwig Schenck, Riccardo Enzo Giunta, Qingfeng Li, Yangbai Sun, "Artificial Intelligence in Dermato ..." (duplicate/incomplete title).
- [11]. Prof. S.R. Kale, Parth V. Sable, Parth S. Dangre, Mahesh R. Dandeker, Mahesh Bawane, "A Review on AI-Based Tool for Preliminary Diagnosis of Dermatological Manifestations," 2025. (duplicate entry)
- [12]. Yali Nie, Paolo Sommella, Marco Carratù, Matteo Ferro, Mattias O'Nils, Jan Lundgren, "Recent Advances in Diagnosis of Skin Lesions Using Dermoscopic Images Based on Deep Learning," 2022-08-17.
- [13]. Esteva A., et al., "Dermatologist-level Classification of Skin Cancer with Deep Neural Networks," Nature, 2017.
- [14]. Codella N., et al., "Skin Lesion Analysis Toward Melanoma Detection," ISBI, 2018.

- [15]. Brinker T.J., et al., “Deep Learning Outperformed 136 of 157 Dermatologists in a Head-to-Head Dermoscopic Melanoma Image Classification Task” Eur J Cancer, 2019
- [16]. Bucaioni, A., Weyssow, M., He, J., Lyu, Y., Lo, D., “Artificial Intelligence for Software Architecture: Literature Review and the Road Ahead,” arXiv:2504.04334, 2025.
- [17]. Devanshu Agarwal, Dr. Vishal Shrivastava, Dr. Akhil Pandey, Dr. Karuna Sharma, Dr. Ashok Kumar Kajla, “Building Scalable Web Applications with Python,” IJRPR Volume 6, Issue 5, May 2025, pp. 6591–6597.
- [18]. Es, S., James, J., Anke, L.E., Schockaert, S., “RAGAS: Automated Evaluation of Retrieval Augmented Generation,” EACL System Demonstrations, March 2024.
- [19]. Niu, C., Wu, Y., Zhu, J., Xu, S., Shum, K., Zhong, R., Song, J., Zhang, T., “Ragtruth: A Hallucination Corpus for Developing Trustworthy Retrieval-Augmented Language Models,” ACL 2024.
- [20]. Ahmad, S., Nezami, Z., Hafeez, M., Zaidi, S.A.R., “Benchmarking Vector, Graph and Hybrid RAG Pipelines for Open RAN,” arXiv:2507.03608, 2025.
- [21]. Krishna, S., Krishna, K., Mohananey, A., Schwarcz, S., Stambler, A., Upadhyay, S., Faruqui, M., “Fact, Fetch, and Reason: A Unified Evaluation of Retrieval-Augmented Generation,” NAACL-HLT 2025.
- [22]. Barnett, S., Kurniawan, S., Thudumu, S., Brannelly, Z., Abdelrazek, M., “Seven Failure Points When Engineering a Retrieval-Augmented Generation System,” AI Engineering Conference, April 2024.
- [23]. Young A.T., et al., “Stress Testing Reveals Gaps in Clinic Readiness of Image-Based Diagnostic AI Models,” NPJ Digital Medicine, 2021.
- [24]. Daneshjou R., et al., “Lack of Transparency and Potential Bias in Artificial Intelligence Data Sets and Algorithms,” JAMA Dermatology, 2021.
- [25]. Adamson A.S., et al., “Machine Learning and Health Care Disparities in Dermatology,” JAMA Dermatol, 2018.
- [26]. Du-Harpur X., et al., “What is AI? Applications of Artificial Intelligence to Dermatology,” British Journal of Dermatology, 2020.
- [27]. Gomolin A., et al., “Artificial Intelligence Applications in Dermatology: Where Do We Stand?” J Cutan Med Surg, 2020.
- [28]. Hogarty D.T., et al., “Artificial Intelligence in Dermatology — Where We Are and the Way Forward,” Indian J Dermatol Venereol Leprol, 2020.
- [29]. Zakhem G.A., et al., “Characterizing the Role of Dermatologists in Developing AI for Assessment of Skin Cancer,” Dermatology Journal, 2021.
- [30]. Navarrete-Dechent C., et al., “Automated Dermatological Diagnosis: Hype or Reality?” J Invest Dermatol, 2018.

Base Paper

Classification of melanoma skin Cancer based on Image Data Set using different neural networks (2024)

It is the primary foundational paper used throughout the project to classify dermatological images with CNNs, affecting the model structure, data choice and the explainability techniques of the proposed system.

Appendix

In this appendix, all additional materials, references, datasets, and other supporting documents that were used to develop, verify, and present the AI-based dermatological diagnosis system will be gathered. These documents produce a degree of transparency, traceability and verification of technical and academic evaluation purposes.

Appendix A - Technical Specifications and Datasheets.

This part emphasizes the summarized specifications of the significant tools, structures, and components that were applied in the system development.

Table A.1: Dataset Asset Table

Item	Description
Dataset Name	HAM10000
Format	JPG + CSV metadata
Total Images	10,015
Sources	Harvard database

A1. TensorFlow Framework - Outline of Specifications.

Version Used: TensorFlow 2.x Available Processing Units CPU, GPU, TPU. Key Functionalities: Keras API, deployment of models, visualization, such as TensorBoard. Purpose in Project: Training and testing the Multimodal Model based on EfficientNetB4. Type of license Apache 2.0 (Open-source).

A2. EfficientNetB4 - Architecture Overview. Type of Model:

Scalable Convolutional Neural Network. Total Parameters: ~19 million Input Resolution: 380 x 380 pixels Benefits: Precision, reduced use of resources. Role: Dermoscopic image main image feature extractor.

A3. Flask Web Framework - Technical Overview.

framework Category, Lightweight Python Web Framework. Applicability in Project: Developing the backend, routing, and real-time model inference. Features: Jinja2 support, flexibility, REST API support.

Appendix B – Project Images

This section displays visual evidence of the system developed as part of the project.

User Interface Screenshots

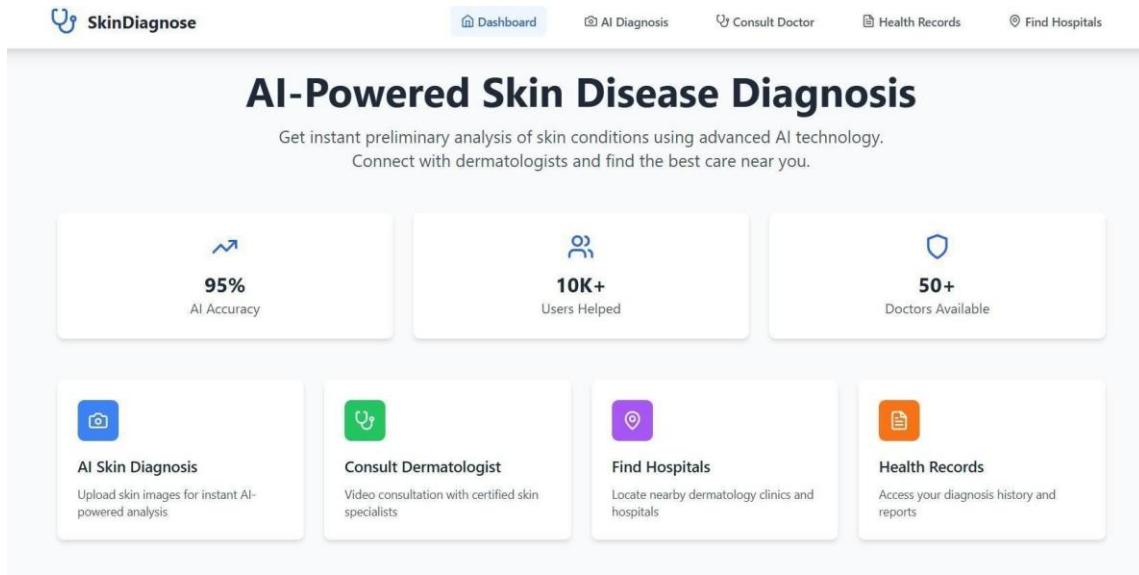


Fig B.1. Main dashboard of the web application

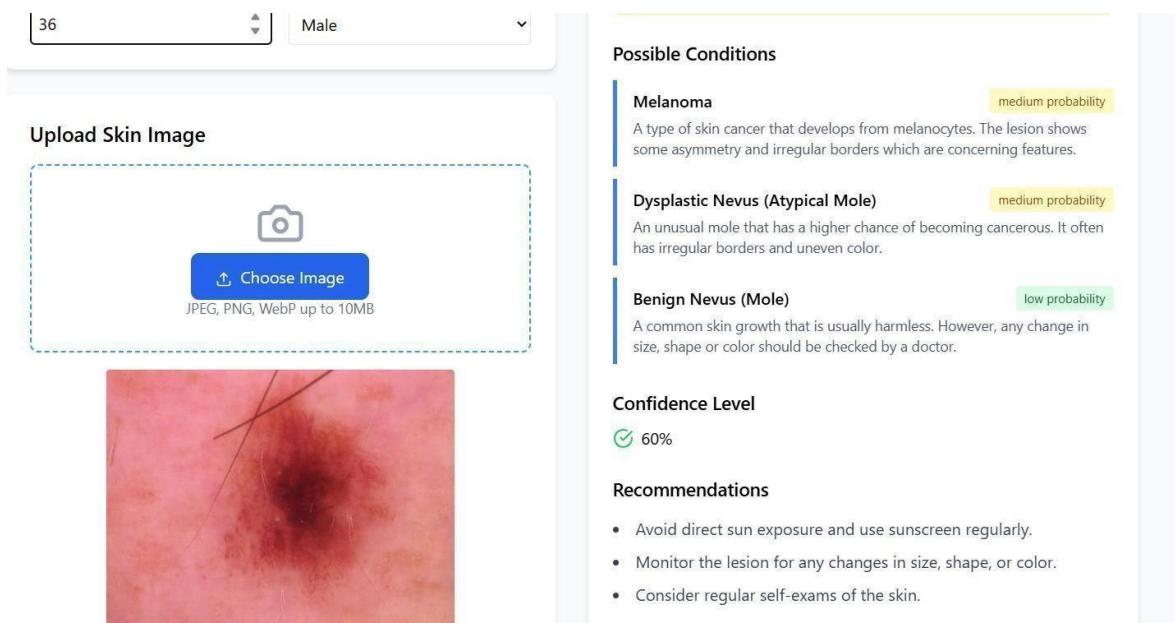


Fig B.2. Image upload section

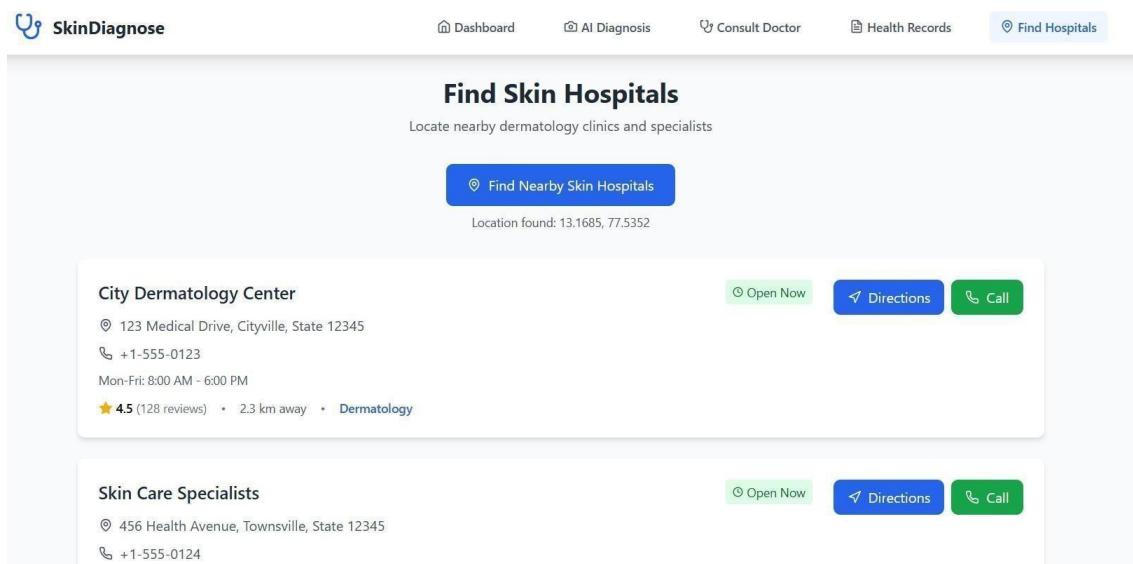


Fig B.3. Clinic/location recommendation screen

