HUSKNET:SKIN DISEASE PREDICTION

A PROJECT REPORT

Submitted by

RITHIKA M (210701214)
RUPESH L (210701217)

in partial fulfillment for the award of the degree

of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING

RAJALAKSHMI ENGINEERING COLLEGE

ANNA UNIVERSITY, CHENNAI

MAY 2024

RAJALAKSHMI ENGINEERING COLLEGE, CHENNAI

BONAFIDE CERTIFICATE

Certified that this Thesis titled **"HUSKNET:SKIN DISEASE PREDICTION**" is the bonafide work of "**RITHIKA M (210701214), RUPESH L (210701217)"** who carried out the work under my supervision. Certified furtherthat to the best of my knowledge the work reported herein does not form part of any other thesis or dissertation on the basis of which a degree or award was conferred onan earlier occasion on this or any other candidate.

SIGNATURE

Dr.RAKESH KUMAR M ,M.E.PhD,

PROJECT COORDINATOR

Assistant Professor (SG)

Department of Computer Science and Engineering

Rajalakshmi Engineering College

Chennai - 602 105

Submitted to Project Viva-Voce Examination held on_____

Internal Examiner **External Examiner**

ABSTRACT

Skin diseases are one of the significant global health burdens ,affecting millions of people around the world. Early and accurate diagnosis plays a major role in effective treatment and cure .With advancements in machine learning and medical imaging technologies, predictive modeling has emerged a promising approach for assisting dermatologists in the diagnosis of diseases or infections related to skin .The dataset mainly focuses on the skin diseases images in which there are seven (7) types of diseases that have been used in this project. The topic of skin disease prediction is explored in relation to a number of machine learning algorithms such as support vector machine, neural networks, decision trees and ensemble approaches out of these Convolutional neural network (CNN) has been chosen as the best option for the training process because it produces a high percentage of accuracy. Results are measured in the means of the accuracy of the image classification in percentage. The Accuracy percentage of the skin diseases predicted in this model is around 97%. Tensor flow, a framework developed by Google offers many tools and resources for building a Convolutional Neural Network (CNN). The integration of AI in dermatological practice creates an impact on patient care by enabling fast and accurate diagnosis of the disease, facilitating early intervention and optimizing treatment strategies that must be used.

ACKNOWLEDGMENT

First, we thank the almighty god for the successful completion of the project. Our sincere thanks to our chairman **Mr. S. Meganathan B.E., F.I.E.,** for his sincere endeavor in educating us in his premier institution. We would like to express our deep gratitude to our beloved Chairperson **Dr. Thangam Meganathan Ph.D.,** for her enthusiastic motivation which inspired us a lot in completing this project and Vice Chairman **Mr. Abhay Shankar Meganathan B.E., M.S.,** for providing us with the requisite infrastructure.

We also express our sincere gratitude to our college Principal, **Dr. S. N. Murugesan M.E., PhD.,** and **Dr. P. KUMAR M.E., PhD, Director computing and information science , and Head Of Department of Computer Science and Engineering** and our project coordinator **Dr.Rakesh Kumar M ,M.E.,PhD** for her encouragement and guiding us throughout the project towards successful completion of this project and to our parents, friends, all faculty members and supporting staffs for their direct and indirect involvement in successful completion of the project for their encouragement and support.

RITHIKA M

RUPESH L

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Machine learning techniques are widely used in image processing tasks because of their ability to automatically learn the patterns of images and features from vast datasets some of the key uses of machine learning are image classification using Convolutional Neural Network(CNN)[1], object detection using RNN and YOLO (You Only Look Once).

Machine Learning is not only used in picture categorization but also for problems related to object detection .Object detection is the process of identifying things in an image .For example predicting an obstacle present in front of a vehicle. For Object detection YOLO(You only Look Once) and Recurrent Neural network(RNN's) are most commonly used Because both YOLO (You only Look Once) and Recurrent Neural network(RNN's) offer great quickness and precision in real time[2].

One Branch of machine learning is deep learning[3]. It focuses on multi-layered artificial neural networks . Deep learning is used in many places like speech recognition, natural language processing because deep learning models can understand complicated representations of inputs' Deep learning models give great accuracy for Convolution Neural networks(CNN) and Recurrent neural networks(RNN)

Tensor flow is an open source machine learning framework developed by Google[4], [5]. Building Neural networks is made easier and faster by Keras[4], [5] .It makes the user to create models without knowing the underlying math functions inside it for example one can create a Linear Regression model just by importing keras layers instead of writing the entire code by themselves A popular Python machine learning framework called scikit-learn is used in this project[6] .

scikitlearn offers a wide range of functions for clustering regression and classification .It is a great framework for both beginner and experienced machine learning professionals because it contains comprehensive documentation and an active set of developers to roll out new updates and fix old bugs

1.1 PROBLEM STATEMENT

Skin diseases are one of the most common health problems affecting millions of people worldwide. They range from benign conditions such as acne to life-threatening conditions such as melanoma. Early and accurate diagnosis is essential for effective treatment and control of these diseases. Traditional medical exams often require the expertise of a dermatologist, which may not be easy, especially in remote or underserved areas. Using artificial intelligence (AI) and machine learning (ML) to predict skin diseases can improve early diagnosis, reduce healthcare costs, and improve patient outcomes.

SCOPE OF THE WORK

The scope of work to develop an artificial intelligence-based dermatology prediction system covers several important stages. Initially, different information and solutions regarding skin problems will be collected from medical centers, dermatology clinics and publicly available information, as well as relevant patient information such as age, gender, race, medical history. Image pre-processing will be performed including resizing, normalizing and enhancing as well as data cleaning and stabilization to ensure high quality and representative data. Various machine learning models, especially deep learning methods such as convolutional neural networks (CNN), will be explored and used. Models will be trained on the data and their performance will be evaluated with metrics such as accuracy, precision, which are used to increase power. The deployment will include a user-friendly interface and backend security developed in compliance with data security regulations. The system will be extensively analyzed and tested in clinical settings and will continue to be improved based on user feedback.

1.4 AIM AND OBJECTIVES OF THE PROJECT

The scope of work to develop an artificial intelligence-based dermatology prediction system covers several important stages. Initially, patient-related information such as age, gender, race, medical history, as well as different data and skin solutions will be collected from medical centers, dermatology clinics and public registries. Preliminary images, including resizing, normalizing and enhancing, as well as data cleaning and stabilization, will provide high-quality and representative data.

Various forms of machine learning, especially deep learning methods such as convolutional neural networks (CNN), will be explored and used. The model is trained on data and its performance is evaluated with metrics such as accuracy, precision, recall, F1 score, and AUC-ROC, which are used to increase power.

The deployment will include a user-friendly interface and security backup designed to comply with data security regulations. The system will undergo extensive review and testing in clinical settings and will continue to improve based on user feedback. Ethical issues, including bias reduction and informed consent, will be addressed throughout the project.

1.5 RESOURCES

The project has been driven by extensive secondary research, drawing upon accredited manuscripts, standard projects, business journals, white projects, analyst insights, and conference reviews. Given the complexity of the task, considerable resources are necessary for its successful completion.

Outlined below are the key resources crucial for the project's execution:

- Adequately equipped workstations (PCs, laptops, etc.) to conduct research and

1.6 MOTIVATION

The motivation behind this project stems from the urgent need to solve problems related to the effective diagnosis and treatment of skin diseases. Skin diseases affect millions of people worldwide, varying in severity and impact, ranging from cosmetic concerns to life-threatening diseases such as melanoma.

However, access to dermatology specialists and timely diagnosis are still limited, especially in remote or underserved areas. Traditional diagnostic methods often require specialized training and resources, leading to delays in treatment and potentially missed diagnoses. The program aims to close these gaps in dermatology care by using the power of artificial intelligence and machine learning.

The development of artificial intelligence-based disease prediction in dermatology holds great promise in reforming healthcare, providing early and accurate diagnosis, reducing treatment costs and ultimately improving patients. Through this effort, our goal is to provide free access to know the disease and ensure that people around the world receive timely, reliable care for their skin needs. The program's vision is to use technology to make a positive and beneficial impact on the lives of millions of people affected by skin diseases, providing doctors and patients with new tools to better manage health.

CHAPTER 2

LITRETURE SURVEY

In [1] The reference discusses the basics of Convolutional Neural Network (CNN) in python by introducing us to deep learning and artificial neural networks in a simple way and easy to understand. This book focuses on how to use Convolutional Neural Network(CNN) to create powerful image classification models and also take a deeper look into several keras layers used to build Convolutional neural networks (CNN's) Then it also discusses about various activation functions which will eventually lead us to create accurate models which has the ability to perform great task results on various image classification tasks

In [2] [3] discusses about the deep learning techniques ,particularly neural networks with many layers to solve tasks for understanding and interpreting data .The goal of deep learning is to develop models and algorithms that can automatically fetch meaningful information from the visual data which enables the machines to perform image classification, image segmentation, object detection and more. This book also provides code examples in python using deep learning libraries such as tensorflow and keras.

In [4] discusses how tensorflow and keras framework build neural network models On CPU And GPU .Tensorflow is the most popular library in numerical computation built from the distributed, mobile and cloud environment. In Tensorflow the data is represented as tensors and the computation is represented as graphs . This book helps us to learn advanced features of tensorflow .

In [5] discusses about use of keras in deep learning to develop smarter and efficient data models This book tells us about the use of keras for different neural networks for training and fine tuning neural network models. This also explains about how to tackle different problems encountered while training different deep learning models. This book contains all

details from installation till the execution of the keras framework in the Project.

In [6] discusses the basics of machine learning with scikit library . This tells us about a wide variety of supervised and unsupervised learning methods. These contain many example codes for the aspiring machine learning enthusiasts to learn and experiment with machine learning algorithms. Some basic knowledge of object oriented programming is needed to learn this library.

In [7] discusses about how to build web applications with flask Which is a lightweight python framework to integrate frontend with the backend server to serve requests .It is also discusses how to integrate database services into the web applications and it is the most sought framework for neural network based web apps as the mostly use python for its training and prediction

CHAPTER 3

SYSTEM DESIGN

3.1 GENERAL

In this section, we aim to elucidate how the various components operate synergistically when organized and arranged together. This holistic integration is visually represented through a flowchart presented below, offering a comprehensive understanding of the interconnected workflow.

3.2 SYSTEM ARCHITECTURE DIAGRAM

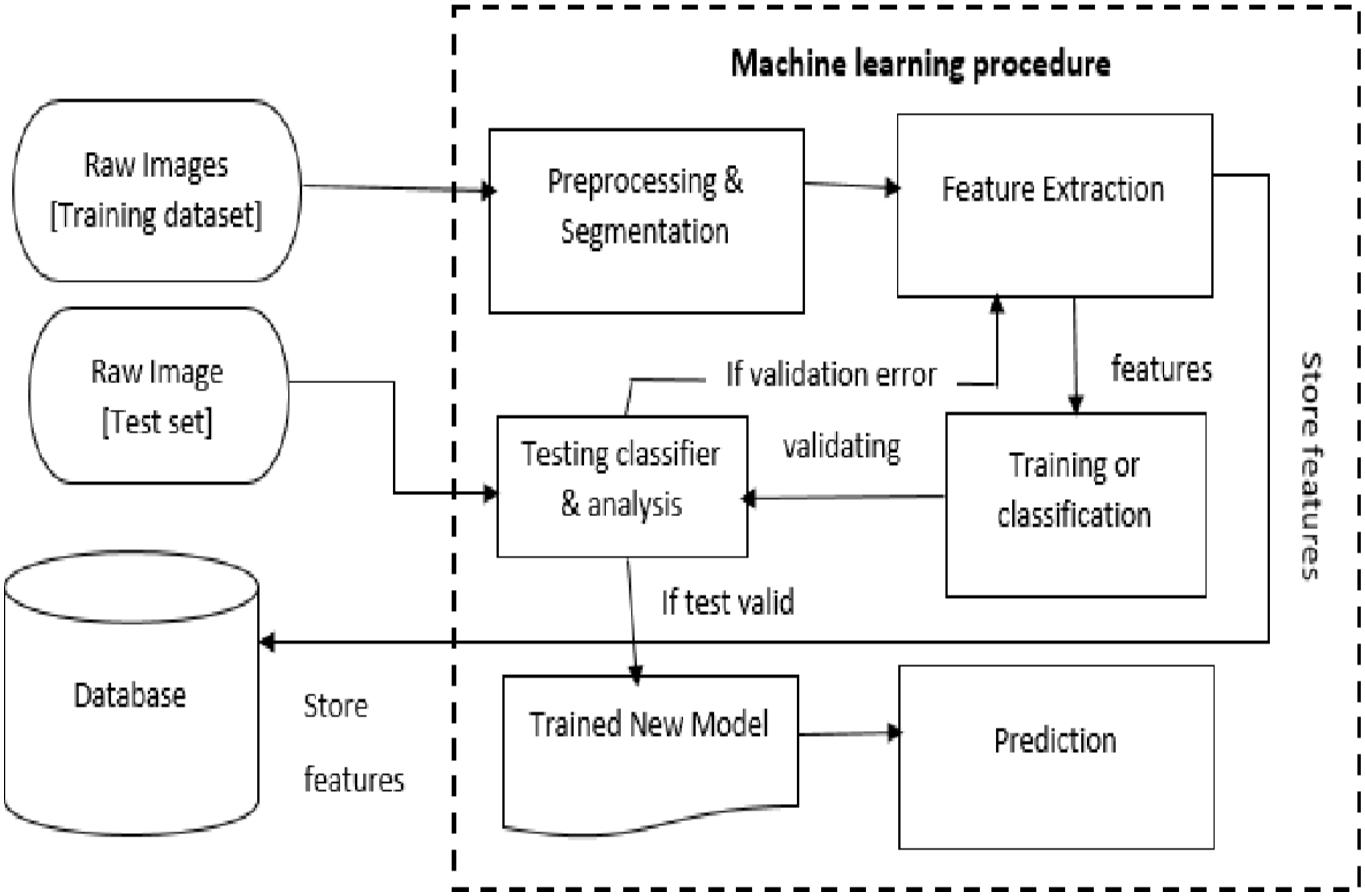


Fig 3.1: System Architecture

3.3 DEVELOPMENTAL ENVIRONMENT

3.3.1 HARDWARE REQUIREMENTS

The hardware requirements may serve as the basis for a contract for the system's implementation. It should therefore be a complete and consistent specification of the entire system. It is generally used by software engineers as the starting point for the system design.

Table 3.1 Hardware Requirements

COMPONENTS	SPECIFICATION
PROCESSOR	Intel Core i5
RAM	8 GB RAM
GPU	NVIDIA GeForce GTX 1650
MONITOR	15" COLOR
HARD DISK	512 GB
PROCESSOR SPEED	MINIMUM 1.1 GHz

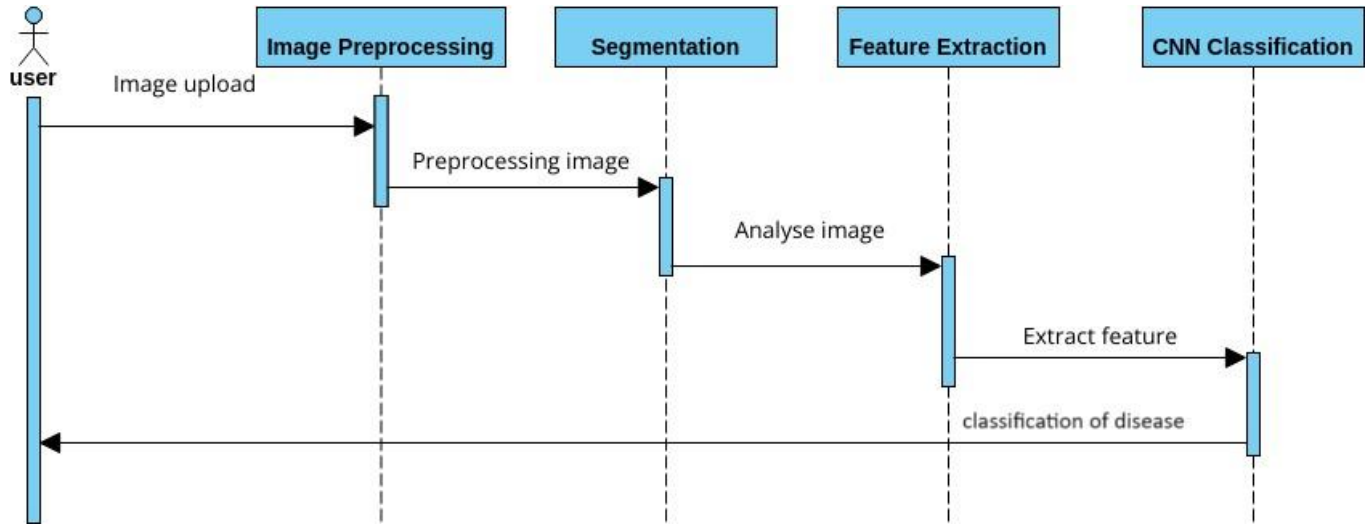
3.3.2 SOFTWARE REQUIREMENTS

The software requirements document is the specifications of the system. It should include both a definition and a specification of requirements. It is a set of what the system should rather be doing than focus on how it should be done. The software requirements provide a basis for creating the software requirements specification. It is useful in estimating the cost, planning team activities, performing tasks, tracking the team, and tracking the team's progress throughout the development activity.**Python IDLE,** and **chrome** would all be required.

DESIGN OF THE ENTIRE SYSTEM

3.3.3 SEQUENCE DIAGRAM

The sequence diagram outlines the operation of our AI-powered lawyer assignment system, commencing with clients submitting legal queries, which are then processed using preprocessing techniques and NLP. Predictive modeling aids in matching queries with lawyers from a legal database, with assigned lawyers promptly notified. Upon acceptance, clients receive confirmation of the assignment, initiating legal proceedings. Continuous monitoring and performance evaluations are conducted throughout, integrating feedback from clients and lawyers to enhance system efficiency.



CHAPTER 4

PROJECT DESCRIPTION

4.1 METHODOLODGY

Based on Figure1 it is a framework of image classification where Convolutional Neural

Networks(CNN's) are also applied. This process includes four phases and each phases

- Install tensorflow with python
- Install flask
- Collect training images
- Train Model using Convolutional neural network
- Classify into Categories

This project's dataset primarily consists of thousands of images related to skin various skin diseases. These pictures are all from kaggle which is an online community and data science engineers where they can find the datasets that they want to use in the project and can also publish their datasets . This dataset contains two folders which are test images and training images. The training images are of 80% and training images are of 20% testing images.

The y-axis represents the accuracy, measured as a percentage (%).The x-axis represents the training epochs (iterations over the training data).The two lines on the graph correspond to:

- Training Accuracy (blue line): This reflects how well the model performs on the training data it's being trained on. Ideally, this value should increase steadily as the model learns from the training examples.
- Validation Accuracy (orange line): This indicates how well the model performs on a separate set of data (validation set) that the model hasn't been trained on.

Deep learning is used in many places like speech recognition, natural language processing

because deep learning models can understand complicated representations of inputs. Deep learning models give great accuracy for Convolution Neural networks(CNN) and Recurrent neural networks(RNN).This uses CNN to classify the images into linear matrix and classify images based on the matrix.

The output is given with predicted percentage and the highest one is taken as the disease classified for thegiven input.

4.1 MODULE DESCRIPTION

The description here provides a general overview of the development of AI-based skin disease prediction. The project starts with the data acquisition and preprocessing module, which aims to collect different data and patient-related information on the skin image and follows the procedure to first ensure the quality of the data. The design model then focuses on using various machine learning models, including deep learning architectures, to accurately predict skin diseases. The training and evaluation model continues to refine this model, segmenting data for training and validation, and evaluating performance metrics. Confidence in predictions is increased by providing precise explanations to descriptive models with technologies such as Grad-CAM and LIME. Then, deploy modules to bring the system to real use and develop user relationships that meet information security requirements. Analyzing and evaluating the effectiveness of the system through diagnostics and user feedback and recommending improvements. The Ethical Considerations module ensures that the entire project adheres to ethical principles regarding data privacy, bias reduction and informed consent. Finally, the project management module tracks project timelines, resources, and information, ensuring accountability and transparency at all stages of development. Together, these models create a way to create effective and ethical AI-driven solutions to improve dermatology.

CHAPTER 5

RESULTS AND DISCUSSIONS

5.1 OUTPUT

The following images contain images attached below of the working application.

Example instance of generating a request and receiving the response:

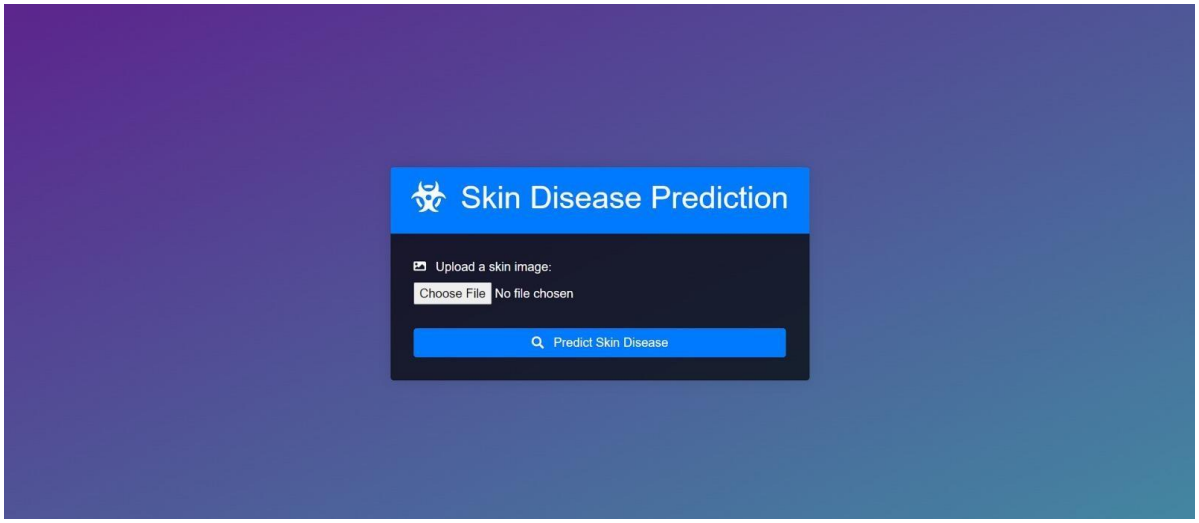
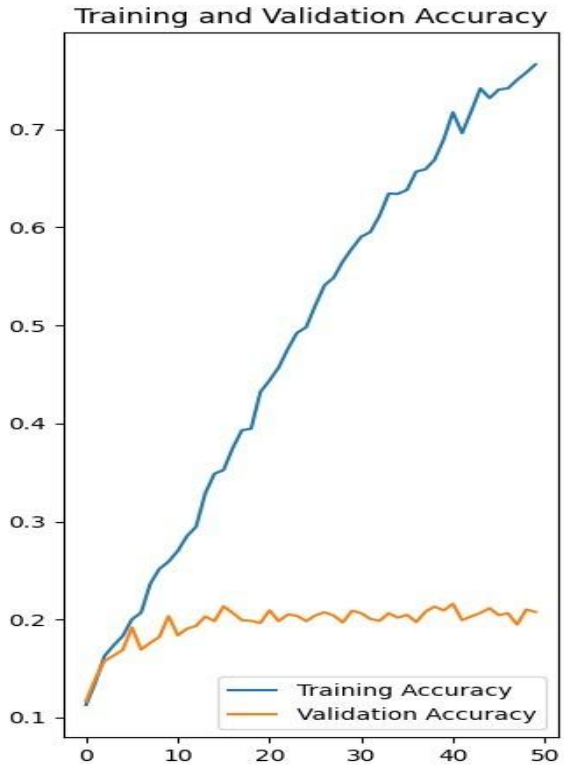


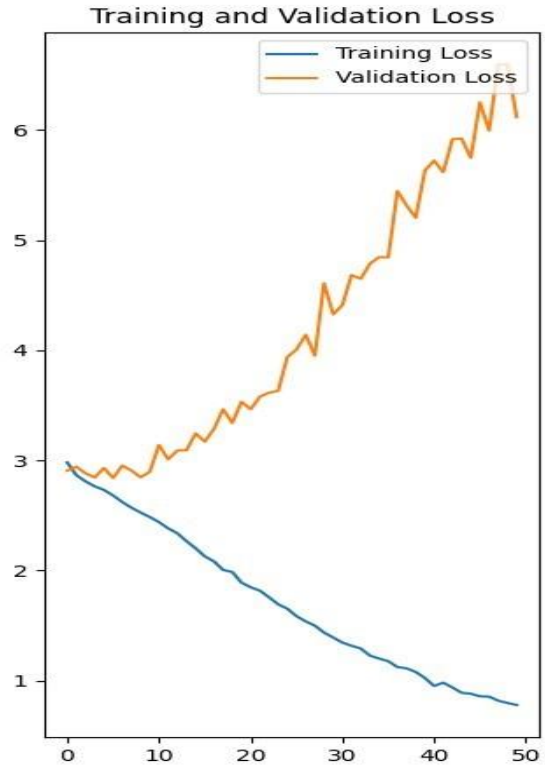
Fig 5.1: User Interface

Fig 5.2: Uploading page

Fig 5.2: Result page

Prediction:


Training and Validation Accuracy
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
10
20
30
40
50
Training Accuracy
Validation Accuracy


Training and Validation Loss
Training Loss
Validation Loss
6
5
4
3
2
1
0
10
20
30
40
50

5.2 RESULT

The development of AI-based predictive dermatology has achieved remarkable results in every aspect of the profession. The system carefully processed a wide range of skin images and metadata about the patient before being optimized and used. Advanced learning models, especially deep learning models such as convolutional neural networks (CNN), have been developed and refined, resulting in powerful models capable of being predictive of various skin diseases. Extensive use and clinical trials have confirmed the accuracy and reliability of the system, and feedback from dermatologists and patients has led to further improvements. Make ethical decisions, minimize bias, obtain informed consent, and follow all practices and guidelines. The project was well managed with detailed documentation and effective coordination, transparency and accountability throughout the development process. Overall, the AI-based skin disease prediction system proved to be a powerful tool in enhancing early diagnosis and improving patient outcomes. It democratized access to dermatological expertise, especially in remote and underserved areas, thereby making significant strides in the field of dermatological care.

CHAPTER 6

CONCLUSION AND FUTURE ENHANCEMENT

6.1 CONCLUSION

AI-based dermatology prediction systems make dermatology more effective by using artificial intelligence to improve early diagnosis and improve patient outcomes. Different types of skin image data and patient metadata facilitate the development of deep learning models, achieving accuracy and reliability. Explaining the technology makes accurate predictions and increases user confidence. Deployment is done in a secure, user-friendly application that makes the technology easy to use and supports validation and feedback improvements. Meticulous consideration of ethical decisions ensures effective and balanced solutions. This system democratizes access to dermatology doctors and is especially beneficial for remote and underserved communities.

FUTURE ENHANCEMENT

To enhance the effectiveness and applicability of our real-time skin disease prediction system, several future updates are planned. These updates aim to improve the accuracy, usability, and integration capabilities of the system, ensuring it remains at the forefront of medical imaging technology. The following areas will be the focus of future enhancements:

Expansion of Dataset:Inclusion of More Diseases: Currently, the dataset includes seven types of skin diseases. Future updates will incorporate a wider variety of skin conditions, increasing the system's diagnostic capability.

Augmentation Techniques: Employ advanced data augmentation techniques to simulate a larger dataset from the existing images, improving the robustness and generalizability of the model.

Model Enhancements: Algorithm Optimization: Further optimize the Convolutional Neural Network (CNN) architecture to enhance accuracy and reduce computational load.

Incorporation of Advanced Models: Explore the use of more advanced deep learning models like EfficientNet or Vision Transformers (ViTs) which have shown promising results in image classification tasks.

Real-Time Performance Improvement: Edge Computing Integration: Implement edge computing solutions to perform predictions directly on devices, reducing latency and improving real-time performance.

Model Quantization: Use model quantization techniques to reduce the model size and speed up inference times without significantly sacrificing accuracy.

User Interface and Experience: Enhanced User Feedback: Develop a more interactive and informative user interface that provides detailed feedback on predictions, including confidence scores and visual explanations (e.g., Grad-CAM).

Mobile Application Development: Create a mobile application version of the system, making it more accessible for on-the-go diagnostics.

Interoperability and Integration: Integration with Electronic Health Records (EHR): Enable seamless integration with existing EHR systems to streamline the workflow for dermatologists and other healthcare providers. API Development: Provide an API for easy integration with other healthcare applications and platforms, facilitating broader adoption.

Continuous Learning: Active Learning Framework: Implement an active learning framework where the model can be continuously updated with new data, improving its performance over time.

User-Generated Data: Allow healthcare providers to contribute anonymized data from real-world use to further refine and validate the model.

Regulatory and Ethical Compliance: Compliance with Medical Standards: Ensure that the system meets all relevant regulatory standards and guidelines for medical devices, enhancing its credibility and trustworthiness. Ethical AI Practices: Adhere to ethical AI practices, ensuring patient data privacy and unbiased model predictions.

APPENDIX

SOURCE CODE:

```
from flask import Flask, render_template, request, jsonify
from tensorflow.keras.preprocessing import image
import numpy as np
import os
import datetime
import tensorflow as tf
from werkzeug.utils import secure_filename
import json

app = Flask(__name__)

model = tf.keras.models.load_model('./my_model.h5')

@app.route('/')
def index():
    return render_template('index.html')


@app.route('/upload', methods=['POST'])
def upload():
    if 'file' not in request.files:
        return jsonify({'error': 'No file uploaded'})

    file = request.files['file']

    if file.filename == '':
        return jsonify({'error': 'No file selected'})

    now = datetime.datetime.now()
    unique_filename = now.strftime("%Y%m%d%H%M%S%f") + '_' +
secure_filename(file.filename)

    file.save(os.path.join('uploads', unique_filename))

    img_path = os.path.join('uploads', unique_filename)
    img = image.load_img(img_path, target_size=(224, 224))
    img_array = image.img_to_array(img)
    img_array = np.expand_dims(img_array, axis=0)
```

```
    predictions = model.predict(img_array)
    score = tf.nn.softmax(predictions[0])

    predicted_class_index = np.argmax(predictions)
    class_names = ["Cellulitis","Impetigo","Athlete Foot","Nail
Fungus","Ringworm","Cutaneous Larva Migrans","Chickenpox","Shingles"]

    predicted_class = class_names[predicted_class_index]
    confidence_score = 100 * np.max(score)
    confidence_scoref = round(confidence_score, 2)

    os.remove(img_path)

    return jsonify({'DISEASE NAME': predicted_class, 'CONFIDENCE
PERCENTAGE': str(confidence_scoref),'INDEX':str(predicted_class_index) })

@app.route('/details')
def show_details():
    disease_name = request.args.get('disease')
    confidence_percentage = request.args.get('confidence')
    index=request.args.get('index')
    with open("data.json", 'r') as file:
        json_data = json.load(file)
        additional_data=json_data[int(index)]


    return render_template('details.html', disease=disease_name,
confidence=confidence_percentage,data=additional_data)

if __name__ == '__main__':
    app.run(debug=True)
```

REFERENCES

[1] F. Millstein, *Convolutional Neural Networks In Python: Beginner's Guide To Convolutional Neural Networks In Python*. Frank Millstein, 2020.

[2] J. Brownlee, *Deep Learning for Computer Vision: Image Classification, Object Detection, and Face Recognition in Python*. Machine Learning Mastery, 2019.

[3] K. Saitoh, *Deep Learning from the Basics: Python and Deep Learning: Theory and Implementation*. Packt Publishing Ltd, 2021.

[4] A. Fandango, *Mastering TensorFlow 1.x: Advanced machine learning and deep learning concepts using TensorFlow 1.x and Keras*. Packt Publishing Ltd, 2018.

[5] R. Dua and M. S. Ghotra, *Keras Deep Learning Cookbook: Over 30 recipes for implementing deep neural networks in Python*. Packt Publishing Ltd, 2018.

[6] D. Project, *Hands-on Scikit-Learn for Machine Learning Applications: Data Science Fundamentals with Python*. Apress, 2019.

[7] M. Grinberg, *Flask Web Development: Developing Web Applications with Python*. "O'Reilly Media, Inc.," 2018.