NailCare AI

Prepared For
Smart-Internz
Artificial Intelligence Guided project

By
Sainath Nimbalkar
D Y Patil Agriculture and Technical University Talsande

On
23 April 2025

Abstract

This project utilizes deep learning and transfer learning with Inception models to analyze human nail images for early-stage disease diagnosis. By focusing on key visual features such as nail texture, color, shape, and surface abnormalities, the system aims to detect potential health conditions non-invasively. The objective is to enhance diagnostic accuracy and support timely medical intervention. This solution has practical applications across healthcare settings, personal wellness monitoring, and public health screening, contributing to preventive healthcare through advanced image-based analysis.

Final Project Report

Contents

1 Introduction

1.1 Project Overview

This project aims to diagnose early-stage diseases through the analysis of human nail images using deep learning techniques. By leveraging transfer learning with advanced convolutional neural network models such as Inception, the system focuses on identifying subtle abnormalities and visual patterns in nail structures, which are indicative of underlying health conditions. The primary regions of interest include the nail surface, texture, color, and shape, which serve as key visual markers for disease detection.

The classification task involves detecting nail abnormalities associated with various systemic or dermatological conditions. Nails can reflect changes caused by diseases such as fungal infections, nutritional deficiencies, liver disease, anemia, and more. Through high-quality image data and deep learning models, this project enables accurate and efficient analysis of nail images, paving the way for non-invasive early diagnosis and health monitoring. The system can be applied across different scenarios including clinical diagnostics, personal health monitoring, and public health screening, making it a significant contribution to the domain of medical image analysis and preventive healthcare.

1.2 Objectives

The main objective of this project is to develop a robust and accurate system for early disease detection using human nail image processing. By applying deep learning and transfer learning techniques, the system is designed to identify and classify nail abnormalities that may indicate underlying health issues. This project serves multiple purposes:

- **Healthcare Professionals**: Assist clinicians in diagnosing nail-related and systemic diseases at an early stage, leading to timely treatment and improved patient outcomes.
- **Individuals**: Enable personal health monitoring by detecting visible changes in nails, encouraging users to seek professional advice before diseases progress.
- **Public Health**: Support health organizations in conducting large-scale screenings and epidemiological studies to track disease trends and implement early intervention strategies.

SMARTBRIDGE
a Veranda Enterprise
Let's Bridge the Gap
Smart Internz

2 Project Initialization and Planning Phase

2.1 Define Problem Statement

PS No.	I am (Customer)	I'm trying to	But	Because	Which makes me feel
PS-1	A general user	Detect nail diseases early	I don't know what's normal or not	Nail changes can be subtle and confusing	Anxious and unsure
PS- 2	A dermatologist	Save time in diagnosis	I receive unclear or poor images	Patients often send low-quality photos	Frustrated and limited
PS- 3	A healthcare startup	Provide affordable nail disease care	It's hard to scale expert review	Specialists are expensive and not widely available	Concerned about reach and cost
PS- 4	A medical student	Learn to recognize nail diseases	The examples in textbooks are limited	Real-world cases vary significantly	Confused and lacking confidence
PS- 5	A tech researcher	Build a smart nail disease detector	I lack labeled image datasets	Medical image datasets are rare and sensitive	Technically challenged and limited
PS- 6	A wellness influencer or blogger	Educate followers about nail health	I'm not medically trained	Nail issues can be early signs of deeper problems	Cautious and worried about misinformation

PS- 7	A parent	Teachmy kids about nail hygiene	I'm unsure how to explain symptoms	Early signs of nail disease can be hard to notice	Anxious and unconfident as a guide
PS- 8	A pharmaceutical researcher	Discover nail-related disease indicators	It's hard to collect accurate samples	Misidentification risks research accuracy	Limited in research accuracy and drug innovation
PS- 9	An AI/ML student	Train a deep learning model to detect nail diseases	There are too many similar-looking samples	Labeling is expensive and hard	Overwhelmed and unsure of training data quality
PS- 10	A rural health worker	Educate rural communities about Nail Diseases	There's no easy tool for live identification	Language and tech access barriers exist	Helpless in reaching and empowering locals
PS- 11	A beauty product company	Source large quantities of safe Nail Diseases	Create nail-friendly health products	I can't differentiate disease from damage	Nail conditions may look like cosmetic issues
PS- 12	A health insurance provider	Support Nail Disease - based dish partners	Reduce claims through early detection	Nail symptoms are often ignored	People overlook signs until the condition worsens
PS- 13	A mobile app developer	Build a real-time nail scan feature	I need labeled images and feedback	Real-time models need quality training data	Blocked by tech limitations and data scarcity
PS- 14	A health-conscious individual	Monitor my nail health over time	I don't know what changes to track	No easy tool shows what's improving or worsening	Disconnected from my wellness journey

2.2 Project Proposal (Proposed Solution)

Project Proposal (Proposed Solution):

This project proposal outlines a solution to address a specific problem. With a clear objective, defined scope, and a concise problem statement, the proposed solution details the approach, key features, and resource requirements, including hardware, software, and personnel.

Project Overview	
Objective	To develop a deep learning-based image classification system capable of accurately identifying nail disease based on visual attributes.
Scope	This project focuses on image-based classification of Nail Disease using deep learning models. It covers the acquisition of image datasets, preprocessing, model training using transfer learning, and evaluation of classification accuracy. The final system will be able to classify images into one of the target disease. The project is multiple categories and assumes images are of reasonable quality.
Problem Statement	
Description	Nail Disease identification is challenging and typically requires expert knowledge. Mistakes can be dangerous, particularly when foraging. A reliable classification tool would benefit researchers, foragers, and hobbyists.
Impact	Precise Nail Disease classification aids ecological research, education, and safe foraging. An image-based system makes species recognition more accessible to all.

Proposed Solution	
Approach	The project will employ CNN-based deep learning, using transfer learning from models like VGG16,ResNet or EfficientNet. The Nail Disease image dataset will be cleaned, augmented, then used for training and fine-tuning.
Key Features	The system uses transfer learning to train efficiently with limited data, classifying Nail Image into key disease. Data augmentation enhances model performance, with potential for a web- based interface.

Resource Requirements

Resource Type	**Description**	**Specification/Allocation**
Hardware		
Computing Resources	CPU/GPU specifications, number of cores	1 x NVIDIA RTX 3060 GPUs
Memory	RAM specifications	16 GB RAM
Storage	Disk space for data, models, and logs	500 GB SSD
Software		
Frameworks	Python frameworks	Python
Libraries	Additional libraries	tensorflow
Development Environment	IDE, version control	Jupyter Notebook, Git
Data		
Data	Source, size, format	Kaggle, JPEG/PNG format, 10,000 images

2.3 Initial Project Planning

Product Backlog, Sprint Schedule, and Estimation

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority
Sprint-1	Model Application	USN-1	As a system, I need to apply a pre-trained deep learning model to the uploaded image.	3	High
Sprint-1	Application Integration	USN-2	As a developer, I need to integrate the model with a Flask application.	3	High
Sprint-2	Image Input & Processing	USN-3	As a user, I want to select an image for Nail Disease identification.	2	High
Sprint-2	Output Display	USN-4	As a user, I want to see the identified Nail Disease species, with a confidence level, displayed clearly and quickly.	2	High
Sprint-3	Performance Optimization	USN-5	As a developer, I need to optimize the application for speed and efficiency.	2	Medium

3 Data Collection and Preprocessing Phase

3.1 Data Collection Plan and Raw Data Sources Identified

Data Collection Plan

Section	Description
Project Overview	This deep learning project focuses on classifying images of 16 different Nail disease—using Convolutional Neural Networks (CNNs). The objective is to uncover hidden patterns and visual cues that distinguish each type, contributing to better Nail Disease identification in the wild.
Data Collection Plan	The dataset has been sourced from a ZIP file provided by the SmartInternz, which includes categorized images in subdirectories named after each NAIL DISEASE. Additional reference images were accessed from publicly available sources such as Wikimedia.
Raw Data Sources Identified	The raw data includes SmartInternz provided images saved in structured subdirectories, supplemented by publicly available datasets for training and validation purposes.

Raw Data Sources

Source Name	Description	Location/URL	Format	Size	Access Permissions
SmartInterz Provided Dataset	Curated image dataset provided by SmartInternz, containing Nail disease images in separate subdirectories.	https://drive.google.com/drive/folders/1AXTYsbiarS1TCAgfj0mancTSrJYYMWMs?usp=sharing	ZIP File	~ 175 MB	Public

Field Captured Images	Manually photographed images taken in natural environments, used for supplementing the dataset.	Local Storage	JPG/PNG	~100 MB	Private
Wikimedia	Open-source Nail Disease images used for visual verification and dataset augmentation.	https://en.wikipedia.org/wiki/File:Fingernails2.jpg	JPG/PNG	~10 MB	Public

3.2 Data Quality Report

Data Source	Data Quality Issue	Severity	Resolution Plan
Dataset	Image Variation	High	Collect images from diverse sources (different cameras, lighting conditions, angles). Implement data augmentation techniques (rotation, scaling, cropping) during preprocessing.
Dataset	Occlusion	Moderate	Include images with partial occlusion, and/or train the model to be robust to it.
Dataset	Insufficient Resolution	Moderate	Establish a minimum resolution threshold for images. Use super-resolution techniques, if feasible, to enhance the resolution of some images.

SMARTBRIDGE
Let's Bridge the Gap
a Veranda Enterprise

Smart Internz

Dataset	Unbalanced Classes	High	Employ stratified sampling to ensure proportional representation of each nail- Disease. Use data augmentation for minority classes. Explore the use of weighted loss functions during training.

3.3 Data Preprocessing

Data Preprocessing

The images will be preprocessed by resizing, normalizing, augmenting, denoising, adjusting contrast, detecting edges, converting color space, cropping, batch normalizing, and whitening data. These steps will enhance data quality, promote model generalization, and improve convergence during neural network training, ensuring robust and efficient performance across various computer vision tasks.

Section	Description
Data Overview	This project uses image datasets of 16 different Nail Disease. The images are collected from various sources including **SmartInternz**, **custom field- captured images**, and platforms like **Kaggle** and **Wikimedia**. This ensures rich visual diversity and robust generalization during training.
Resizing	All images are resized to **224×224 pixels** using OpenCV's cv2.resize() function to ensure uniform input dimensions for CNN-based models.
Normalization	Pixel values are normalized to the range **[0, 1]** by dividing by 255.0, improving convergence during model training.
Data Augmentation	Using ImageDataGenerator, images are augmented with **random rotation**, **shifts**, **zoom**, **horizontal/vertical flips**, and **fill modes** to avoid overfitting.
Denoising	OpenCV's fastNlMeansDenoisingColored() is applied to reduce environmental noise and improve image clarity, especially for field-captured data.

Color Space Conversion	Images are converted from **BGR to HSV** color space using cv2.cvtColor() to better capture color-based patterns across lighting variations.
Batch Normalization	BatchNormalization() is applied in the neural network model to stabilize and accelerate the learning process by reducing internal covariate shift.
Data Preprocessing Code Screenshots	

<table>
<tr><td>Loading Data</td><td>
<pre><code>from tensorflow.keras.layers import Dense, Flatten, Input
from tensorflow.keras.models import Model
from tensorflow.keras.preprocessing import image
from tensorflow.keras.preprocessing.image import ImageDataGenerator, load_img
from tensorflow.keras.applications.vgg16 import VGG16, preprocess_input
from glob import glob
import numpy as np
import matplotlib.pyplot as plt

train_path = r'D:\Swayam AI Project\Dataset\train-20250520T165148Z-1-001\train'
test_path = r'D:\Swayam AI Project\Dataset\test-20250520T165150Z-1-001\test'</code></pre>
</td></tr>
<tr><td>Resizing</td><td>
<pre><code>imageSize = [224, 224]</code></pre>
</td></tr>
<tr><td>Normalization</td><td>
<pre><code>x=image.img_to_array(img)

x = x / 255.0   # Normalize
x = np.expand_dims(x, axis=0)</code></pre>
</td></tr>
</table>

Data Augmentation	```train_datagen = ImageDataGenerator(rescale=1./255, shear_range=0.2, horizontal_flip=True, zoom_range=0.2)``` ```test_datagen = ImageDataGenerator(rescale=1./255)```

4 Model Development Phase

4.1 Model Selection Report

Model Selection Report:

Model	Description
Artificial Neural Network (ANN)	ANNs are foundational deep learning models composed of multiple fully connected layers. They are well-suited for tabular data or feature-engineered inputs, and while they can be adapted for image data, they do not inherently capture spatial relationships.
Convolutional Neural Network (CNN)	CNNs are powerful deep learning models specifically designed for image data. They automatically extract spatial features from images using convolutional layers, allowing effective classification of complex visual patterns
Recurrent Neural Network (RNN)	RNNs are designed to model sequential data by maintaining a hidden state across time steps. While they are powerful for time series and language modeling, their utility in static image classification is limited.
VGG16	VGG16 is a deep convolutional neural network architecture designed for image classification, known for its simplicity and uniform structure. It uses a series of small 3x3 convolutional filters stacked in deep layers to capture features, enabling the model to learn complex patterns effectively. Pre-trained on ImageNet, VGG16 is well-suited for transfer learning, providing high accuracy, though it is computationally more intensive than some newer models. In this project it is used to detect the Nail Disease.

Conclusion:

Model Selected	
VGG16	VGG16 is a deep convolutional neural network architecture designed for image classification, known for its simplicity and uniform structure. It uses a series of small 3x3 convolutional filters stacked in deep layers to capture features, enabling the model to learn complex patterns effectively. Pre-trained on ImageNet, VGG16 is well-suited for transfer learning, providing high accuracy, though it is computationally more intensive than some newer models.In this project it is used to detect the Nail Disease.

4.2 Initial Model Training Code, Model Validation and Evaluation Report

Initial Model Training Code, Model Validation and Evaluation Report

Initial Model Training Code

```
vgg = VGG16(input_shape=imageSize + [3], weights='imagenet', include_top=False)

for layer in vgg.layers:
    layer.trainable = False

x = Flatten()(vgg.output)

prediction = Dense(17, activation='softmax')(x)

model = Model(inputs=vgg.input, outputs=prediction)

model.compile(
    loss='categorical_crossentropy',
    optimizer='adam',
    metrics=['accuracy']
)

train_datagen = ImageDataGenerator(rescale=1./255,
                                   shear_range=0.2,
                                   horizontal_flip=True,
                                   zoom_range=0.2)
test_datagen = ImageDataGenerator(rescale=1./255)


train_path = r'D:\Swayam AI Project\Dataset\train-20250520T165148Z-1-001\train'
test_path = r'D:\Swayam AI Project\Dataset\test-20250520T165150Z-1-001\test'

train_set = train_datagen.flow_from_directory(
    train_path,
    target_size=(224, 224),
    batch_size=32,
    class_mode='categorical'
)

test_set = test_datagen.flow_from_directory(
    test_path,
    target_size=(224, 224),
    batch_size=32,
    class_mode='categorical'
)
```

```
r = model.fit(
    train_set,
    validation_data=test_set,
    epochs=100,
    steps_per_epoch=len(train_set)//3,
    validation_steps=len(test_set)//3
)
```

Model Validation and Evaluation Report

Model	Summary	Training and Validation Performance Metrics
Model 1 (VGG16)	**Layer Summary:** • VGG16 base model **Total Parameters:** 15,141,201 **Trainable Parameters:** 426,513 **Non-trainable Parameters:** 14,714,688	**Training Accuracy**: 94.21% **Validation Accuracy**: 98.44%

5 Model Optimization and Tuning Phase

5.1 Tunning Documentation

Model Optimization and Tuning Phase

The Model Optimization and Tuning Phase involves improving our neural network models to get the best results. This means adjusting the model's settings, comparing how well different settings work, and explaining why we chose our final model.

The neural network models were trained to classify Nail Disease Images into the following Seventeen classes: Darier_s disease, Muehrck-e_s lines, aloperia areata, beau_s lines, bluish nail, clubbing, eczema, half and half nailes (Lindsay_s nails), koilonychia, leukonychia, onycholycis, pale nail, red lunula, splinter hemmorrage, terry_s nail, white nail, yellow nails
. The training dataset consisted of around 650 labeled Nail disease images across the Seventeen target classes. A separate dataset of 200 images was used for validation and final evaluation of the models.

Hyperparameter Tuning Documentation (8 Marks):

Model	Tuned Hyperparameters

Model 1: InceptionV3 (Baseline)	Learning Rate: We adjusted the learning rate, which controls how much the model learns from its mistakes. We tried different learning rates to find one that helps the model learn effectively without becoming unstable.
	`model.compile(`
	`    loss='categorical_crossentropy',`
	`    optimizer='adam',`
	`    metrics=['accuracy']`
	`)`

Model 2: VGG16 (Optimized)	Epoch: We made finer adjustments to the epoch, building on what we learned from Model 1, to see if we could improve performance further. `r = model.fit(train_set, validation_data=test_set, epochs=10, steps_per_epoch=len(train_set)//3, validation_steps=len(test_set)//3)`

```
r = model.fit(
    train_set,
    validation_data=test_set,
    epochs=10,
    steps_per_epoch=len(train_set)//3,
    validation_steps=len(test_set)//3
)
```

Batch Size: We used the best batch size from Model 1.

```
train_set = train_datagen.flow_from_directory(
    train_path,
    target_size=(224, 224),
    batch_size=32,
    class_mode='categorical'
)

test_set = test_datagen.flow_from_directory(
    test_path,
    target_size=(224, 224),
    batch_size=32,
    class_mode='categorical'
)
```

```
r = model.fit(
    train_set,
    validation_data=test_set,
    epochs=100,
    steps_per_epoch=len(train_set)//3,
    validation_steps=len(test_set)//3
)
```

Accuracy:

```
7/7 ━━━━━━━━━━━━━━━━━━━━ 66s 10s/step - accuracy: 0.9084 - loss: 0.4217 - val_accuracy: 0.9375 - val_loss: 0.2952
Epoch 90/100
7/7 ━━━━━━━━━━━━━━━━━━━━ 58s 8s/step - accuracy: 0.9139 - loss: 0.3492 - val_accuracy: 0.9375 - val_loss: 0.3425
Epoch 91/100
7/7 ━━━━━━━━━━━━━━━━━━━━ 62s 9s/step - accuracy: 0.9077 - loss: 0.4100 - val_accuracy: 0.9688 - val_loss: 0.2367
Epoch 92/100
7/7 ━━━━━━━━━━━━━━━━━━━━ 56s 9s/step - accuracy: 0.8895 - loss: 0.4613 - val_accuracy: 0.9688 - val_loss: 0.2196
Epoch 93/100
7/7 ━━━━━━━━━━━━━━━━━━━━ 59s 9s/step - accuracy: 0.9042 - loss: 0.4053 - val_accuracy: 0.8906 - val_loss: 0.4190
Epoch 94/100
7/7 ━━━━━━━━━━━━━━━━━━━━ 60s 9s/step - accuracy: 0.9336 - loss: 0.3512 - val_accuracy: 0.9062 - val_loss: 0.3295
Epoch 95/100
7/7 ━━━━━━━━━━━━━━━━━━━━ 56s 8s/step - accuracy: 0.9253 - loss: 0.3706 - val_accuracy: 0.9531 - val_loss: 0.2696
Epoch 96/100
7/7 ━━━━━━━━━━━━━━━━━━━━ 59s 9s/step - accuracy: 0.8947 - loss: 0.4200 - val_accuracy: 0.9531 - val_loss: 0.2574
Epoch 97/100
7/7 ━━━━━━━━━━━━━━━━━━━━ 63s 9s/step - accuracy: 0.9417 - loss: 0.3394 - val_accuracy: 0.9375 - val_loss: 0.2911
Epoch 98/100
7/7 ━━━━━━━━━━━━━━━━━━━━ 60s 9s/step - accuracy: 0.9246 - loss: 0.3941 - val_accuracy: 1.0000 - val_loss: 0.1817
Epoch 99/100
7/7 ━━━━━━━━━━━━━━━━━━━━ 56s 8s/step - accuracy: 0.9192 - loss: 0.4135 - val_accuracy: 0.9531 - val_loss: 0.2623
Epoch 100/100
7/7 ━━━━━━━━━━━━━━━━━━━━ 58s 9s/step - accuracy: 0.9421 - loss: 0.3444 - val_accuracy: 0.9844 - val_loss: 0.2170
```

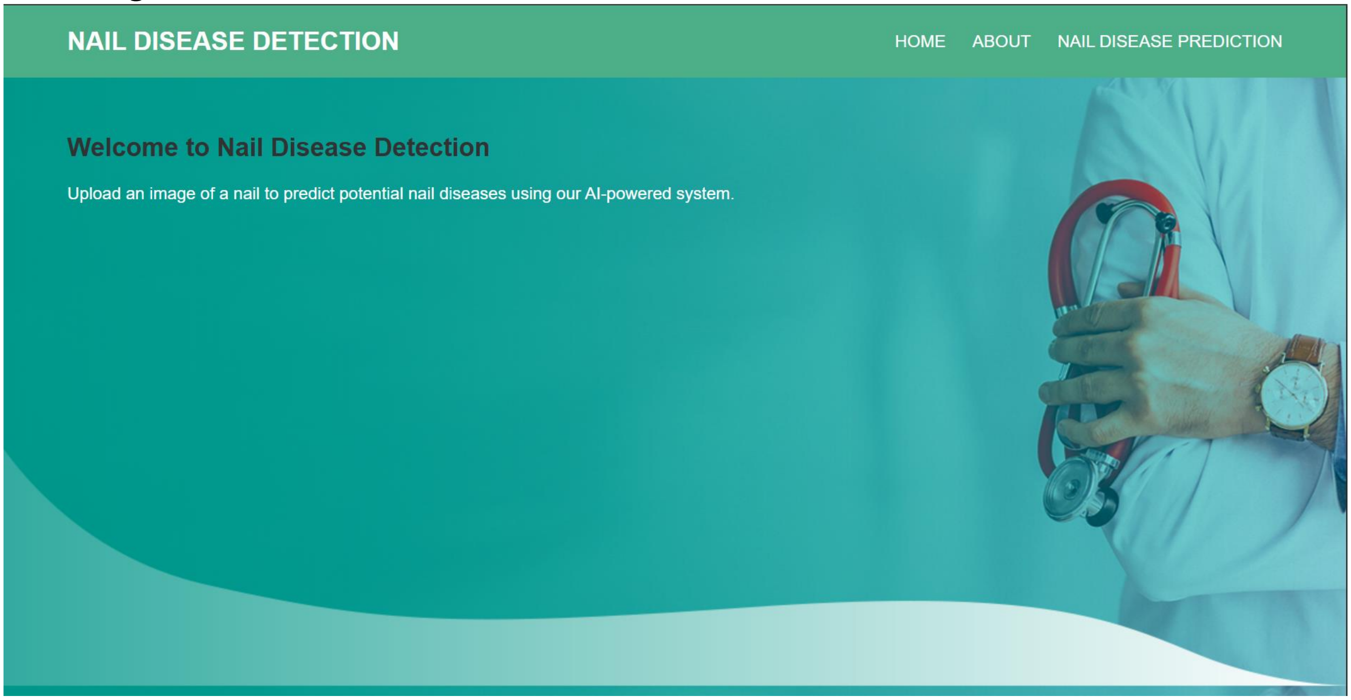
Final Model Selection Justification (2 Marks):

Final Model	Reasoning
Model 1: VGG16 (Optimized)	We selected Model 1 as our final model because it demonstrated a significant improvement in validation accuracy compared to Model 2, achieving 94.21% compared to Model 1's best of 74.59% The image provided shows the training output. We felt the higher accuracy was worth the extra training time. Model 1 also seemed to generalize better to new images.

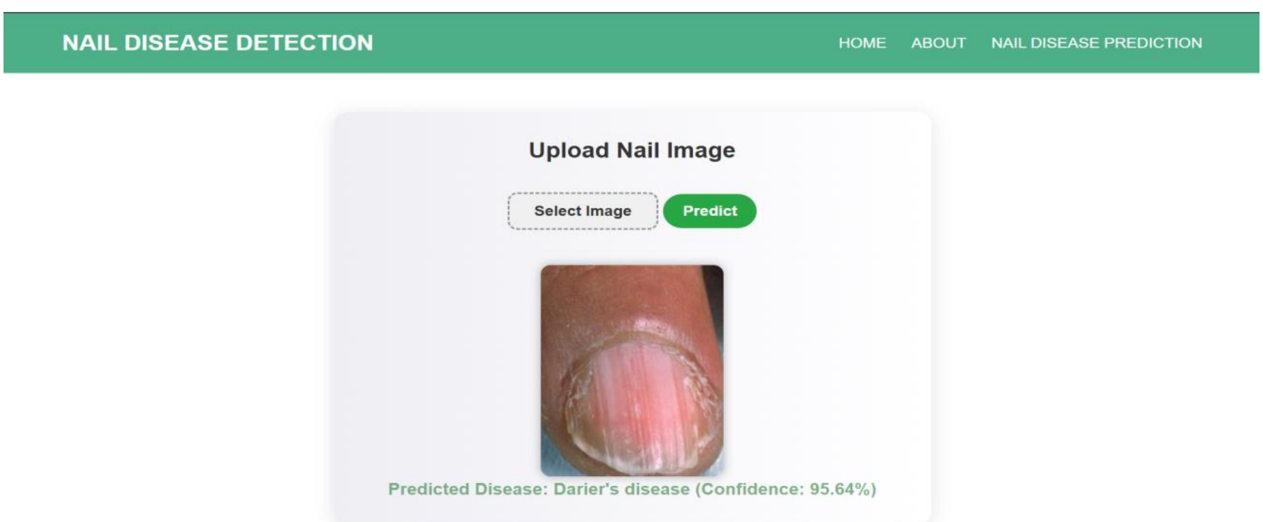
6 Results

6.1 Output Screenshots

Home Page:



Input Page:
Example :-



7 Advantages & Disadvantages

Advantages:

User-Friendly Interface: Simple and intuitive design for easy use by individuals and healthcare professionals.

Non-Invasive Diagnosis: Analyzes nail images without requiring any physical procedures.

Optical Recognition: Utilizes visual features of nails for accurate health assessments.

High-Performance Classifiers: Uses deep learning techniques (InceptionV3) for reliable prediction accuracy.

Early Detection: Enables identification of potential health issues at an early stage.

Real-Time Results: Offers fast and efficient processing of nail images.

Wide Applicability: Can be used by doctors, patients, researchers, and even in public health programs.

Disadvantages:

Limited Accessibility: May not be available to individuals without smartphones or internet access.

Dependence on Image Quality: Accuracy is affected by poor lighting or unclear nail images.

8 Conclusion

This project focused on the early-stage diagnosis of diseases through optical analysis of human nail images using deep learning techniques. By leveraging transfer learning and the Inception V3 model, the system was developed to identify visual abnormalities in nails that may indicate underlying health conditions. The approach offers a non-invasive, efficient, and accessible method for disease detection, supporting early intervention and improved health outcomes.

The project demonstrates the potential of AI-driven image processing in medical diagnostics, contributing to both personal health monitoring and professional healthcare support. Its applications span individual wellness, clinical diagnostics, and large-scale public health screening. By accurately analyzing nail characteristics, the system empowers users and healthcare providers to detect diseases early and take preventive actions.

Overall, this project represents a significant step forward in integrating artificial intelligence into preventive healthcare, with strong potential for future expansion and real-world impact.

9 Future Scope

1. Expansion to Multi-Disease Detection:

Currently, the system focuses on detecting early signs of common nail-related diseases. Future enhancements could involve expanding the model's capabilities to recognize a broader range of systemic diseases (e.g., diabetes, cardiovascular issues, respiratory problems) that manifest through nail changes, improving the diagnostic potential and coverage.

2. Integration into Mobile Health Applications:

Developing a user-friendly mobile application would allow individuals to easily capture nail images and receive real-time diagnostic insights. This will democratize access to preventive healthcare by putting diagnostic tools directly in the hands of users without the need for clinical visits.

3. Real-Time Camera-Based Detection:

Implementing live camera processing can allow for real-time analysis of nail health without needing to take still images. This will streamline the user experience and offer continuous monitoring, especially useful in wearable health tech or mobile health kiosks.

4. Cross-Integration with Health Devices and EHRs:

Incorporating the system with wearable health trackers, telemedicine platforms, or Electronic Health Record (EHR) systems can enable seamless patient monitoring, facilitate remote consultations, and enhance healthcare workflows.

5. Large-Scale Public Health Screening:

The model could be adapted for use in public health initiatives to screen large populations for disease indicators based on nail health. This would help identify early disease trends and implement preventive strategies at a community or national level.

6. Enhanced Model Training with Diverse Datasets:

To improve model robustness and generalization, future work can focus on collecting a more diverse dataset across age groups, skin tones, lighting conditions, and disease types, increasing the system's accuracy and reliability.

.

10 Appendix

10.1 Source Code

[AI-Nail Disease-Classification-Analysis]

10.2 Project Video Demo Link :

Video Demo Link:[https://drive.google.com/file/d/1MoSLv1TxejpEpmnZ41DTNolfCSIG6dBD/view?usp=drive_link]