Data Collection and Preprocessing Phase

Date	20-06-2025
Team ID	SWDTID1749906902
Project Title	Early Stage Disease Diagnosis System Using Human Nail Image Processing
Maximum Marks	6 Marks

Data Exploration and Preprocessing Report

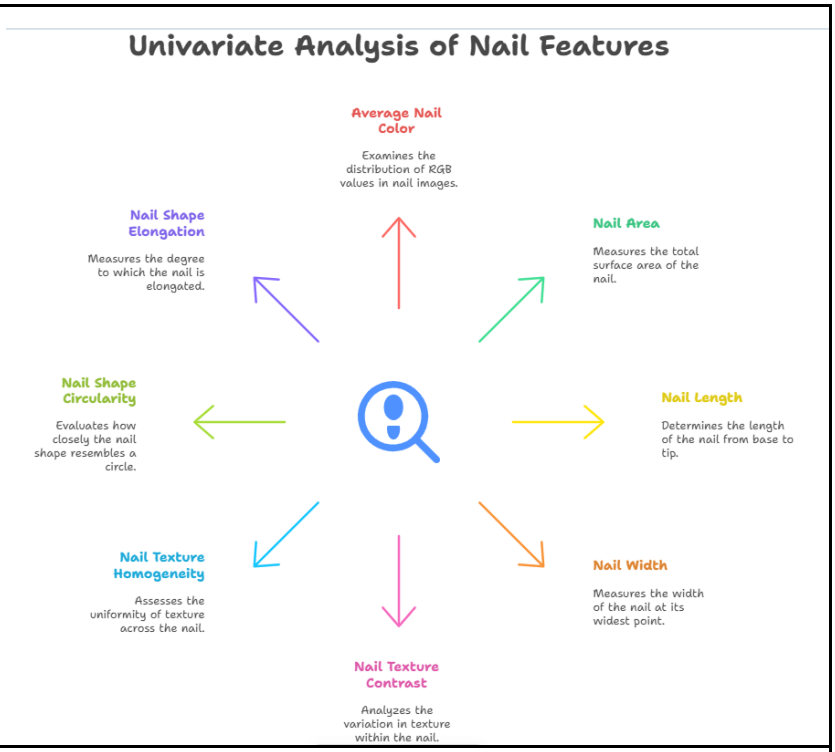
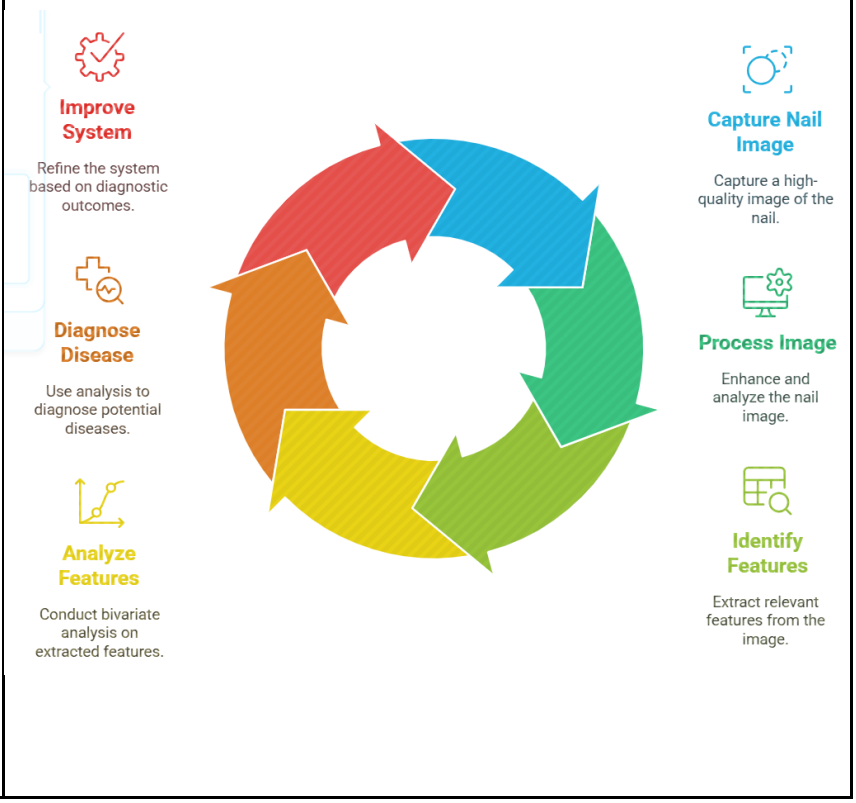
Dataset variables will be statistically analyzed to identify patterns and outliers, with Python employed for preprocessing tasks like normalization and feature engineering. Data cleaning will address missing values and outliers, ensuring quality for subsequent analysis and modeling, and forming a strong foundation for insights and predictions.

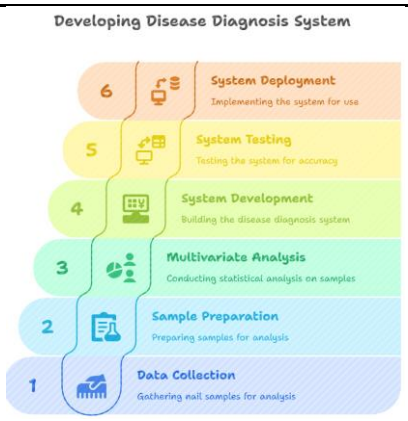
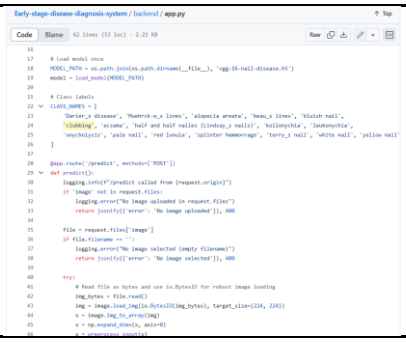
Section	Description
Data Overview	<u>Dimension:</u> **Total params:** 15,141,201 (57.76 MB) **Trainable params:** 426,513 (1.63 MB) **Non-trainable params:** 14,714,688 (56.13 MB) <u>Descriptive Statistics:</u>

Model: "functional"

Layer (type)	Output Shape	Param #
input_layer (InputLayer)	(None, 224, 224, 3)	0
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1,792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36,928
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73,856
block2_conv2 (Conv2D)	(None, 112, 112, 128)	147,584
block2_pool (MaxPooling2D)	(None, 56, 56, 128)	0
block3_conv1 (Conv2D)	(None, 56, 56, 256)	295,168
block3_conv2 (Conv2D)	(None, 56, 56, 256)	590,080
block3_conv3 (Conv2D)	(None, 56, 56, 256)	590,080
block3_pool (MaxPooling2D)	(None, 28, 28, 256)	0
block4_conv1 (Conv2D)	(None, 28, 28, 512)	1,180,160
block4_conv2 (Conv2D)	(None, 28, 28, 512)	2,359,808

Univariate Analysis	

	
Bivariate Analysis	

Multivariate Analysis	
Outliers and Anomalies	-
Data Preprocessing Code Screenshots	
Loading Data	
Handling Missing Data	(If applicable, for metadata associated with images or if images have corrupted sections) Screenshots of code for addressing missing values.

Data Transformation	Screenshots of code for image transformations (e.g., resizing, normalization, color space conversion).
Feature Engineering	Attached the codes in final submission. (e.g., extracting texture features, shape features, or color histograms from nail images).
Save Processed Data	-