Infosys SpringBoard Virtual Internship 6.0

DermalScan:AI_Facial Skin Aging Detection App

Submitted by

Meghana sandya Vallabharaju

Under the guidance of Mentor

Mr. Praveen Bhargav

Introduction:

AI-DermalScan is an end-to-end AI/Deep Learning project designed to analyze facial skin conditions from images and classify them into different categories. The system leverages image processing and convolutional neural networks (CNNs) to preprocess, augment, and learn patterns from skin images, enabling accurate classification and diagnosis support.The main goal of this project is to assist dermatologists and skincare professionals by providing an automated, intelligent system that can identify common skin issues such as:

Wrinkles

Dark Spots

Clear Skin

Puffy Eyes

By combining image preprocessing, augmentation, and deep learning models, AI-Dermal Scan ensures that the model generalizes well on real-world images and is robust to variations like lighting, orientation, and facial expressions.

Project Statement:

The objective is to develop a deep learning-based system that can detect and classify facial aging signs—such as wrinkles, dark spots, puffy eyes, and clear skin—using a pretrained EfficientNetB0 model. The pipeline includes face detection using Haar Cascades, custom preprocessing and data augmentation, and classification with percentage predictions. A web-based frontend will enable users to upload images and visualize aging signs with annotated bounding boxes and labels.

Objectives:

● Detecting and locating facial characteristics which indicates the presence of age.

● Categorise the characteristics into 4 specific features like wrinkles, dark spots, puffy eyes, and clear skin using a trained CNN model.

● Train and evaluate an EfficientNetB0 model for robust classification.

● Build a web-based frontend for uploading facial images and viewing annotated outputs.

● Integrate a backend pipeline that processes input images and returns annotated results.

● Export annotated outputs and logs for documentation or analysis.

Milestone 1

Module 1: Dataset Setup and Image Labelling

In this module, the dataset is prepared and organized for further processing. The dataset is downloaded and stored locally within the project directory, ensuring easy access during model development. Python's os module is used to handle the dataset structure, allowing directory traversal, file listing, and automated reading of class folders.

Classes

Each class in the dataset represents a specific facial aging category. The project includes four primary classes are Clear Skin, Dark spots, Puffy Eyes, and Wrinkles.

Importance in Deep Learning

Proper categorical labelling is crucial because deep learning models rely on accurately labelled data to learn distinguishing features for each category. Mislabelled or mixed images can lead to confusion during training and result in poor classification performance. By organizing images into separate folders corresponding to each facial aging sign, the model can effectively learn the unique characteristics of each class.

Dataset Verification

After organizing and labelling the dataset, it is essential to ensure class balance by checking the number of samples in each category. Imbalanced datasets can lead to model bias, where the network may favor classes with more samples, negatively affecting classification accuracy.

Class Distribution Visualization

To understand the dataset composition, a bar graph is used to visualize the number of images present in each facial aging category. This graphical representation provides a clear comparison of class-wise image distribution and helps in identifying any imbalance among the classes.

The bar chart is generated using the matplotlib library, where:

The x-axis represents the facial aging categories

The y-axis represents the number of images in each category

This visualization plays a crucial role in determining whether data augmentation techniques are required in later stages to balance the dataset.

Training Class Distribution

Output of Module 1

- Verified dataset structure
- Clear class definitions
- Class-wise image count analysis
- Bar graph visualization of class distribution

Module 2: Image Preprocessing & Augmentation

In this module, the images in the dataset are prepared so that they can be used effectively by the deep learning model. Since neural networks require uniform input size, all images are resized to **224 × 224 pixels**. After resizing, the pixel values are normalized by scaling them between 0 and 1 using a rescale factor of **1/255.** This normalization helps the model train faster and improves stability during learning.

To make the model more robust and to reduce overfitting, image augmentation techniques are applied to the training data. Augmentation includes operations such as horizontal flipping, rotation, zooming, and shifting. These techniques create variations of the original images and

help the model learn generalized patterns instead of memorizing the training data. Important facial features like wrinkles, dark spots, and puffy eyes are preserved during augmentation.

The dataset is divided into training and validation sets, where 80% of the images are used for training and 20% are used for validation. This split allows the model to be evaluated on unseen data and helps in monitoring overfitting during training.

Additionally, class labels are converted into **one-hot encoded vectors**. One-hot encoding transforms each class label into a binary format suitable for multi-class classification, allowing the model to process categorical data effectively.

Outcome of Module 2

By completing this module, the dataset is fully prepared for model training. All images are standardized in size, properly normalized, and augmented to increase data diversity. The training and validation datasets are now ready with correctly encoded labels, ensuring smooth and effective model learning in the next module.