Age Classifier Model

# 1. Introduction

This project implements an age classification system using facial images. The dataset is organized into folders labeled by age (20, 30, ..., 80). Each folder contains 10 images of the same 10 individuals observed at different ages. For example, image1.jpg in folder 20 and image1.jpg in folder 50 represent the same person at different ages. The goal is to classify an image into its correct age group.

# 2. Project Structure

The project is divided into two main scripts:

* 1. data\_processing.py - Handles data loading, augmentation, feature extraction, and saving NumPy arrays.
* 2. train\_model.py - Loads preprocessed data, defines the model, and evaluates it using accuracy, precision, recall, and F1 score.

# 3. data\_processing.py

This script performs the following tasks:

• Loads the dataset, where each age folder contains 10 images of 10 different individuals.

• Augments each image using Keras ImageDataGenerator to expand dataset size (rotation, zoom, shift, etc.).

• Encodes age labels using LabelEncoder.

• Splits data into train, validation, and test sets using stratified sampling.

• Extracts features from images using MobileNetV2 (pre-trained on ImageNet, pooling='avg').

• Saves the resulting features and labels into .npy files for later training.

# 4. train\_model.py

This script performs the following tasks:

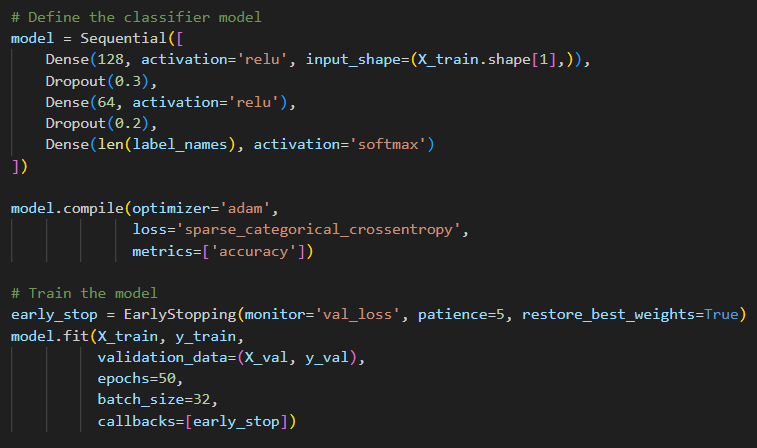
• Loads the saved NumPy arrays for features and labels.

• Defines a dense neural network classifier with dropout for regularization.

• Compiles the model using Adam optimizer and sparse categorical crossentropy loss.

• Trains the model with early stopping based on validation loss.

• Evaluates the model using accuracy, precision, recall, and F1 score on the test set.

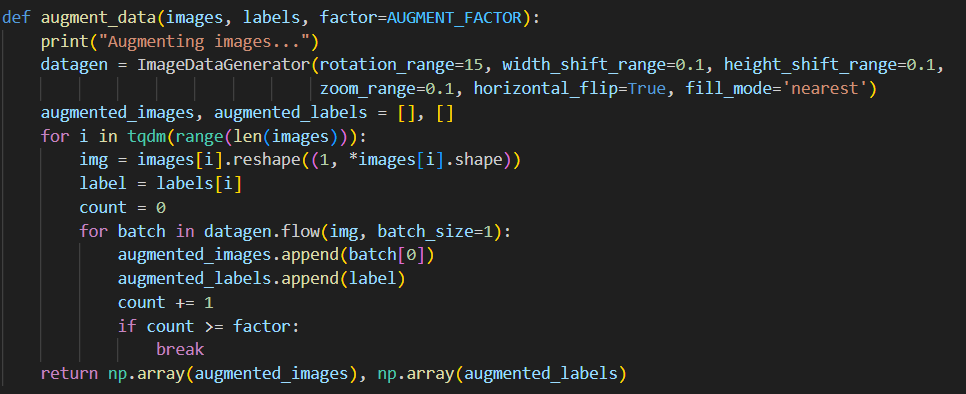


# 5. Problems

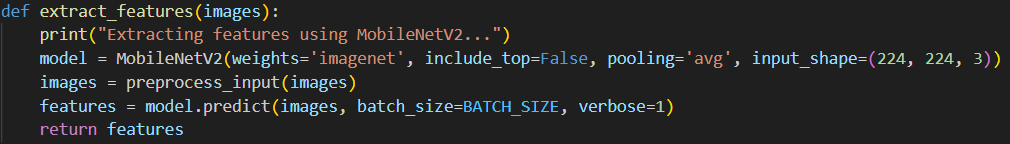
* Major problem faced was the small size of the data set which was causing overfitting.
* How to extract the factors affected by aging on a person’s face.

# 6. Approaching the Problems

* The overfitting of the model due to small dataset size was tackled by augmentation. I used ImageDataGenerator model from Keras which applies: Rotation, Shifting, Zooming and Flipping to the original image to generate its 5 duplicates hence increasing the size of the data.

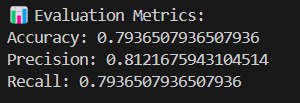


* The major changes on a person’s face due to aging are: wrinkles, skin texture, sagging. These were the features required to extract from the data. To do so I used the MobileNetV2 model which is trained on a large ImageNet data for these features.



# 7. Conclusion

* When the model was trained on the features extracted from the image dataset provided instead of images directly gave better results.



## 