Skin Vision

Reference : <https://www.kaggle.com/code/malakalaabiad/skin-tone-classification#4.-Model-Creation-Using-Transfer-Learning-(MobileNetV2)>

Data Augmentation and Preprocessing

Crucial for training a more robust and generalized model.

Why augment data?

Prevents overfitting by making the model see slightly modified versions of images.

Helps the model generalize better by learning invariant features (like skin tones from different angles or lighting).

train\_datagen = ImageDataGenerator(

rescale=1./255,

rotation\_range=20,

zoom\_range=0.2,

width\_shift\_range=0.2,

height\_shift\_range=0.2,

horizontal\_flip=True

)

rescale=1./255 - Normalizes pixel values from 0–255 to 0–1 (MobileNetV2 expects normalized input).

rotation\_range=20 - Randomly rotates images by up to 20 degrees.

zoom\_range=0.2 - Randomly zooms in or out by up to 20%.

width\_shift\_range=0.2 - Horizontally shifts images by up to 20%.

height\_shift\_range=0.2 - Vertically shifts images by up to 20%.

horizontal\_flip=True - Randomly flips images horizontally.

Validation Data

valid\_datagen = ImageDataGenerator(rescale=1./255)

Only rescaling is applied — no augmentation.

Why?

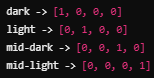
Validation data should reflect real-world, unaltered images.

It tests how well the model performs on untouched data.

Data loaded as batches. Batch size = 32 images

Class Mode — class\_mode='categorical' because:

Your labels are one-hot encoded — like this,



Model creation using MobileNetV2

base\_model = MobileNetV2(weights='imagenet', include\_top=False, input\_shape=(224, 224, 3))

MobileNetV2 — loads a pre-trained MobileNetV2 model.

weights='imagenet' — uses weights trained on the ImageNet dataset (1,000 classes like dogs, cats, etc.).

include\_top=False — removes the final classification layer (the one for 1,000 ImageNet classes).

input\_shape=(224, 224, 3) — sets the input size to 224x224 RGB images (3 channels for color).

Why remove the top layer?

ImageNet’s final layer outputs 1,000 classes (like dogs, cats, etc.).

Your task only has 4 classes (skin tones), so you remove the top layer and add your own custom layers.

Freeze base layers

for layer in base\_model.layers:

layer.trainable = False

freezing layers means their weights won’t be updated during training.

The convolutional layers of MobileNetV2 already learned to detect basic features (edges, textures, shapes) during ImageNet training.

This lets you reuse those features for skin tone classification without re-learning from scratch.

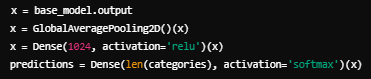
Why freeze layers?

Transfer learning works best when you reuse low-level features from pre-trained models.

Early layers detect edges, shapes, and patterns — useful for any image task (like skin textures).

You only train the new layers (which learn skin tone-specific patterns).

New layers add by me.



The output layer will give something like this for a given image:

A screen shot of a computer

AI-generated content may be incorrect.

The class with the highest probability is the predicted class.

Then we want to define the **full model** with connecting them.

[Input (224, 224, 3)] → [MobileNetV2 layers (frozen)] → [GlobalAveragePooling2D] → [Dense 1024 ReLU] → [Dense 4 Softmax]

model.compile(optimizer=Adam(learning\_rate=0.001), loss='categorical\_crossentropy', metrics=['accuracy'])

Adam optimizer — updates the model’s weights during training using the Adam algorithm (adaptive learning rate optimization).

Learning rate (lr=0.001) — controls how fast/slow the model updates weights.

categorical\_crossentropy — loss function for multi-class classification with one-hot labels (like [1, 0, 0, 0] for "dark").

metrics=['accuracy'] tracks accuracy during training.