Image Classification of 5 F1 Drivers

1. Business Context:

This project was born as a combination of my interest in ML applications and the F1 motorsport. The aim was to identify and classify any given image as one of my 5 favourite F1 drivers. Business uses may include: sorting news articles (segregating based on cues from images), cross-cultural fashion elements can be easily searched on e-commerce websites (the model has to be altered and trained for fashion items), searching for and identifying a particular person in new images (like in Instagram, though they use CNN and deep learning).

2. Problem Statement:

Given images of various Formula 1 drivers; classify those as belonging to the right personality (among Charles Leclerc, Daniel Ricciardo, George Russell, Max Verstappen and Sebastian Vettel).

3. Solution Developed:

Though modern image classification algorithms use CNN (Deep Learning), classic ML algorithms can also be effective in image classification. To this end, the main idea behind the solution is to use libraries in OpenCV to extract the facial features of an individual (their pre-trained models are open-sourced in their website); and using these features along with the original image, train a model. The process also involves significant data cleaning and web scrapping. In short, we use the horizontal, vertical and diagonal details extracted using wavelet-transform. The model initially used was SVM.

4. Improvements to the Solution:

The classification score was significantly improved upon using hyperparameter tuning; using GridSearchCV. Increasing the number of images used for training by 46 (145 total to 191 total, 37 used for testing), raised the validation score by 2% for Logistic Regression model.

3 different models; SVM, Random Forest and Logistic Regression were trained and the best results were obtained from Logistic Regression (though SVM had close score).

5. GitHub Link to the Jupyter Notebook:

https://github.com/aswathp/ML_Projects/blob/main/F1%20Drivers%20.ipynb