Machine Learning Engineer Nanodegree

Capstone Proposal

Rashid Kamran
May 14, 2018

Proposal

Domain Background

I recently came across an interesting study "A Diverse Benchmark Dataset for Multi-Paradigm Facial Beauty Prediction (FBP)"[1]. The study mainly assessed the facial attractiveness that is consistent with human perception.

In my project I want to replicate the results of the study and take it fun step further to deploy the trained model to an iOS app that would help people capture best of their faces (aka selfies) using an indicator on the screen.

[1]https://arxiv.org/pdf/1801.06345.pdf

Problem Statement

This is a classic image analysis problem but rather than classifying we are assigning a numerical score between 0-5 to measure attractiveness based on the faces in the dataset. Higher score perceived to be more attractive based on the score provided in the dataset.

The utility of the original regression seems limited as it would be unwise to put a score on a face but by using the technique I will find a way to help people capture their best-looking faces.

Datasets and Inputs

For this project I am relying on the original dataset[2] of 5500 face images (Asians and Caucasian male/female). Every face has given a score between 0-5 of perceived attractiveness. All images are 350x350 in shape. Out of 5500 images 4000 faces are Asian(male/female) so it would be interesting at least anecdotally to see the effect on other faces not represented in the dataset.
[2] **https://github.com/HCIILAB/SCUT-FBP5500-Database-Release**

Solution Statement

Like the original authors I am planning to use a pre-trained feature extractor for the final solution like AlexNet or ResNet and a dense layer on top for regression. The trained model will be converted to CoreML model to be used in the iOS app.

It's assumed for this solution that pre-trained Resnet50 is one of the state of the art feature extractor and will be sufficient for simple regression on top.

Benchmark Model

Since the dataset is small so as a bench mark I am also planning to use Support Vector Regression and compare the performance with the targeted solution using ResNet50 as a feature extractor and a Dense layer on top.

Evaluation Metrics

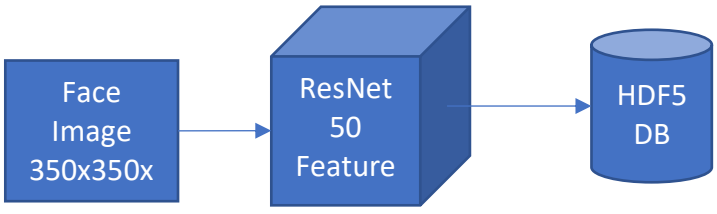
For a simple regression albeit on image data we will be relying on RMSE evaluation metrics. I will try to compare against study's results and will do both 5-Fold training and validation and also 90/10 split validation.

For anecdotal proofs I will also visualize the results of the model using some celebrities and personal, family and friend's photos.

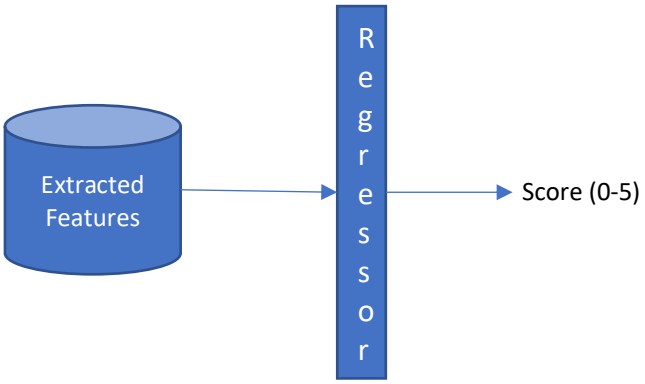
Project Design

As stated above I am going to split the problem into three steps and demonstrate my understanding of transfer learning and utilization of deep learning in a real-world situation.

1. Using pretrained ResNet50 to extract features from entire dataset and store into a HDF5 database. This will help us train fast for regression and perform other experiments without loading the huge ReseNet50 again.



2. Use a dense layer on top of features to perform regression.



3. Covert to CoreML model and build an iPhone app on top. On iOS capturing video stream and running an ML model is a very trivial task using Vision APIs so the details of which I am not documenting here.