

STAT 231 / CS 276A - Project 3

1. Objectives

This project is an exercise for studying the social attributes of human faces using Support Vector Machines. See examples in the figure below. The goal of this project is 1) to train classifiers that can automatically infer the perceived human traits from facial photographs and 2) to apply the model to analyze the outcomes of real-world political elections. We have collected a set of face images of US politicians, and in some cases we have a pair of images for the two competing candidates running for public offices: Congressmen, Senators or Governors.

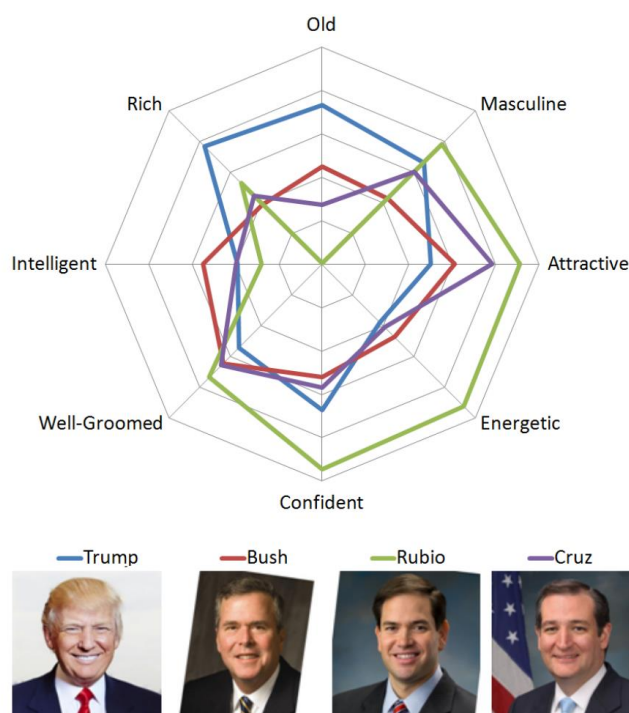


Figure 1. The inferred social dimensions of faces of a few Republican politicians who may run for 2016 presidency, predicted by our learned model.

You will train various models using the provided images as well as the corresponding annotations (traits or voting shares). The first step is to build trait classifiers from the facial photographs. We used Mechanical Turk to obtain the human judgments on the perceived trait dimensions such as 'honest' or 'intelligent.' Your classifiers should be trained to automatically infer these annotated traits using both geometric and appearance features. Then we use these trained classifiers to predict the traits of novel faces using the second part of dataset which contains the pairs of rival candidates who ran against each other.

2. Data

In the enclosed zip file, you will see two image folders:

- a. img/ # This is the set of images that you use for training trait classifiers. Use this for the Part I.
- b. img-elec/ # This folder contains two sub-folders (senator and governor). Use this for the Part II.

Provided Codes:

./demo.m - an example script
./HoGfeatures.cc - a mex implementation of HOG extraction.
./HoGfeatures.mexw64 - a pre-compiled mex file
./libsvm_matlab/ - libsvm

Annotations:

train-anno.mat: This contains the perceived trait annotations (491 x 14 matrix). The nine variables correspond to {**Old, Masculine, Baby-faced, Competent, Attractive, Energetic, Well-groomed, Intelligent, Honest, Generous, Trustworthy, Confident, Rich, Dominant**}

Note: The trait annotations are obtained in the form of ranking and you will see **real-valued** scores instead of binary classes. Consider these options: a) Set an arbitrary decision threshold (e.g., 0) and divide the whole set into a positive and a negative set, or b) Use regression (e.g., SVR or RankingSVM).

In addition, we provide you with pre-computed facial landmark coordinates (491 x 160 matrix). Each row represents 80 keypoint locations [x1, x2, ..., x80, y1, y2, ... , y80]

stat-sen.mat & stat-gov.mat : These files contain the pre-computed facial landmarks for the Part II images and the actual voting share differences between the candidate pairs.

In this project, you will do the following steps.

Part 1, Design facial features for predicting the social attributes.

The goal of the part 1 is to train binary SVMs (or SVRs) to predict the perceived traits (social attributes) from facial photographs. You can use the pre-computed facial keypoint locations and extract HoG (histogram of oriented gradient) features using the enclosed matlab function. You can further try your own favorite features.

We do not explicitly divide the image set into train/test sets. Therefore, you need to perform k-fold cross-validation and report the obtained accuracy.

Part 2: Using SVM to predict the election results based on the social attributes.

Now we use the learned classifiers to analyze the outcomes of real-world elections. Run your trained classifiers on the images in "img-elec" directory and obtain the predicted trait scores. Then train an

“election winner” classifier based on the trait features of size 14. You should also perform cross-validation to measure accuracy and see if you can get accuracy better than chance.

Since each race comprises two candidates, binary classification is not suitable. A simple trick is to define a pair of politicians as one data point by subtracting a trait feature vector A from another vector B, and train a binary classifier: $f_{ab} = f_a - f_b$. Do not include a bias term.

Also, show the correlations between the facial traits and the election outcomes. What are the facial attributes that lead to the electoral success?