Adrian Marinovich
Springboard Data Science Career Track
Capstone Project #1 Proposal
September 23, 2018

Detection of smiles in images of faces

What is the problem you want to solve?
The problem is to detect smiles in images of faces.

Who is your client and why do they care about this problem? In other words, what will your client do or decide based on your analysis that they wouldn't have done otherwise?
The detector developed here may find eventual implementation in a human-machine interface, as one step in developing emotional communication tools to control such things as musical instruments via MIDI, or to allow for more safe and reliable human-robot interactions, so that robots can better infer human intentions and thereby enhance their response decisions. The clients for such applications may be found in the music, robotics, and physical programming fields. A range of additional clients may be interested in smile detection as a step towards emotional classification from human facial expression, for example in market research, to better gauge interest in and reaction to products in order to guide decisions on product design, or gaming, to support decisions in creating more immersive and exciting player experiences.

What data are you using? How will you acquire the data?
The data will be obtained from a labelled subset of the cropped version of the Labeled Faces in the Wild (LFW) dataset (LFWcrop: https://conradsanderson.id.au/lfwcrop/), in which faces are centered on the image with the background largely omitted. The LFWcrop dataset consists of 13,233 images, available as both 3-color and grayscale. The list of face images labelled as smiles comprises 600 images, and the list of face images labelled as non-smiles consists of 603 images (lists available at: https://data.mendeley.com/datasets/yz4v8tb3tp/5). This is a balanced dataset with ~600 images per target class. The cropped images have a resolution of 64x64 pixels.

Briefly outline how you'll solve this problem. Your approach may change later, but this is a good first step to get you thinking about a method and solution.
Machine learning techniques, including support vector machines and neural networks, will be used to classify the images into smile and non-smile images.

The maximum dimensionality of each image is 12,288 with 3 colors (64x64x3). Limiting analysis to grayscale images would yield a reduced dimensionality, D, of approximately 4,096. This gives a D/N ratio of 6.8 (4,096/600) at the outset. Dimensionality reduction techniques, such as principal components analysis, are expected to further reduce the D/N ratio. Regularization

techniques will be used to reduce overfitting. For neural networks random dropout layers will help ensure a generalized model that should work on an untrained subsample of the dataset.

Working within the Amazon Web Services and other GPU-based environments such as Google Colaboratory, the analysis will be conducted using Python tools such as OpenCV, Scikit-Learn, TensorFlow and Keras.

What are your deliverables? Typically, this includes code, a paper, or a slide deck.
The deliverables will include the machine learning code, a brief paper describing the findings of the analysis, and a slide deck, to be made available on GitHub.

References:
Arigbabu, Olasimbo Ayodeji, et al. "Smile detection using hybrid face representation." Journal of Ambient Intelligence and Humanized Computing (2016): 1-12.

Huang GB, Mattar M, Berg T, Learned-Miller E (2007) Labeled faces in the wild: a database for studying face recognition in unconstrained environments. University of Massachusetts, Amherst, Technical Report.