CS 221 Project Proposal

Task Definition

Several platforms have attempted to draw information from pictures. These systems range from those capable of identifying an image's geolocation to those that can recognize different faces within a picture. For our project, we aim to build a program that can accurately discern a person's facial expression within an image and classify their sentiment into a finite list of categories.

Input/Output Data

Our input data will be taken from this kaggle challenge: https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge/data . The data consists of 28,709 grayscale images of faces that have been standardized for size/scale. The images have been translated in to a string of space separated pixel values in row major order. Each image has been classified in to one of seven categories (0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral).

The output of our system will similarly be a classification of the expression in a given image.

For example:

input: 70 80 82 72 58 58 60 63 54 58 60 48 89 115 121 119 115 110 98 91 84 84 90 99 … 154 140 78 21 11 61 144 168 173 157 138 150 148 132 159 182 183 136 106 116

output: 0

Evaluation

Because this is a kaggle challenge, there is test data that is formatted in the same way as the training data. We will use this as the primary evaluation criteria for our system. If we test our project with images that are not part of the kaggle data we will use human discretion (the average of several opinions) to determine if the classifier is correct.

Baseline and Oracle

For our baseline assessment, we plan to classify sample input data using linear classification with the feature set being the value for each pixel. We expect that this will be somewhat effective, because our data has been standardized in terms of location of facial features so there should be similarities in the pixel values of each image. However, this will not work very well given the large variety of shapes that may all be classified as "sad," "happy," etc. by a human viewer, and the differences in the sizes and colorings of human facial features. Furthermore, we do hope to use our system on real images, instead of just standardized kaggle data and this system does not generalize to images that are not scaled perfectly.

Challenges we anticipate facing with this project include differing picture quality and hues (as we go from pictures to strings of pixel values), the complexity of recognizing shapes within pictures, and the wide range of ways that humans can express each emotion.

Our oracle is a human's assessment of a person's emotion given a picture of that person's face. Studies in the past have been done upon the complex factors that determine a person's ability to predict emotion given a picture of a face, and variety of factors people take into account when deciding upon the emotion in a picture.

Related Work:
The task of deriving information from pictures of faces has been explored in several contexts. One common example of this is face recognition software, a variation of which is used by Facebook to "guess" which of your friends you should tag in your picture. Outside of face recognition, other platforms use image analysis to explore information that can be derived from facial expressions. For example, a paper from University of California San Diego analyzed how a student's facial expression before an exam can be used to predict the student's exam score. A separate paper from the National Institute of Mental Health, used identification of fear in a person's facial expression to predict prosocial behavior. From these diverse uses of expression analysis, it is clear that this topic can be taken in many different directions.
There are also a variety of approaches to actually solving the problem of expression analysis. For instance, the UCSD study used features such as brow lowering intensity, and mouth dimpling. The NIMH study used 3D morphable models and Elastic Bunch Graph modeling. Perhaps more relevant to our analysis, other approaches that have been attempted include geometric feature-based methods and appearance based methods. In the geometric methods, facial feature points are extracted to form a feature vector that represents face geometry. In appearance based methods image filters are applied to the face to extract a feature vector. Some specific techniques that can be used to extract features include the Rough Contour Estimation Routine to get features of eyebrows, eyes, and mouth, the Point Contour Detection method to improve the precision. Approaches have also varied in whether they classify the entire image or attempt to break the face into small parts that can then be classified, with the former tending to provide better results.

References
Competition Overview: https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge
Automatically Recognizing Facial Expression: Predicting Engagement and Frustration
http://www.educationaldatamining.org/EDM2013/papers/rn_paper_09.pdf
Computer software analyzing facial expressions accurately predicts student test performance
http://www.sciencedaily.com/releases/2014/04/140416101335.htm
Accurate identification of Fear Facial Expressions Predicts Prosocial Behavior
http://www.ncbi.nlm.nih.gov/pubmed/17516803
Image Analysis for Face Recognition
http://www.face-rec.org/interesting-papers/general/imana4facrcg_lu.pdf
Geometric and Appearance Based Methods
http://link.springer.com/chapter/10.1007/0-387-27257-7_12#page-1
literature review on the topic
http://www.academia.edu/6761678/A_literature_survey_on_Facial_Expression_Recognition_using_Global_Features

Team
Chelsea Lewis, Katie Redmond, Rohisha Adke