

Capstone Proposal

Omar Ezzat

June 24, 2017

Domain Background Computer Vision

Using the Laptop webcam to classify the users facial expression.

Facial expression are a good indicator for the emotional state. They are also a strong form of non verbal communication

Application with an basic understanding for those emotional states can improve the user experience.

This problem is particular interesting for me to improve my portfolio. It held the possibility to build a web application which is entertaining, scientific advanced and easy to use.

Similar project: [memeoji](#)

Interesting paper describing a solution using convnet: [Stanford Report](#)

Problem Statement

We want to map the user snapshot to one of 7 emotions Angry, Disgust, Fear, Happy, Sad, Surprise, Neutral.

Let $X := [0, 255]^{48 \times 48}$ be a set of images and $Y := 0, 1, 2, 3, 4, 5, 6$ the label. We need to find a function $f : X \mapsto Y$. This mapping will result to an emotional state.

The input $x \in X$ will result to an emotion $y \in Y$

Datasets and Inputs

I'm using the kaggle dataset fer2013 [dataset](#)

"The data consists of 48x48 pixel greyscale images of faces. The faces have been automatically registered so that the face is more or less centered and occupies about the same amount of space in each image. The task is to categorize each face based on the emotion shown in the facial expression in to one of

seven categories (0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral)."

Training set: 28709 examples
Test set: 3589 examples
Validation set: 3589 examples

Solution Statement

Convolutional nets will be used to solve this problem.
The solution should result into a web application, which classifies user snapshots.
We need to pre process the image, resize them to 48x48 and remove the color channels.

Benchmark Model

As a benchmark model I will use a out of the box SVM. Using the validation set.

Example: [link](#)

Note: The kaggle winner scored an accuracy of 0.71162.

Evaluation Metrics

The evaluation metric will take use of the confusion matrix.

Calculating:

Precision: $\text{True Positive} / (\text{True Positive} + \text{False Negative})$

Recall: $\text{True Positive} / (\text{True Positive} + \text{False Positive})$

F1 score: $F1 = 2 * (\text{precision} * \text{recall}) / (\text{precision} + \text{recall})$

Having an unbalanced dataset, especially the class disgust with only 113 samples, makes it hard to find a good evaluation metric.

Comparing all three of them for each class will give an better understanding of the performance.

Project Design

Steps 1:

Pre process the data. Understand the faces, the emotions, looking for outlier.

Step 2:

Run the benchmark model.

Step 3:

Construct my model. Using the techniques I learned from the deep learning

chapter. (Inputlayer, convnets, dropout, pooling, fully connected layer, softmax, stochastic gradient descent, learning rate decay, regularization)

For example:

Inputlayer - 3-5 convnet-layer - 2 fully connected layer - output layer

Step 4:

Evaluate the different models.

Step 5:

Decide which model is the best.

Step 6:

Implement the trained model into the webpage.