Project 5 Proposal

Jessica Hudiono

Overview

Emotion detection with machine vision, also known as "affective computing", is an emerging field with many expected applications but still under development. I would like to explore this subject by taking an existing face recognition neural network and adding new layers.

Accuracy is my primary metric.

MVP

On my first pass I will focus on classifying one emotion at a time, such as "happy" vs. "not happy". If this is successful I will try multiclass classification with the emotions considered standard in FACS (happiness, sadness, surprise, fear, anger, disgust, neutral).

Data

I found a few possible candidates for my training and testing datasets. IMFDB has the most volume but uneven quality. I don't know yet if my model will be able to perform with low resolution photos, so if I can't get good results I will try combining KDEF and SOF.

Datasets

	# of images	Labeled?	Color or grayscale?	Position?	Notes
IMFDB	30k+	FACS	color	Frontal, Left, Right, Up, Down	Some are very bad resolution, no "neutral" label
KDEF	4900	FACS	color	Straight, full profile, half profile	
JAFFE	213	FACS	grayscale	Frontal	Good for toy example
CK+	~1000?	FACS	grayscale	Frontal	Person not always in center
SOF	2400	?	color	Mostly frontal, mix of control setting	More challenging. Labels are in .mat

				and "in the wild"	format.
[ChaLearn](#)	video	FACS, fake or real	color	Frontal	Honest vs dishonest expressions

Features

Variable	Type	Description	Used for Model
Image Pixels R channel	int	Red component of pixel	Y
Image Pixels G channel	int	Green component of pixel	Y
Image Pixels B channel	int	Blue component of pixel	Y
Gender	Categorical	Male, Female	Y
Position	Categorical	Frontal, Left, Right, Up, Down	Y

Known Unknowns

It is not guaranteed that the photo subjects gave an authentic performance of the emotion. I'm not sure yet if this will make my model less accurate, either on cleaned test data or a more realistic setting.

A quick test run of GradientBoostingClassifier on JAFFE with 256x256 resolution took about 12 minutes on my local machine. If I want to use larger datasets, I might have to use AWS.

I don't know yet how training on color images will affect performance on grayscale, and vice versa.

Further goals

- More emotions
  - Clustering to find expression groupings than FACS
  - Regression to rate expressions on two or more axes, for example the circumplex model of affect (pleasure vs arousal).
- Investigate if training on a combination of still images and video data shows improvement.
- Create a live demo with a webcam.
- Realistic data: different poses, lighting, occlusion

- Fake vs authentic emotion.