CV Pipeline for Facial Expression Recognition

Kamil Akhmetov
Vadim Solovov

https://github.com/kamilkoduo/cv-fer

Human Interaction Systems

“We create intelligent and emotionally aware interactions between people and machines”

— Dr. Mark Sagar,
Soul Machines, New Zealand

Why Facial Expression Recognition?

- The gap between machines and humans becomes apparent with time.
- Reading social signals is one of the ground points in Human Interaction Systems.
- Understanding the relationship between human Facial Expressions and Social Signals allows to decrease the gap.
- Computers can learn empathy and understand customer needs better.

Problem Statement

- Deep Learning - standard approach. End-to-end models.
- CNNs are opaque. Expressive models, but hardly interpreted
- Health-care, psychology, sociology, driver safety, virtual reality, cognitive sciences require both **effective and interpretable** models
- Traditional Machine learning - more transparent, comprehensible results, however they require additional stages

Project goals

- Learning CV and ML in the field of Facial Expression Recognition
- Building a pipeline with preprocessing, feature extraction and classification stages

Existing methods

- Classifiers - SVM, AdaBoosted Random Forest
- Feature Extractors - Histogram of Oriented Gradients, Local Binary Pattern
- Facial Landmarks Extraction - Ensemble of Regression Trees

We combine these methods together.

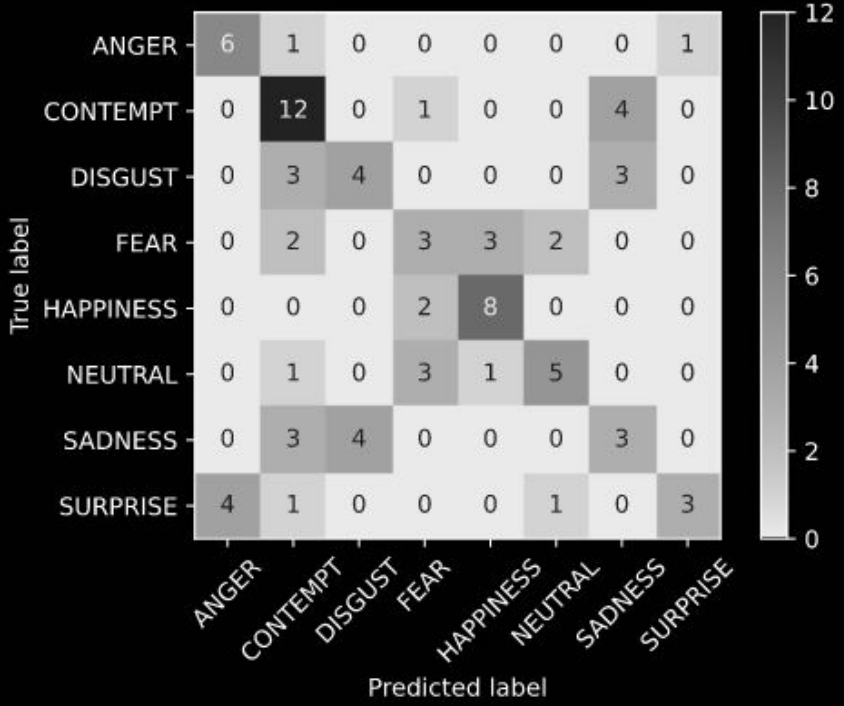
Dataset

- Taiwanese Facial Expression Images Database (TFEID)
- relatively small: 336 samples
- 20 female and 20 male models
- 8 facial expressions

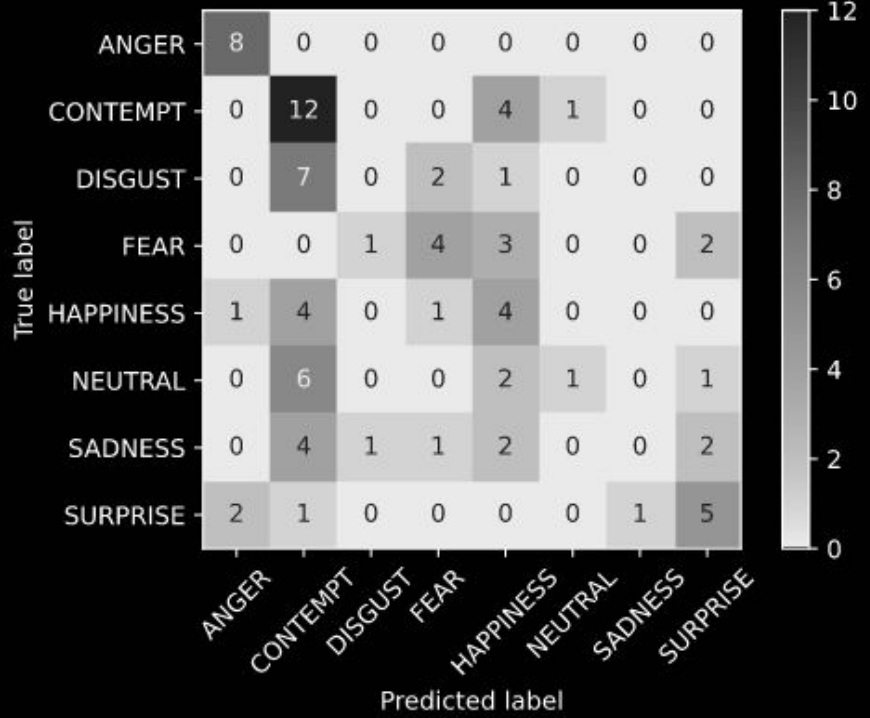
Experiments

HOG and LBP

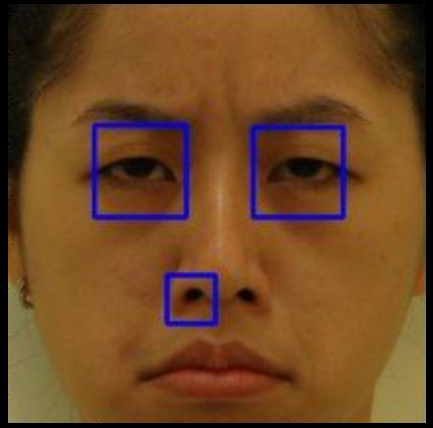
HOG + AdaBoost: 52.38%

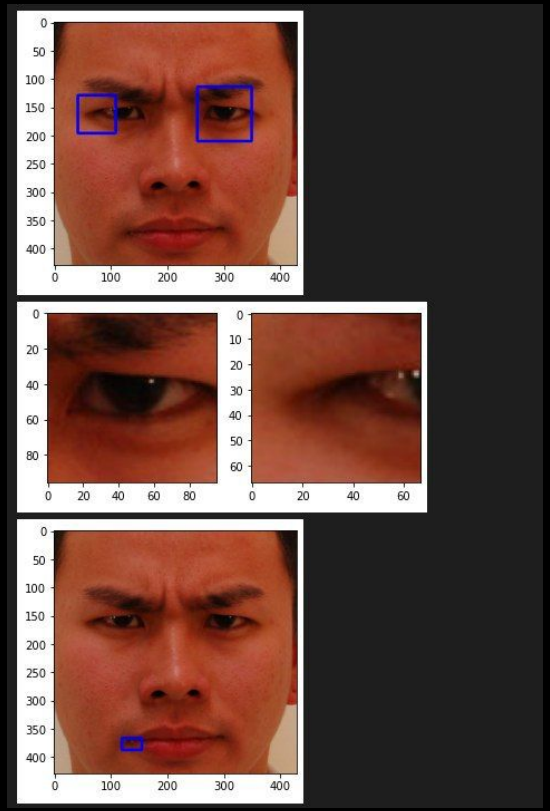


LBP + SVC: 40.47%



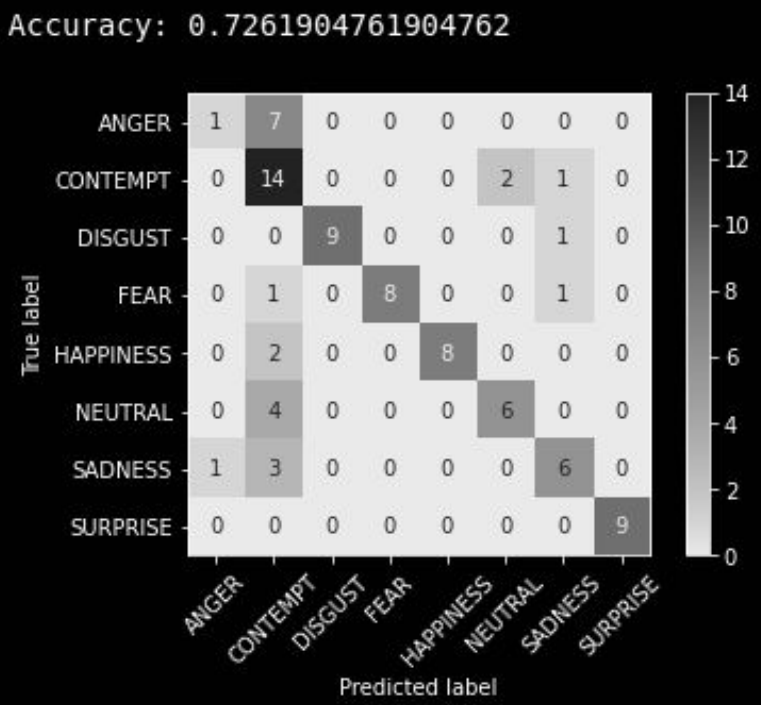
First try with Viola



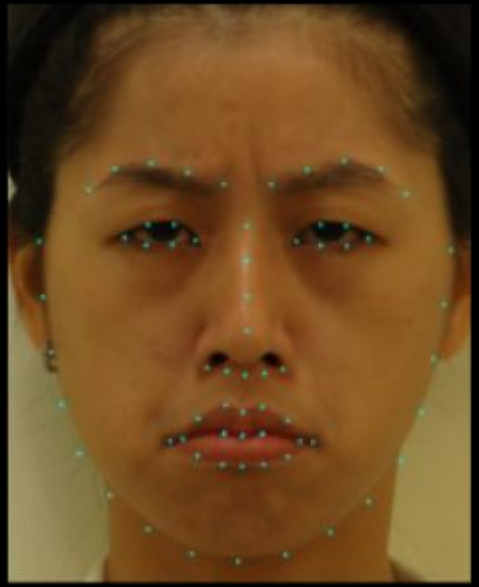


Features as Images

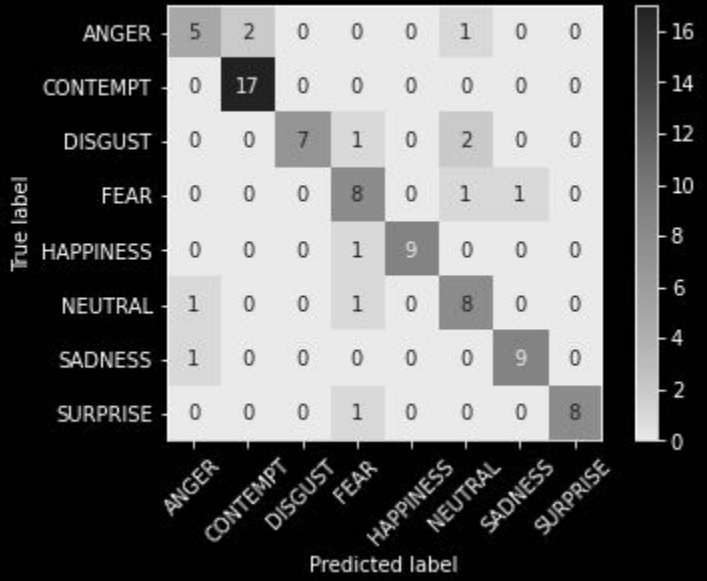
Features as Images



Features as Points



Accuracy: 0.8452380952380952



Results

- FL **points** + AdaBoost + HOG : 84.52 % accuracy on TFEID
- FL **images** + AdaBoost + HOG : 72.61 %

Possible improvements

- **More Feature Extractors**: There may exist better feature extraction algorithms.
- **Better data**: TFEID is still very limited (there are only 40 Asian models) and may introduce bias to the model.

References

- One millisecond face alignment with an ensemble of regression trees
  DOI:10.1109/CVPR.2014.241

- Github: https://github.com/kamilkoduo/cv-fer