FACIAL EXPRESSION RECOGNITION

Rahul Sridhar

CS 216 - Project

June 2017

OUTLINE

- Problem
- Data
- Initial Solution Design
- Final Solution Design
- Implementation, Results and Benchmarks
- Future Work

PROBLEM

- Facial Expression Recognition (FER)
  - Categorize faces based on emotion

Angry Disgust Fear Happy Sad Surprise Neutral

DATA

- FER dataset
  - Source: Kaggle
  - 48x48 grayscale images

	Training	Public Test	Private Test
# Images	28,709	3,589	3,589

  - Class labels: 0-6 (Angry…Neutral)

INITIAL SOLUTION DESIGN

3-stage pipeline:

1. Baseline classifier using either:
    a. Convolutional Neural Networks (CNN), or
    b. K-nearest neighbors (K-NN)
2. Facial expression generation using InfoGAN
3. Feature augmentation to step 1 and build new classifier

FINAL SOLUTION DESIGN

Implemented:

- Gradient Boosted (GB) ensemble using CNN **and** K-NN (on top *n* components from Principal Component Analysis)

Implemented, but unused:

- Facial expression generation using InfoGAN

Not implemented:

- Feature augmentation to step 1 and build new classifier

IMPLEMENTATION, RESULTS AND BENCHMARKS…

Classifier		Prediction Accuracy*		
		Training	Public Test	Private Test
PCA + K-NN		-	37%	37%
CNN		48.5%	47%	48%
GB Ensemble		-	49%	49%
Random		14%		
Human		65%		
State-of-the-art	Softmax	65%		
	SVM loss	-	70%	71%

Ensemble still better than ~25 Kaggle submissions

** Numbers are rounded*

IMPLEMENTATION, RESULTS AND BENCHMARKS…

Correctly classified

Surprise

Happy

Fear

IMPLEMENTATION, RESULTS AND BENCHMARKS…

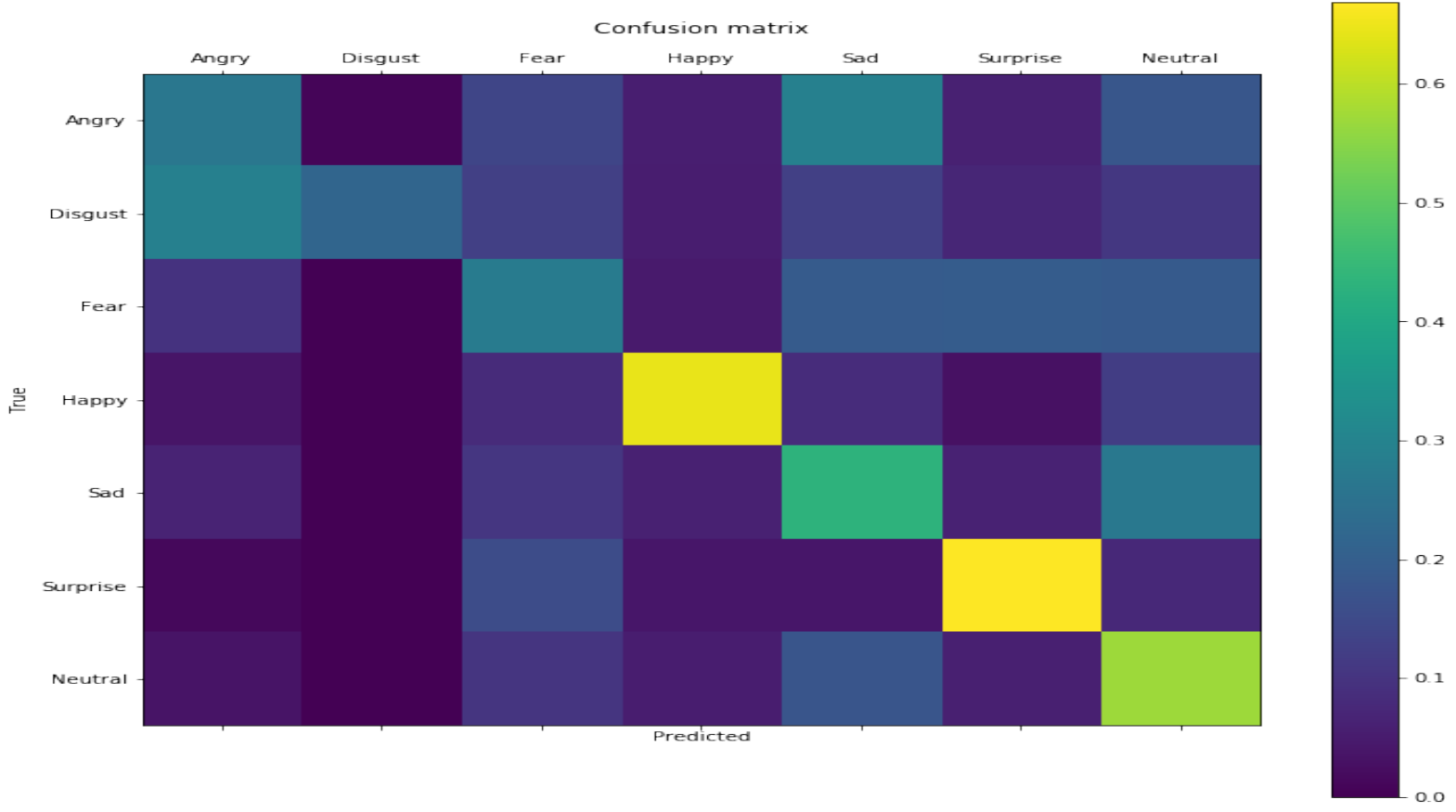
Incorrectly classified - Actual (Predicted)

Sad (Neutral)

Sad (Neutral)

Happy (Sad)

CONFUSION MATRIX



IMPLEMENTATION, RESULTS AND BENCHMARKS

Implementation details:

- Data preprocessing: Standardization + Mean Centering
- CNN:
  - #Layers; #Neurons; Filter sizes; Activation; Dropout; Pooling; Regularization; Batch normalization; Horizontal flips; #Iterations; Weight initialization; Batch size; Optimizer
- PCA + K-NN:
  - Components; Number of neighbors; Distance function
- GB Ensemble:
  - # Trees; Maximum depth

IMPLEMENTATION, RESULTS AND BENCHMARKS

Technology:

- Python (Jupyter notebook)
  - TensorFlow (not very fast, but highly flexible)
  - MXNet (fast, but not very flexible)
  - scikit-learn
  - GraphLab

- MacBook Air (i5 processor, 8 GB RAM)

CHALLENGES AND FUTURE WORK

- Computationally expensive models
  - Better hardware (GPU)
- Test two hypotheses with the state-of-the-art models:
  1. Feature augmentation using InfoGAN
  2. Ensemble of classifiers

REFERENCES

1. Kaggle - "*Challenges in Representation Learning: Facial Expression Recognition Challenge*", 2013
2. Chen et al., "*InfoGAN: Interpretable Representation Learning by Information Maximizing Generative Adversarial Nets*", 2016
3. Y. Tang, "*Deep Learning using Linear Support Vector Machines*", 2013
4. Goodfellow et al., "*Challenges in Representation Learning: A report on three machine learning contests*", 2013

THANK YOU