Facial Emotion Recognition using Ensembles of Deep Neural Networks

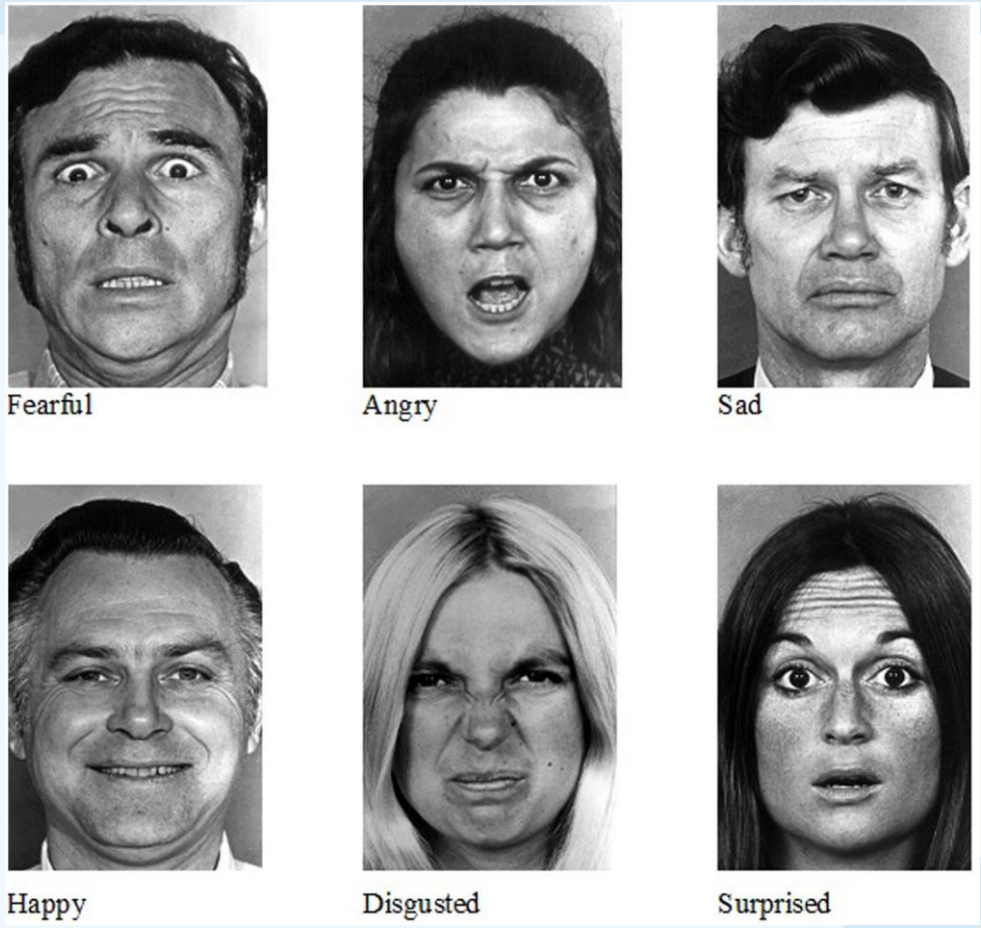
Deep Learning in Practice with Python and LUA
VITMAV45

Csaba Bokányi
Tamás Mészégető

05.01.2021.

Project and motivation

- A team of Deep Learning rookies
- A good start:
  - Classification
  - Challenging for humans
  - Relatively small models



Common FER workflow

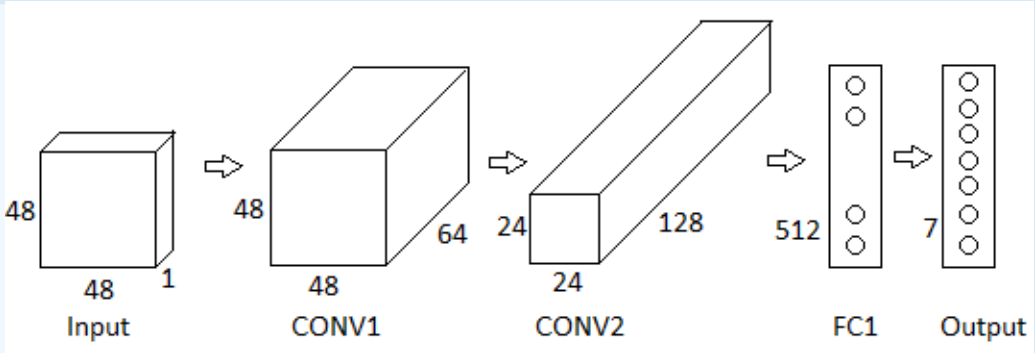
- Preprocessing:
  - Create more data, enhance variance
  - Eliminate certain variances by normalizing pictures (pose/illumination etc.)
- Network design
  - All common achitectures (CNN, DBN, RNN, GAN, etc.)
  - Usually shallower networks as in other image processing tasks
  - Lately many deep CNNs (VGG, Inception, ResNet)
- Decision making
  - Ensembles of a few/thousands of different networks

The FER 2013 database

- ~35 000 images
- Train, valid, test split
- 48x48, grayscale
- 7 emotions (including neutral)
- One of the main static FER databases

Trained Models 1

- Shallow (figure), deep model
- 2(4) CNN, 1(2) FC layer
- Batch normalization, Dropout, Early stopping
- CNN: Max pooling
- Test accuracies: 0.606 (S), 0.665 (D)
- Tuning:
  - Learning rate (0.01)
  - Weight decay (1e-7)
  - Dropout rate (0.3)



Trained Models 2

- Agrawal et al. (2020)
- Overall accuracy: 63.36%

Input data (64 × 64) grayscale image
Data augmentation
CONV 8 × 8 × 32,BATCH NORM
CONV 8 × 8 × 32, RELU, STRIDE (2 × 2)
CONV 8 × 8 × 32, BATCH NORM
CONV 8 × 8 × 32, RELU, STRIDE (2 × 2)
CONV 8 × 8 × 32, BATCH NORM
CONV 8 × 8 × 32, RELU, STRIDE (2 × 2)
CONV 8 × 8 × 16, BATCH NORM
CONV 8 × 8 × 16, RELU, STRIDE (2 × 2)
CONV 8 × 8 × 16, BATCH NORM
CONV 8 × 8 × 16, RELU, STRIDE (2 × 2)
CONV 8 × 8 × 16, BATCH NORM
CONV 8 × 8 × 16, RELU, STRIDE (2 × 2)
CONV 8 × 8 × 8, BATCH NORM
CONV 8 × 8 × 8, RELU, STRIDE (2 × 2)
CONV 8 × 8 × 8, BATCH NORM
CONV 7 × 7 × 7, RELU, STRIDE (1 × 1)
SOFTMAX

Ensembles

- Decision level ensembles
  - Simple averaging
  - Overall accuracy weighted averaging
  - Categorical cofidence weighted averaging
- A total of 9 models, accuracies between 55-66.48%
- Differences in the order of most confident classes

- Best accuracy weighted ensemble: 69.01%

Summary

- Reconstructed recent works and reached over human-level accuracies
- Combined different architectures and showed that even small ensembles can yield significant performance improvements

Thank you for watching!

Csaba Bokányi
Tamás Mészégető