Recognizing Facial Expressions Using Deep Learning

Alexandru Savoiu (savoiu), James Wong (jwhwong)
Department of Computer Science, Stanford University

Objective

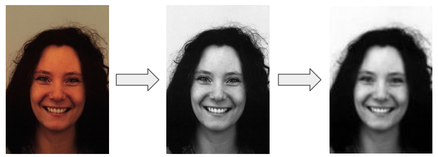
Use various deep learning models and techniques to identify the 7 key main human emotions: anger, disgust, fear, happiness, sadness, surprise, and neutrality.

Data

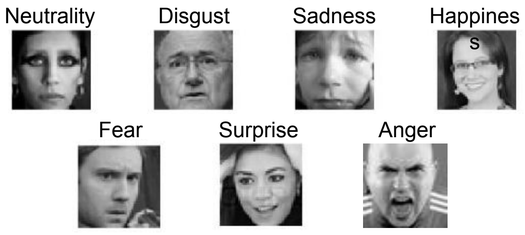
Databases

- Kaggle (Facial Expression Recognition Challenge)
  - 35,887 images
  - 48 x 48 pixels (8-bit grayscale)
  - Various individuals at various angles
- Karolinska Directed Emotional Faces (KDEF)
  - 4900 images
  - 562 x 762 pixels (32-bit RGB)
  - 70 individuals, each displaying 7 different emotional expressions, and each expression is photographed twice from 5 different angles
- Both databases store images representative of the human species, with various age, race, ethnicity, gender, etc

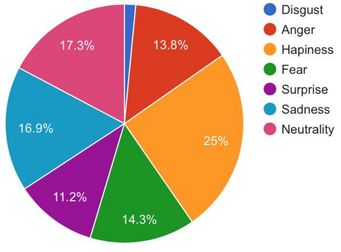
Data Processing: Mean centering, colorimetric (luminance-preserving) conversion to grayscale and resizing (KDEF)



Emotions Represented: anger, disgust, fear, happiness, sadness, surprise, and neutrality



Emotions Distribution



Methodology

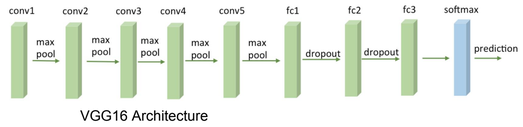
SVM (Baseline)

$$f(x_i, W, b) = Wx_i + b$$

$$L_i = \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + \Delta)$$

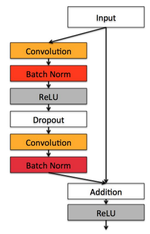
$$L = \frac{1}{N} \sum_{i=1}^{N} L_i + \lambda \sum_{j=1}^{C} \sum_{k=1}^{D} W_{j,k}^2$$

VGG-16

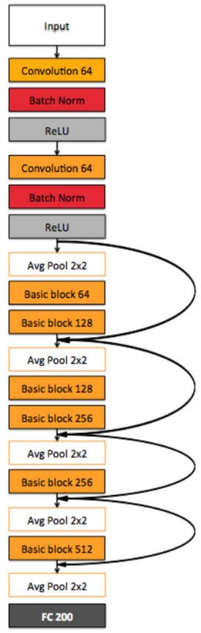


VGG16 Architecture

ResNet50

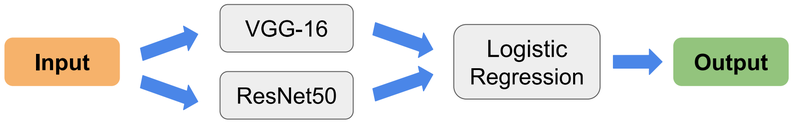


ResNet- Basic Bloc with a dropout to reduce overfitting



ResNet- Sample Architecture

Ensemble Learning



Transfer Learning

Results (overall comparison)

Kaggle Dataset

	Accuracy	Precision	Recall
SVM (baseline)	31.8%	43.7%	54.2%
VGG-16	59.2%	70.1%	69.5%
ResNet	65.1%	76.5%	74.8%
Ensemble	67.2%	79.4%	78.2%

Results (overall comparison)

Karolinska Directed Emotional Faces (KDEF)

	Accuracy	Precision	Recall
SVM (baseline)	37.9%	50.1%	54.9%
VGG-16	71.4%	81.9%	79.4%
ResNet	73.8%	83.3%	80.7%
Ensemble	75.8%	85.0%	82.3%

Transfer Learning (Kaggle -> KDEF)

	Accuracy	Precision	Recall
SVM (baseline)	*37.9%*	*50.1%*	*54.9%*
VGG-16	73.6%	84.2%	81.1%
ResNet	76.0%	86.1%	82.5%
Ensemble	78.3%	87.3%	84.3%

Results (by emotion)

Kaggle dataset, ensemble learning