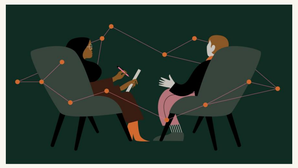
EmoVision: A Facial Emotion Recognition Model

Ashley Chon, Christine Han, Jean Yoo

Motivation

- AI systems with emotional intelligence



Automatic facial expression analysis can facilitate sociable robotics, driver fatigue surveillance, and assistive technologies for people with ASD (autism spectrum disorder).

Problem

Emotion Classification
Predict the emotion conveyed by a facial expression.

FER-2013 Dataset
Train on 48x48 grayscale images of facial expressions, experimenting with different pre-processing methods.

Goal

1. Given an image of a facial expression, accurately predict the emotion conveyed by it.
2. Achieve around 55% accuracy (human performance on FER-2013 is estimated to be around 65.5% [1]).

Insights from Experimentation

Fine-Tuning with VGG

- Direct training of deep networks on relatively small datasets tend to overfit.
- Fine-tuning on pre-trained models like VGG can help mitigate this problem.

VGG architecture from exiting literature which achieved 84% accuracy on the FER+ dataset.

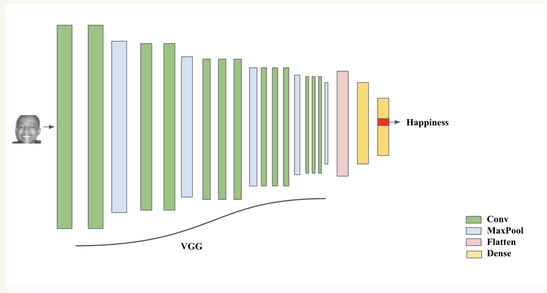
Data Pre-Processing – Cropping around the Eyes

- Eyes and eyebrows are often the most important parts of communicating emotions.
- Experimented with cropping each input image and using only the eyes to train the model:

Example training image after cropping.

- Led to only marginally better performance than baseline

Final Model Architecture



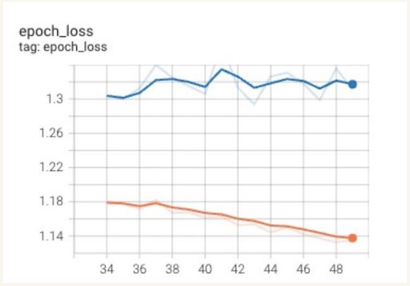
- 5 blocks of convolutional layers, each followed by a max pooling layer
- Filter size of 3x3, number of filters increase from 64 to 512
- Feedforward network that includes a dropout layer, fully connected layer with ReLU activation, and a final output layer with a softmax activation.

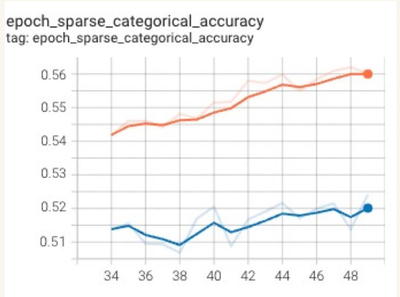
Results

Model	Test Set Accuracy
Baseline Model	24.71%
Eye Cropping Model	27.47%
Draft VGG Model	41.54%
Final VGG Model	52.08%

Table 1. Test Set Accuracy of Different Models.

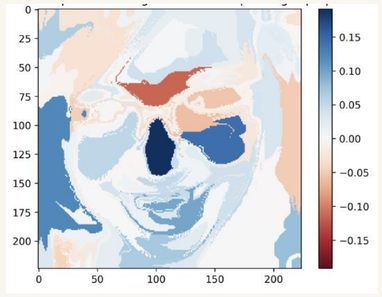
More Results

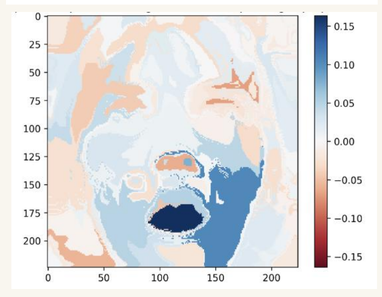




Left: Loss and accuracy graphs over epochs (for epochs 34-50). Training (blue) and testing (orange).

Right: LIME explainer images for "happy" and "surprised".





References

[1] I. J. Goodfellow et al., "Challenges in representation learning: A report on three machine learning contests," Neural Networks, vol. 64, 2015.
[2] M. Zhang K. Liu and Z. Pan. Facial expression recognition with cnn ensemble. Proceedings - 2016 International Conference on Cyberworlds, CW 2016.
[3] Shan Li and Weihong Deng. Deep facial expression recognition: A survey. IEEE Transactions on Affective Computing, vol. 13, no. 3, pp. 1195-1215, 2022.
[4] Rodolfo Pavez, Jaime Diaze, Jeferson Arango-Lopez, Danay Ahumada, Carolina Mendez-Sandoval, and Fernando Moreiera. Emo-mirror: a proposal to support emotion recognition in children with autism spectrum disorders. Neural Comput Applic 35, 7913-7924, 2023.