EmoRec: An iOS Mobile App for Emotion Detection and Gamification to Improve Social Skills

Katherine Chen Flora Huang

Department of Computer Science, Stanford University

We developed ***EmoRec***, an iOS app that **recognizes human emotions** from user-provided images and helps users **practice** expressing specific emotions, designed for people with autism, social anxiety, or related disorders.

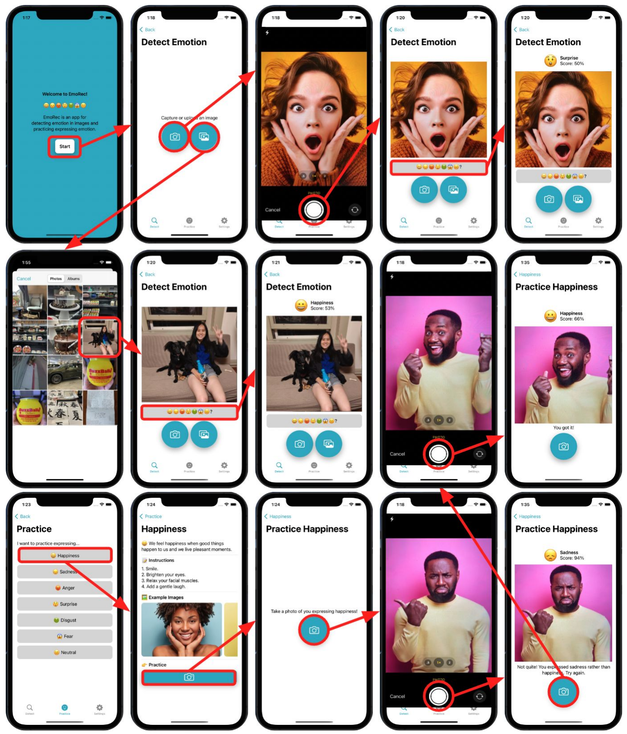

Figure 1. EmoRec UI (detection mode & practice mode).

Background

- Emotions are vital for social interaction, but individuals with autism, social anxiety, etc. struggle with recognition and expression
- Existing solutions have limitations in accurately detecting emotions
- EmoRec is designed to overcome these limitations and provide an effective tool for emotion recognition & expression practice

Problem Statement

- *EmoRec* aims to recognize human emotions from provided images
- **Inputs**: Images from the user's camera or photo library
- **Outputs**: Emotion categories (e.g., happy, sad, angry)
- **Evaluation metrics**: Accuracy, precision, recall, and F1 score

Methods

8 Layer AlexNet (Krizhevsky et al.)

- Proven effective in general image classification tasks
- **5 conv layers** (with max pooling & batch normalization) + **3 FC layers**
- **Higher** storage requirements

Data Augmentation

- Made up for imbalance in FER-2013
- **Horizontal flips** and **translation**

Figure 2. Left: original FER- 2013 image. Right: augmented image.

Regularization

- **Early Stopping**: Stopped training before our model overfit on our dataset
- **L2 Regularization**: Penalty term in the model's loss function kept weights small
- **Dropout**: Enhanced robust feature learning by varying neuron reliance

Hyperparameter Tuning

- Grid Search for **Learning Rates (lr)** and **Batch Sizes (bs)**
- Further tuned for the **L2 Regularization Constant (λ)**

Mobile App Integration

- Used **Core ML** to make TensorFlow model compatible with iOS devices
- Used **Xcode** and **Swift** to build the UI and implement the functionality of the app

3 Layer DenseNet (Huang et al.)

- Addresses vanishing gradients and encourages feature reuse
- Each layer receives feature maps from all preceding layers
- **3 dense blocks** with dropout layers in between
- **Lower** storage requirements

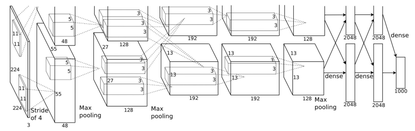

Figure 3. AlexNet model architecture.

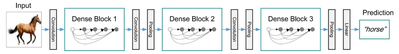

Figure 4. DenseNet model architecture.

Dataset

- **FER-2013**: 35,887 facial images labeled with 7 emotion categories
- **Preprocessing**: Resizing, normalization

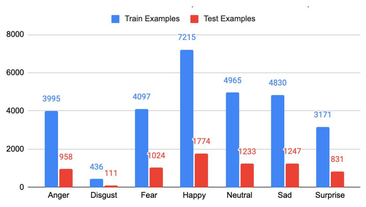

Figure 5. Distribution of emotions in the FER-2013 dataset.

Figure 6. Images from the FER-2013 dataset, labeled by emotion category

Experiments & Analysis

Data Augmentation

- Did **not** improve accuracy
- Introduced too much **variation**
- Our DenseNet model (~2 million parameters) may be **too shallow**

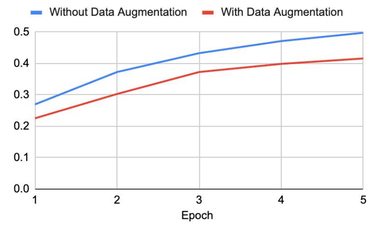

Figure 7. Training accuracies with & without augmentation (lr=0.005, bs=32).

Hyperparameter Tuning

- **Results**: lr=0.001, bs=64, λ = 0.01

bs \ lr	0.001	0.005	0.01
16	0.2513	0.5711	0.2513
32	0.6452	0.6065	0.6225
64	0.6570	0.6361	0.6445

λ	Training Accuracy
0.005	0.6298
0.01	0.6361
0.05	0.6107

Figure 8. Left: Training accuracies for various learning rates, batch sizes, and L2 regularization constants.

Mobile App Testing

- Skew towards happiness
- Visually appealing, intuitive to use
- **Detection mode**: confusion due to inaccuracies in emotion predictions
- **Practice mode**: engaging & helpful in improving expression skills

Final Model

- **72.81% accuracy** on test set
- Good at **happiness & surprise**
- Bad at **disgust & fear**
- **Imbalanced** dataset
- **Subjectivity** of emotions → difficult to differentiate

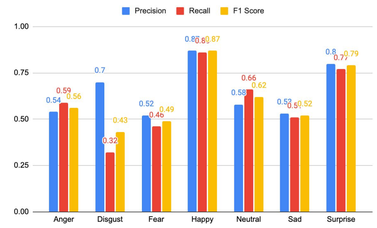

Figure 9. Performance metrics.

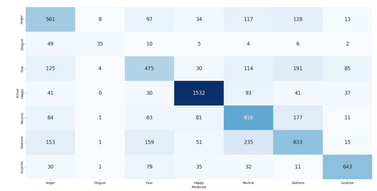

Figure 10. Confusion matrix.

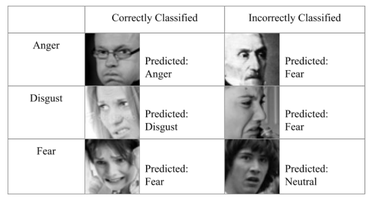

Figure 11. Misclassified examples.

Conclusion & Future Work

- *EmoRec* is effective for emotion recognition and expression practice
- Underlying DenseNet model achieved **72.81% accuracy** on our test set
- Future work:
  - Exploring more **CNN architectures**, such as VGGNet, ResNet, or InceptionNet
  - Augmenting FER-2013 with **larger, more diverse facial emotion datasets**
  - User study assessing EmoRec's effectiveness for our **target users**