**Motivation**

Emotion detection within a mobile platform may be used by content developers to provide a more fulfilling user experience, such as, updating an interface and/or experience in real-time based on user emotional feedback.

**Objective**

To create a facial expression recognition (FER) system that can accurately and efficiently process depth sensor data from a smartphone in order to elicit users' emotional state, specifically the six Ekman emotions.

**Methods**

1. Facial Expression Recognition
2. Emotion Recognition
3. Convert to Mobile Platform

![Fig. 1: Flowchart of system](Image)

1. **Facial Expression Recognition**

An iPhone X smartphone, which has an integrated depth sensor, was used to capture the facial expressions of users. The iOS SDK was utilized to access a facial geometry mesh generated from the front sensors. Using the BU-3DFE database as reference, 83 landmark vertices were selected from the mesh to act as the feature points necessary to process an emotion.

![Fig. 2: 83 Facial landmarks used to detect facial expressions (left) and the facial mesh provided by iOS SDK (center).](Image)

2. **Emotion Recognition**

A feed-forward NN was developed with TensorFlow (TF) and trained on data consisting of 2400 labeled facial data, each containing 83 vertex (X, Y, Z) landmarks. The model was defined by a 249-feature (83 x 3 axes) input layer and two hidden layers, consisting of 127 nodes each, excluding biases.

**NN Hyper-Parameters:**
- Adam optimizer
- ReLu activation (hidden layers)
- Softmax activation (output layer)
- Cross-entropy loss function
- 10,000 epochs for minimum loss

![Fig. 3: NN architecture](Image)

3. **Mobile Platform**

The NN model was converted to a CoreML model using built-in TF tools. This allowed the system to run on an Apple device, specifically the iPhone X, while utilizing the integrated depth sensor within the device.

![Fig. 4: iOS application that recognized emotions using 83 identified landmarks](Image)

**Conclusions**

A mobile application has been created that utilizes 83 facial landmarks as inputs to a NN which identifies user emotional state. The application can consistently determine four of the Ekman emotions: happy, angry, disgust and surprise. The system has difficulty capturing sadness and fearful. This limitation is due to the similarities between these emotions’ training landmarks. Moreover, because the landmarks were manually selected, there is opportunity to increase the accuracy of the overall system.

**Future Work**

Increasing overall accuracy of the NN Model would allow this system to be effectively used in enterprise software. This may be achieved through additional hyper-parameter configuration testing while examining the effect of each change on the overall system. Furthermore, the NN model could be converted to a predictive system, in which emotions are determined prior to the visual cue occurring.

**References**

2. Lijun Yin; Xiaozhou Wei; Yi Sun; Jun Wang; Matthew J. Rosato. "A 3D Facial Expression Database For Facial Behavior Research". 7th International Conference on Automatic Face and Gesture Recognition, 10-12 April 2006 S.