Brains In Jars

Machine Learning Term Project

Hemachandra Kolisetty	-14MA20015
Sarath NVP	-14ME10039
Sri Keerthi B	-14IE10007
Nayan Raju Vysyaraju	-14MA20049
Nikitha Kondapalli	-14IE10017
Sri Lekha D	-14EE10014
Manasa Gogineni	-14EC10018
Sharath Chandra Vemula	-14EC10060

Title:
Facial Emotion
Recognition

Problem Statement

The goal of this project is to predict, from the grayscale picture of a person's face, which emotion the facial expression conveys. Our evaluation metric will be the accuracy for the emotions (fraction of correctly classified images). In short,

Input : 48 x 48 grayscale image of a face

Output : Emotion conveyed by facial expression

Methodology

Data Set

- We used a dataset **FER2013** provided by Kaggle website.
- **Type**: csv file comprising label followed by pixel values of grayscale images(48x48).
- **Emotion Categories**: Anger, Disgust, Fear, Happy, Sad, Surprise and Neutral.
- **Training Set**: 28,709 examples   **Validation Set**: 3,589 examples   **Test Set**: 3,589 examples
- **Labels**: 0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral

Pre Processing

- We performed central cropping to the images to get the target width and height(32x32), so that they form a bounding box around the face region.
- We accounted for variation in lighting of the input images by adjusting brightness and contrast to fixed band values.
- We converted the image pixel values to zero mean and unit norm with the per image whitening function.

CNN Architecture

The final architecture retained can be described as follows :

- 9 x 9 Conv (stride 1) - ReLU - 3 x 3 Max Pool (stride 2) - LR_Norm with 64 filters
- 5 x 5 Conv (stride 1) - ReLU - 3 x 3 Max Pool (stride 2) - LR_Norm with 128 filters
- 3 FC hidden layers with ReLU activation
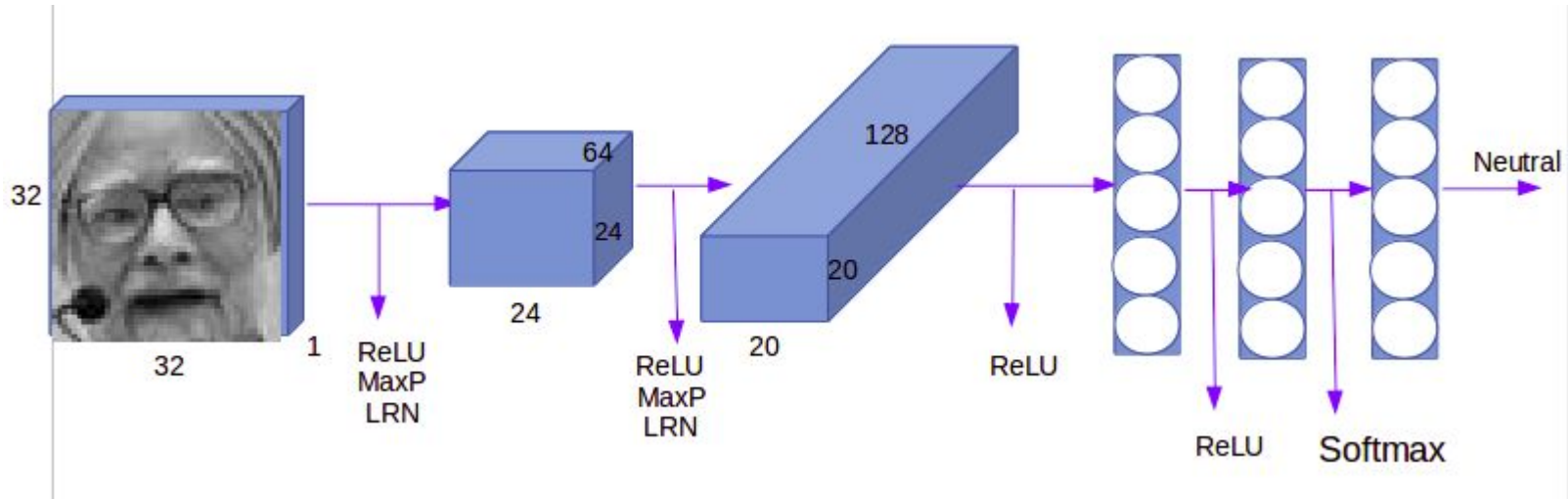
Architecture Characteristics

- Exponentially decayed learning rate, L2 regularization

- Local Response normalization after each layer

- Gradient Descent optimiser

- ReLU nonlinearity activation

- 9x9 convolution filter with stride 1 for first convolutional layer and 5x5 convolution filter with stride 1 for second convolutional layer

- 3x3 max pools with a stride of 2 for both the convolutional layers
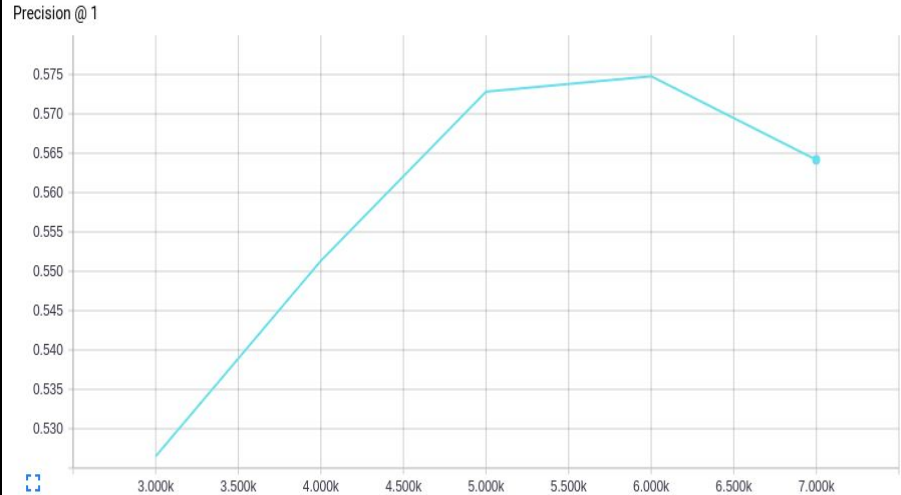
- Softmax cross-entropy to compute cost
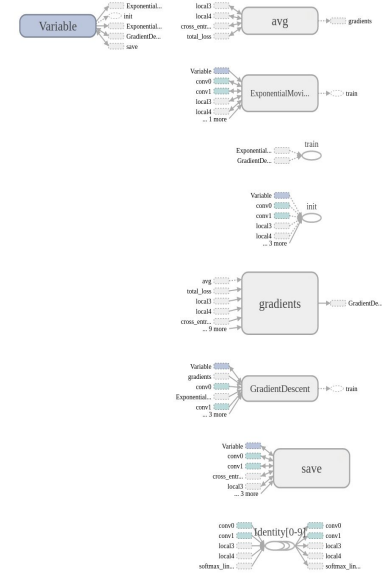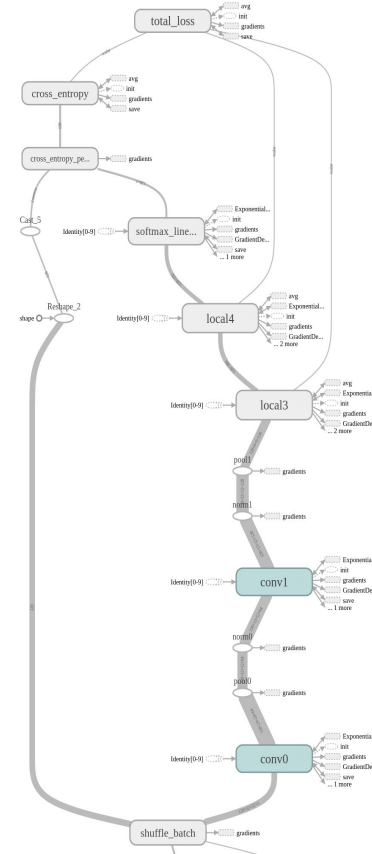
Results

- We achieved final validation accuracy of 57% while the state of the art test accuracy for 7 emotion categories using deep networks is 61%, and the top Kaggle implementation received an accuracy of 71%.

Train Accuracy : 0.997
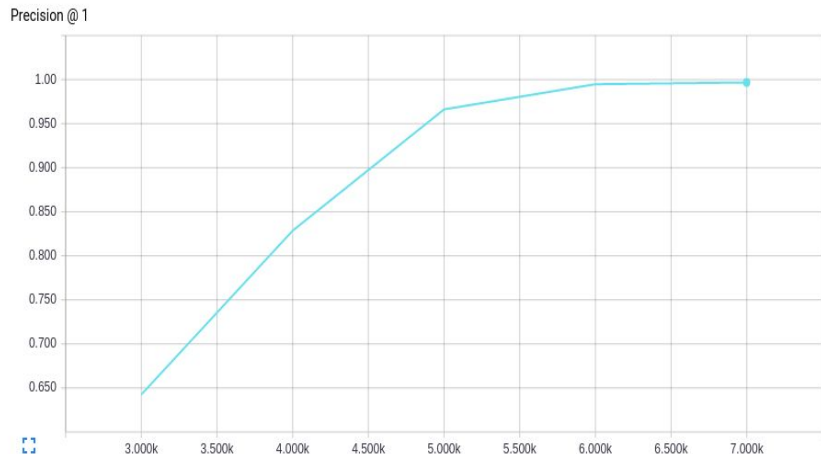
Test Accuracy : 0.57

Loss Function