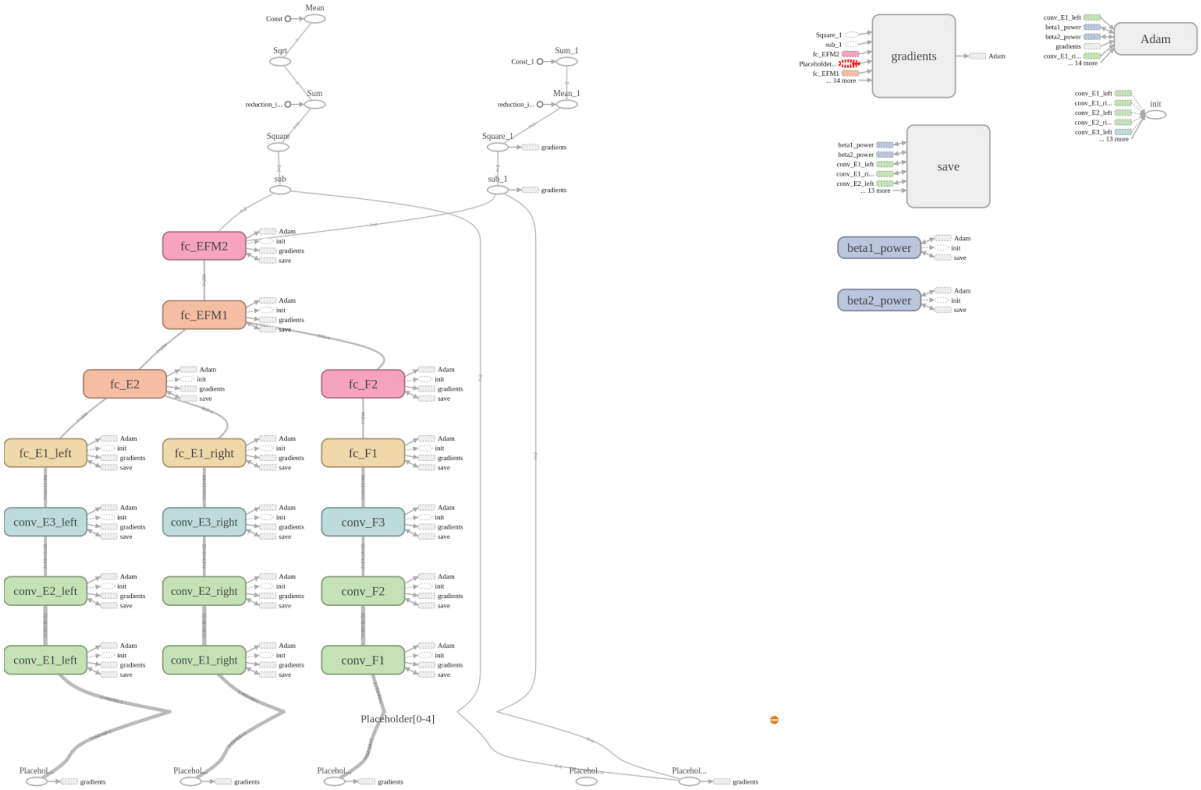
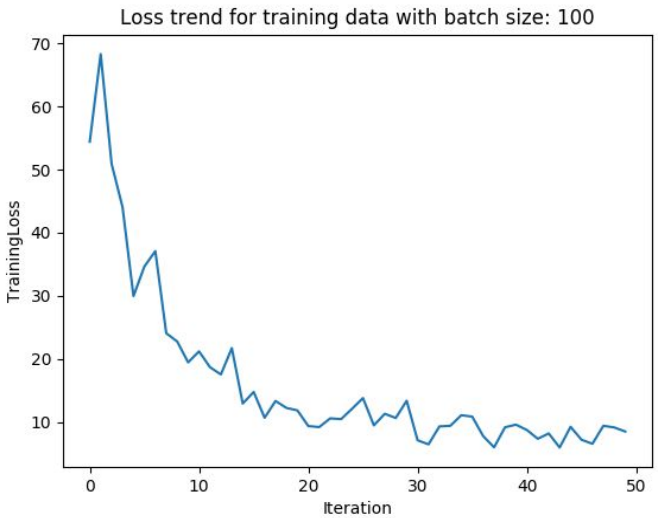
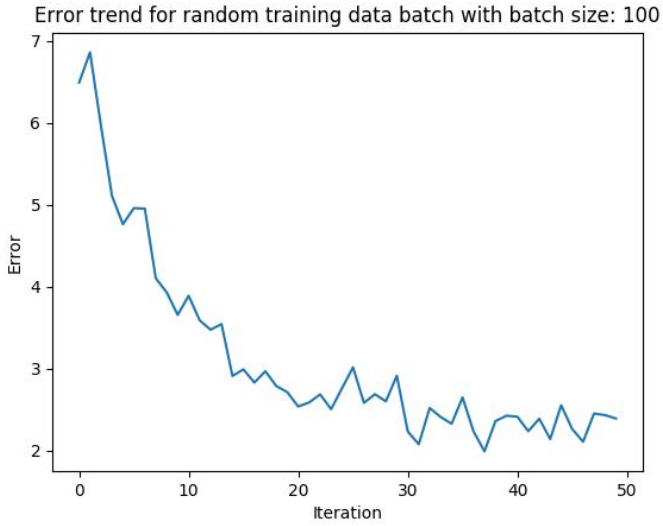
ECSE 6965: Mobile Eye Gaze Estimation with Deep Learning
Yogish Didgi 661538970

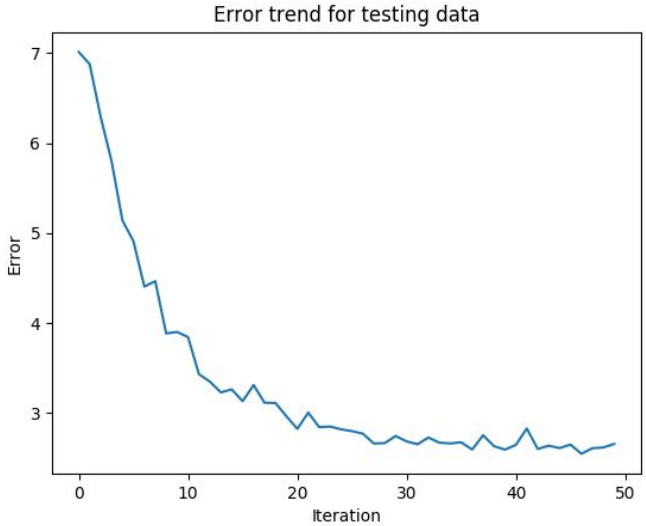
Model architecture:



Loss & Error plots:





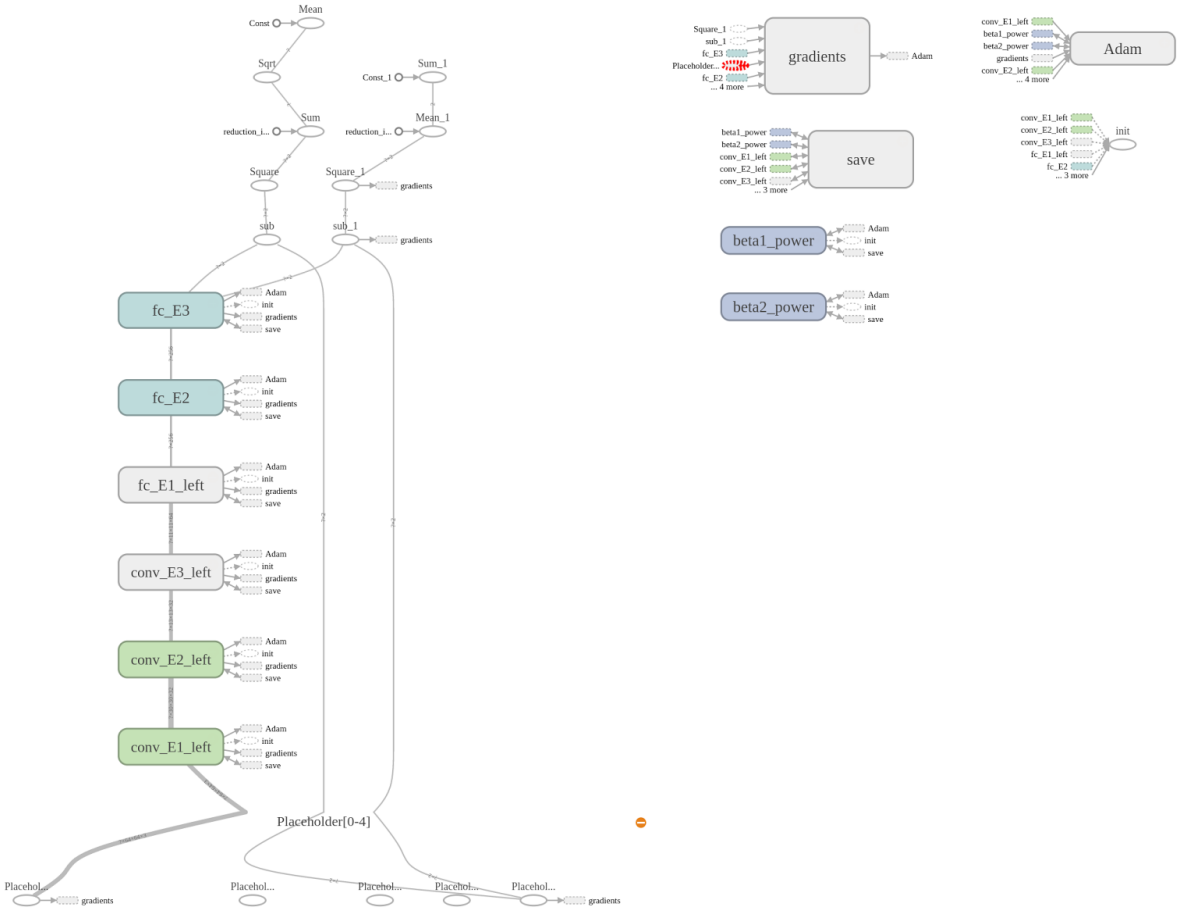


The architecture chosen gave a training error of around 2.5 cm and validation error of around 2.4 cm after 5000 iterations. The design of the architecture was driven mainly based on the original paper. The computation limitations of my laptop also made it difficult to choose a more complex architecture. The model makes use of the left, right eye and face data. The mask information was not used.

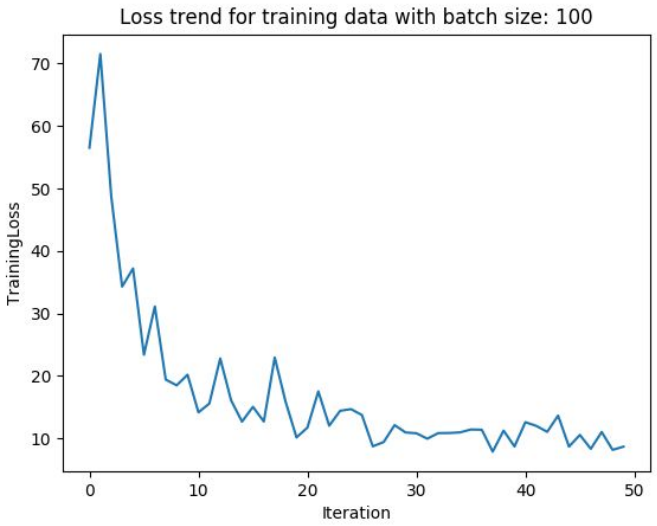
Each of the eyes, face data is passed through three convolutional layers. The first two of those are followed by a max pooling layer. The size is reduced by a factor of 2 in both the cases. The parameters of the filters were also kept simple to reduce the computations. So, the filter size chosen was either 5x5 or 3x3. The number of filters were also limited to 32 or 64. ReLu was used as the activation function. Note that the weights are not shared between the eyes layers.

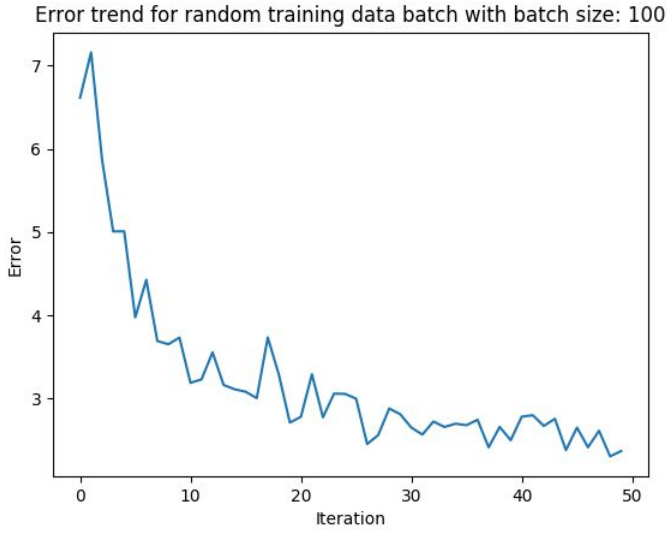
After the convolution, each of the three pathways is passed through 2 layers of fully connected layers with 256 or 512 hidden nodes. Then each of the outputs from the eyes and face is concatenated and passed through two more fully connected layers. The second of those is the final layer with the regression output. The loss and error are variations of mean square error.

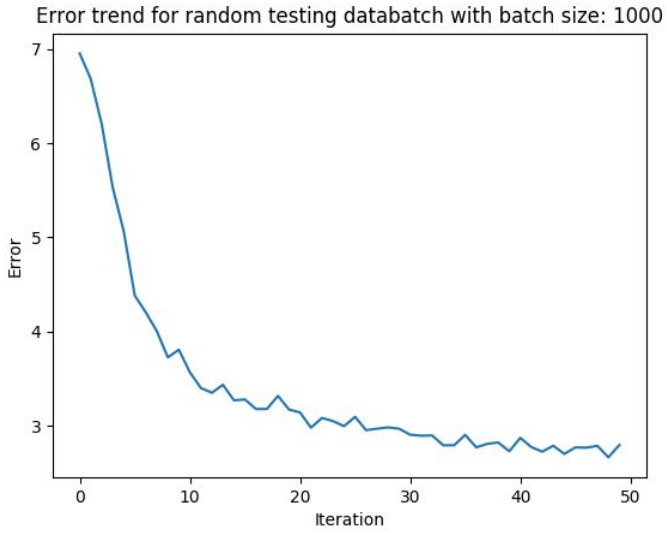
Comparison with a simpler model:



Loss & Error plots:







In the second model, only the left eye data is used to train the network. As can be seen from the plots, the error value also reaches a value of around 2.3cm and 2.7cm for training and validation set respectively for the same number of optimization iterations (5000). We can observe that although training data reduces to a low value, the validation accuracy remains lower than the original architecture. This strengthens the argument that the gaze is not only dependent on the eyes but also on the face orientation.