

Pranesh Navarathna
ECSE 6965 Deep Learning
Final Project
5/10/2017

Overview

For this project, two models were tested for accuracy. Both models input the normalized eye data into a CNN that has 3 convolutional layers and 2 pooling layers. The data was normalized by dividing by 255. The first two convolutional layers have 32 filters each of size 5x5. The last convolutional layer has 64 filters each of size 3x3. After each convolutional layer, there is ReLu activation. Each pooling layer is max pooling of size 2x2 and stride of 2. No padding is implemented. The optimizer used is AdamOptimizer with a learning rate of 0.001. The weights and biases were initialized using Xavier initialization. The loss function used for optimization is the mean Euclidean distance between the predicted coordinates and the labels.

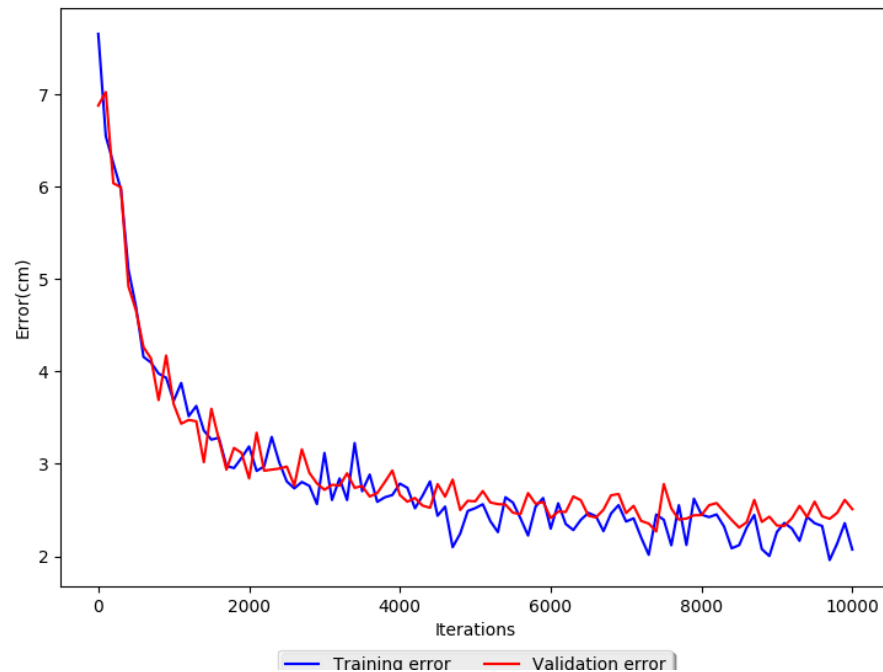
Model 1 uses only the left eye data whereas Model 2 uses both left and right eye data. In Model 2, the outputs of the CNN for left and right eyes are concatenated and fed into a feedforward neural network with 2 ReLu activated hidden layers of 20 nodes each. The output nodes of the feedforward network have linear activation for regression purposes. A batch size of 200 was used for training and validation for model 1. Model 2 uses a batch size of 100. Model 1 achieves approximately 2.3 cm batch validation accuracy whereas Model 2 achieves 1.74 cm batch validation accuracy.

Upon running the validation script, Model 2 achieves approximately 2.05 cm overall validation accuracy.

Model 2 has been saved and submitted for this project.

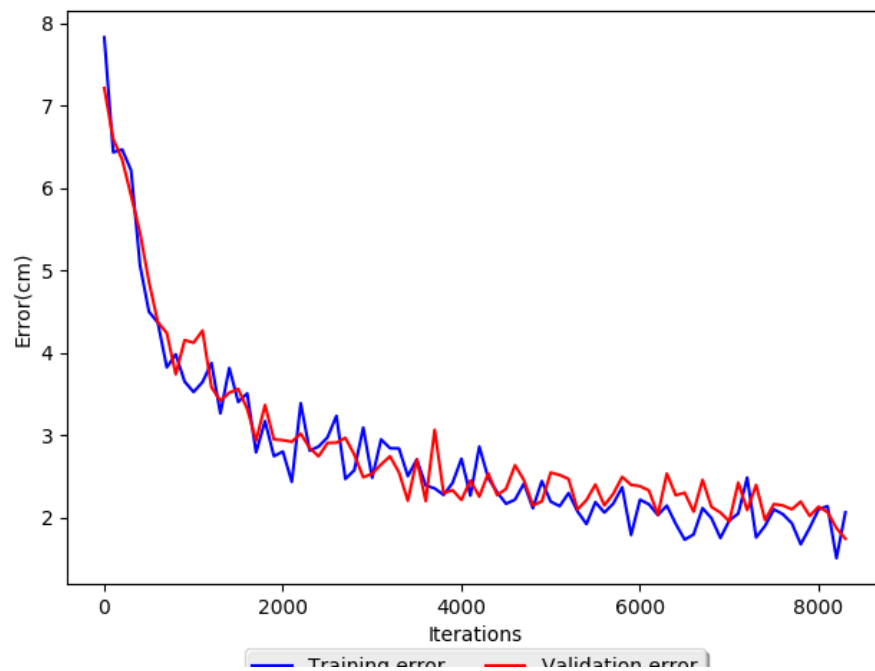
Error Plots

Model 1:



As can be seen in the figure above, model 1 is trained for about 10000 iterations which equates to about 21 epochs. Slight overtraining is observed after about 9000 iterations. The lowest observed error in this model was approximately 2.3 cm.

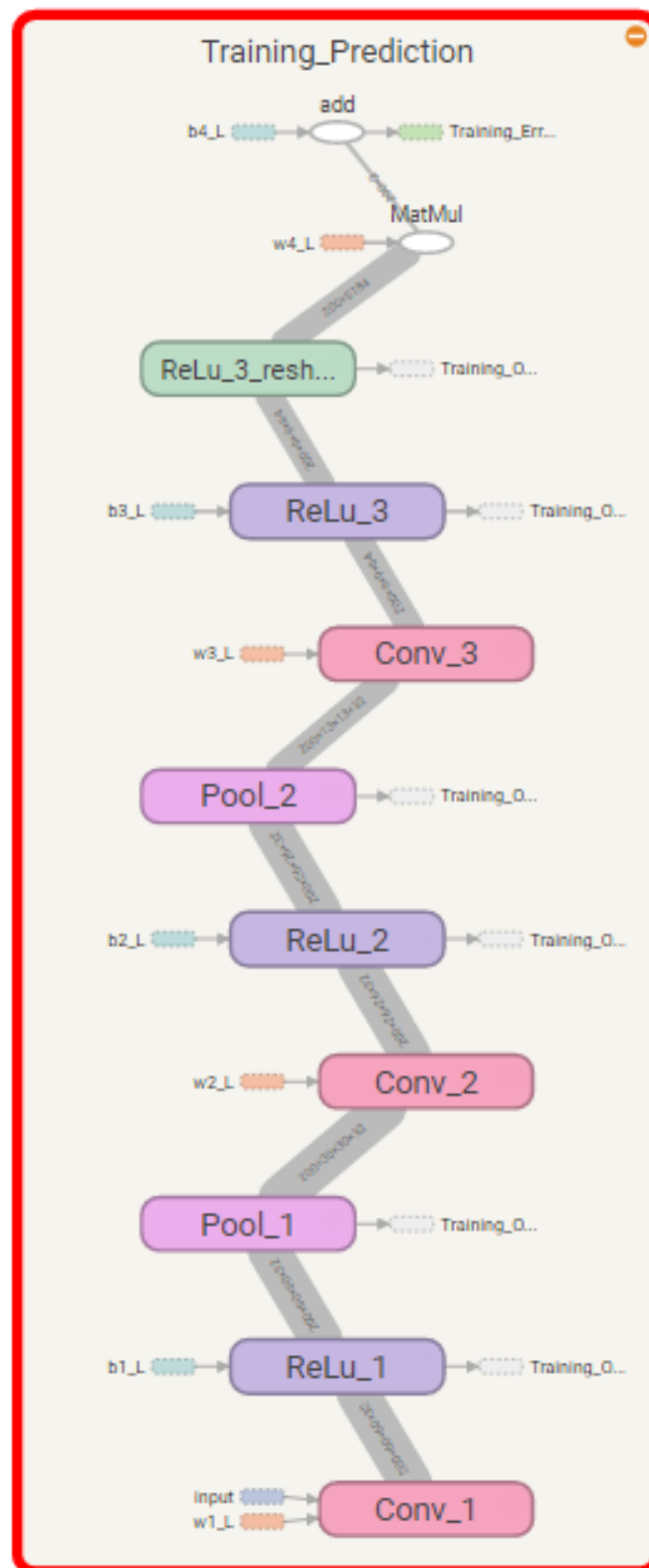
Model 2:



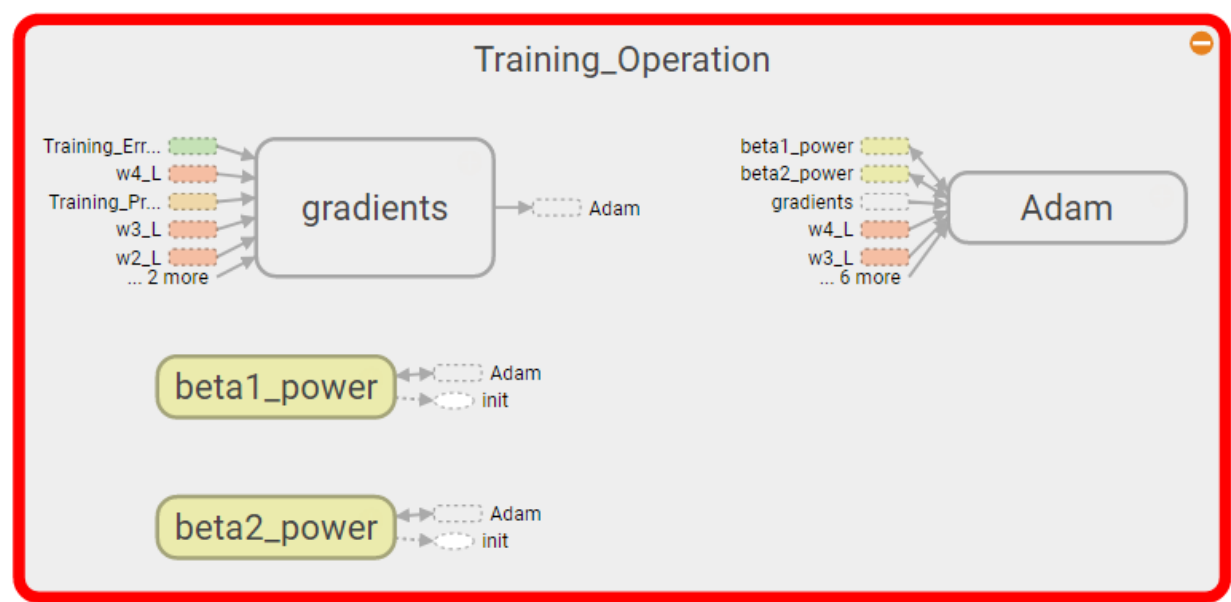
Model 2 is observed to hit 1.74 cm batch validation accuracy in about 8300 iterations or 17.3 epochs.

Tensorboard Visualization – Model 1

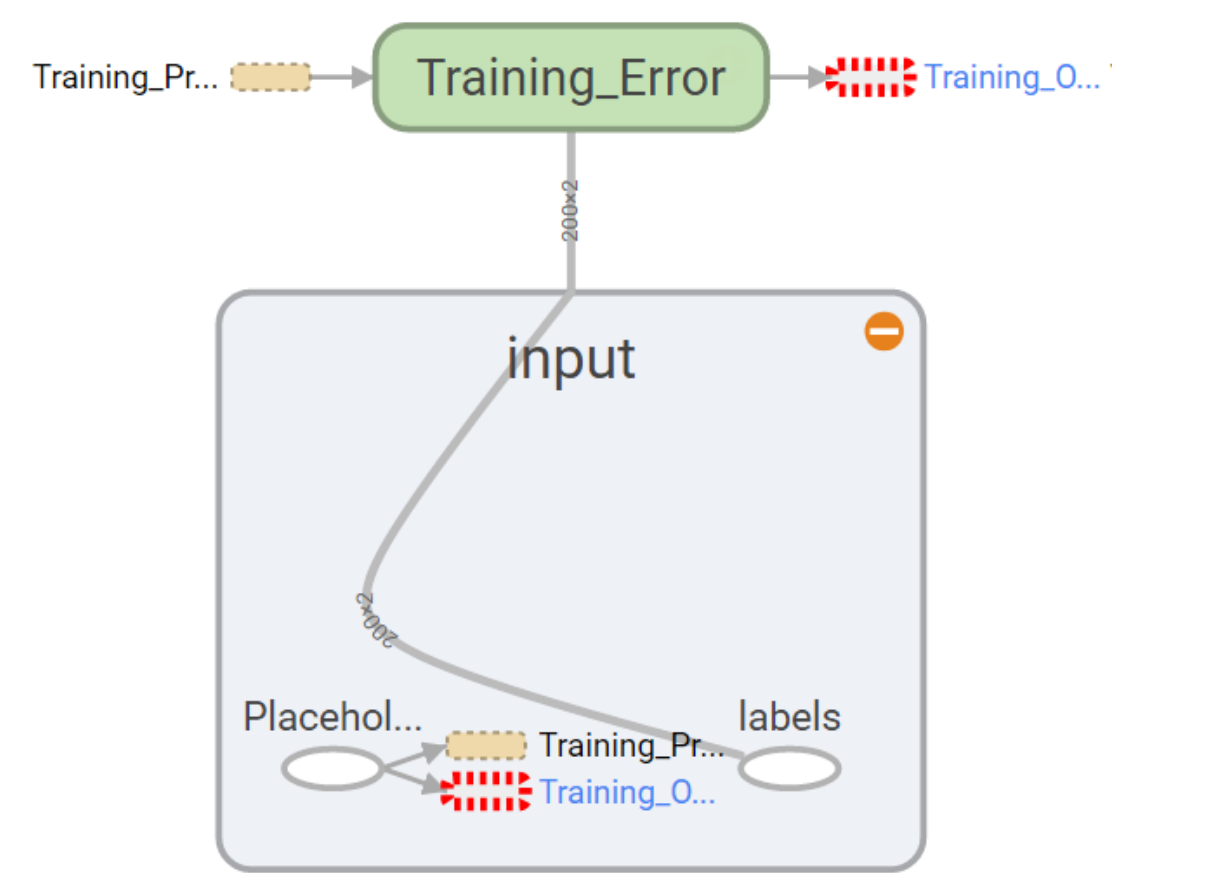
The pathway that the left eye input is fed into is shown below.



Overview of the Training Operation for Model 1:

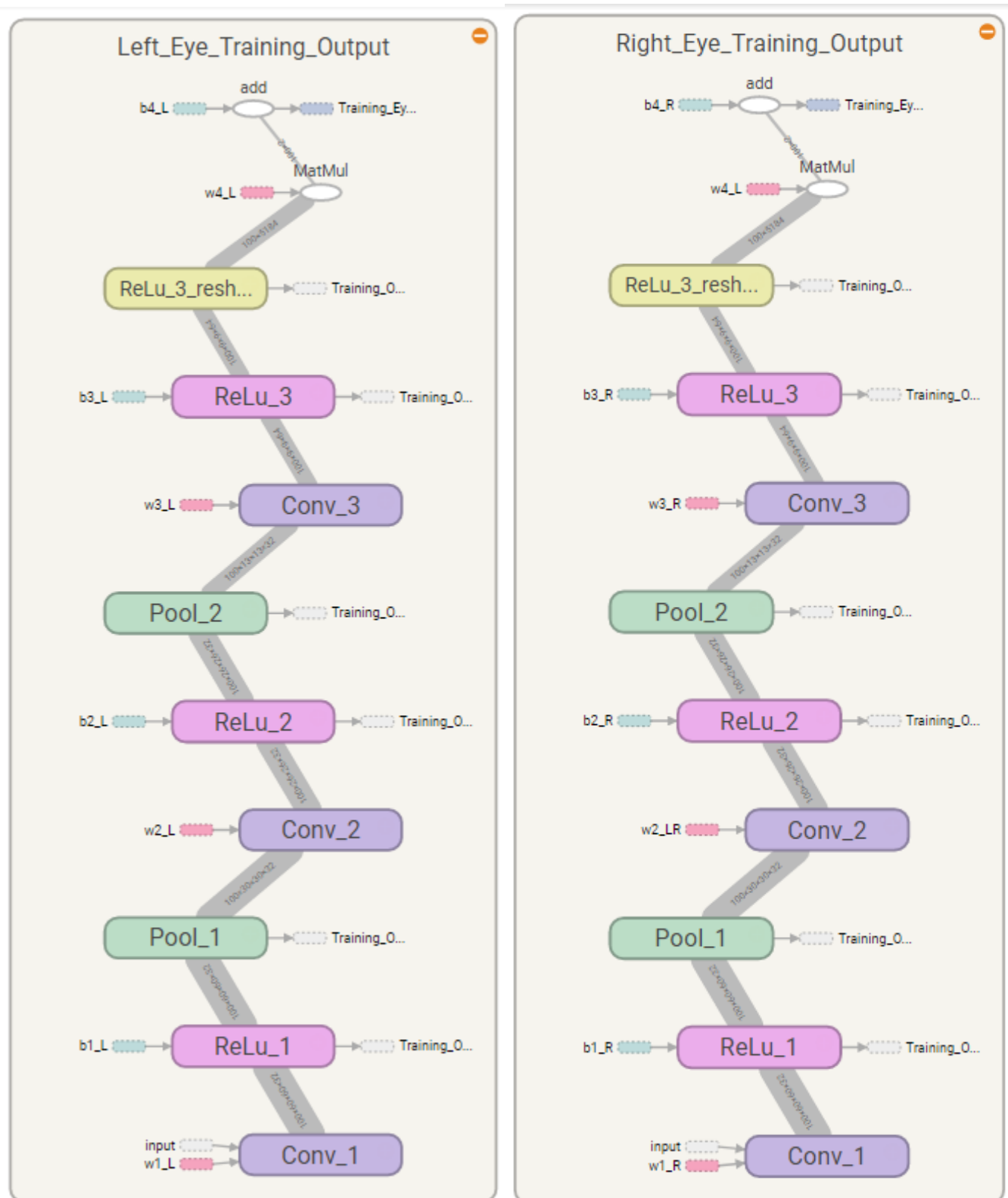


Overall Model Structure [Model 1]:



Tensorboard Visualization – Model 2

The left eye and right eye pathways are shown below.

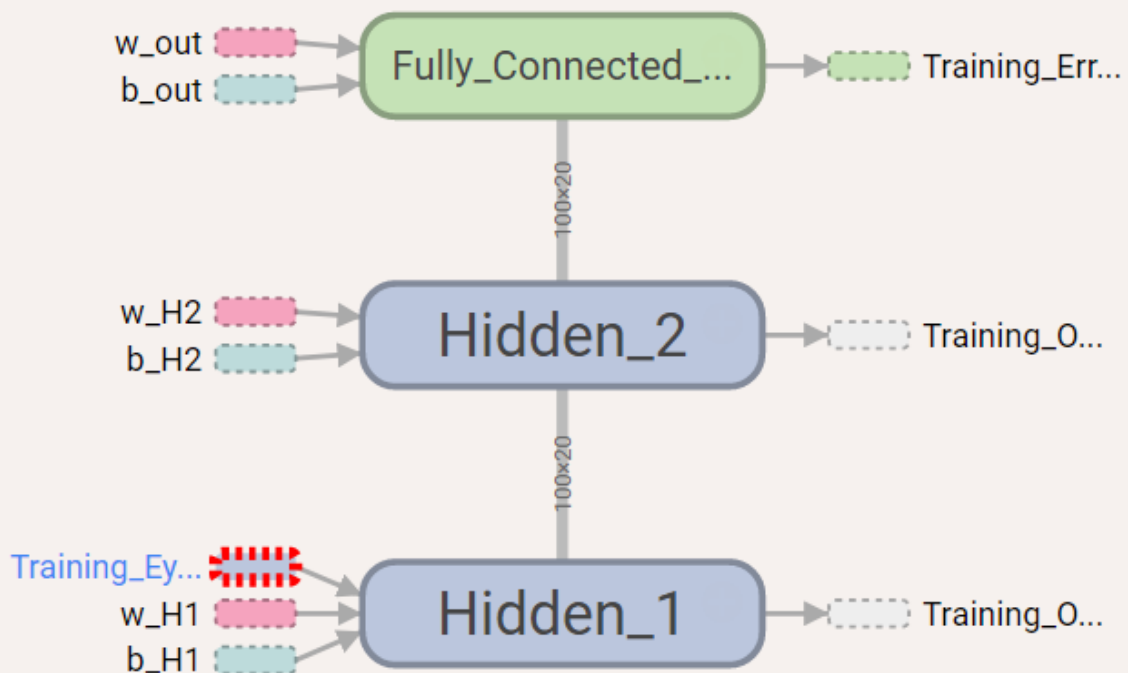


The outputs from the above two pathways is concatenated and the result is fed into a feedforward network with two hidden layers, each with ReLu activations. The output has a linear activation for regression purposes.

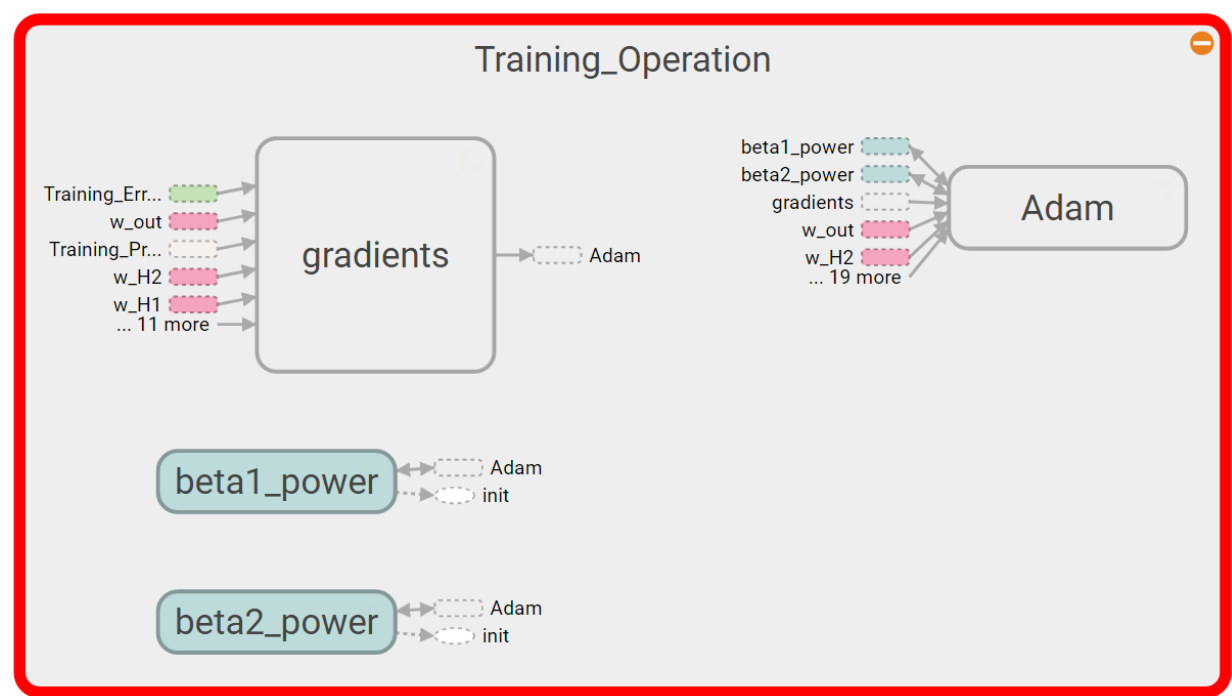
Training_Eyes_Concatenate



Training_Prediction



Overview of the Training Operation for Model 2:



Overall Model Structure [Model 2]:

