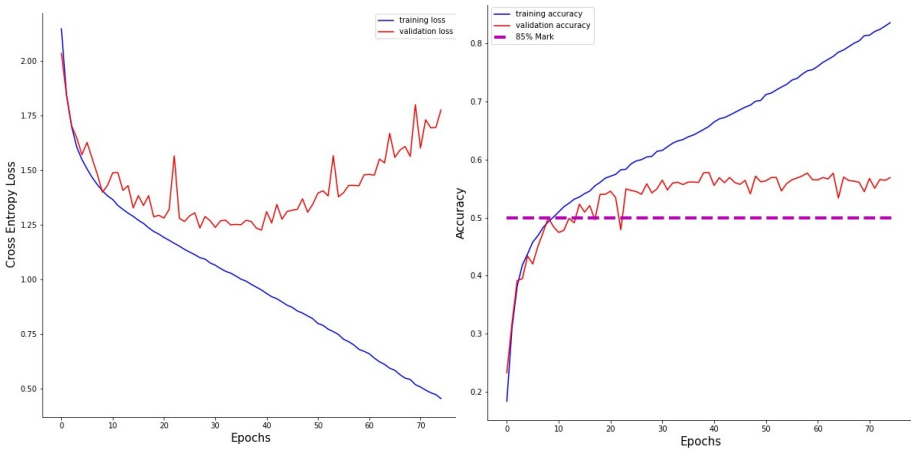
Michael Potter
UID: 705644667
mlpotter124@g.ucla.edu

Fully Connected Network

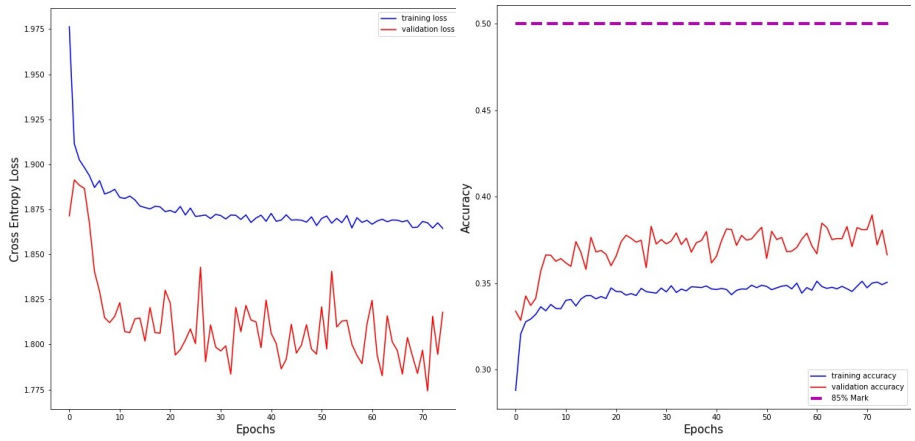
A seven layer fully connected (FC) neural network (NN) with ReLU activation functions (architecture details in Table 2) was trained and evaluated on the CIFAR-10 dataset. The cross-entropy objective function (Figure 1a) was minimized using stochastic gradient descent with momentum (SGD-M) set to 0.9 and a learning rate of 0.05. Early stopping was applied by saving the model at the best cross-validation accuracy during the 75 epochs.



(a) FCNN Cross Entropy (b) FCNN Accuracy

Figure 1: Fully Connected Network Learning Curves with ReLU

The activation functions were removed from the FCNN and the same objective function, optimizer, optimizer settings, early stopping, and number of epochs were used to train the FCNN (Figure 2)

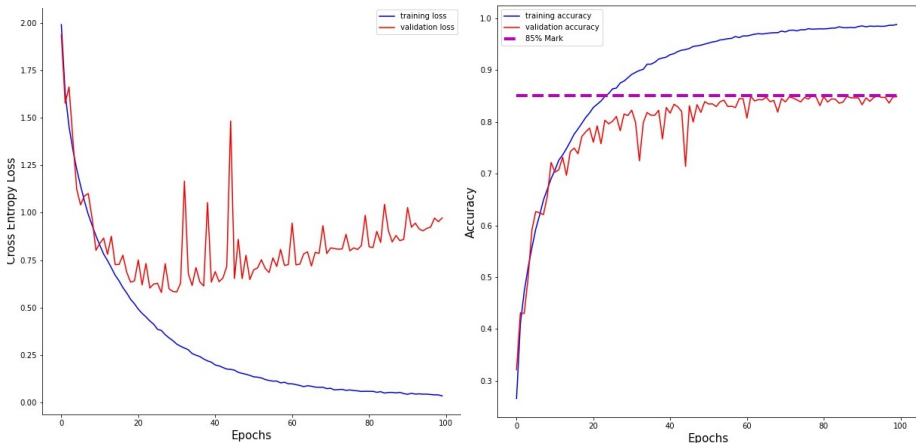


(a) FCNN Cross Entropy (b) FCNN Accuracy

Figure 2: Fully Connected Network Learning Curves without ReLU

Convolution Neural Network

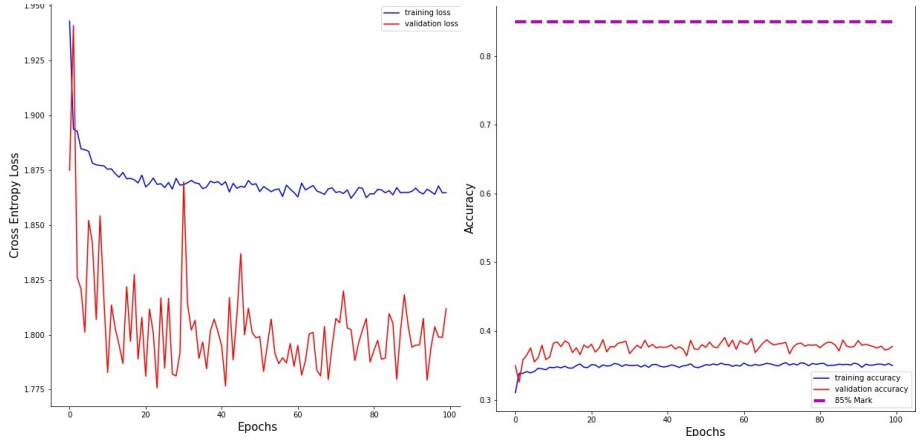
A four layer convolutional neural network (CNN) feature extractor combined with a three layer FCNN with ReLU activation functions (architecture details in Table 2) was trained on the CIFAR-10 dataset. The cross-entropy objective function (Figure 3a) was minimized using stochastic gradient descent with momentum (SGD-M) set to 0.9 and a learning rate of 0.05. Early stopping was applied by saving the model at the best cross-validation accuracy during the 100 epochs.



(a) CNN+FCNN Cross Entropy (b) CNN+FCNN Accuracy

Figure 3: CNN + FCNN learning Curves with ReLU

The activation functions were removed from the FCNN and the same objective function, optimizer, optimizer settings, early stopping, and number of epochs were used to train the FCNN (Figure 4)



(a) CNN+FCNN Cross Entropy (b) CNN+FCNN Accuracy

Figure 4: CNN + FCNN learning Curves without ReLU

Final Test Accuracy

Model Accuracy		
Model	Activations ON	Activations ON
FCNN	57.79%	38.94%
CNN	85.08%	39.04%

Table 1: The Final Test Accuracy for Trained Model

The final test accuracy of the FCNN trained model with ReLU activations reached over 50% accuracy and the CNN+FCNN trained model with ReLU activations reached over 85% accuracy (Table. 1, Figure 1b, Figure 3b)). There is a difference of approximately 28% accuracy between the FCNN and the CNN+FCNN, showing that the CNN is better for computer vision related objectives/tasks. This may be explained by various aspected such as the biological motivation for CNN being closer to the human visual system, the sparse connectivity of CNN parameters which leads to better memory and computation efficiency, and the spatial locality assumptions and spatial invariance of CNN. Removing the ReLU activation functions from both the FCNN model and the CNN+FCNN model greatly reduced the accuracy to 39% (Table. 1, Figure 2b, Figure 4b)), which is to be expected. When no non-linear activation functions are used, the model will act as a single affine layer (with no non-linearity). For example, it is easy to show with the FCNN that the multiplication of multiple weight matrices of parameters is equivalent to a single weight matrix of parameters (and the same logic applies for the CNN+FCNN even though there are convolutions) due to all the model computations being linear functions

(and linear combinations) without non-linear components in the networks.

$$\boldsymbol{W_7 W_6 W_5 W_4 W_3 W_2 W_1 x = W_0 x}$$

Neural Network Architecture Details

Following the homework details, a 7 layer FCNN with ReLU activation functions (and without ReLU activation functions) was constructed (Table 2). The architecture for the FCNN in Table 2 without ReLUs is the same but does not include the ReLU activation function after each fully-connected layer (excluding the output layer). Similarly, following the homework details, a 7 layer CNN+FCNN with 4 convolutional layers and 3 fully connected layers, with ReLU activation functions (and without ReLU activation functions) was constructed (Table 2). No maxpool layers, batch normalization layers, etc were used.

Convolutional Neural Network Architecture	
Layer	**Structure**
Input	Input Size: 32x32x3
Conv2d_1	In Channels: 3 Out Channels: 64 Stride: (2,2) Kernel Size: 3x3x3 Padding: (1,1) Activation: ReLU
Conv2d_2	In Channels: 64 Out Channels: 124 Stride: (1,1) Kernel Size: 3x3x64 Padding: (1,1) Activation: ReLU
Conv2d_3	In Channels: 124 Out Channels: 256 Stride: (1,1) Kernel Size: 3x3x124 Padding: (1,1) Activation: ReLU
Conv2d_4	In Channels: 256 Out Channels: 64 Stride: (2,2) Kernel Size: 3x3x256 Padding: (1,1) Activation: ReLU
FC_1	Input Dimension: 4096 Output Dimension: 4096 Activation: ReLU
FC_2	Input Dimension: 4096 Output Dimension: 2048 Activation: ReLU
FC_3	Input Dimension: 2048 Output Dimension: 10

Fully Connected Neural Network Architecture	
Layer	**Structure**
Input	Input Size: 32x32x3 = 3072x1
FC_1	Input Dimension: 3072 Output Dimension: 1536 Activation: ReLU
FC_2	Input Dimension: 1536 Output Dimension: 1536 Activation: ReLU
FC_3	Input Dimension: 1536 Output Dimension: 1536 Activation: ReLU
FC_4	Input Dimension: 1536 Output Dimension: 1536 Activation: ReLU
FC_5	Input Dimension: 1536 Output Dimension: 768 Activation: ReLU
FC_6	Input Dimension: 768 Output Dimension: 384 Activation: ReLU
FC_7	Input Dimension: 384 Output Dimension: 10

Table 2: Neural Network Architecture Descriptions