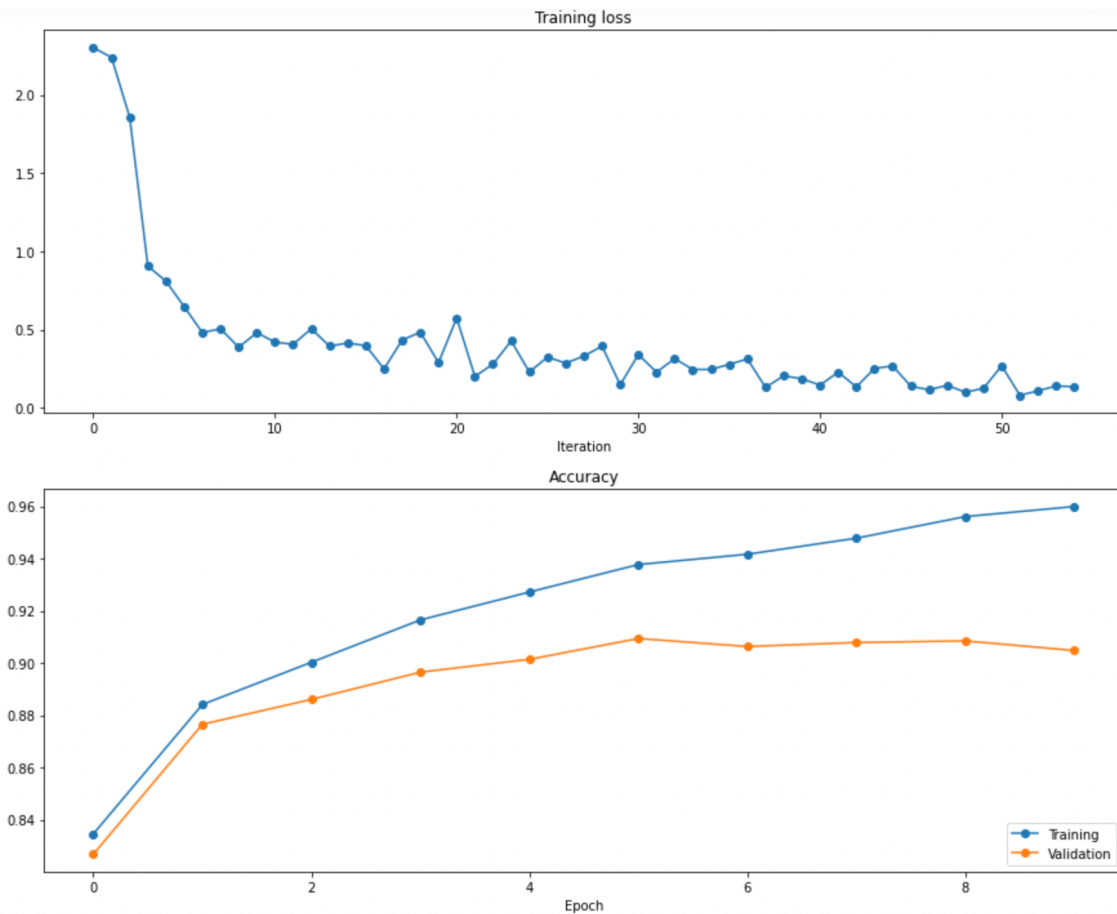


Problem 2

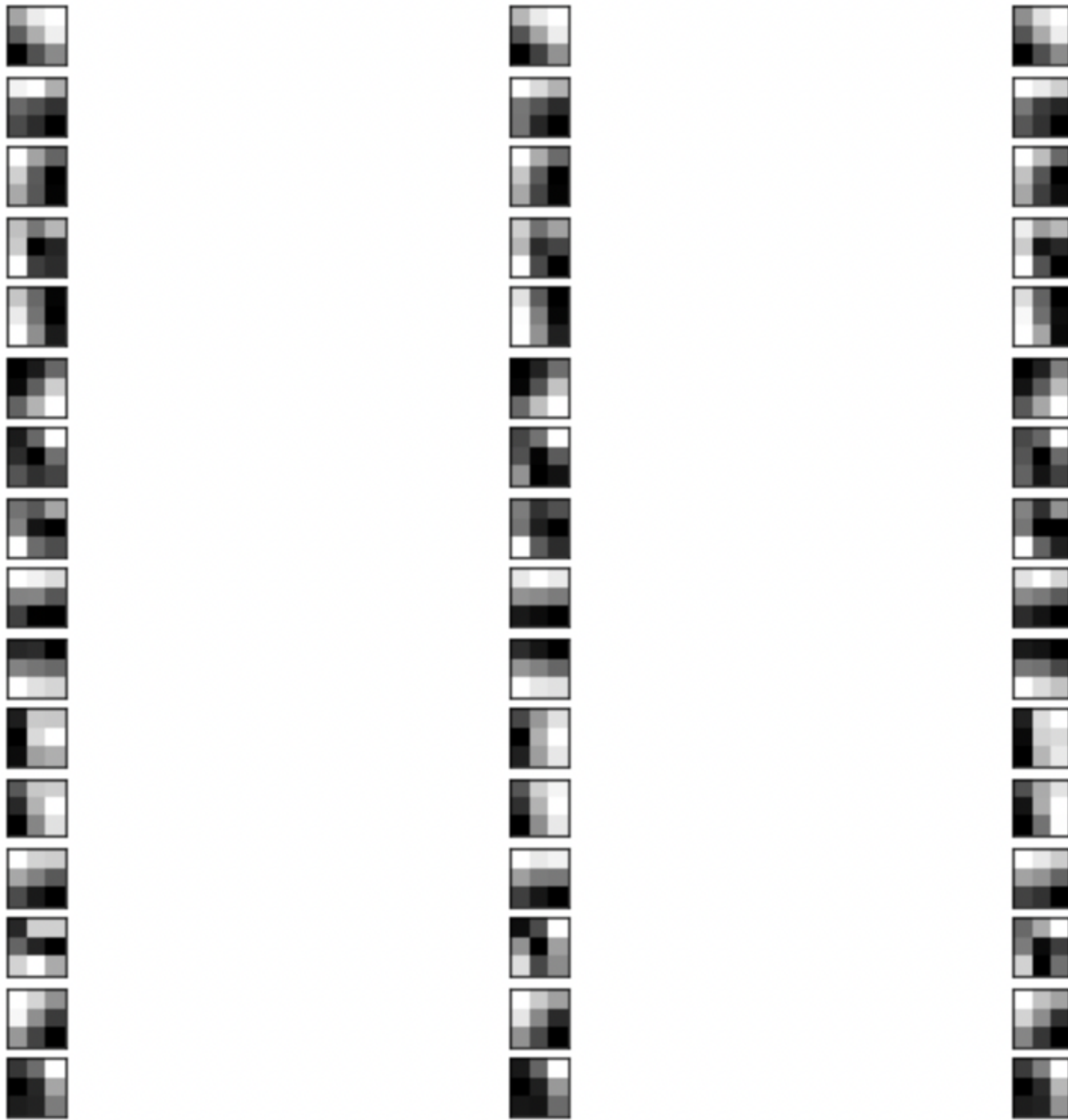
1. Training loss / accuracy curves for CNN training



```
(Iteration 1 / 5460) loss: 2.3025502718695856
(Iteration 501 / 5460) loss: 0.7147962761183483
(Epoch 1 / 10) Training Accuracy: 0.8346, Validation Accuracy: 0.8268345102855389
(Iteration 1001 / 5460) loss: 0.40271941116270277
(Epoch 2 / 10) Training Accuracy: 0.8841428571428571, Validation Accuracy: 0.8765735339269266
Decaying learning rate of the optimizer to 0.00085
(Iteration 1501 / 5460) loss: 0.2690541597691176
(Epoch 3 / 10) Training Accuracy: 0.9003142857142857, Validation Accuracy: 0.8860914952410194
(Iteration 2001 / 5460) loss: 0.29811615074083453
(Epoch 4 / 10) Training Accuracy: 0.9165285714285715, Validation Accuracy: 0.8965305495855082
Decaying learning rate of the optimizer to 0.0007224999999999999
(Iteration 2501 / 5460) loss: 0.2915215436052228
(Epoch 5 / 10) Training Accuracy: 0.9272142857142858, Validation Accuracy: 0.9014430457476205
(Iteration 3001 / 5460) loss: 0.3225101393867451
(Epoch 6 / 10) Training Accuracy: 0.9377428571428571, Validation Accuracy: 0.9094258520110531
Decaying learning rate of the optimizer to 0.000614125
(Iteration 3501 / 5460) loss: 0.22207987608865906
(Epoch 7 / 10) Training Accuracy: 0.9416857142857142, Validation Accuracy: 0.9063555419097329
(Iteration 4001 / 5460) loss: 0.34173404186256484
(Epoch 8 / 10) Training Accuracy: 0.9478285714285715, Validation Accuracy: 0.907890696960393
Decaying learning rate of the optimizer to 0.00052200625
(Iteration 4501 / 5460) loss: 0.2756848690225722
(Epoch 9 / 10) Training Accuracy: 0.9561142857142857, Validation Accuracy: 0.9085047589806571
(Iteration 5001 / 5460) loss: 0.17560006214442503
(Epoch 10 / 10) Training Accuracy: 0.9599571428571428, Validation Accuracy: 0.9048203868590727
```

Run the code below to generate the training plots.

2. Visualization of convolutional filters



3. Answers to inline questions about convolutional filters

There are 16 filters and each filter on one row has three channels for red, green and blue. I see the filters with different colors, some grids have dark color and some grids have bright color. The dark color shows a small weight and the light color shows a large weight in the filters. However, the patterns that the filters learn are very general so that it may be used to detect the shapes and edge in images, and it can be applied for later layers to learn in more details about specific objects. It is reasonable since they are filters from the first convolutional layer and at a very early stage of model training.