CSCI 611 Assignment 2 Report

Repository Link: https://github.com/FlansaasSoft/CSCI611_Logan_Flansaas

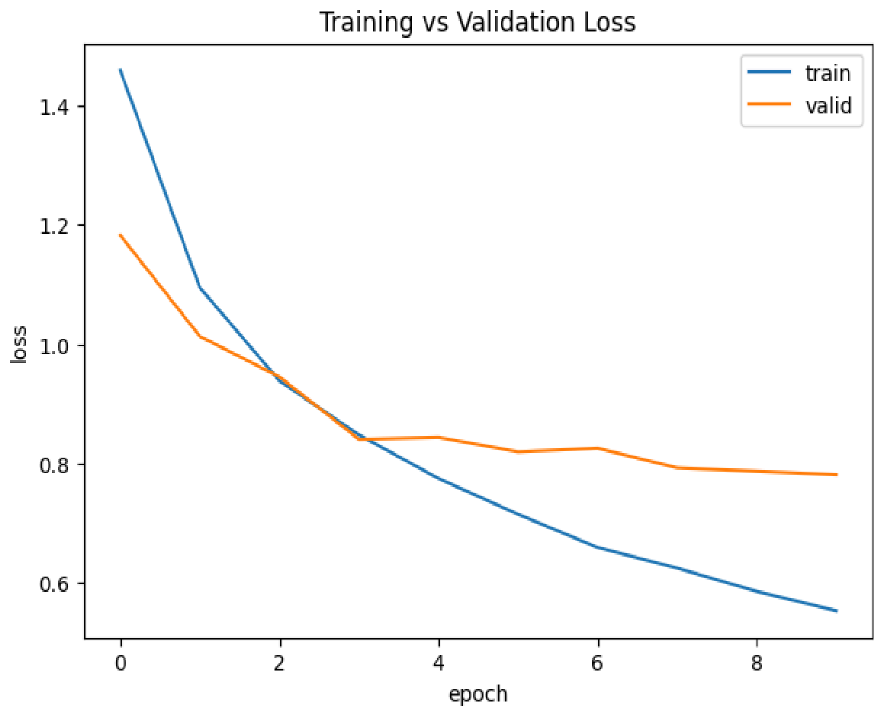
Model Architecture and Training Setup

My Cifar-10 CNN model consists of 3 convolutional blocks of progressively increasing channel depth (16 -> 32 -> 64) with a kernel size of 3x3 + ReLU, and a 2x2 max pooling layer after each block. The information from these blocks is then flattened and then fed into 2 each of alternating dropout and fully connected linear classification + ReLU layers. The model is trained using the CrossEntropyLoss loss function, the Adam optimizer with a learning rate of 0.001, a batch size of 20, and 10 epochs.

Training Results

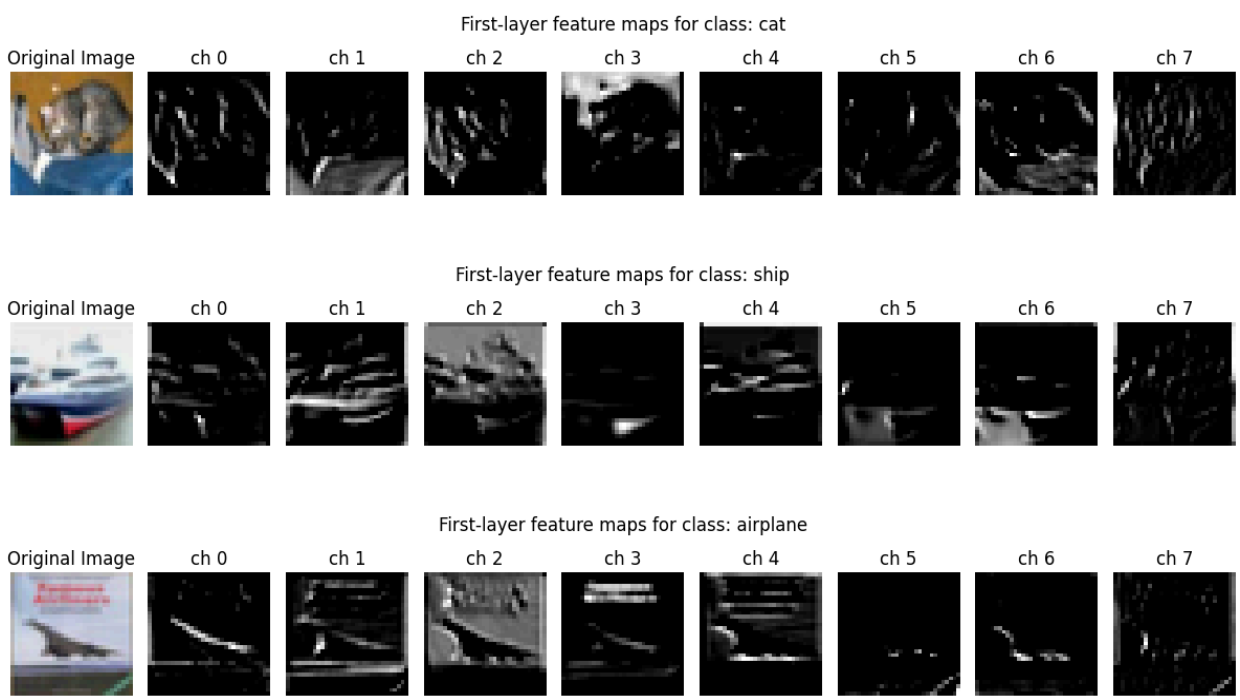
The model achieved an overall test accuracy of 73.56% with a test loss of 0.779072. Class accuracy varied significantly with a gap from 60% to 82.7%. The model achieved a final training loss of 0.577204 and a final validation loss of 0.750366. The loss curves suggest a case of mild overfitting to the training data. While the training loss continues to steadily decrease, the validation loss begins to mostly plateau at epoch 3, which is a classic signal of overfitting.

Class	Accuracy
Airplane	81.5%
Automobile	82.7%
Bird	60.1%
Cat	60.6%
Deer	68.4%
Dog	60%
Frog	80.2%
Horse	77%
Ship	82.7%
Truck	82.4%



Feature Map Visualization (Early Layer)

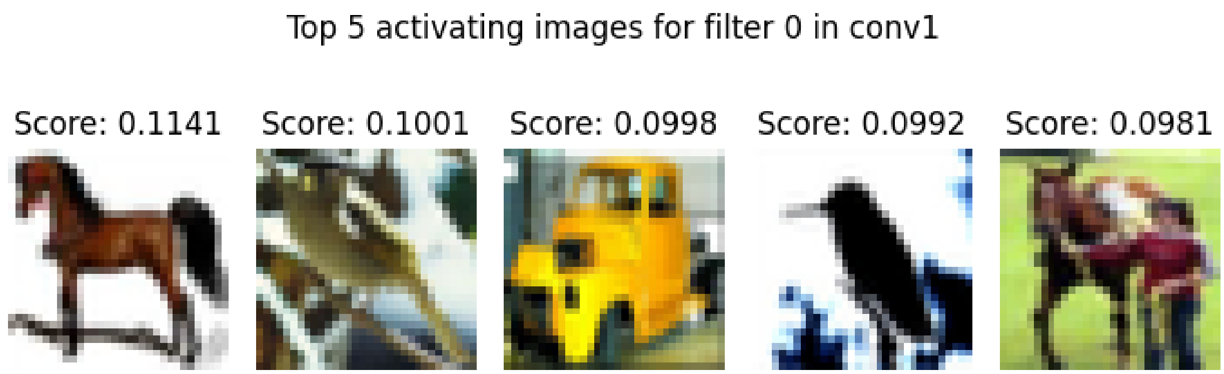
The feature maps show the various patterns that the filters are detecting and measuring in the images. Bright regions indicate strong matching to that filter's pattern and dark regions indicate weak matching to that filter's pattern. Early layer feature maps typically represent lower level patterns such as edge detection and direction, simple textures, color and brightness contrasts, and rough silhouettes of objects. For example, channels 0 and 1 appear to be responding to either vertical and horizontal edges respectively, or basic silhouettes and deep blues respectively. Channel 2 appears to respond strongly to the backgrounds of the images, and channel 3 respond strongly to the red and brown colors in the images. Channels 4, 5, and 6 appear to respond to silhouettes and and contrasting regions of the images, while channel 7 suggests a filter for basic textures in the images. These feature maps all demonstrate the various different patterns that can be identified from the same input image.

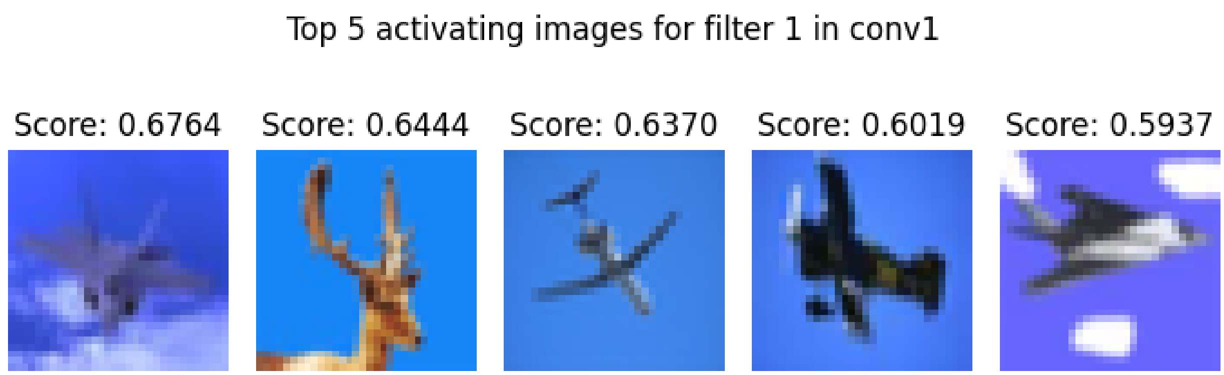


Maximally Activating Images

The maximally activating images are the images that produce the strongest activation or score for a particular pattern that a filter is looking and measuring for. I measure the score for these images using the mean values of the underlying feature maps created by the filters. Using the mean value for measurement of activation score means that filters that look for small, local, and infrequent patterns will generally have a low overall score compared to other filters. Meanwhile, filters that activate on broad/global or smooth patterns will generally have a higher overall score compared to other filters. This can be seen in the top 5 images for filters 0, 1, and 2 below. I extracted the filters for convolutional layer 1, so the patterns that activate these filters are generally going to be for general features and low level patterns of the images rather than class related patterns. For filter 0, all scores are rather low and the images do not clearly belong to a single class or similar set of classes. This suggests that filter 0 may be measuring for patterns like the silhouettes and boundaries of objects, which generally do not dominate image spaces and would only see activation in a small portion of the images, which explains the overall low scores. For filter 1, activation scores are high overall and the images are all very clearly dominated by blue backgrounds, with 4 of the top 5 images appearing to be airplanes. Since this filter is pulled from the first convolutional layer, it’s unlikely that this filter is activating on airplanes in particular because early layer filters rarely focus on class specific patterns. So, what this filter is most likely to be activating on are the smooth bright blue backgrounds that dominate the image space, which would explain the high overall scores when measuring with the mean. Finally, filter 2 activation scores are somewhat high overall with the top images largely consisting of smooth bright backgrounds with foreground objects that have long neat edges and contrast clearly with the foreground.

3 of the images are airplanes, but again this an early layer filter, so this filter should not be activating on class specific patterns. This suggests that filter 2 is activating on the contrasting areas of images, which again would explain the high mean scores.





Discussion and Reflection

Overall, my model works reasonably well with a 73% overall accuracy, which seems to be pretty good for such a relatively small CNN. Among the most interesting results of my models training and testing is the wide gap in classification accuracy between some of the classes. The bird, cat, deer, and dog had the lowest accuracies with them all in the 60% range, with bird, cat, and dog at the bottom of that range. Meanwhile all of the vehicle classes scored at the top of the accuracy range, with horse and frog being not far behind. The reason for this discrepancy may be because the lower scoring classes often have more cluttered image spaces with finer details distinguishing them and have high levels of similarity like with cat and dog. Meanwhile classes like the vehicles typically have more straight and clear forms with significantly less variation between individuals of those classes, while horse and frog have much more distinct colors and shapes compared to the other animal classes. Some ways to potentially improve the accuracy of the model that I did not explore in this assignment would be data augmentation like rotating, flipping, and cropping, and adding more channels or additional convolutional layers. Adding more channels and layers would increase the number and complexity of patterns detected, but would require more compute and training time as a result