

CSE849 Project 2

Javen W. Zamojcin
zamojci1@msu.edu

2025, March 21

Report the validation accuracy in your report

Validation accuracy = 70.52%

Visualize all the filters from the first layer of your best model. Do these filters remind you of any patterns related to any of these CIFAR-10 classes?

Figures 2, 3 depict all 16 of the first layer filter visualizations from Q2. Interestingly, filter-9 appears to have a strong linear pattern (diagonal) corresponding to an activation emphasis on vehicles (planes in particular). This makes sense as airplanes and ships often have straight edges. Additionally, filter-2 has a "bulging" pattern that captures and corresponds well to the roundish features of automobiles.

Include a few bar plots corresponding to filters from the first and the final convolutional layers.

Figures 4, 5 depict all 16 of the first layer filters, while Figure 6 depicts 8 selected final layer filters.

Explain what difference you note between the bar plots from these two layers. How can you use them to explain what your model has learned?

In the 16 first layer filters, we see much more uniform activation and less variance among the classes. While they have relatively uniform activations, we do observe a noticeably higher activation emphasis on the vehicle related classes (e.g., airplane, truck, automobile). This suggests that common, low-level features such as edges and corners are being detected in this earlier layer. It's also worth noting that several "dead" filters exist for the the first layer, implying near-zero weights and no discernible patterns. This could be due to vanishing gradients during training.

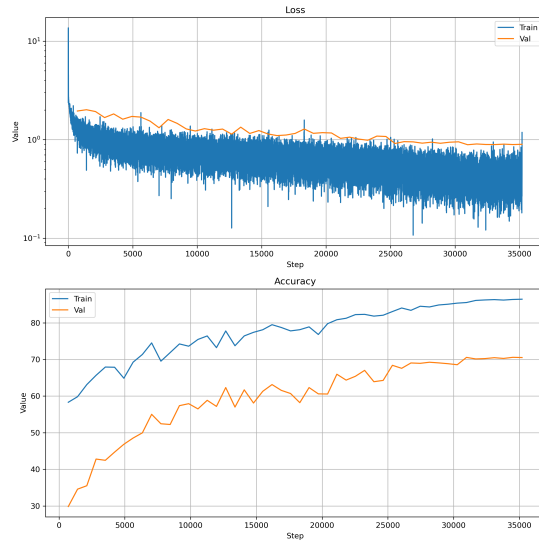


Figure 1: Q1 Training History

In the deeper fifth layer, we observe less uniform activation and higher variance among the filters. This suggests higher class specialization, or that the filters learn to detect features specific to one or few classes and only react strongly to these classes.

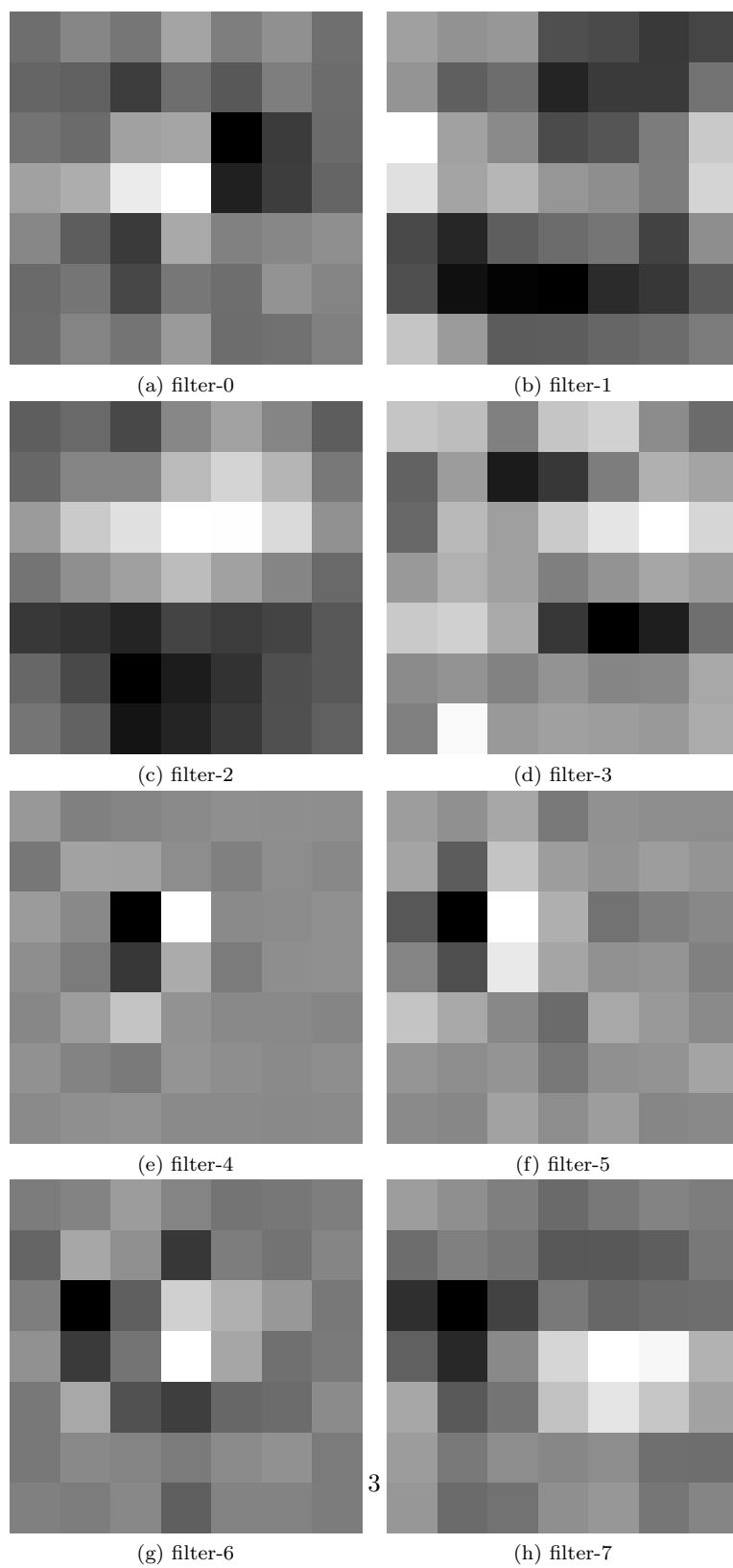


Figure 2: Visualized Layer-1 Filters 0-7

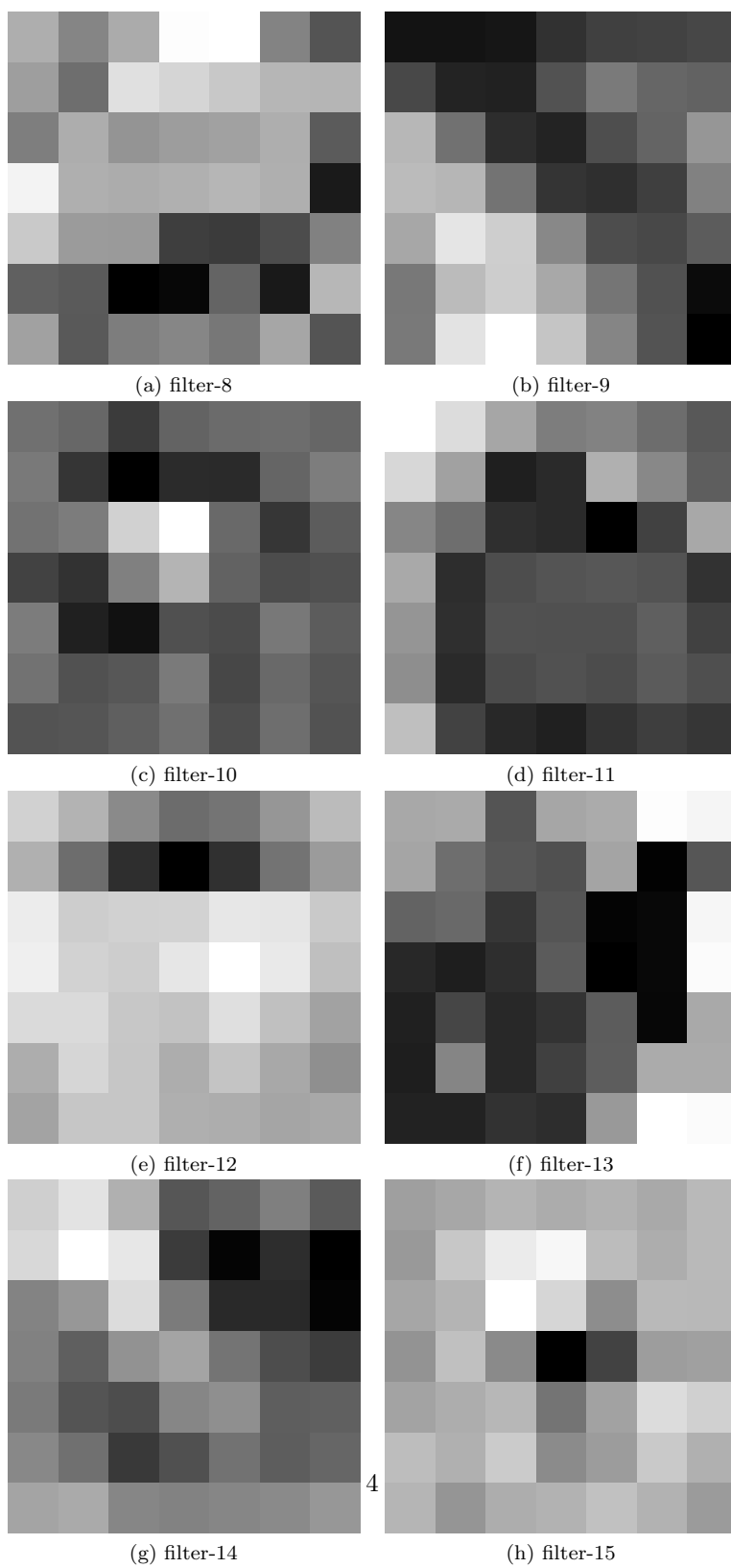
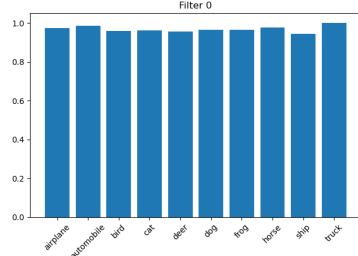
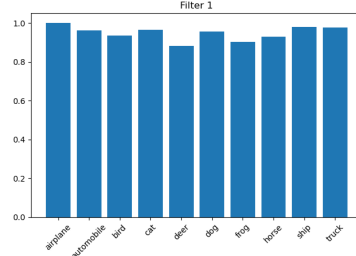


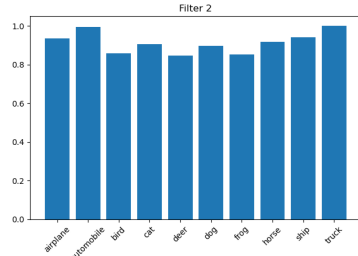
Figure 3: Visualized Layer-1 Filters 8-15



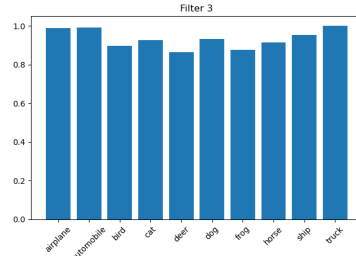
(a) filter-0 var=0.0003



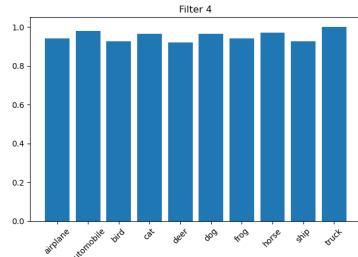
(b) filter-1 var=0.0013



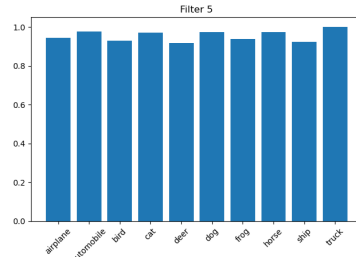
(c) filter-2 var=0.0030



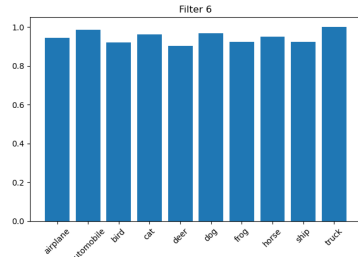
(d) filter-3 var=0.0023



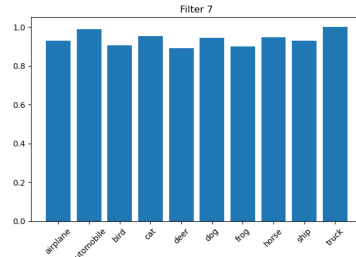
(e) filter-4 var=0.0007



(f) filter-5 var=0.0008

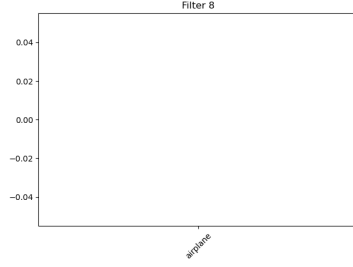


(g) filter-6 var=0.0009

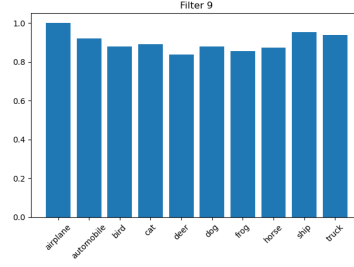


(h) filter-7 var=0.0013

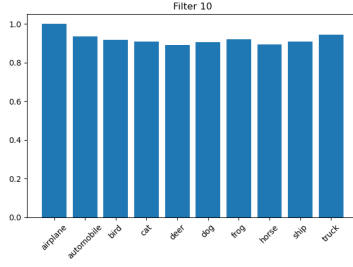
Figure 4: Average Classwise Norm Layer-1 Filters 0-7



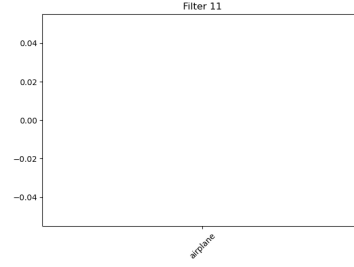
(a) filter-8 var=0.0



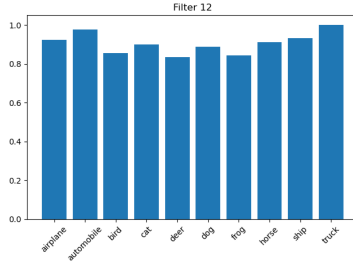
(b) filter-9 var=0.0025



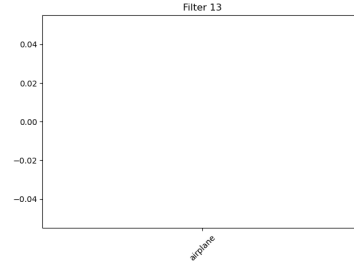
(c) filter-10 var=0.0010



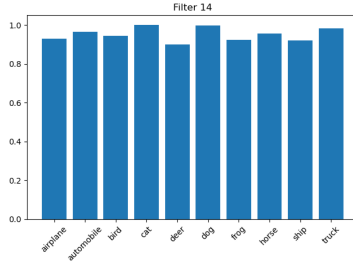
(d) filter-11 var=0.0



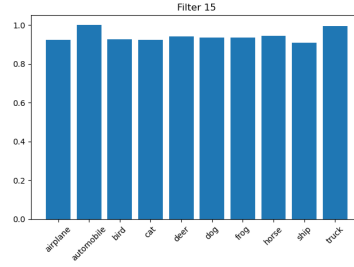
(e) filter-12 var=0.0030



(f) filter-13 var=0.0

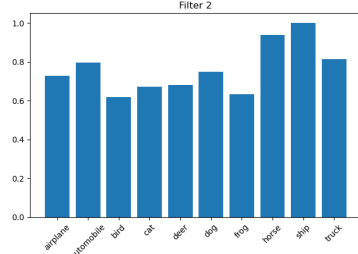


(g) filter-14 var=0.0012

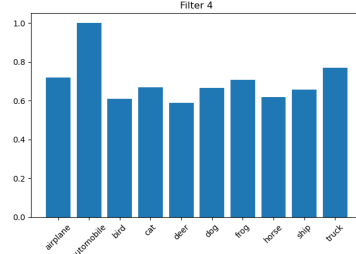


(h) filter-15 var=0.0009

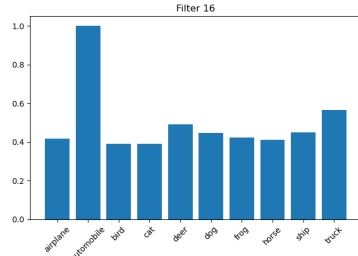
Figure 5: Average Classwise Norm Layer-1 Filters 8-15



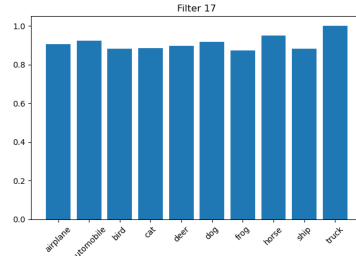
(a) filter-2 var=0.0160



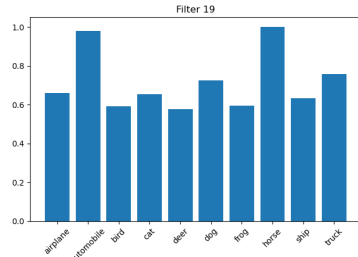
(b) filter-4 var=0.0141



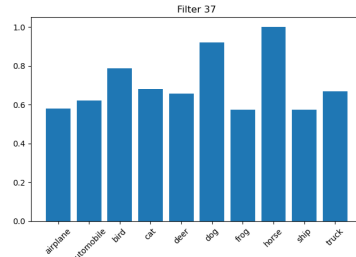
(c) filter-16 var=0.0338



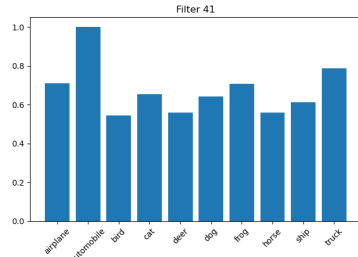
(d) filter-17 var=0.0015



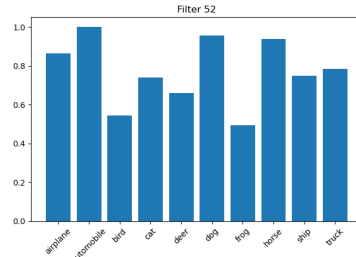
(e) filter-19 var=0.0239



(f) filter-37 var=0.0223



(g) filter-41 var=0.0118



(h) filter-52 var=0.0296

Figure 6: Average Classwise Norm Layer-5 Filters