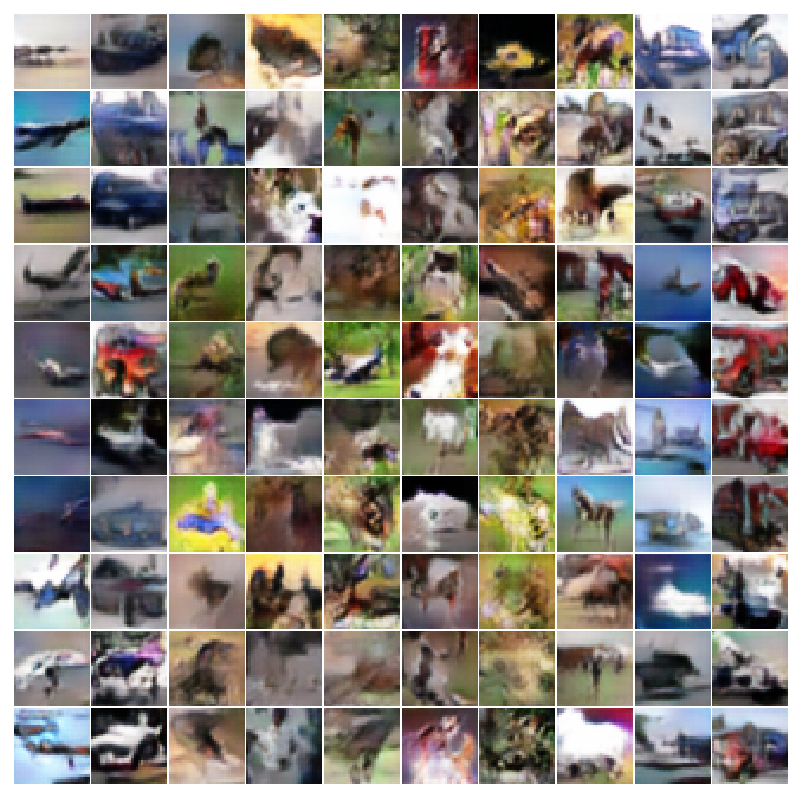
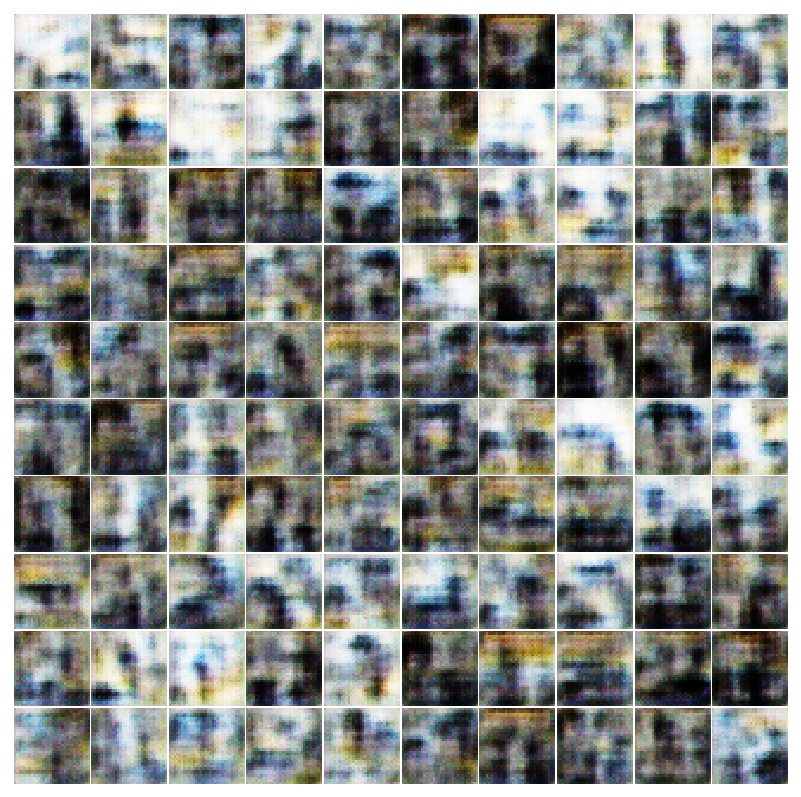
CS598 DL HW6 Report

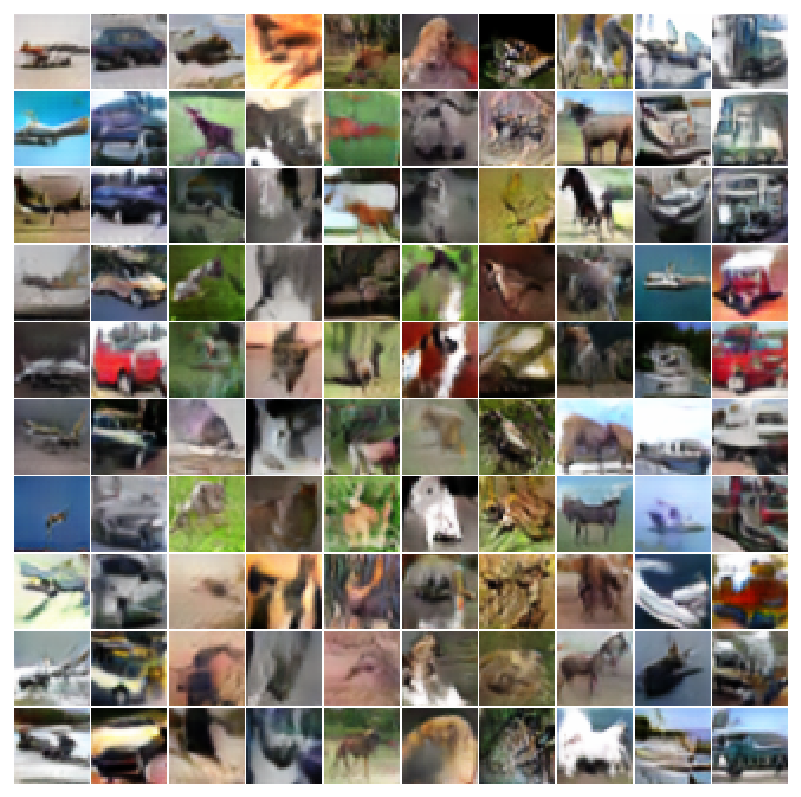
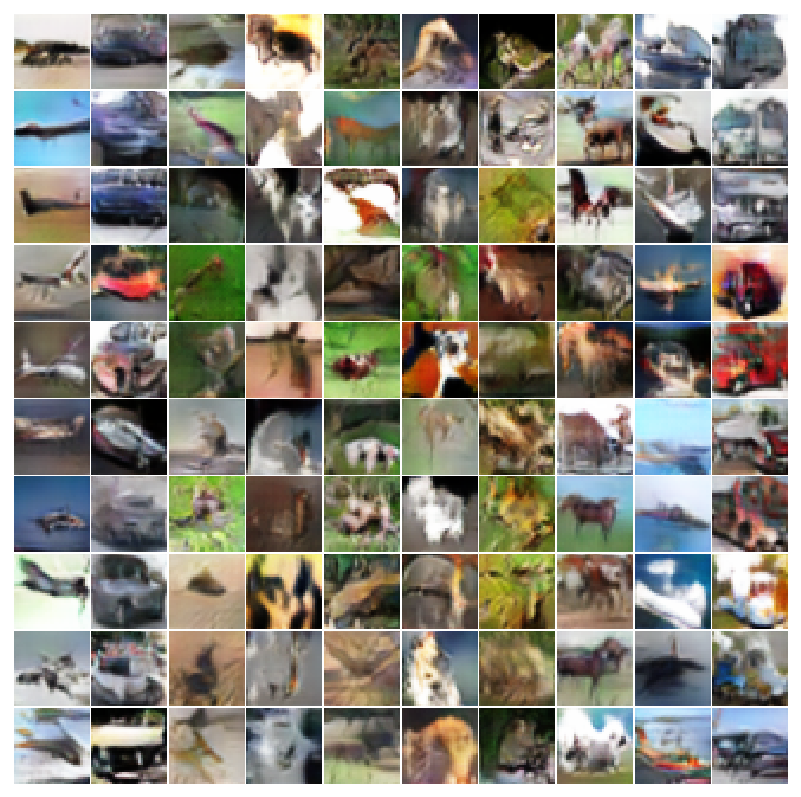
Zhenbang Wang(zw11)

1. Training Process:

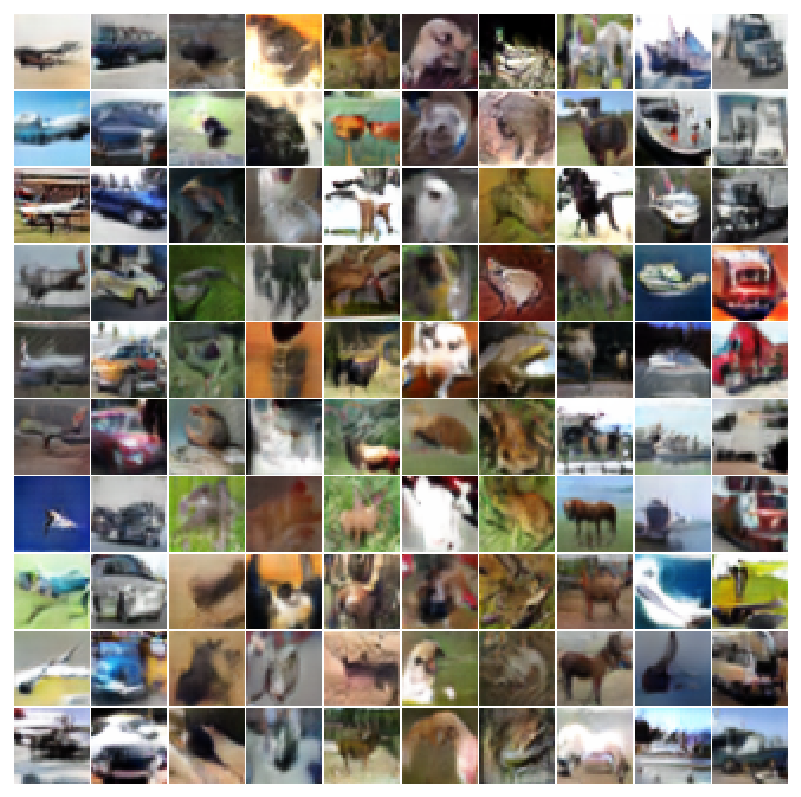
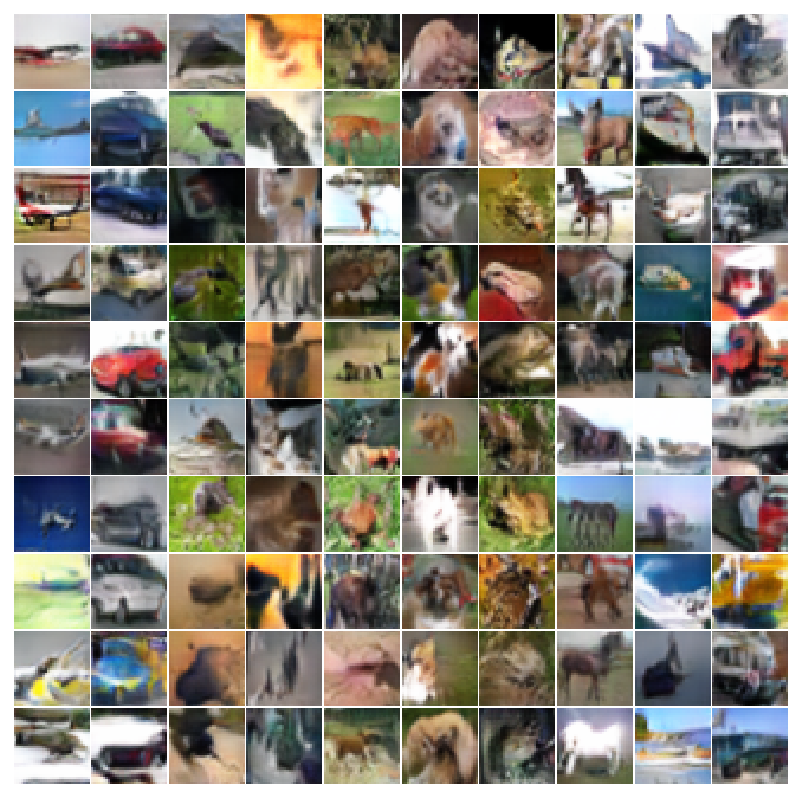
epoch 0 epoch 40



epoch 80 epoch 120



epoch 160 epoch 200

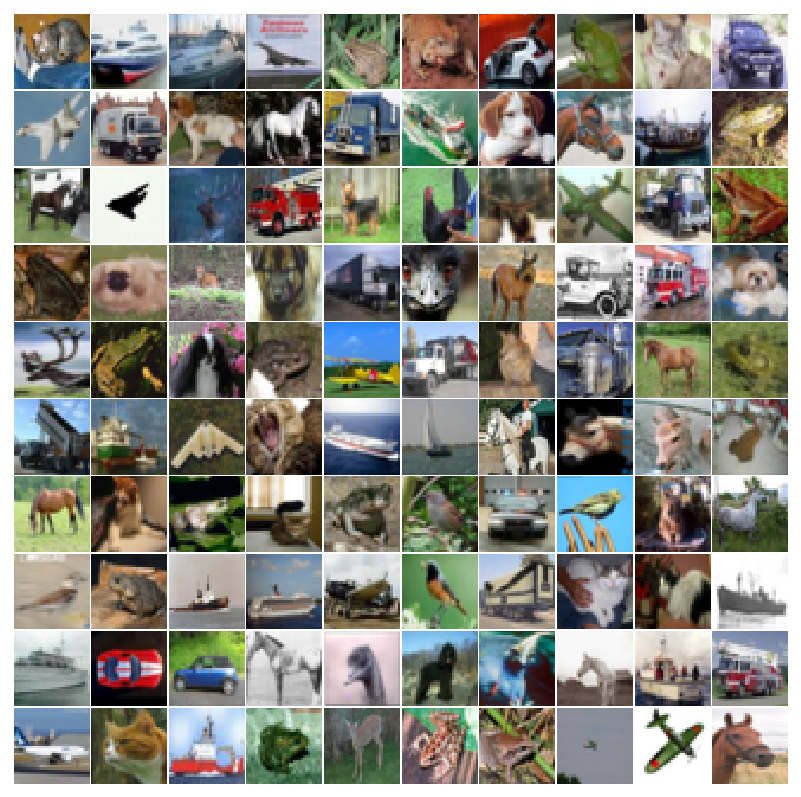


2. Fast Gradient Sign Method:

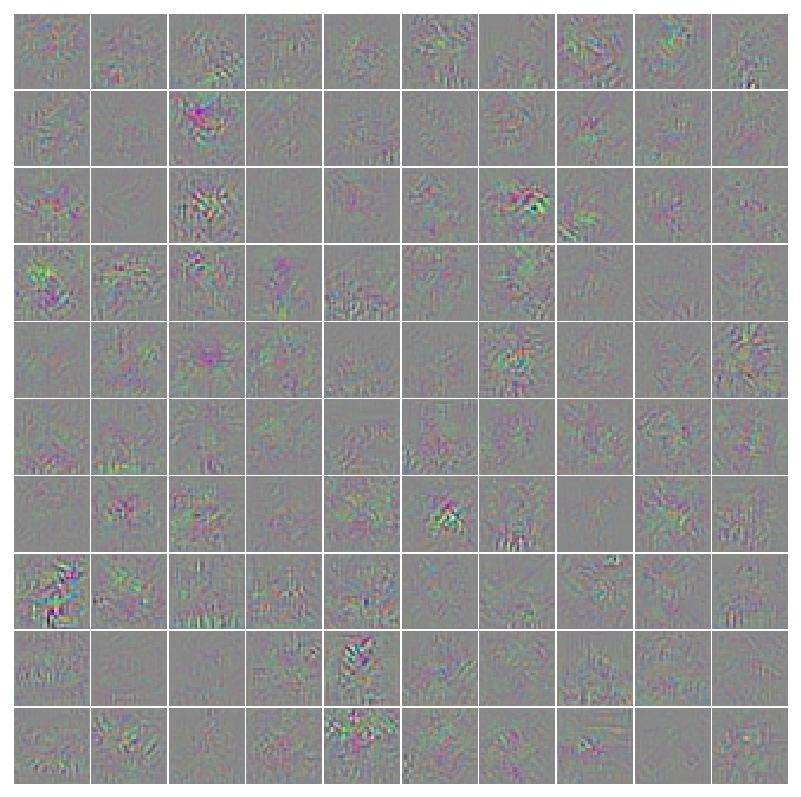
Original Accuracy: 90.625%

Adversarial Accuracy: 13.28125%

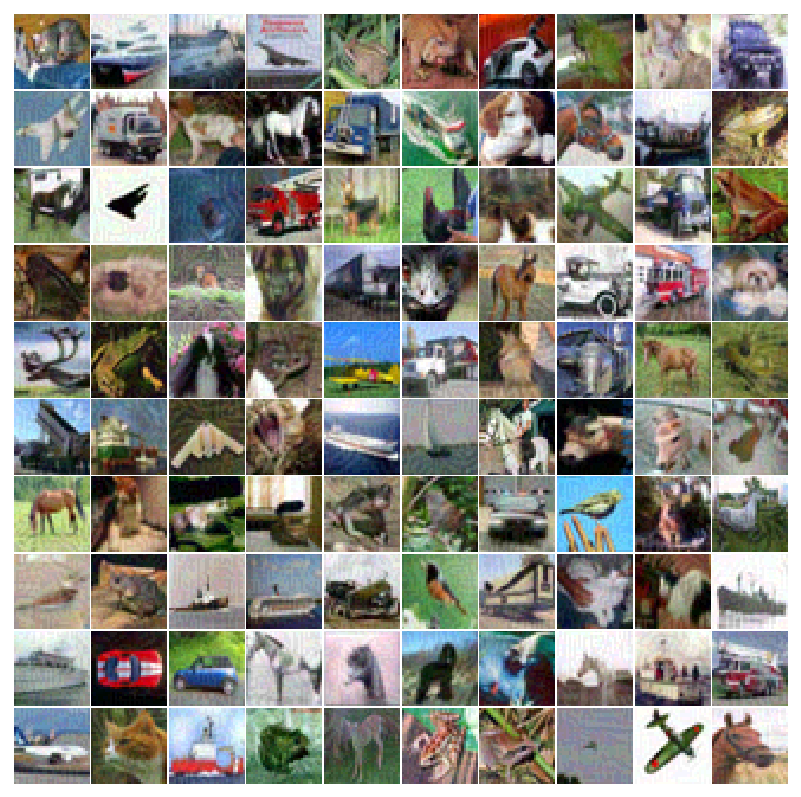
real image:



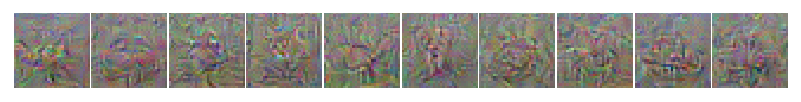
gradient:



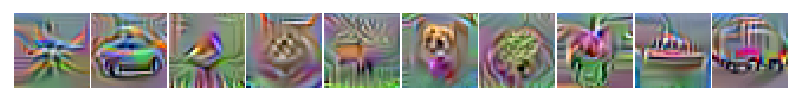
Adversarial images:



3. Synthetic images maximizing the class output

discriminator trained without the generator:

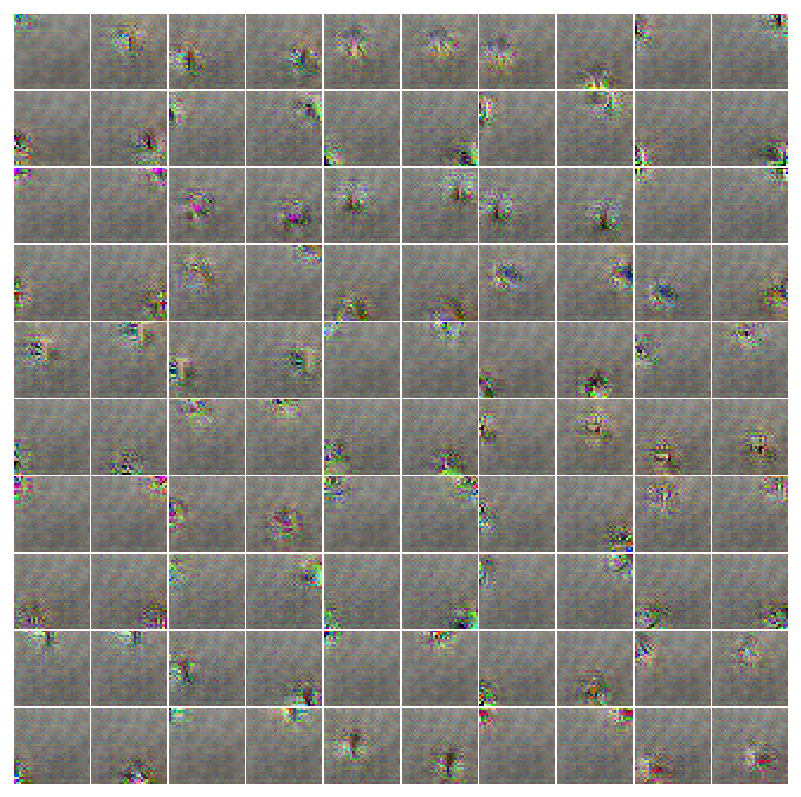
discriminator trained with the generator:



4. Synthetic images maximizing a particular layer of features

discriminator trained without the generator:

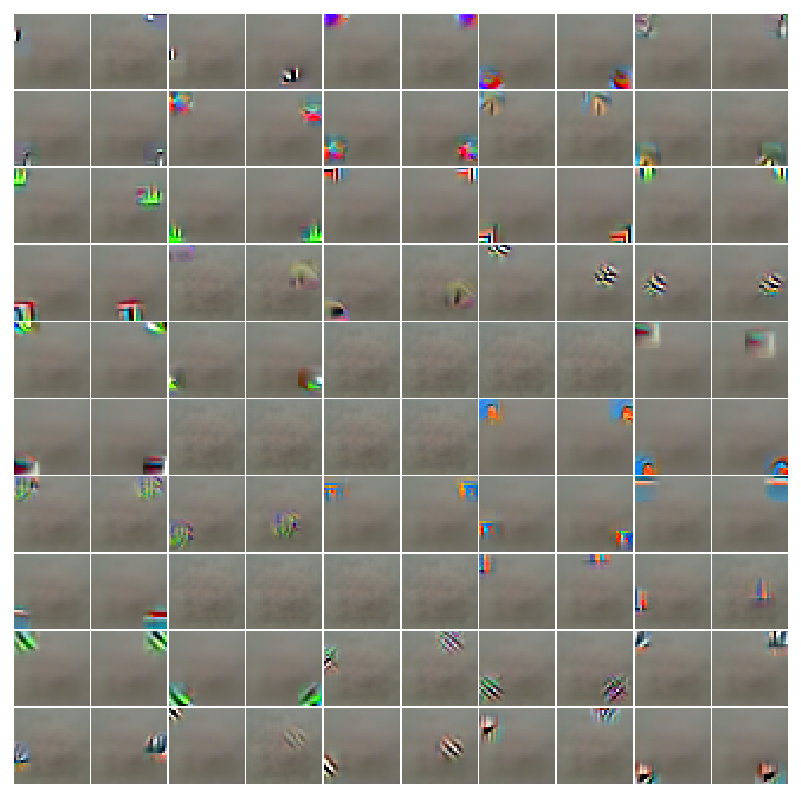
layer 4:



layer 8:



discriminator trained with the generator:

layer 4: 

layer 8:



 5. Test Accuracy

 Discriminator trained without the generator: 89.07%

 Discriminator trained with the generator: 84.09533227848101%