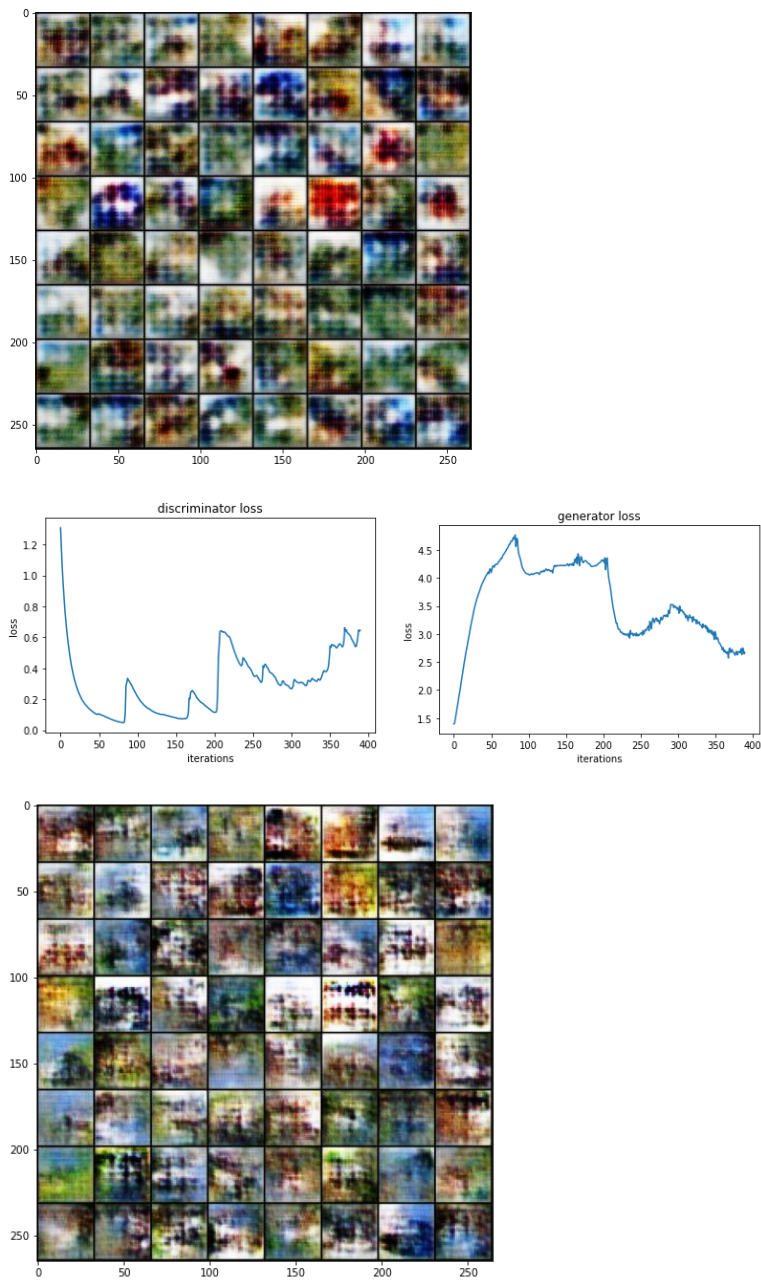
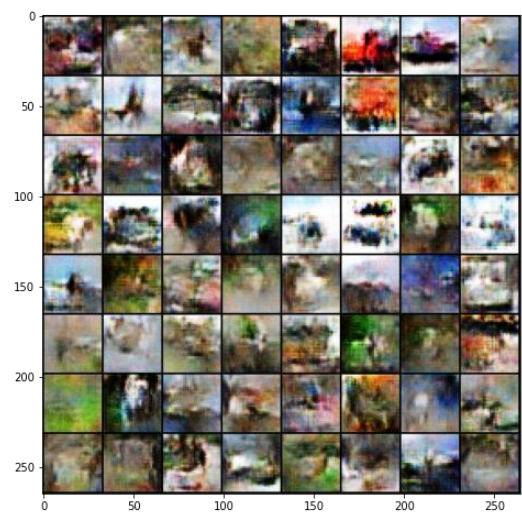
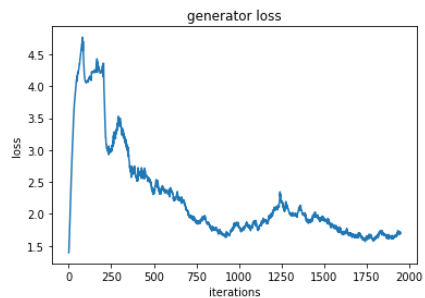
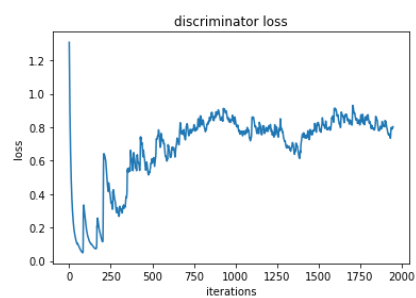
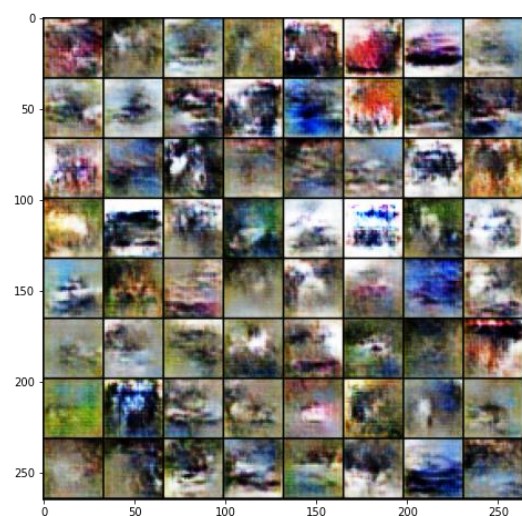
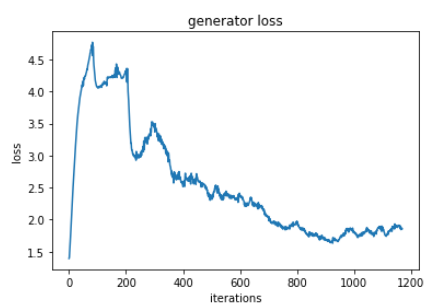
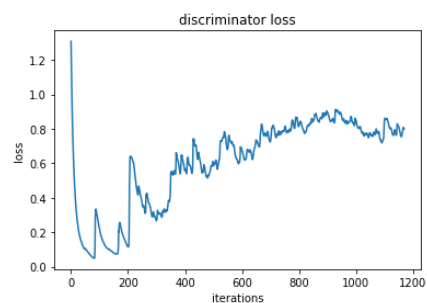


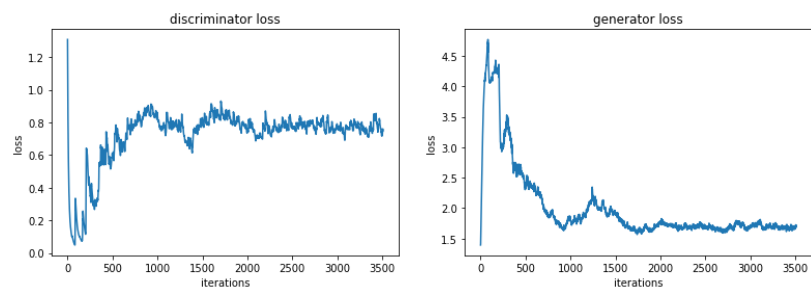
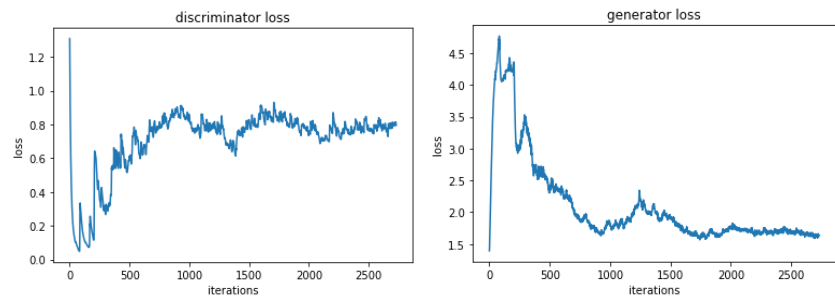
PROBLEM 2 SOLUTION

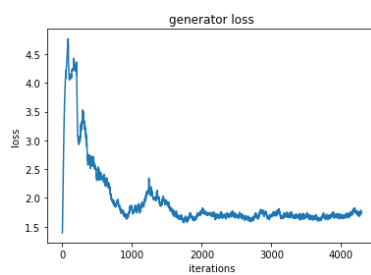
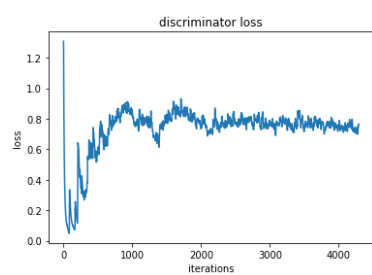
MEHAK PIPLANI

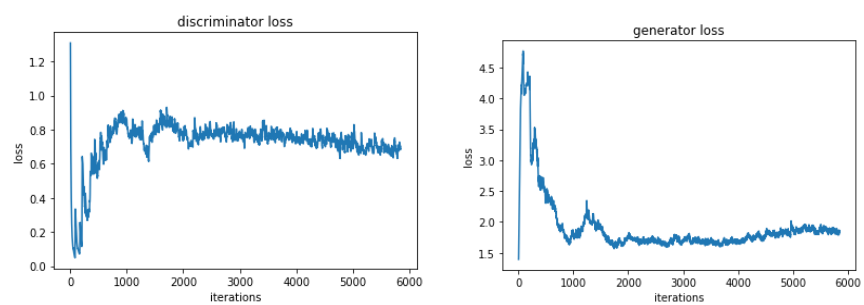
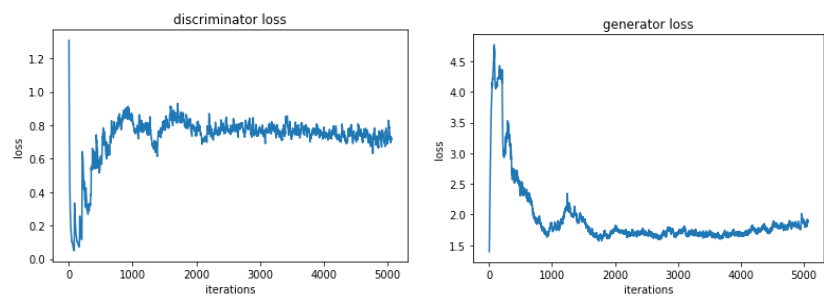
Training curves and generated images for the GAN

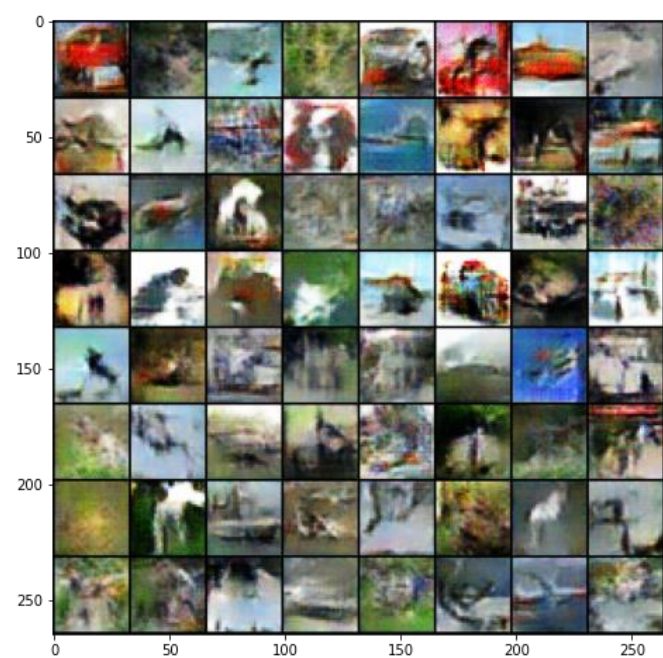
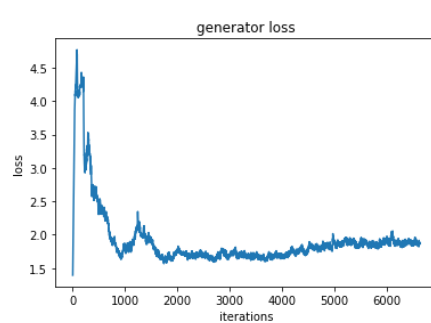
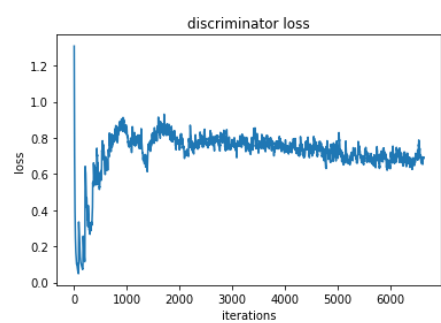


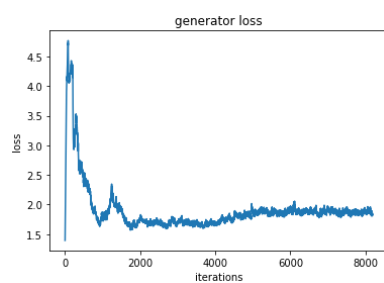
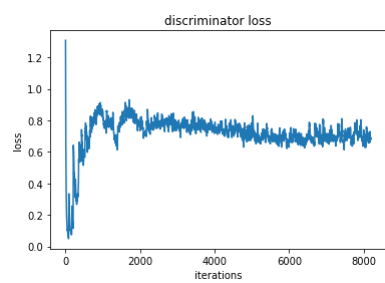
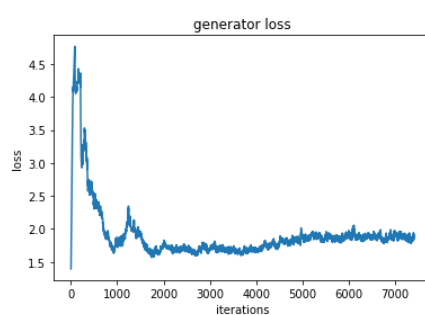
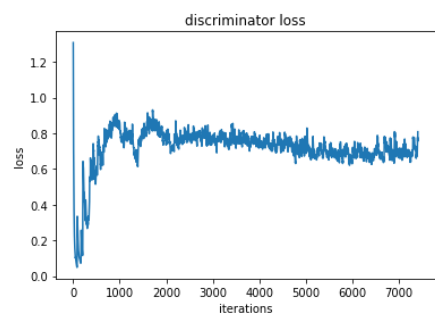


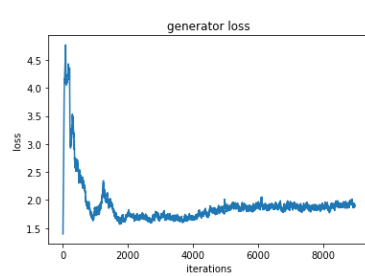
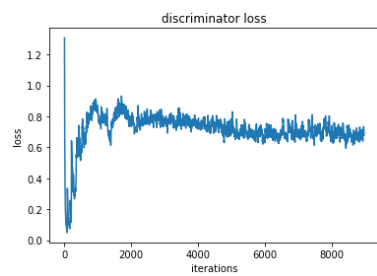


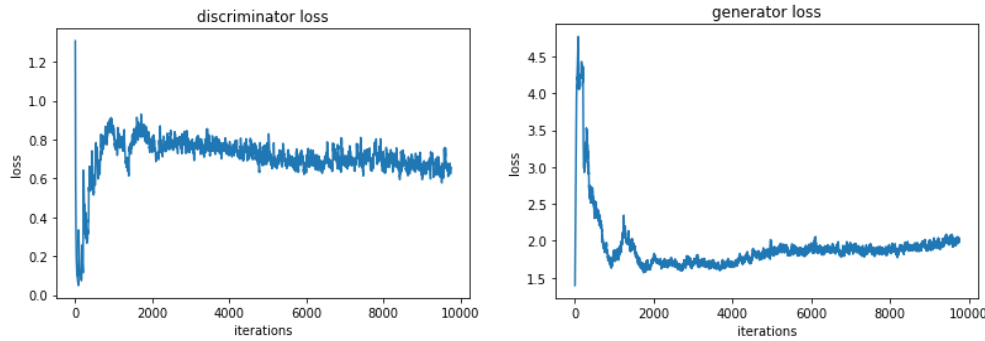












Problem 2-2: The forger versus the police, revisited

Question: In the forger versus police story, we made part of it hand-wavy to hide a flaw that makes the story improbable to actually happen and makes it a bad analogy of how the training works in a GAN. Now that you have implemented a GAN, can you spot the flaw? Specifically, when we consider one of the two parties, the other is treated as a black box. They know their opponent's result but not how they works. What is wrong here?

Answer: When training the discriminator, the training samples are detached of the generator's results and hence, The discriminator always uses the generator as a black box never examining its internal parameters. However, when training the generator, we do not detach samples passing the discriminator's parameters to compute its gradient direction. Hence, this the flaw in the story as discriminator is not a black box to the generator.

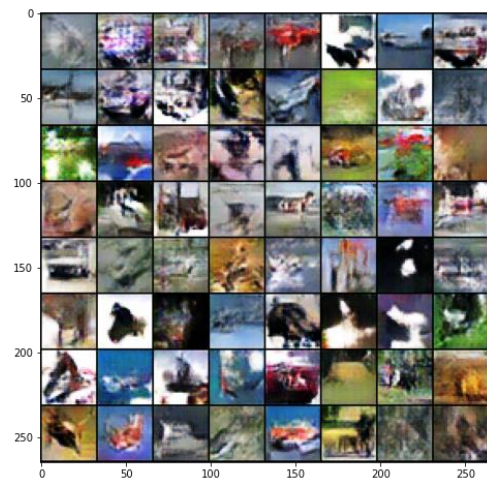
Problem 2-3: The Batch Normalization dilemma

Question: By removing the first batch normalization layer, for two different distributions to get confused with each other they must produce two distributions after `dis_lrelu1` such that one can be obtained by applying an isotropic scaling and a translation to the other. Such a case is still possible but extremely unlikely to happen. Propose a different way of feeding the samples to solve the problem in the second question without omitting any batch normalization layers or changing their mode of operation.

Answer:

One way of feeding the samples is transforming the real batch samples and fake batch samples differently before feeding it into the Discriminator so that the two distributions are different. One way to do such a transformation is to apply on non-linear filter like adding noise (speckle noise). This can help to directly apply batch norm to all the layers and not result in sample oscillation or model instability as mentioned in the DCGAN paper.

Activation Maximization



TEST SAMPLES



RECONSTRUCTIONS (Reconstruction loss = 0.0135)

