Jiaming Li
CSI 5340 Intro to Deep Learning and Reinforcement Learning
Homework Exercise 4
Nov 14, 2021

Part I: Introduction

This assignment uses two datasets, which are MNIST (has a training set of 60,000 examples, and a test set of 10,000 examples) and CIFAR10 (consists of 60000 color images in 10 classes, with 6000 images per class), to implement and train the three models, they are VAE, GAN and WGAN. For VAE model, it contains encoder and decoder components; for GAN model it contains generator and discriminator components; for WGAN model, it contains generator and critic components. The parameters are set as below in each model. Also, the plots of loss values and dataset images after training are shown in this report.

The Figure 1 shows below is the original 8 * 8 MNIST and CIFAR10 Dataset image displayed before the three models training. From the figure shown, we can see that the handwritten digits and 10-class objects are clear.

Left: MNIST Dataset **Right:** CIFAR10 Dataset

Figure 1: Two Datasets Display before Model Training

Part II: Models for MNIST Dataset

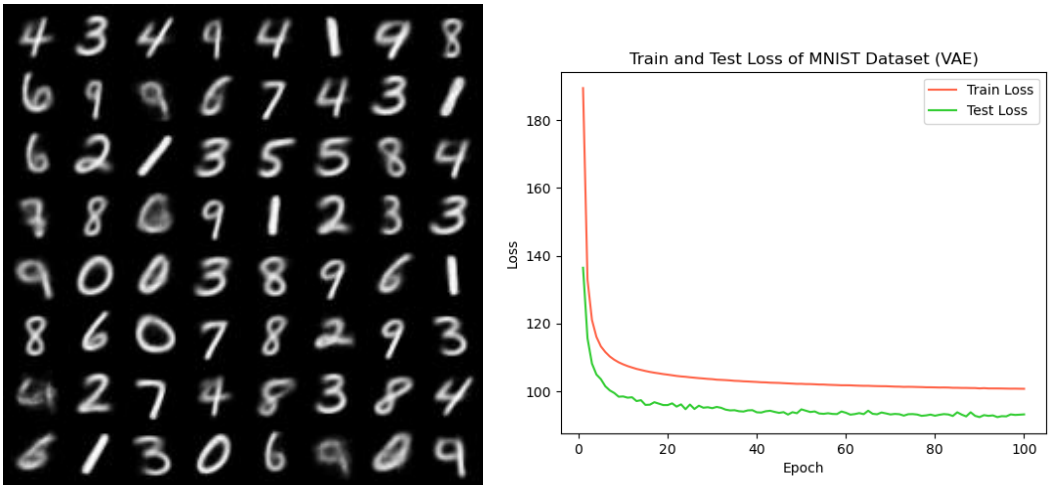
1. Model VAE

1.1 Parameters Settings

Loss Function: Binary Cross Entropy Loss.

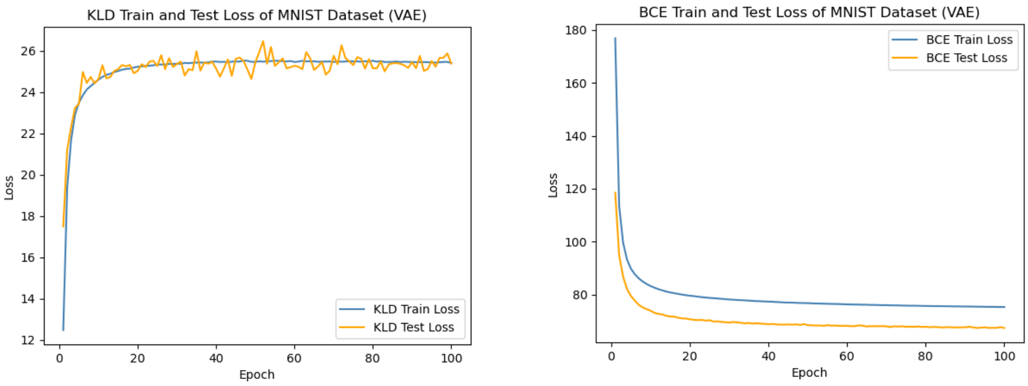
Reconstruction + KL divergence losses summed over all elements and batch.
Optimizer: Adam
d = 20
Batch Size = 256
Learning Rate = 0.001
Epoch = 100

1.2 Images and Result Plot Show



Left: MNIST Display After 100 Epochs VAE Training **Right:** Train and Test Loss of MNIST (VAE)

Figure 2: Image Display and Plot Result of MNIST for Model VAE



Left: KLD Train and Test Loss of MNIST (VAE) **Right:** BCE Train and Test Loss of MNIST (VAE)

Figure 3: KLD & BCE Plot Result of MNIST for Model VAE

1.3 Results and Conclusions

The model was trained for 100 epochs, with the loss function Binary Cross Entropy loss. It is used to calculate the reconstruction loss, which is a part of the whole loss. For another part of loss, KL divergence is mentioned in this model. From the Figure 2 shows above, it can be easily concluded that the test loss is always lower than the train loss curve and they start from different loss value, which has a big difference. But they all decrease drastically when epoch is at around 7 or 8. Then they become relatively stable when epoch is 20. Figure 3 shows that the KLD and BCE losses of VAE model changeover 100 epochs.

2. Model GAN

2.1 Parameters Settings

Loss Function: Binary Cross Entropy Loss

Optimizer: Adam

Input Dim = 100

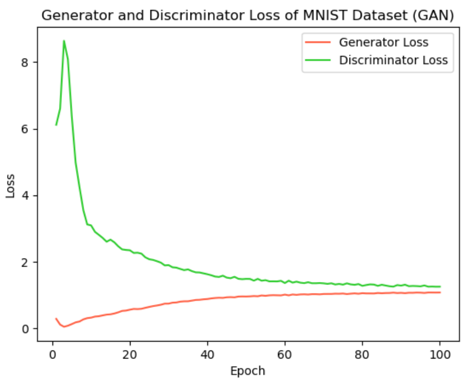
Batch Size = 100

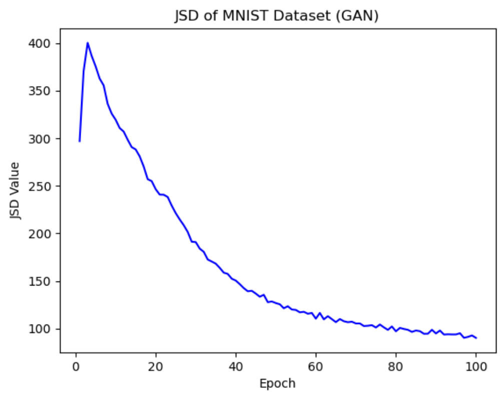
MNIST Dim = 1 * 28 * 28

Learning Rate = 0.0002

Epoch = 100

2.2 Images and Result Plot Show





Left: Generator and Discriminator Loss of MNIST (GAN)

Right: JSD Value of MNIST (GAN)

Figure 4: Plot Results of MNIST for Model GAN

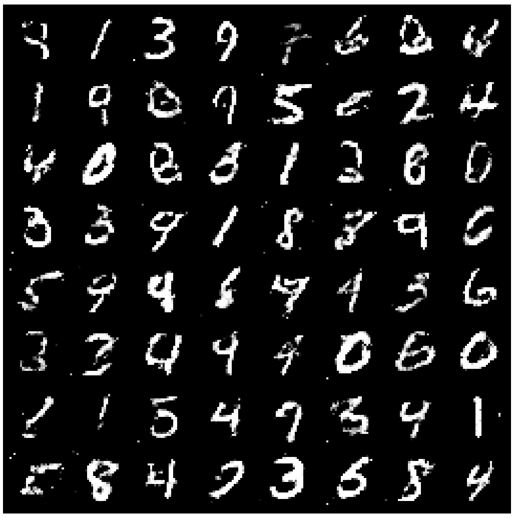


Figure 5: MNIST Display After 100 Epochs GAN Training

2.3 Results and Conclusions

From Figure 4(left) shown above, the generator loss curve seems to increase in general trend whereas the discriminator loss curve decreases rapidly before 20th epoch. However, at the beginning of training, discriminator loss would rise, and generator loss would decrease, but after several epochs, generator loss and discriminator loss would gradually increase and decrease stably. The Figure 4 (right) shows that JSD Value plot before the first five to eight epochs, the curve increases, but then it keeps going down and stop lower than 100 when it reaches the 100th epoch.

3. Model WGAN

3.1 Parameters Settings

WGAN

Optimizer: RMSprop

Batch Size = 64

n_critic = 2

n_noise = 100

Learning Rate = 0.0005

Epoch = 50

WGAN-GP

Optimizer: Adam

Batch Size = 64

n_critic = 5

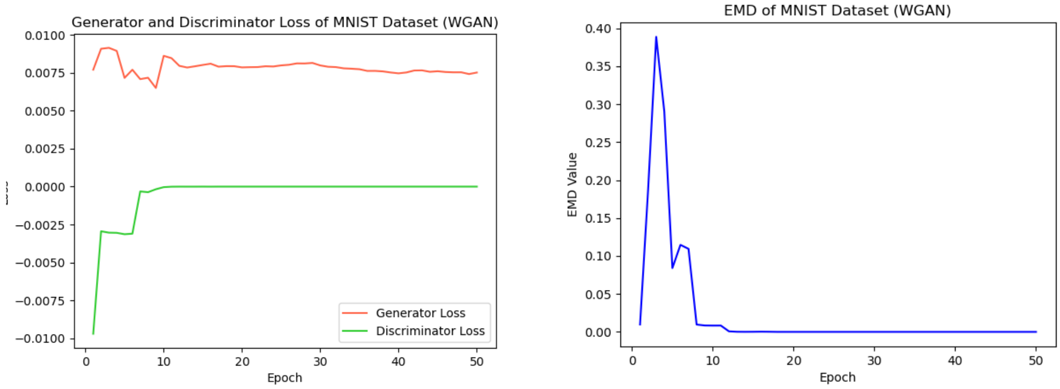
n_noise = 100

p_coeff = 10

Learning Rate = 0.0001

Epoch = 50

3.2 Images and Result Plot Show

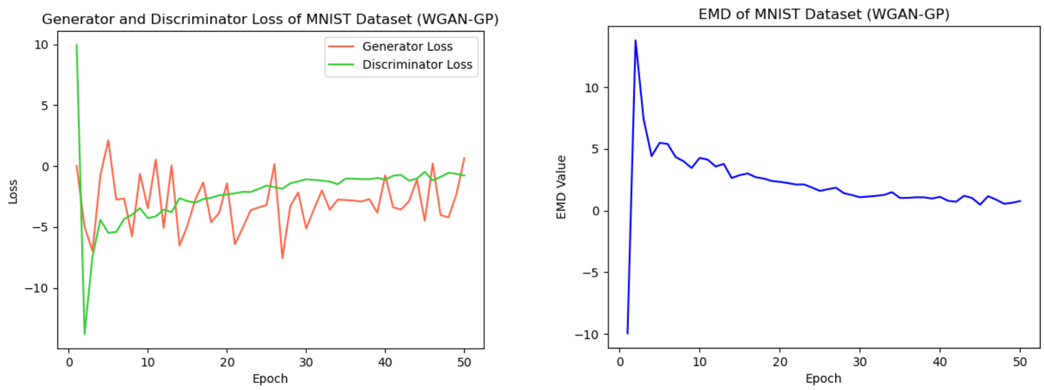


Left: Generator and Discriminator Loss of MNIST (WGAN)

Right: EM Distance Value of MNIST (WGAN)

Figure 6: Plot Results of MNIST for Model WGAN

Figure 7: MNIST Display After 50 Epochs WGAN Training



Left: Generator and Discriminator Loss of MNIST (WGAN-GP)

Right: EM Distance Value of MNIST (WGAN-GP)

Figure 8: Plot Result of MNIST for Model WGAN-GP

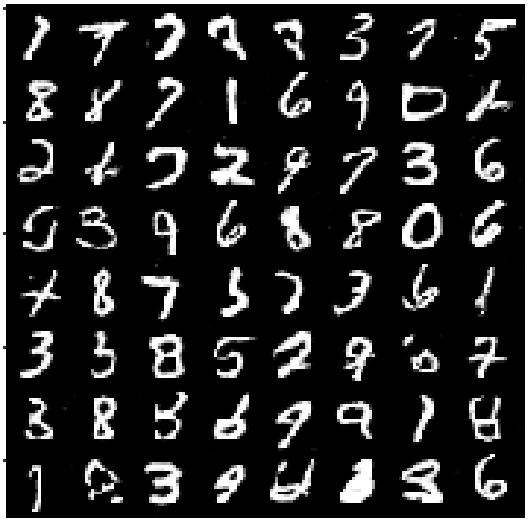


Figure 9: MNIST Display After 50 Epochs WGAN-GP Training

3.3 Results and Conclusions

The WGAN model has been training for a while now, but the output is essentially noise, and it seems to be a case of convergence failure as Figure 6 shows above. Over time, my generator loss gets more and more negative while my discriminator loss remains around -0.003. I guess it is probably because the discriminator is not improving enough, the generator does not get improve enough.

WGAN-GP enhances training stability. As Figure 8 shown below, when the model design is less optimal, WGAN-GP can still create good results while the original GAN cost function fails.

Part III: Models for CIFAR10 Dataset

1. Model VAE

1.1 Parameters Settings

Loss Function: Binary Cross Entropy Loss.

Reconstruction + KL divergence losses summed over all elements and batch.

Optimizer: Adam

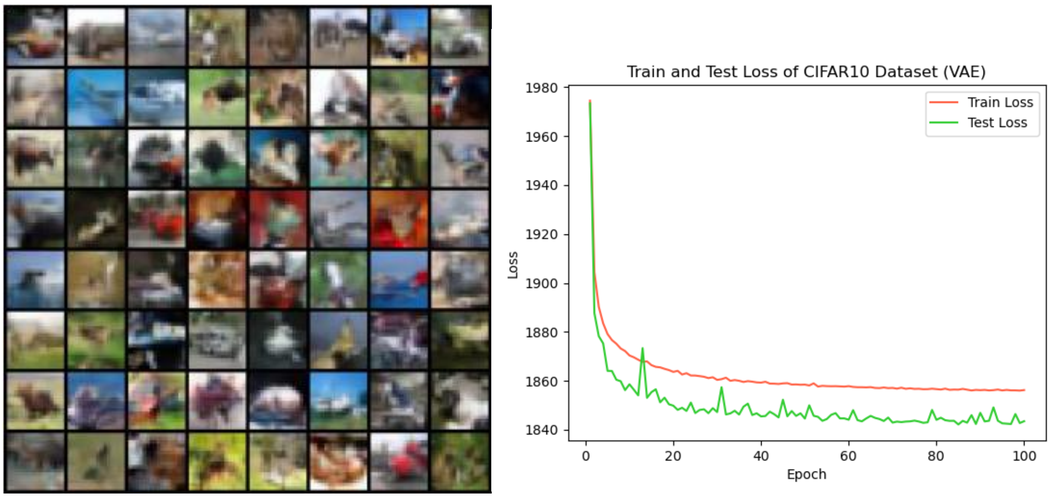
d = 20

Batch Size = 256

Learning Rate = 0.001

Epoch = 100

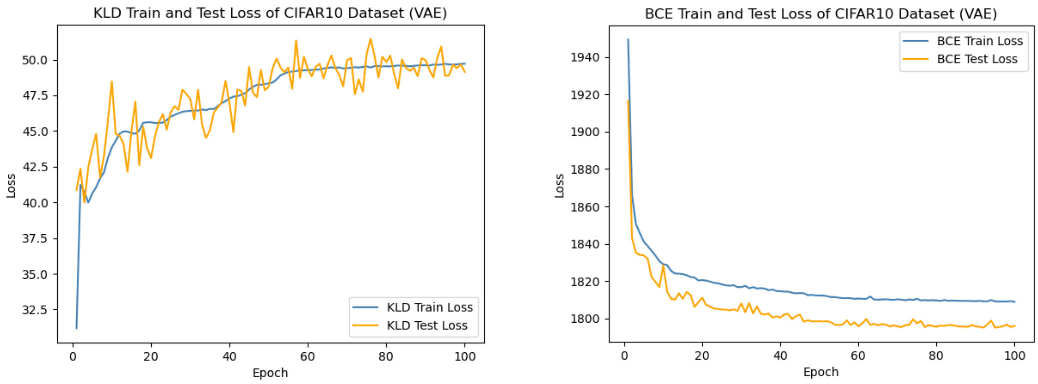
1.2 Images and Result Plot Show



Left: CIFAR10 Display After 100 Epochs VAE Training

Right: Train and Test Loss of CIFAR10 (VAE)

Figure 10: Image Display and Plot Result of CIFAR10 for Model VAE



Left: KLD Train and Test Loss of CIFAR10 (VAE)

Right: BCE Train and Test Loss of CIFAR10 (VAE)

Figure 11: KLD & BCE Plot Result of CIFAR10 for Model VAE

1.3 Results and Conclusions

As Figure 10 and 11 displayed above, different from MNIST dataset, the train loss and test loss of VAE for CIFAR10 dataset start form the similar loss value. They both decrease sharply at the first 15 epochs, then keep stable for the rest epochs. But the test loss is relatively lower than the train loss. Although the test loss is low, it fluctuates a lot, even each fluctuation is small. The train loss is much more stable.

2. Model GAN

2.1 Parameters Settings

Loss Function: Binary Cross Entropy Loss

Optimizer: Adam

Input Dim = 100

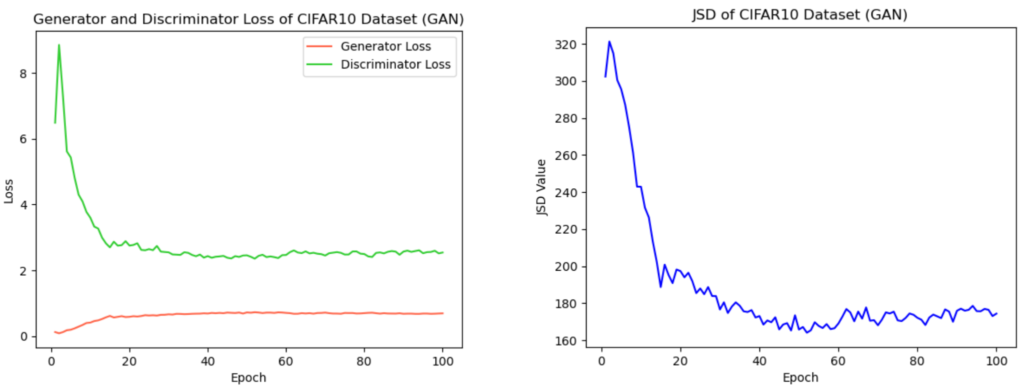
Batch Size = 100

CIFAR10 Dim = 3 * 32 * 32

Learning Rate = 0.0002

Epoch = 100

2.2 Images and Result Plot Show



Left: Generator and Discriminator Loss of MNIST (GAN) **Right:** JSD Value of CIFAR10 (GAN)

Figure 12: Plot Results of CIFAR10 for Model GAN

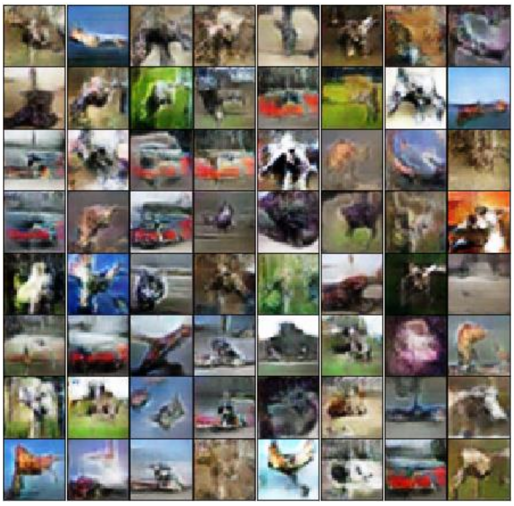


Figure 13: CIFAR10 Display After 100 Epochs GAN Training

2.3 Results and Conclusions

In the GAN model, from the figure 12 shown above, the two loss curves seem more stable compared to MNIST dataset. The discriminator loss goes up to the climax and then keep going down and remain stable, it remains stable. For generator loss, the curve does not too fluctuated, it seems like always remains under 1. But the curve on the right is not as smooth as the JSD curve of MNIST and it has been fluctuating slightly from epoch 20 to 100.

3. Model WGAN

3.1 Parameters Settings

Optimizer: RMSprop

Batch Size = 128

Latent Vector Length = 128 and 512

Sample Size = 64

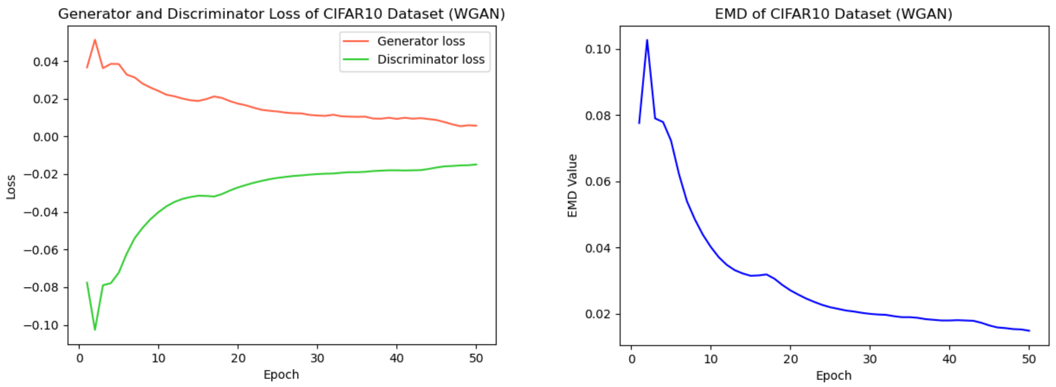
K = 5

Learning Rate = 0.0005

Weight Clip = 0.01

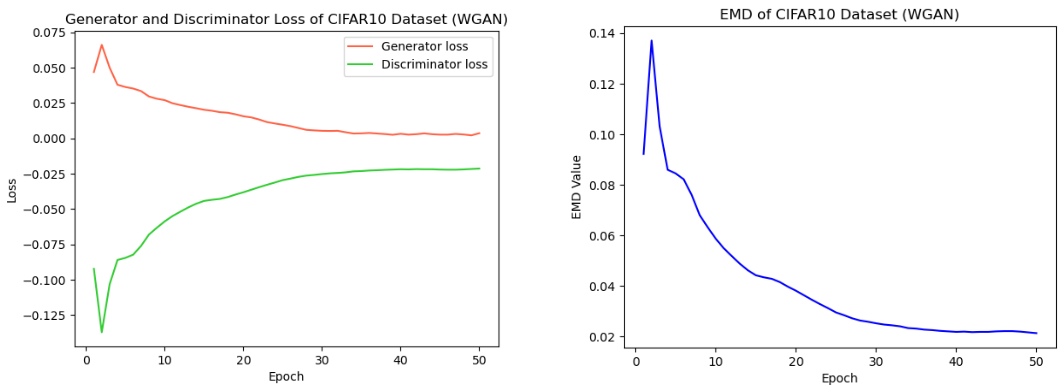
Epoch = 50

3.2 Images and Result Plot Show



Left: Generator and Discriminator Loss of CIFAR10 (WGAN) **Right:** EM Distance Value of CIFAR10 (WGAN)

Figure 14: Plot Results of CIFAR10 for WGAN with Latent Vector Dimension 128



Left: Generator and Discriminator Loss of CIFAR10 (WGAN) **Right:** EM Distance Value of CIFAR10 (WGAN)

Figure 15: Plot Results of CIFAR10 for WGAN with Latent Vector Dimension 512

Figure 16: CIFAR10 Display After 50 Epochs WGAN Training

3.3 Results and Conclusions

For WGAN model of CIFAR10 dataset, the latent vector dimension is 128 and 512 as shown above, also, the Figure 14 and 15 present the generator and discrimination losses change over 100 epochs with training the critic 5 times. The discriminator and generator losses are quite different from dataset MNIST. The curve in CIFAR10 dataset looks more rounded, while the curve in the MNIST dataset looks straighter. The EM Distance increases a little before the first three of five epochs, then it keeps decreasing as Figure 14 and 15 on the right shown.