CSI 5138
Introduction: Deep Learning & RL

Homework 4

Date: 2020/12/23

NAME: Yunfeng Li
Student Number: 300139547

Experiment Description

The purpose of this experiment is to implement three models for MNIST and CIFAR-10 datasets. The experiment is carried out per the procedures below.

1. Design the VAE model for MNIST and CIFAR-10 datasets.
2. Design the GAN model for MNIST and CIFAR-10 datasets.
3. Design the WGAN model for MNIST and CIFAR-10 datasets.
4. Summarize results.

Results Analysis

This section will be separated into three parts by different generative models. Since the MNIST dataset is more efficient than the CIFAR-10 dataset in the training process, this report will talk more in the MNIST dataset. In order to optimize results, each model will tune two hyperparameters which are latent dimensions and model complexity.

VAE

In the following content, the VAE model is used to generate images by the MNIST and CIFAR-10 dataset. In the VAE, the model complexity is defined as Dense layers in the input of encoder and decoder.

MNIST

The structure of the VAE model based on the MNIST dataset is shown below. And the detail layers content of encoder and decoder is described in the files with the report.

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
input_1 (InputLayer)         [(None, 784)]             0
_________________________________________________________________
model (Functional)           [(None, 2), (None, 2), (N 403972
_________________________________________________________________
model_1 (Functional)         (None, 784)               403728
=================================================================
Total params: 807,700
Trainable params: 807,700
Non-trainable params: 0
```

Figure 1. VAE Layers

First, this model is developed with the latent dimension (2) and the model complexity (512). The loss and images gallery are shown below.

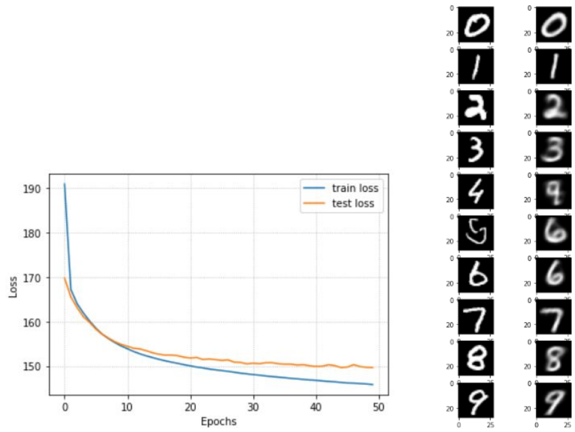
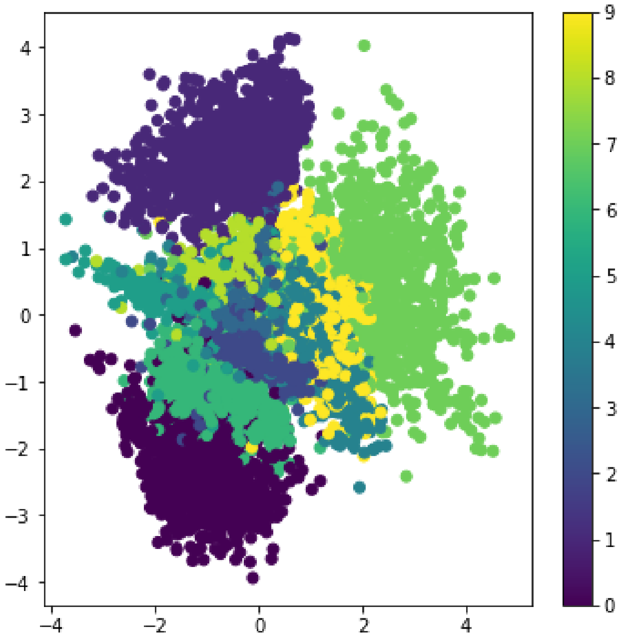


Figure 2. VAE Loss and Images Gallery (latent space:2, model complexity:512)

Tuning hyperparameters

The model complexity and latent dimension will be tuned in the VAE model. First, we will explore the impact of latent dimension and followed by model complexity.

Latent Dimension



To examine the impact of latent space, the model complexity is applied to the same value, which is 512. And we try three latent dimensions, here are the results of 10 and 20 latent dimensions, and the original model above is implemented which the latent dimension is 2.

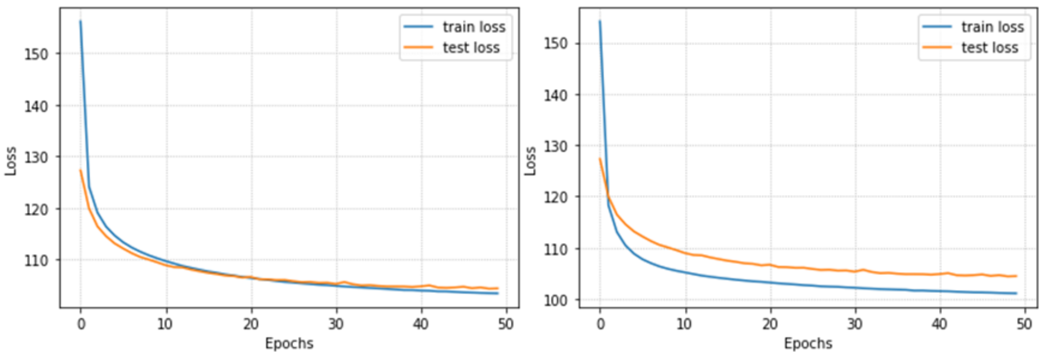


Figure 3. Loss of VAE Latent Dimension: 10(Left), 20(Right)

The loss of different latent dimensions is plotted above. From the image, we can find that the loss will be less with the latent space. And finally, higher latent space will have a lower loss result. The images gallery is also presented below.

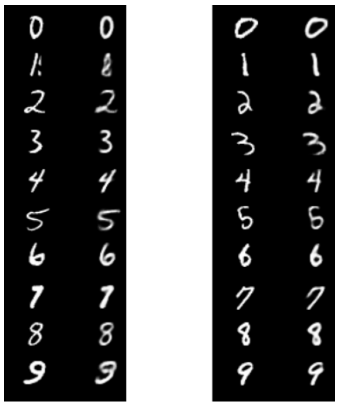


Figure 4. Images Gallery Latent Space:10(Left), 20(Right)

While we only study the value change of loss in the loss diagram, the images gallery can give a direct observation of generative results. According to the gallery, it can be found that high latent space can make images sharper than low latent space, since with 20 latent space, the image gallery is clearer.

Model Complexity

The latent space is fixed in this part, which is 2. The following diagram is from two model complexity values which are 256 and 1024. The above original model's model complexity is 512.

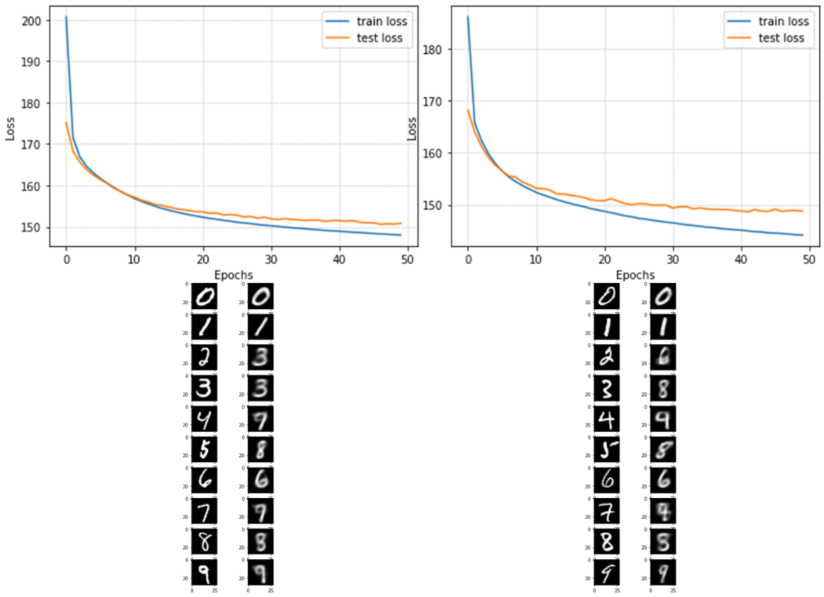


Figure 5. Results of Tuning Model Complexity 256(left) 1024(Right)

Overall, we can find that model complexity affects the result of the VAE model. When the model complexity is 1024, the loss will be lowest among other models. However, increasing the model complexity cannot increase the quality of generated images, and some results are even worse.

CIFAR-10

The structure of the VAE model based on MNIST dataset is shown below. And the detail layers content of encoder and decoder is described in the files with the report.

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
input_1 (InputLayer)         [(None, 3072)]            0
_________________________________________________________________
model (Functional)           [(None, 2), (None, 2), (N 1575428
_________________________________________________________________
model_1 (Functional)         (None, 3072)              1577472
=================================================================
Total params: 3,152,900
Trainable params: 3,152,900
Non-trainable params: 0
_________________________________________________________________
```

Figure 6. VAE Layers

First, this model is developed with the latent dimension (2) and the model complexity (512). The loss and images gallery are shown below.

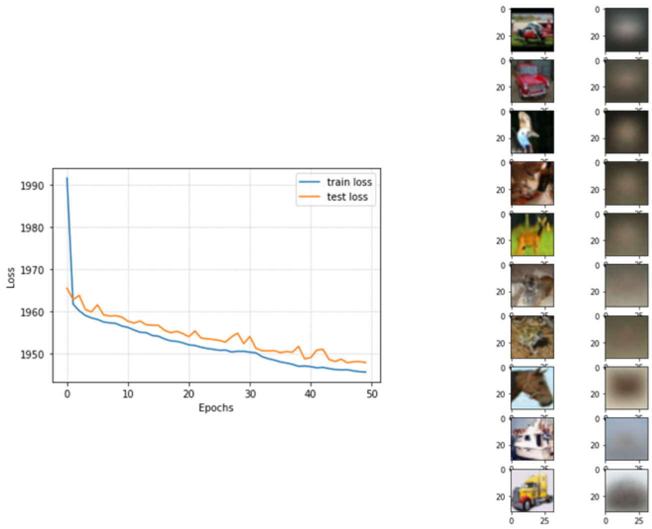


Figure 7. Original Results of VAE model (latent space:2, model complexity:512)

Tuning hyperparameters

The model complexity and latent dimension will be tuned in the VAE model. First, we will explore the impact of latent dimension and followed by model complexity.

Latent Dimension

Like the MNIST dataset, the latent space value will be set with 10 and 20 to observe the results. Meanwhile the model complexity is 512. The results are shown below.

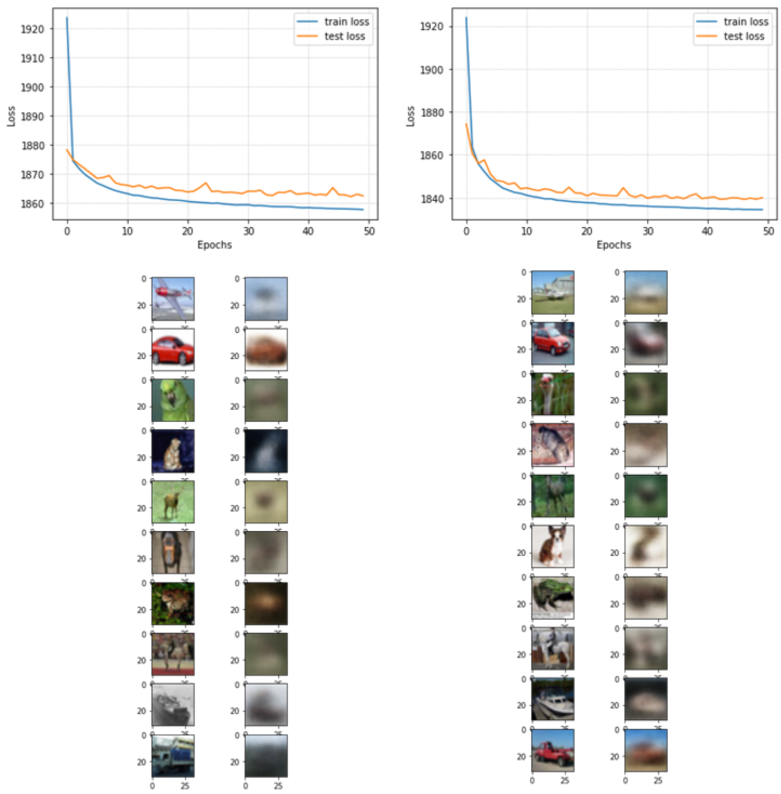


Figure 8. Results of Tuning Latent Space 10(Left) 20(Right)

From the diagram, the results are worse than the MNIST dataset. The loss value is more than 1000. But it is still approved that higher latent space will make the loss lower. And the quality of generated pictures is better when the latent space is larger.

Model Complexity

The model complexity of CIFAR-10 is also similar to MNIST dataset. 256 and 1024 are examined in this part, while keeping the latent space as 2.

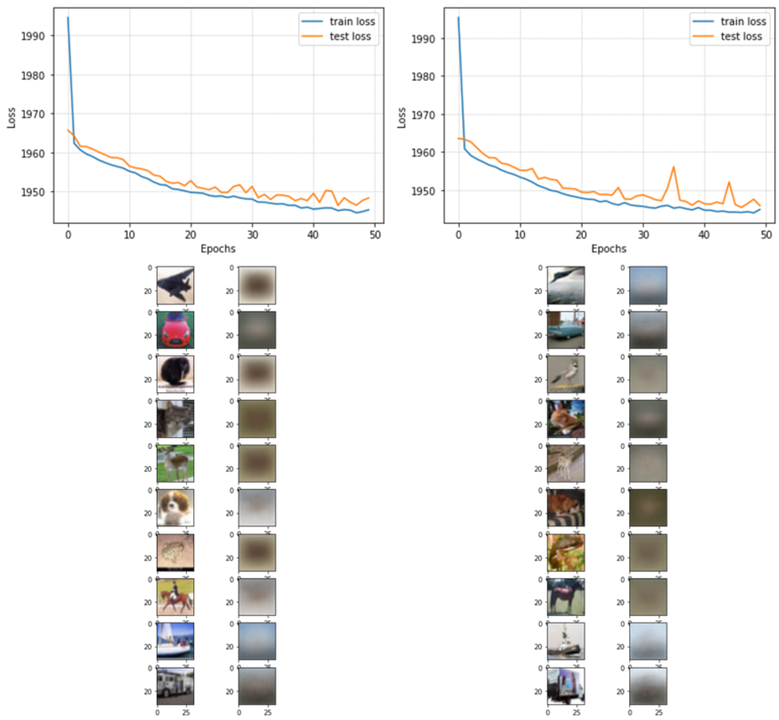


Figure 9. Results of Tuning Model Complexity 256(Left) 1024(Right)

According to the diagram, the model complexity only has few impacts on the loss and images' quality. These results are different from MNIST dataset. The image format may be the main reason.

GAN

For the GAN model, in this experiment, the JSD loss and image quality are studied. First, varying the latent dimensions will be explored and then we also observe how model complexity influences the model. In the GAN, the model complexity is defined as the dimension of the input in the dense layer of the generator.

MNIST Tuning Hyperparameters

Latent Dimension

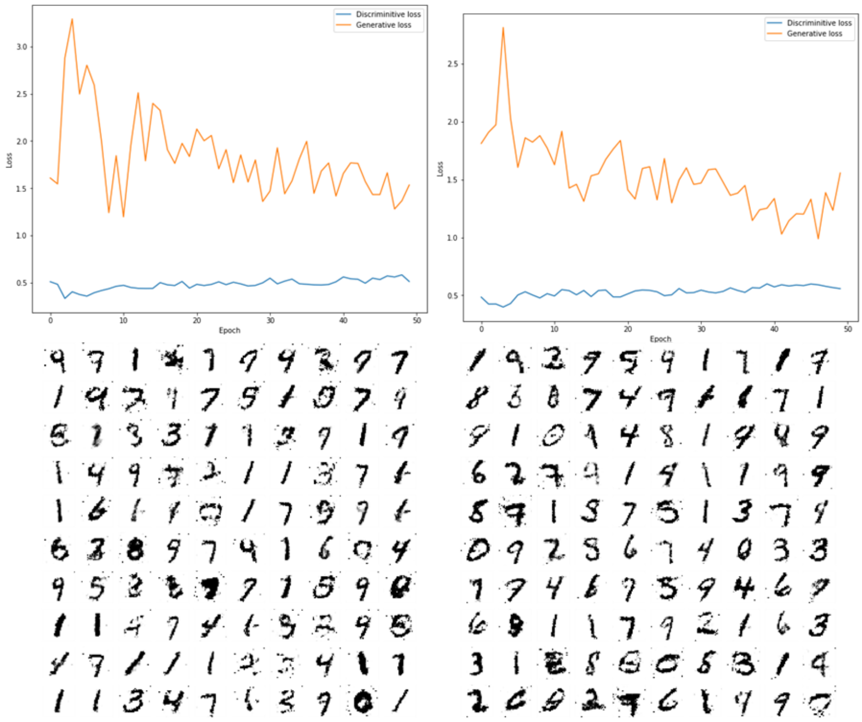


Figure 10. Latent Dimension 10(Left), 100(Right)

The diagram above is the results varying the latent dimensions. The images on the left are 10 latent dimensions and the right are 100.

Model Complexity

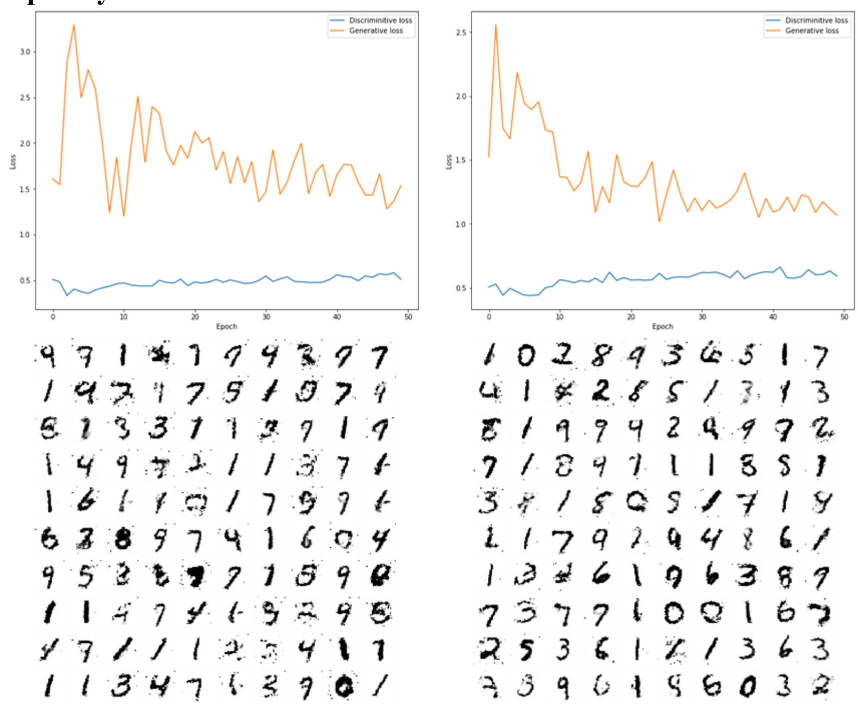


Figure 11. Model Complexity 256(Left) 1024(Right)

The images above are results varying the model complexity. The latent space is fixed to 10, the left is the complexity of 256, the right is the complexity of 1024.

CIFAR-10 Tuning Hyperparameters

Latent Dimension

The latent dimension is studied below, with the same configuration as MNIST dataset.

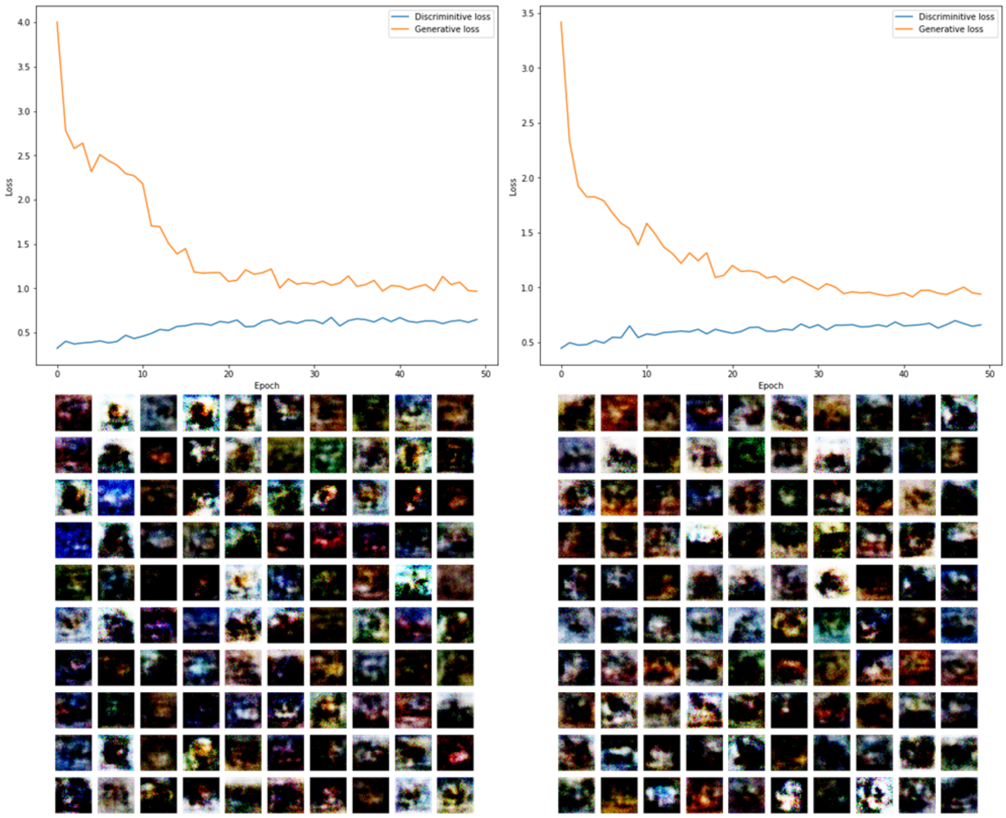


Figure 12. Latent Space 10(Left), 100(Right)

Model Complexity

The model complexity is analyzed below. The latent space is set as 100. We try two model complexity, which are 256 and 1024.

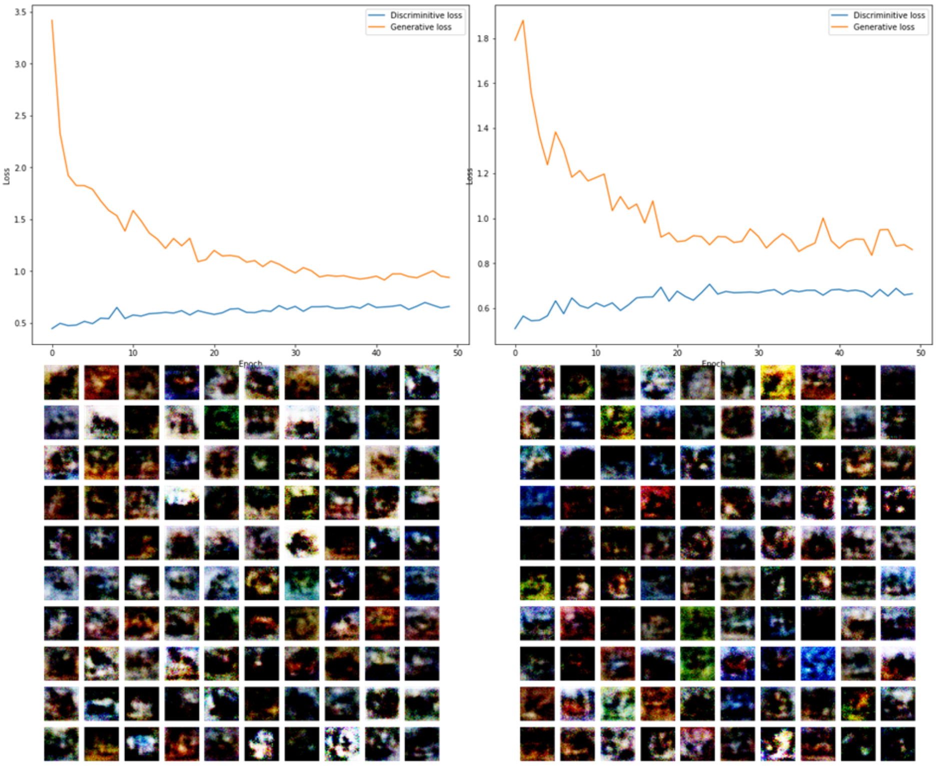


Figure 13. Model Complexity 256(Left) 1024(Right)

WGAN

From the WAGN model, the EMD loss and image quality are studied. First, varying the latent dimensions will be explored and then we also observe how model complexity influences the model. In the WGAN, the model complexity is defined as the dimension of the input in the dense layer of the generator.

MNIST Tuning Hyperparameters

Latent Dimension

The latent dimension is examined with two values which are 10 and 100, while the model complexity is 128.

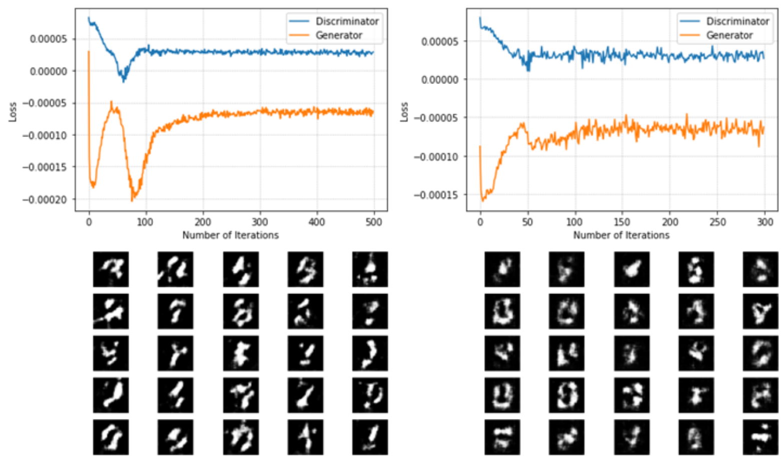


Figure 14. Latent Dimension 10(Left) 100(Right)

Model Complexity
The model complexity is studied with fixed latent space. The results are shown below.

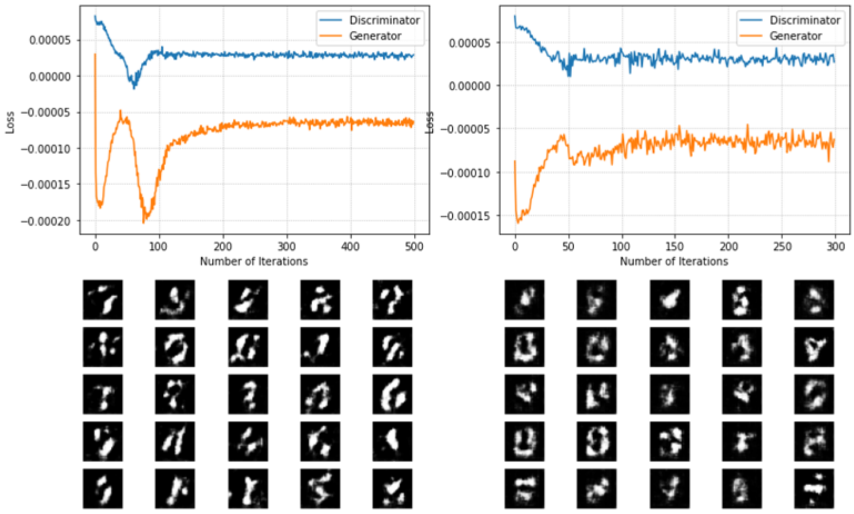


Figure 15. Model Complexity 128(Left) 1024(Right)

CIFAR-10 Tuning Hyperparameters
Latent Dimension
The latent dimension is examined with two values which are 10 and 100, while the model complexity is 128.

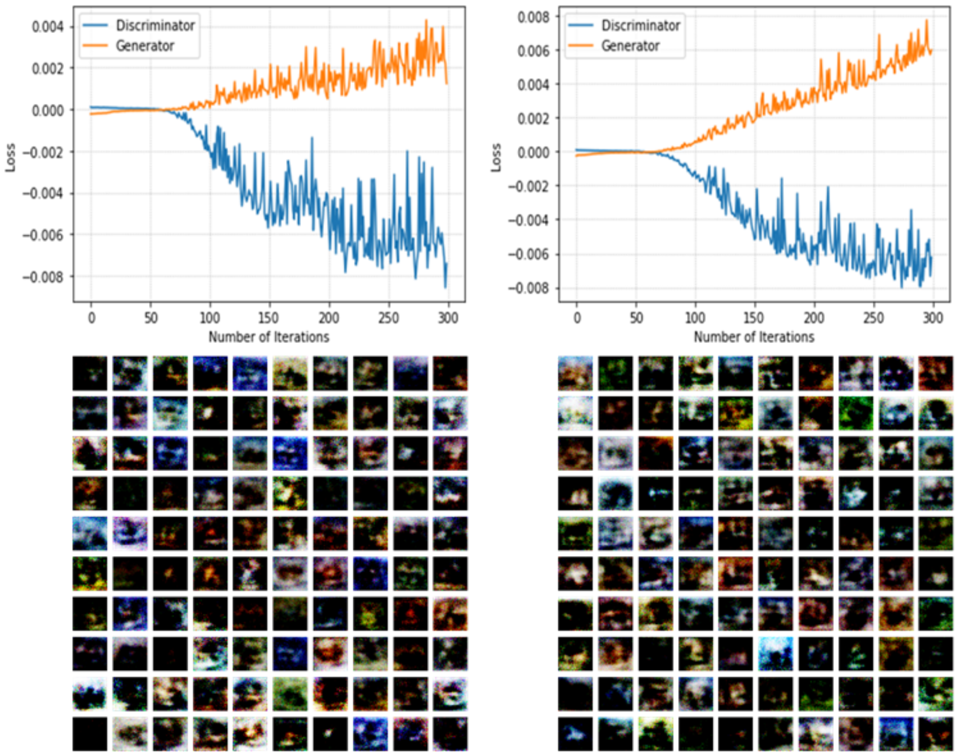


Figure 16. Latent Space 10(Left) 100(Right)

Model Complexity
The model complexity is studied with fixed latent space. The results are shown below.

Figure 17. Model Complexity 128(Left) 512(Right)

Conclusion

To summarize, this experiment is carried out with three models for MNIST and CIFAR-10 dataset. The VAE, GAN, and WGAN are designed by considering different datasets. Overall, some conclusions can be drawn. Firstly, the dataset has a significant impact on models. The basic data format like MNIST which only has one channel, can generate better images. Second, hyperparameters will influence results. This will be described separately below.

1. For VAE models, the latent space and model complexity both have an obvious impact on generated results. It can be observed from loss and image gallery. In the MNIST dataset, increasing the latent space can decrease the loss, which has the same result with increasing the model complexity. Besides, the model with higher latent space can generate shaper images. In contrast, when the model complexity becomes larger, the generated results are worse. Some digits even have wrong results. Same results can also be proved by the CIFAR-10 dataset.
2. For GAN models, from the results generated and the JSD loss, the latent space and model complexity only have minimal impact. Increasing the latent space can have a shape image both in MNIST and CIFAR-10 dataset. However, the loss has not changed significantly. This conclusion can also be applied to model complexity.
3. For WGAN models, models of both MNIST and CIFAR-10 have a satisfactory performance.

It can be observed that increasing the model complexity and latent space cannot make results better. Although the EMD loss is still good from the diagram above. The reason for this may be the training epochs are not enough. Moreover, layers of both generator and discriminator are also too basic. And in the future work, increasing model layers and training epochs will be the main concern.