Course Title: CPSC-8430: Deep Learning - Homework 4

Student: Tharun Kumar Vaspari (tvaspar@clemson.edu)

GitHub Repository:

Access the homework code at this link.

https://github.com/VASPARITHARUNKUMAR/CPSC-8430---Deep-Learning---HW4

A Generative Adversarial Network (GAN) is made up of two main parts: the generator and the discriminator. The goal of training a GAN is to improve the generator's capability to create realistic images that can deceive the discriminator, which is trained to distinguish between authentic and generated images.

The main purpose of GAN training is to produce high-quality images from a given dataset, such as CIFAR-10.

Dataset Overview:

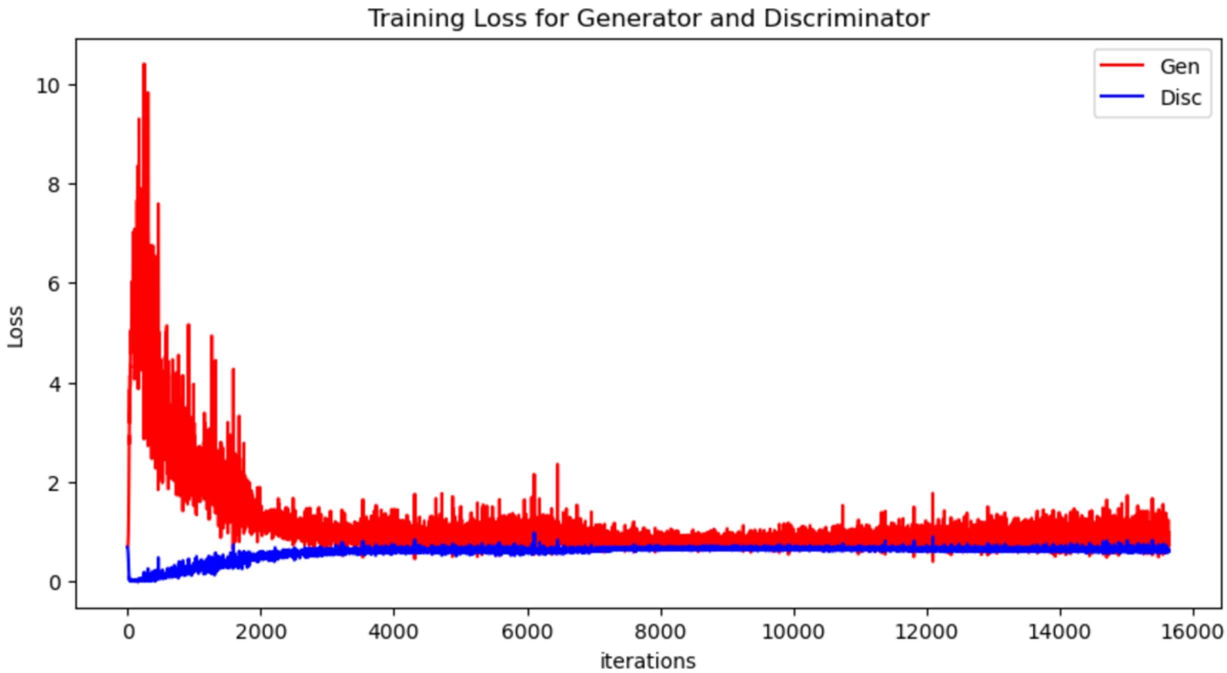
The CIFAR-10 dataset holds 60,000 32x32 color images that have been classified into 10 various categories. There are 6,000 photos in each class, 10,000 of which are used for testing and 50,000 for training. Five training batches and one testing batch make up the dataset's structure. The classes include Dog, frog, truck, cat, ship, bird, horse, airplane, automobile, deer, making it a widely used benchmark for image classification in computer vision research.

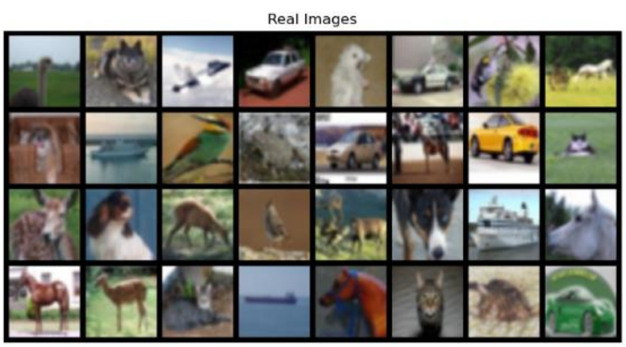
DCGAN:

The Deep Convolutional Generative Adversarial Network (DCGAN) is a well-known and effective variant of GAN, where both the generator and discriminator networks utilize convolutional layers. Unlike traditional models that rely on fully connected layers or pooling, DCGAN uses convolutional strides and transposed convolutions for downsampling and upsampling. This design makes DCGAN highly adaptable and enables the generation of high-quality images without the need for complex generator architectures.

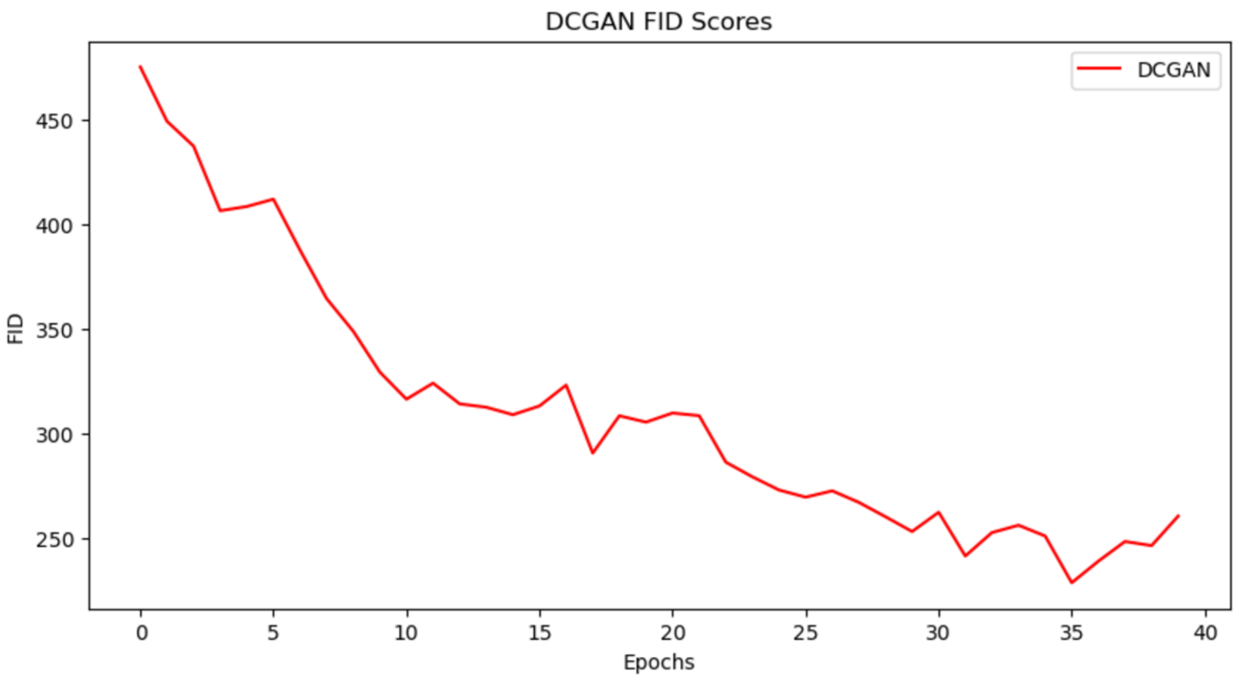
Parameters:

- Num_Epochs = 40
- Noise_Dim = 100
- Channels_Img = 3
- Image_Size = 64
- Batch_Size = 128
- Learning_Rate = 5e-5


Training Loss for Generator and Discriminator
Gen
Disc
Loss
iterations
0
2
4
6
8
10
0
2000
4000
6000
8000
10000
12000
14000
16000


Real Images


Fake Images



Results:

- The generator and discriminator training losses stabilized at 0.746 and 0.627, respectively.
- The generator achieved a minimum loss of 0.401, while the discriminator reached a minimum loss of 0.00194.
- The Fréchet Inception Distance (FID) score varied between 228.508 and 475.327, with an average of 309.993.
- The final FID score of 260.418 reflects a satisfactory level of image quality.

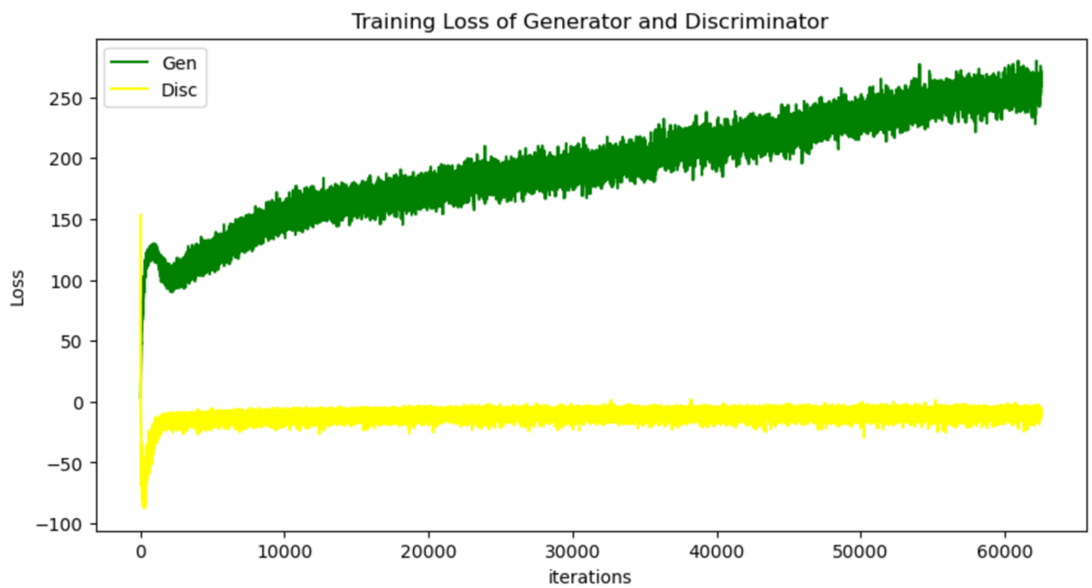
WGAN:

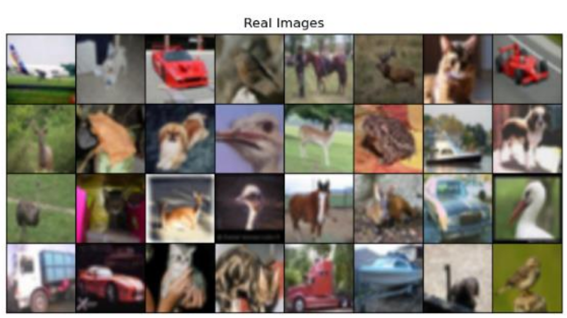
The Wasserstein GAN (WGAN) is a type of generative model that improves training stability by using a loss function closely tied to image quality. WGAN builds on robust mathematical foundations and requires only small adjustments to the DCGAN framework. The discriminator in WGAN uses a linear activation function in its output layer rather than sigmoid, with real images scaled to -1 and generated images to 1. Training typically involves a low learning rate without momentum, and the discriminator is updated more often than the generator. Additionally, WGAN leverages Wasserstein distance for optimizing both the generator and discriminator, which requires applying weight constraints.

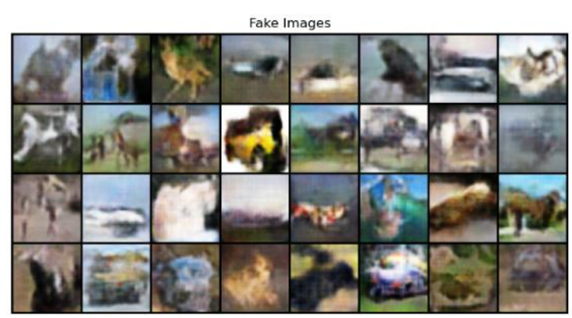
Parameters:

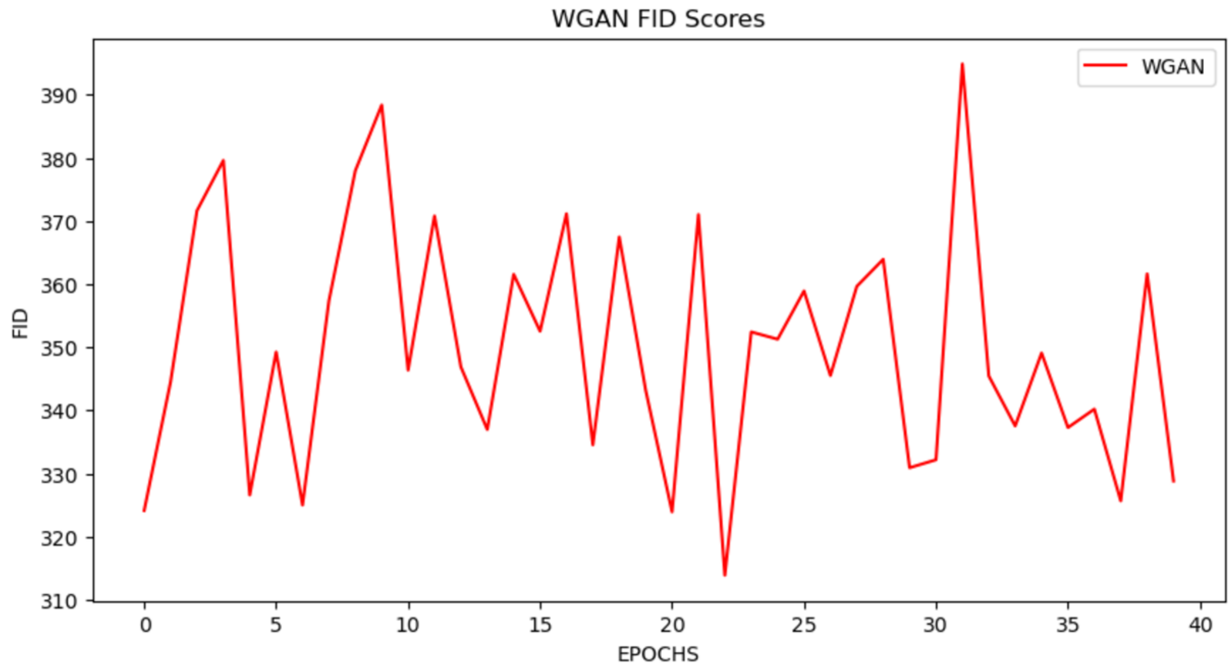
- Num_Epochs = 40
- Noise_Dim = 100

- Channels_Img = 3
- Learning_Rate = 5e-5
- Image_Size = 64
- Batch_Size = 32









Results:

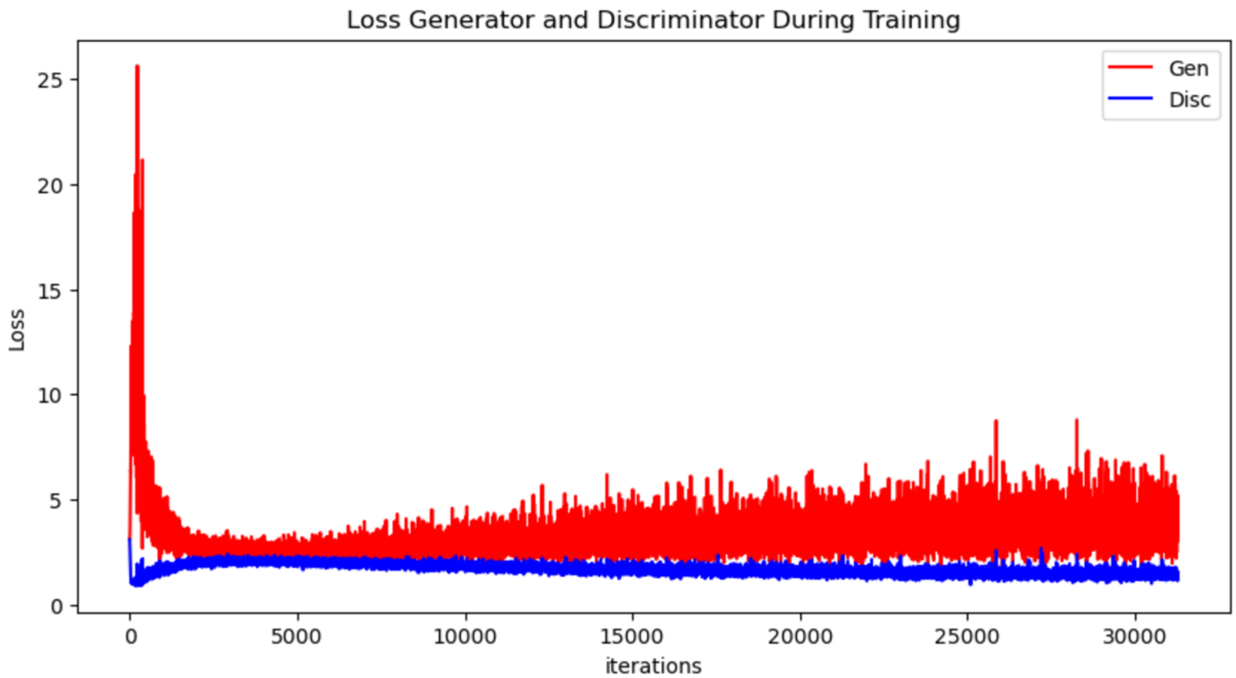
- The generator and discriminator training losses converged to 3.254 and -12.778, respectively.
- The generator's lowest loss was 3.254, while the discriminator's lowest loss was -87.503.
- The Fréchet Inception Distance (FID) score ranged between 316.050 and 387.832, with an average of 346.103.
- The final FID score of 327.162 suggests an acceptable level of image quality.

ACGAN:

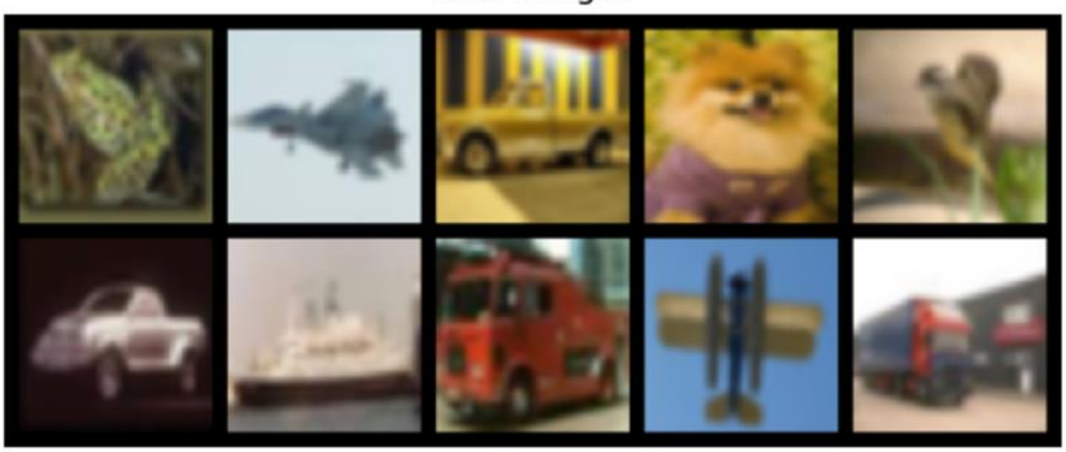
The Auxiliary Classifier GAN (AC-GAN) is a type of conditional GAN where the discriminator predicts the class label of an input image instead of taking it as input. This method enables the generation of large, high-quality images while learning a latent space representation that remains independent of class labels. Additionally, AC-GAN improves the stability of the training process. Similar to conditional GANs, AC-GAN generates images conditionally by feeding the generator a latent space point along with a class label.

Parameters:

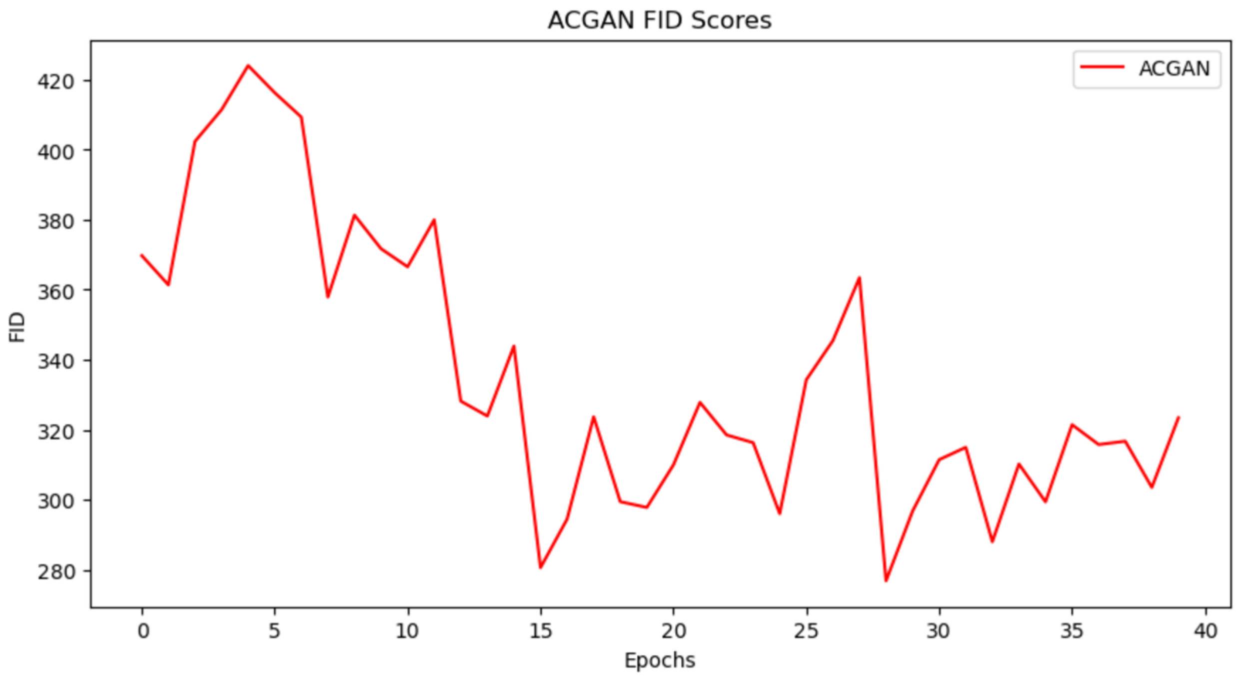
- Num_Epochs = 40
- Noise_Dim = 100
- Channels_Img = 3
- Learning_Rate = 5e-5
- Image_Size = 64
- Batch_Size = 64



Real Images



Fake Images



Results:

- The generator and discriminator training losses stabilized at 3.125 and 1.386, respectively.
- The generator's minimum loss was 1.579, while the discriminator's minimum loss was 0.884.
- The Fréchet Inception Distance (FID) score varied between 276.839 and 424.045, with an average of 335.863.
- The final FID score of 321.444 reflects a decent level of image quality.

Conclusion:

Among the three models, DCGAN outperforms WGAN and ACGAN, achieving the most stable training, the lowest losses, and the best image quality with the lowest FID scores. While ACGAN performs moderately well, WGAN shows instability and lower output quality in comparison.