CPSC – 8430 – Deep Learning

Homework 4 Report

Anupriya Dominic |CU13338590| https://github.com/anuprid/Deep-Learning-HW4

Introduction:

Generative Adversarial Networks (GANs) have emerged as one of the most powerful and influential frameworks for image generation. GANs consist of two neural networks a generator and a discriminator that compete in a minimax game.
The generator attempts to produce realistic images from random noise, while the discriminator tries to distinguish real images from generated ones. Through this adversarial training process, the generator gradually learns the underlying data distribution and becomes capable of producing high-quality synthetic images.

The objective of this homework is to implement, train, and evaluate three major variants of GANs on the CIFAR-10 dataset, a widely used benchmark containing 60,000 32×32 color images across 10 object categories. The models explored in this report represent different stages of GAN evolution:

1. DCGAN (Deep Convolutional GAN)
2. WGAN (Wasserstein GAN with gradient penalty)
3. ACGAN (Auxiliary Classifier GAN)

Models:

DCGAN: It is a popular and influential version of GANs that uses deep convolutional networks to make image generation more stable and realistic. Instead of relying on fully connected layers, DCGAN builds the generator and discriminator using convolutional and transposed-convolutional layers, which helps the model understand and reproduce spatial patterns like edges, textures, and shapes. The generator starts with random noise and gradually up-samples it using BatchNorm and ReLU activations until it forms a full image, while the discriminator does the opposite it down-samples an image to decide whether it looks real or fake. These design choices make DCGAN much more reliable and capable of producing visually coherent images.

The discriminator is a mirrored convolutional network that progressively down-samples the input image using strided Conv2d layers and LeakyReLU activations, ending in a single sigmoid output for prediction. All convolutional and batch-norm weights are initialized from a normal distribution. Both networks are trained from scratch using the Adam optimizer with learning rate and binary cross-entropy loss, batch size 128, and 50 training epochs.

```
Epoch [1/50]Batch 390/391                D-loss: 0.147, D-loss: 3.424 FID:393.684
Epoch [2/50]Batch 390/391                D-loss: 0.284, D-loss: 2.541 FID:425.883
Epoch [3/50]Batch 390/391                D-loss: 0.214, D-loss: 1.907 FID:330.557
Epoch [4/50]Batch 390/391                D-loss: 0.599, D-loss: 0.893 FID:431.928
Epoch [5/50]Batch 390/391                D-loss: 0.448, D-loss: 1.325 FID:423.095
Epoch [6/50]Batch 390/391                D-loss: 0.616, D-loss: 1.231 FID:327.111
Epoch [7/50]Batch 390/391                D-loss: 0.532, D-loss: 1.090 FID:322.796
Epoch [8/50]Batch 390/391                D-loss: 0.601, D-loss: 0.963 FID:295.585
Epoch [9/50]Batch 390/391                D-loss: 0.608, D-loss: 0.808 FID:290.894
Epoch [10/50]Batch 390/391                D-loss: 0.610, D-loss: 0.936 FID:305.616
Epoch [11/50]Batch 390/391                D-loss: 0.690, D-loss: 0.697 FID:283.014
Epoch [12/50]Batch 390/391                D-loss: 0.522, D-loss: 1.103 FID:271.906
Epoch [13/50]Batch 390/391                D-loss: 0.587, D-loss: 1.010 FID:287.855
Epoch [14/50]Batch 390/391                D-loss: 0.606, D-loss: 0.809 FID:297.544
Epoch [15/50]Batch 390/391                D-loss: 0.595, D-loss: 1.023 FID:287.290
Epoch [16/50]Batch 390/391                D-loss: 0.600, D-loss: 0.889 FID:265.352
Epoch [17/50]Batch 390/391                D-loss: 0.667, D-loss: 0.779 FID:270.439
Epoch [18/50]Batch 390/391                D-loss: 0.609, D-loss: 0.895 FID:247.625
Epoch [19/50]Batch 390/391                D-loss: 0.643, D-loss: 0.877 FID:255.110
Epoch [20/50]Batch 390/391                D-loss: 0.586, D-loss: 0.830 FID:283.866
Epoch [21/50]Batch 390/391                D-loss: 0.633, D-loss: 0.868 FID:250.288
Epoch [22/50]Batch 390/391                D-loss: 0.600, D-loss: 0.890 FID:258.808
```

Fig 1: DCGAN training logs.

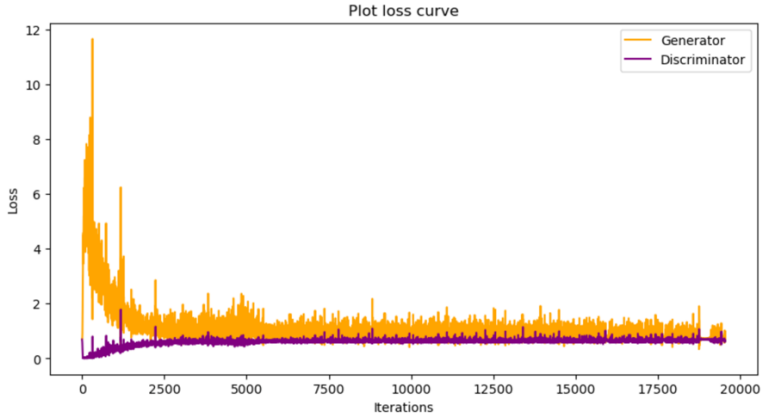

Fig 2: Generator/Discriminator loss curves of DCGAN.

In Fig 2 shows that at the beginning of training, the discriminator loss is low and the generator loss is high, indicating that the discriminator can easily distinguish real from fake images. As training progresses, the generator improves and its loss gradually decreases, while the discriminator loss increases slightly and then stabilizes.

Fig 3: Dataset vs DCGAN Generated images.

WGAN: Wasserstein GAN (WGAN) replaces the standard GAN loss with the Wasserstein distance, which provides smoother gradients and improves training stability. Instead of

predicting probabilities, the discriminator outputs real-valued scores that estimate how real an image looks. The key idea is that minimizing the Wasserstein distance leads to more meaningful updates, reduces mode collapse, and makes training much more stable even when the generator and critic are not balanced. To enforce the Lipschitz constraint required by the Wasserstein formulation, the model uses weight clipping, which prevents the critic from becoming too sharp and destabilizing training. These changes make WGAN one of the most reliable GAN variants for image synthesis.

During training on CIFAR-10, WGAN showed noticeably more stable behavior than DCGAN, even though its losses grow linearly instead of oscillating. This is expected because WGAN losses do not represent classification accuracy, they measure how well the critic separates fake and real distributions. The generated images progressively became clearer and more structured over epochs, and the model captured the dataset's color distribution and object shapes reasonably well.

```
Epoch  [1/50] Batch 781/782          D-loss -24.764, G-loss 108.476, FID Score: 313.902
Epoch  [2/50] Batch 781/782          D-loss -16.173, G-loss 91.110, FID Score: 322.897
Epoch  [3/50] Batch 781/782          D-loss -14.178, G-loss 99.433, FID Score: 367.724
Epoch  [4/50] Batch 781/782          D-loss -14.145, G-loss 123.175, FID Score: 367.914
Epoch  [5/50] Batch 781/782          D-loss -14.736, G-loss 117.928, FID Score: 353.026
Epoch  [6/50] Batch 781/782          D-loss -20.030, G-loss 115.784, FID Score: 348.178
Epoch  [7/50] Batch 781/782          D-loss -21.756, G-loss 131.049, FID Score: 342.336
Epoch  [8/50] Batch 781/782          D-loss -15.719, G-loss 131.475, FID Score: 355.125
Epoch  [9/50] Batch 781/782          D-loss -16.092, G-loss 132.172, FID Score: 388.485
Epoch  [10/50] Batch 781/782          D-loss -16.392, G-loss 143.357, FID Score: 360.140
Epoch  [11/50] Batch 781/782          D-loss -19.116, G-loss 151.700, FID Score: 343.539
Epoch  [12/50] Batch 781/782          D-loss -19.917, G-loss 167.251, FID Score: 381.058
Epoch  [13/50] Batch 781/782          D-loss -28.575, G-loss 142.320, FID Score: 341.536
Epoch  [14/50] Batch 781/782          D-loss -28.177, G-loss 167.054, FID Score: 315.824
Epoch  [15/50] Batch 781/782          D-loss -20.130, G-loss 162.732, FID Score: 338.454
Epoch  [16/50] Batch 781/782          D-loss -26.002, G-loss 185.070, FID Score: 361.365
Epoch  [17/50] Batch 781/782          D-loss -28.830, G-loss 158.303, FID Score: 360.130
Epoch  [18/50] Batch 781/782          D-loss -14.265, G-loss 185.066, FID Score: 324.210
Epoch  [19/50] Batch 781/782          D-loss -21.148, G-loss 178.753, FID Score: 360.528
Epoch  [20/50] Batch 781/782          D-loss -14.365, G-loss 181.058, FID Score: 340.649
Epoch  [21/50] Batch 781/782          D-loss -17.250, G-loss 185.005, FID Score: 318.868
Epoch  [22/50] Batch 781/782          D-loss -16.634, G-loss 190.980, FID Score: 365.729
Epoch  [23/50] Batch 781/782          D-loss -18.983, G-loss 183.375, FID Score: 305.381
Epoch  [24/50] Batch 781/782          D-loss -27.400, G-loss 201.088, FID Score: 360.386
```

Fig 4: WGAN Training Logs.

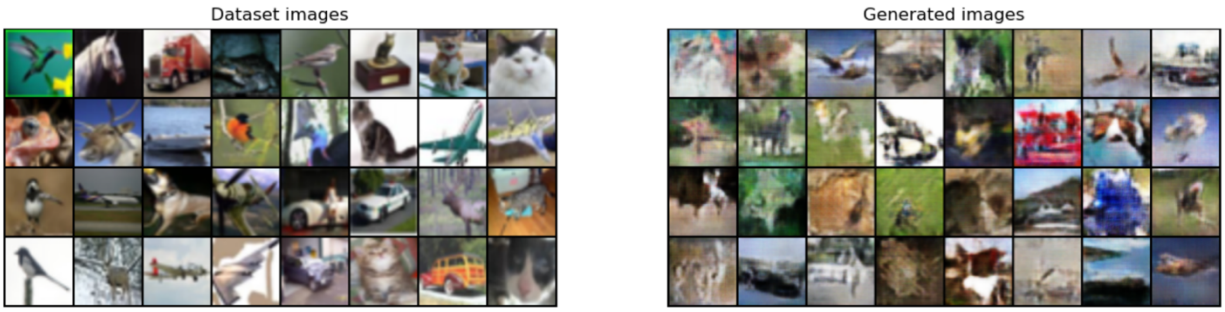

Fig 5: Dataset Images vs. WGAN Generated Images

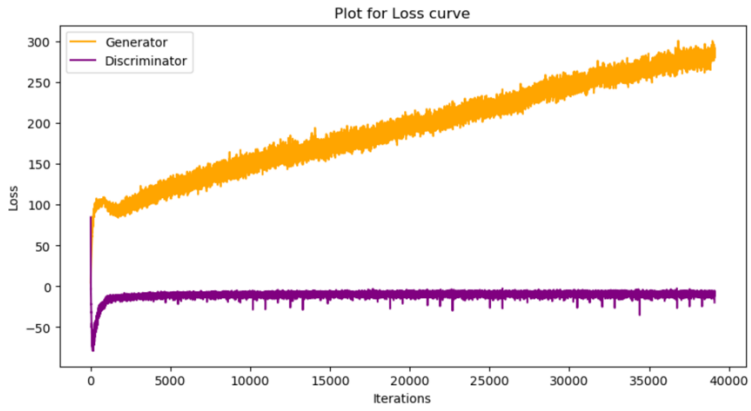

Fig 6: Generator/Discriminator loss curves of WGAN.

The Fig 6 shows that the critic loss becomes more negative as it improves at distinguishing real images from generated ones, while the generator loss increases as the critic applies stronger pressure. The important observation is that the curves are smooth and monotonic rather than unstable and oscillatory.

ACGAN: It extends the standard GAN framework by adding class-conditioning to both the generator and the discriminator. Instead of generating images purely from noise, the ACGAN generator takes both noise and a class label as inputs. This allows the model to intentionally generate images belonging to specific CIFAR-10 categories. The discriminator is also modified, it now produces two outputs, a real or fake prediction and a class label prediction. During training, the discriminator learns to classify real images correctly while also distinguishing generated samplesOverall, ACGAN demonstrates the benefit of conditioning in GANs, especially when dealing with datasets like CIFAR-10 that contain diverse object categories.

```
Epoch [1/50] Batch 781/782          D-loss: 2.271, G-loss: 2.746, FID Score: 455.365
Epoch [2/50] Batch 781/782          D-loss: 2.122, G-loss: 2.920, FID Score: 362.413
Epoch [3/50] Batch 781/782          D-loss: 2.075, G-loss: 2.442, FID Score: 379.845
Epoch [4/50] Batch 781/782          D-loss: 2.107, G-loss: 2.872, FID Score: 383.827
Epoch [5/50] Batch 781/782          D-loss: 1.663, G-loss: 3.238, FID Score: 364.257
Epoch [6/50] Batch 781/782          D-loss: 1.902, G-loss: 2.321, FID Score: 319.671
Epoch [7/50] Batch 781/782          D-loss: 2.155, G-loss: 3.134, FID Score: 337.999
Epoch [8/50] Batch 781/782          D-loss: 2.148, G-loss: 4.123, FID Score: 312.921
Epoch [9/50] Batch 781/782          D-loss: 1.810, G-loss: 3.278, FID Score: 337.809
Epoch [10/50] Batch 781/782          D-loss: 1.789, G-loss: 2.632, FID Score: 385.301
Epoch [11/50] Batch 781/782          D-loss: 1.715, G-loss: 3.033, FID Score: 366.282
Epoch [12/50] Batch 781/782          D-loss: 1.506, G-loss: 4.207, FID Score: 300.242
Epoch [13/50] Batch 781/782          D-loss: 1.575, G-loss: 3.447, FID Score: 336.366
Epoch [14/50] Batch 781/782          D-loss: 1.531, G-loss: 3.012, FID Score: 338.107
Epoch [15/50] Batch 781/782          D-loss: 1.706, G-loss: 3.491, FID Score: 309.404
Epoch [16/50] Batch 781/782          D-loss: 1.480, G-loss: 3.241, FID Score: 345.199
Epoch [17/50] Batch 781/782          D-loss: 1.554, G-loss: 4.330, FID Score: 321.631
Epoch [18/50] Batch 781/782          D-loss: 1.671, G-loss: 4.440, FID Score: 274.167
Epoch [19/50] Batch 781/782          D-loss: 1.706, G-loss: 2.465, FID Score: 321.341
Epoch [20/50] Batch 781/782          D-loss: 1.793, G-loss: 4.193, FID Score: 326.063
Epoch [21/50] Batch 781/782          D-loss: 1.380, G-loss: 3.766, FID Score: 314.368
Epoch [22/50] Batch 781/782          D-loss: 1.393, G-loss: 2.930, FID Score: 302.418
Epoch [23/50] Batch 781/782          D-loss: 1.471, G-loss: 4.803, FID Score: 321.312
```

Fig 7: ACGAN Training Logs.

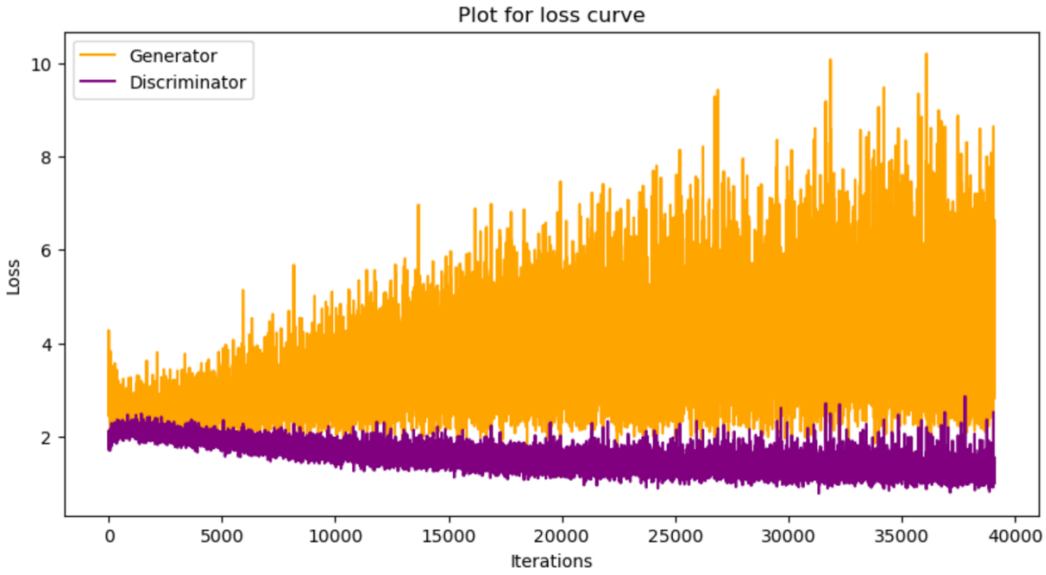

Fig 8: ACGAN Loss Curves.

The Fig 8 shows that the losses do not converge to a fixed value; instead, the generator curve rises and shows variability as the discriminator improves. The discriminator loss remains lower and more stable, suggesting it has a relatively easy time distinguishing real images from generated ones early on.

Fig 9: Real Images vs ACGAN Generated Images.

Comparison:

Training Stability: When comparing the training stability of the three models, each one behaves very differently. DCGAN shows the most unstable behavior, with its generator and discriminator losses constantly going up and down. The early part of training is especially noisy, and the balance between the two networks keeps shifting, which makes DCGAN the least stable overall. WGAN performs the best in terms of stability. Its critic and generator losses move in steady, predictable directions instead of bouncing around, thanks to the Wasserstein loss that provides smoother and more reliable gradients. Throughout training, WGAN shows no signs of collapse and remains consistently well-behaved. ACGAN lands somewhere in the middle. Its discriminator loss stays steady, but the generator loss has

more variation because the model is learning both image realism and class information at the same time. Even with this extra complexity, ACGAN trains reliably and does not collapse. Overall, WGAN is the most stable model, followed by ACGAN, with DCGAN showing the least stable training behavior.

Visual Quality of Generated Images: When looking at the visual quality of the images produced by the three models, clear differences appear. DCGAN generates images that capture the general colors and textures of CIFAR-10, but many samples are still blurry or lack well-defined shapes. Some images look reasonably structured, while others appear noisy or distorted, showing that DCGAN can learn the dataset's style but struggles with fine details. WGAN generally produces cleaner and more consistent images. The textures look smoother, and the samples contain fewer random artifacts, giving the impression of better overall quality, even though the images are still low-resolution. Among the three, ACGAN delivers the strongest sense of structure and meaning. Because it uses class labels during training, many generated images resemble the expected categories more clearly, such as animal-like shapes or vehicle outlines. Overall, ACGAN offers the most semantically meaningful results, WGAN provides the cleanest textures, and DCGAN serves as a solid baseline with decent but less refined outputs.

FID Score Analysis: When comparing the FID scores of the three models, we can see noticeable differences in how well each one matches the statistics of the real CIFAR-10 images. DCGAN achieves the strongest overall performance, with its FID dropping quickly during the early epochs and settling at the lowest value among the three models. This indicates that DCGAN learns the global distribution of the dataset effectively, even though its training is not the most stable. WGAN, despite producing visually consistent images, shows the weakest FID behavior. Its scores fluctuate widely from epoch to epoch and never settle into a clear downward trend. This instability is common with weight-clipped WGANs, where the critic's limited capacity prevents steady improvements in distribution matching. ACGAN starts with the highest FID scores, but they gradually decrease as training progresses and the model becomes better at generating class-meaningful images. By the end of training, ACGAN performs better than WGAN but still does not reach DCGAN's best values. Overall, DCGAN achieves the lowest FID, ACGAN shows moderate improvement, and WGAN performs the weakest according to this metric.

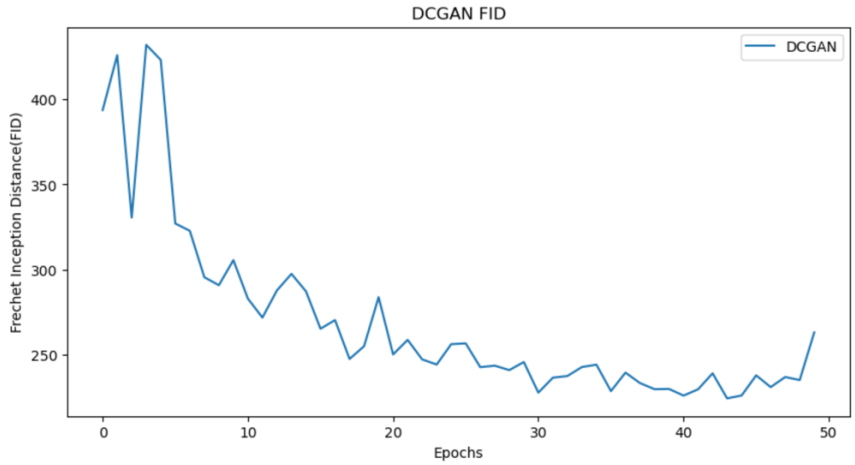

Fig 10: DCGAN FID Score Curve.

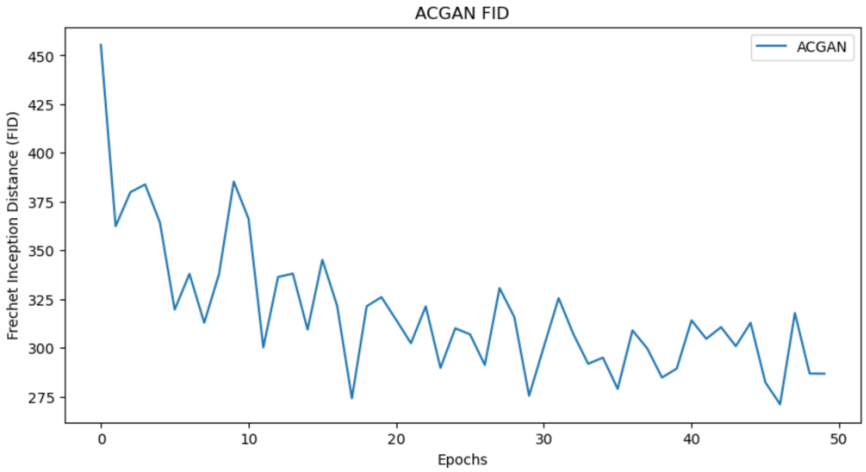

Fig 11: ACGAN FID Score Curve.

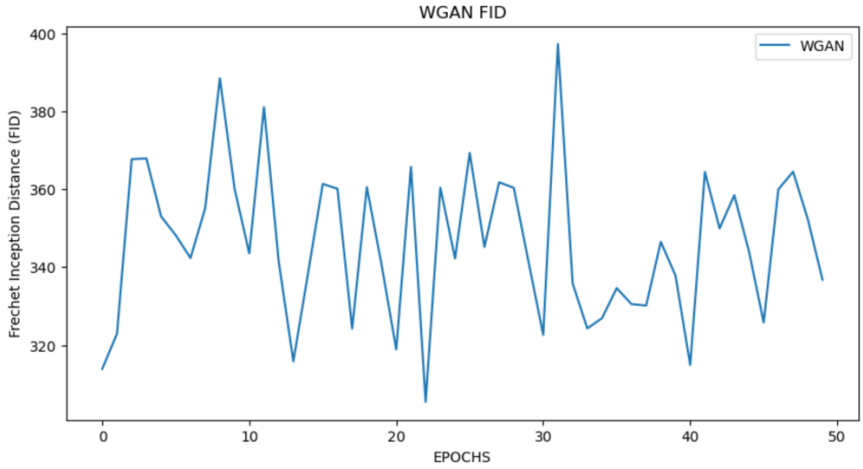

Fig 12: WGAN FID Score Curve