DS-265 DLCV 2025: Assignment 3

Kaniki Bhikshapathi - 24489

Experimental Setup

- **Dataset:** CIFAR-10 (60,000 images, 32x32 pixels, 10 classes, resized to 64x64)
- **Batch size:** 128
- **Latent Vector Size:** 100
- **Learning rate:** 0.0002
- **Beta1:** 0.5 (Adam optimizer)
- **Epochs:** 25
- **Optimization:** Adam optimizer

1 Experiment

Loss Curves Analysis

- **Discriminator Loss** (Blue curve): starting the loss fluctuates because the discriminator tries to differentiate between the real and the fake .Then it stabilizes , it says that high quality real images are generating after some epochs.
- **Generator Loss** (Orange curve):It starts with high fluctuations it continues to fluctuate till end,but the performance is increasing.The loss deecreases as the generator lears more about the real images distribution.

Key Observations

- The **discriminator** loss stabilize as training continues, saying that discriminator is learning to classify more effectively the real and fake images
- The **generator** loss oscillates but decreases gradually, saying that the generator is improving, though still facing challenges in consistently producing realistic images.

Loss Plot

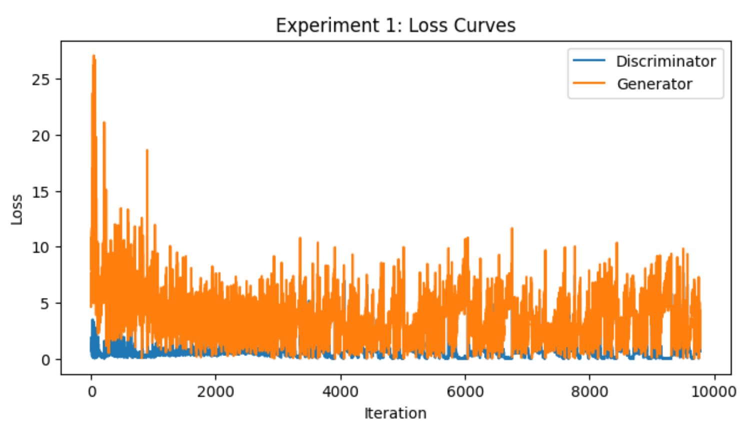

Figure 1: Loss Curves for Generator and Discriminator

2 Experiment

- The generated images are very diverse in texture and colour but not that realistic to look because of lack of sharpness and detail and some are blurry. .
- Many classes(10 clases) are used for training, but some classes are dominat meaning the model struggled with generalising all classes equally well
- Visual artifacts, such as irregular shapes and color blending, are present, which is common in GANs during initial training stages.
- With looking we can say they belong to CIFAR-10 category, but we cannot identify the individual image since lack of sharpness and quality

Figure 2: Generated 10x10 grid of images

3 Experiment

The following table summarizes the observed dynamics and image quality for each configuration:

Configuration	**Generator Loss Behavior**	**Discriminator Loss Behavior**	**Generated Image Quality**
G_steps = 1, D_steps = 1	Fluctuates significantly, slow improvement	Stabilizes faster	Blurry, abstract, minimal detail
G_steps = 2, D_steps = 1	Faster loss reduction, improved quality	More fluctuations due to frequent generator updates	Sharper, more detailed images
G_steps = 1, D_steps = 2	Slow loss reduction, struggles with stability	More fluctuations, challenging for the generator	Noisy, blurry, diverse

Table 1: Summary of observed dynamics and image quality for each configuration.

Loss Curves and Generated Images

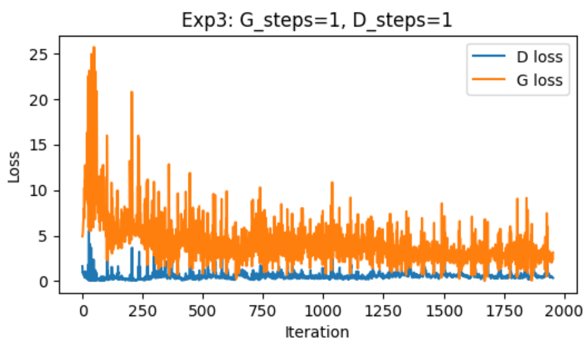

Figure 3: Loss Curves (G_steps=1, D_steps=1)

Figure 4: Generated Images (G_steps=1, D_steps=1)

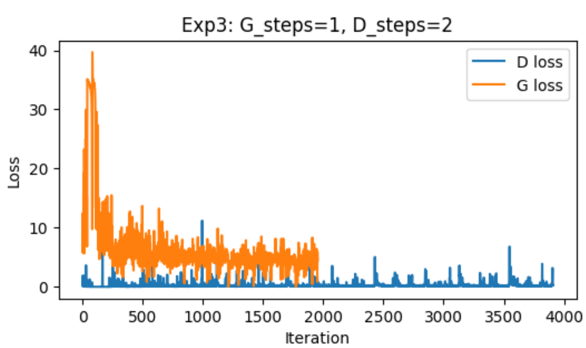

Figure 5: Loss Curves (G_steps=2, D_steps=1)

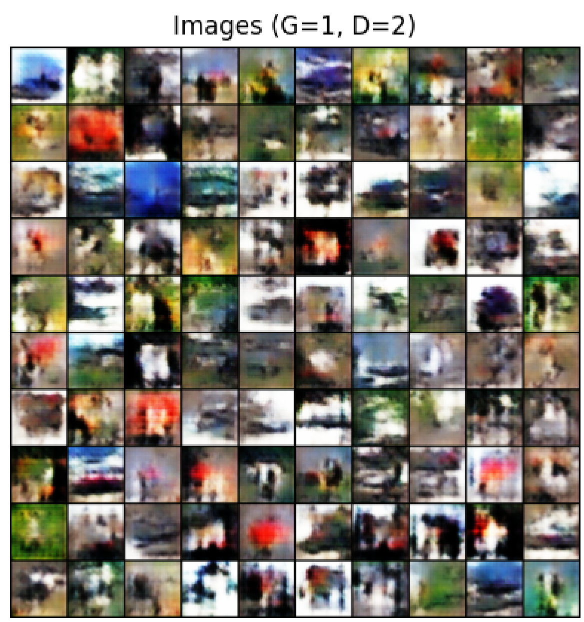

Figure 6: Generated Images (G_steps=2, D_steps=1)

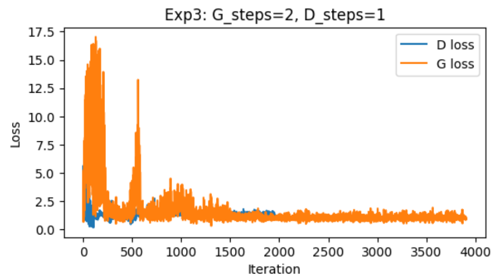

Figure 7: Loss Curves (G_steps=1, D_steps=2)

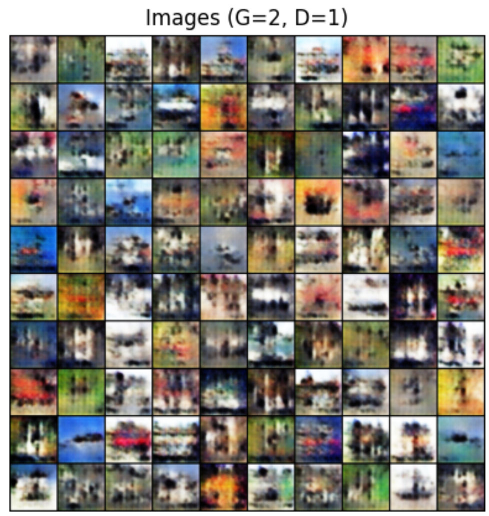

Figure 8: Generated Images (G_steps=1, D_steps=2)

Experiment 4: FID Scores for Different Training Configurations

Configuration	FID Score	Analysis
G_steps = 1, D_steps = 1	139.10	Moderate image quality with blurry and abstract images.
G_steps = 2, D_steps = 1	196.88	Higher FID due to instability, leading to poorer image quality.
G_steps = 1, D_steps = 2	122.04	Best image quality with sharper features and the lowest FID.

Table 2: FID Scores for Different Generator and Discriminator Training Configurations

- For G=2 and D=1 the gernerator update twice but due to lack of training to the descriminator the images generated by the generator lack quality and realistic to eye.So low FID.
- For G=1 and D-2 , even through the generator is update less than descrimator the image quality is good, because we are making the generator to update to correct distribution by making dicriminator more effective. So high FID

Experiment 5:Analysis of Latent Interpolation

- **Smooth Transitions:** We can see a smooth transition from the left to right ,saying the generator network layer learned a continuous latent space,bcz small change in latent vector produced a completely different image transformations
- **Intermediate States:** The feature of the image are visible in the middle stages reflects that generator understandings in latent space
- **Image Quality:** Image Quality is not that good , but more training of generator and discrimator will certainly enhances the image quality

Figure 9: Latent Interpolation

Figure 10: Latent Interpolation

Figure 11: Latent Interpolation

DDPM Training and FID Analysis

FID Comparison

The Fréchet Inception Distance (FID) scores for the final epoch of each model were as follows:

- **500 Timesteps:** FID = 188.2755
- **1000 Timesteps:** FID = 210.2537
- **100 Timesteps:** FID = 237.5298

Analysis

Capacity for Representation

- **500 Timesteps:** Captured key data aspects without overfitting to minor details.
- **100 Timesteps:** Underrepresented data complexity, leading to blurry images and higher FID.
- **1000 Timesteps:** May have over-parameterized, overfitting to noise and artifacts, producing less varied images.

Noise Handling, Regularization, and Generalization

- **500 Timesteps:** Handled noise effectively and generalized well, avoiding overfitting to minor fluctuations.
- **100 Timesteps:** Lacked sufficient noise-variance sampling, leading to blurry images and poor generalization.
- **1000 Timesteps:** Over-refined noise, possibly removing important features and focusing too much on data peculiarities, reducing generalization power.

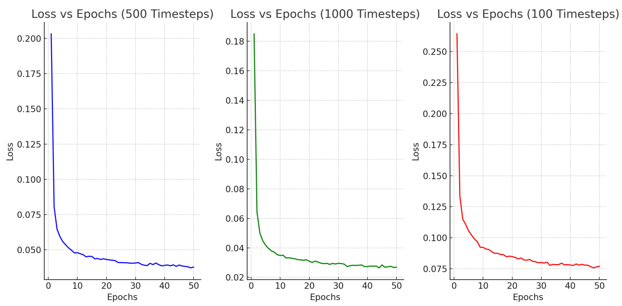

Figure 12: Loss vs Epochs Timesteps)

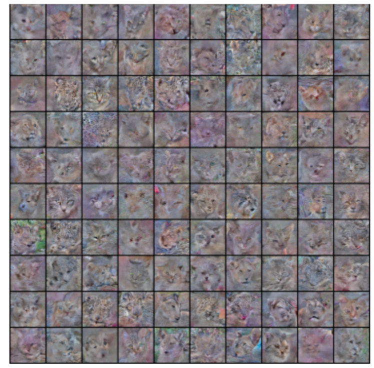

Figure 13: Generated Images with 100 Timesteps

Figure 14: Generated Images with 500 Timesteps

Figure 15: Generated Images with 1000 Timesteps