Luleå University of Technology

D7047E – Advanced Deep Learning

28 Mai 2023

Exercise 5 | Group 1 | Berezin Ilya

Contents

Generative Adversarial Networks (GAN) 3
1.1 Vanilla GAN using a Binary Cross Entropy loss function with 20k iterations: 3
1.3 Vanilla GAN using a Binary Cross Entropy loss function with 100k iterations: 5
1.2 Using a Logistic loss function with 20k iterations: 7
1.4 Using the concept of conditional GANs, generating a particular output class image 8........ 8
Generative Adversarial Networks (GAN). Theory questions. 9
Question 1 9
Question 2 9
Question 3 (optional) 9
Question 4 10
Question 5 10
Question 6 10

Generative Adversarial Networks (GAN)

1.1 Vanilla GAN using a Binary Cross Entropy loss function with 20k iterations:

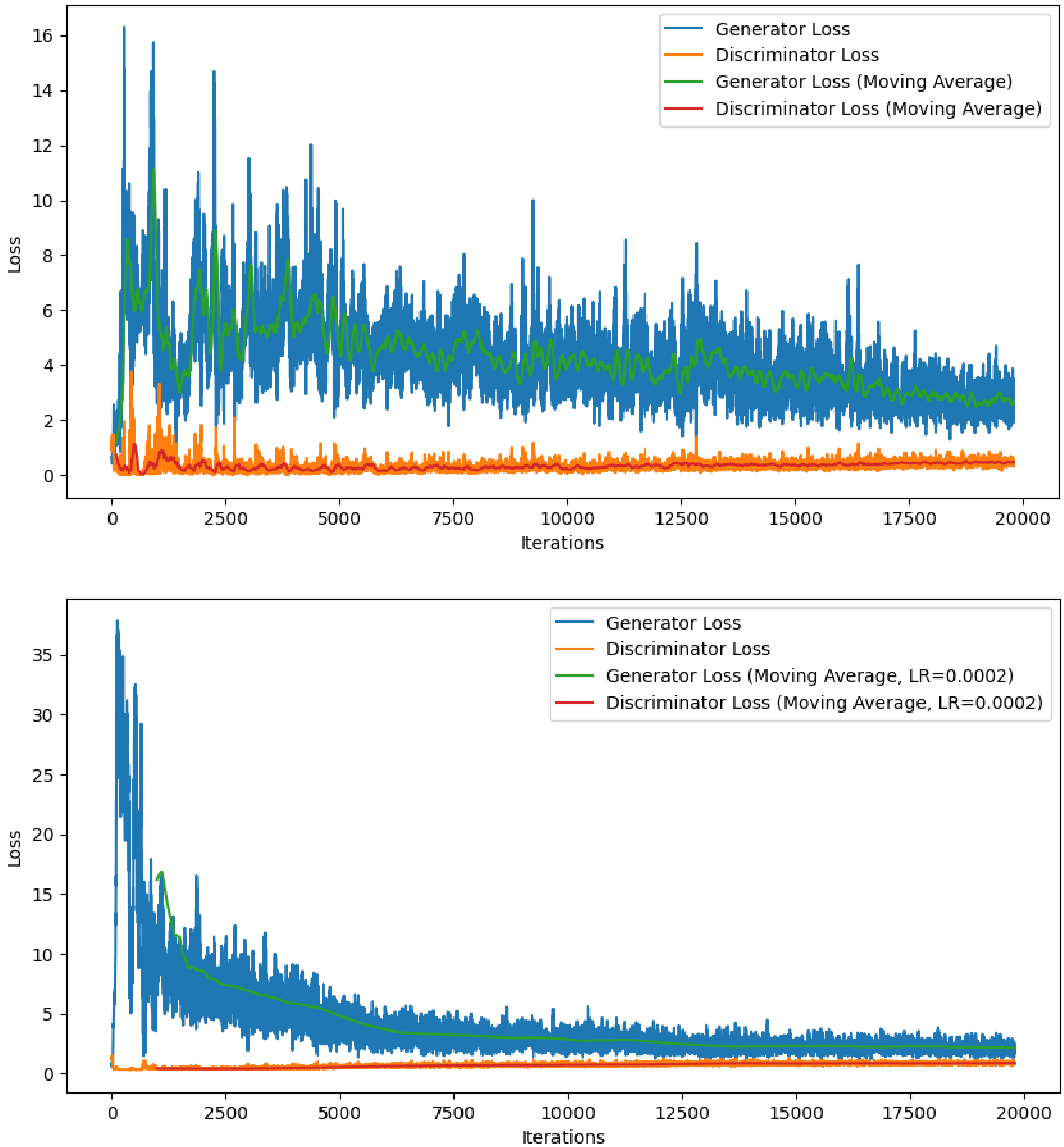


Figure.1.1.1 Models loss, trained on MNIST for 20k iterations with Binary Cross Entropy (BCE) loss.

Figure.1.1.2 MNIST images generated for 20k iterations with Binary Cross Entropy (BCE) loss.

1.3 Vanilla GAN using a Binary Cross Entropy loss function with 100k iterations:

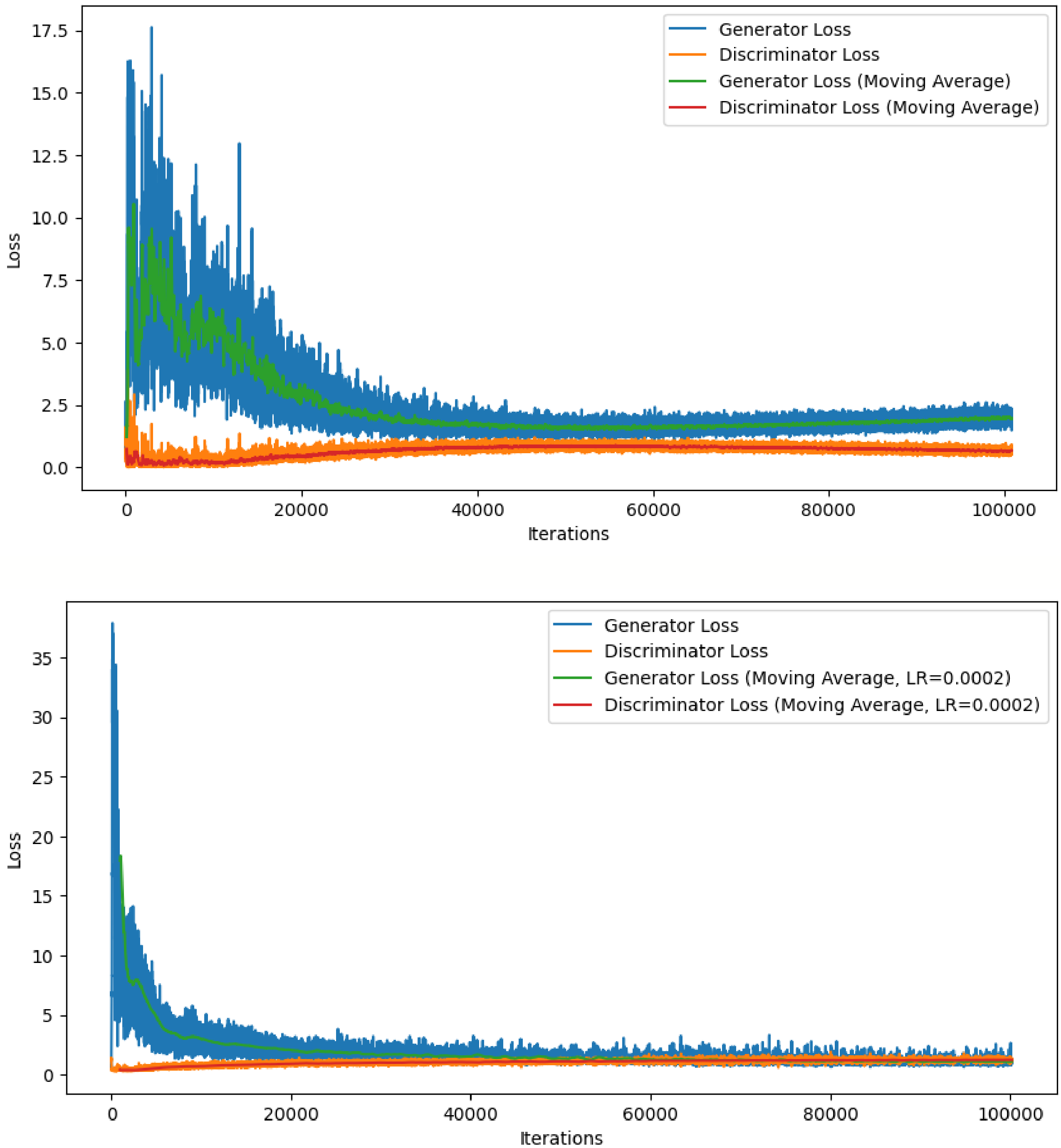


Figure.1.3.1 Models loss, trained on MNIST for 100k iterations with Binary Cross Entropy (BCE) loss.

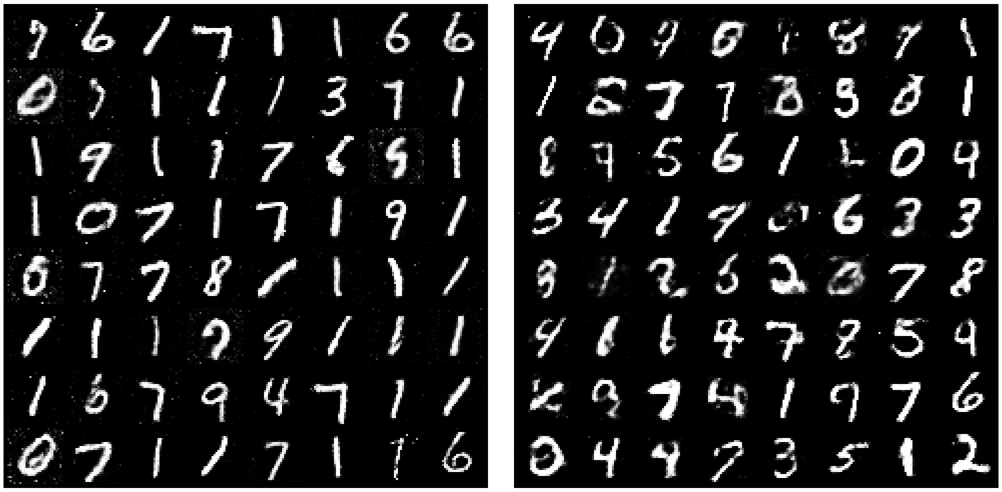


Figure.1.3.2 MNIST images generated for 100k iterations with Binary Cross Entropy (BCE) loss.

1.2 Using a Logistic loss function with 20k iterations:

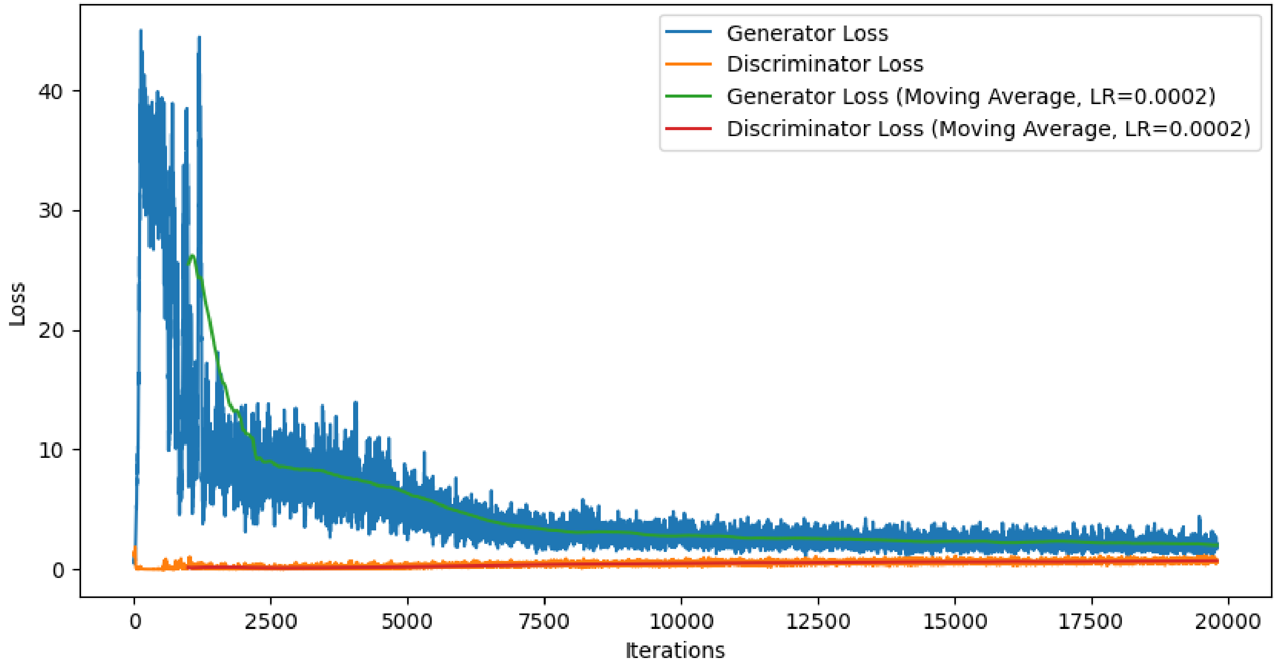
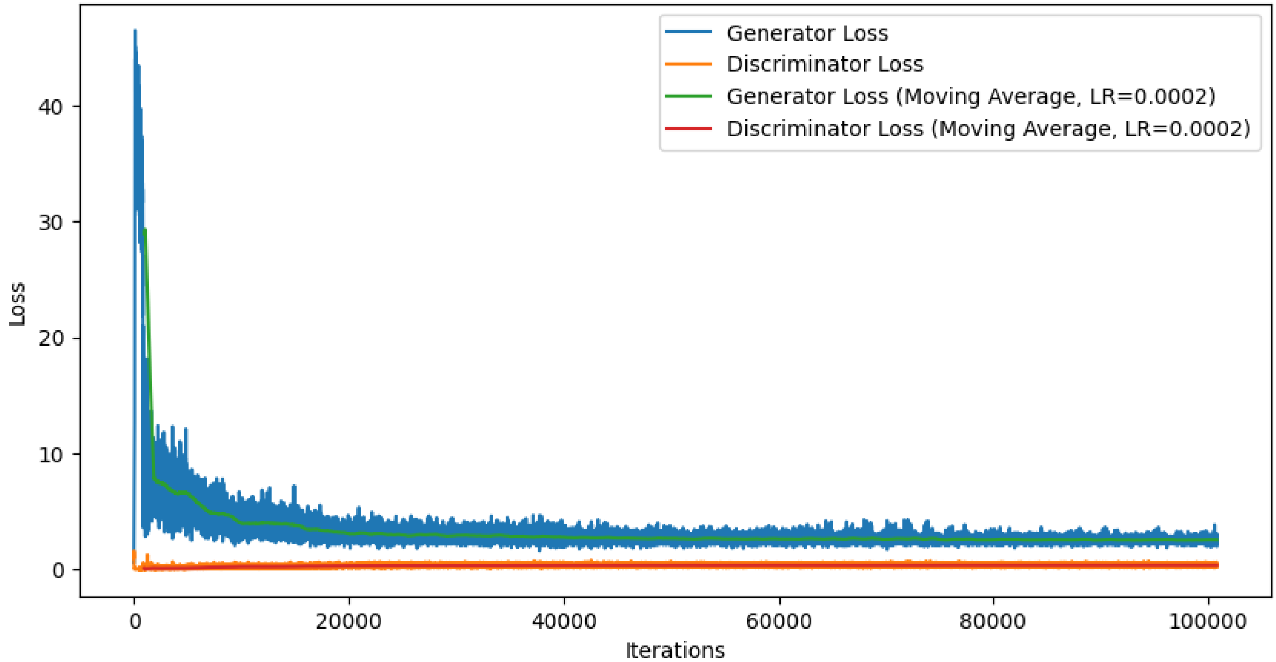


Figure.1.2.1 MNIST images and losses, trained on MNIST for 20k iterations with the logistic loss modifications.

Figure.1.2.2 MNIST images generated for 20k iterations with the logistic loss modifications.

1.4 Using the concept of conditional GANs, generating a particular output class image 8.



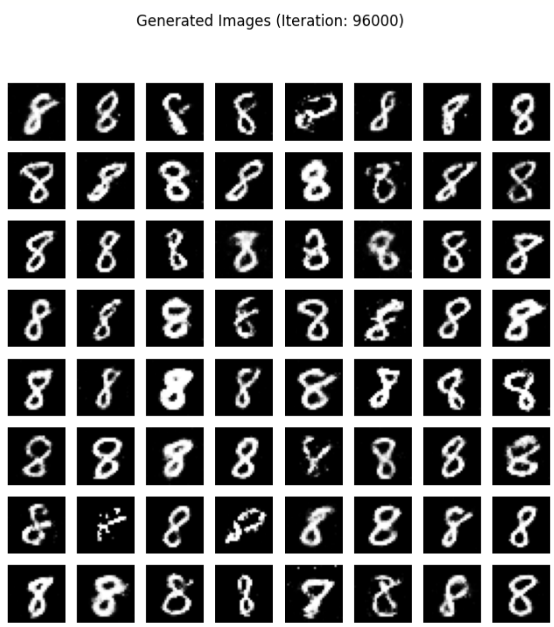


Figure.1.4 MNIST image 8 generated for 100k iterations with Binary Cross Entropy loss and conditional GAN.

Generative Adversarial Networks (GAN). Theory questions.

Question 1

State the objective function of a GAN (the one proposed by Goodfellow et al.)

Formally, the GAN objective can be written as:

$$\min_{\theta} \max_{\phi} V(G_\theta, D_\phi) = \mathbb{E}_{\mathbf{x} \sim \mathbf{p}_{\text{data}}}[\log D_\phi(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p(\mathbf{z})}[\log(1 - D_\phi(G_\theta(\mathbf{z})))]$$

G and D are the generator and discriminator networks.

$\mathbb{E}_{\mathbf{x} \sim \mathbf{p}_{\text{data}}}$ are the expectations over the real data distribution and the input noise distribution.

D(x) is the discriminator's estimate of the probability that real data instance x is real.

D(G(z)) is the discriminator's estimate of the probability that a fake instance is real.

Question 2

Describe each term in the objective function of the GAN.

$\mathbb{E}_{\mathbf{x} \sim \mathbf{p}_{\text{data}}}[\log D_\phi(\mathbf{x})]$ is the expected value of the logarithm of the discriminator outputs for the real examples from the real data distribution. This term encourages the discriminator to recognize real data.

$\mathbb{E}_{\mathbf{z} \sim p(\mathbf{z})}[\log(1 - D_\phi(G_\theta(\mathbf{z})))]$ is the expected value of the logarithm of one minus the discriminator outputs for the fake examples produced by the generator from the noise distribution. This term encourages the discriminator to recognize the generator's data as fake and the generator to fool the discriminator.

The discriminator seeks to maximize this objective (max_D), while the generator seeks to minimize it (min_G). The discriminator's goal is to correctly classify real and fake data, while the generator's goal is to produce data that the discriminator cannot distinguish from real data.

Question 3 (optional)

Can you reformulate the objective function of GAN in terms of Categorical Cross-Entropy? Justify your answer.

Binary cross-entropy loss is equivalent to using the categorical cross-entropy loss for a two-class problem.

The term $\mathbb{E}_{\mathbf{x}\sim\mathbf{p}_{\text{data}}}[\log D_\phi(\mathbf{x})]$ can be considered as binary cross-entropy loss for the real data (labelled as 1), and $\mathbb{E}_{\mathbf{z}\sim p(\mathbf{z})}[\log(1 - D_\phi(G_\theta(\mathbf{z})))]$ as binary cross-entropy loss for the generated data (labelled as 0). The overall objective function is the sum of these two binary cross-entropy losses.

Question 4

Describe the Generator and Discriminator. Be as formal as possible.

Generator (G): The generator is a neural network that learns to create synthetic data by mimicking the true data distribution. It takes a random noise vector z as input and gives a synthetic data instance as output. The objective of the G is to produce data that the discriminator cannot distinguish from the real data.

Discriminator (D): The discriminator is a neural network that learns to distinguish the synthetic data produced by the generator from the real data. It takes a data instance (either real or synthetic) as an input and outputs a probability that the input data is real.

Question 5

What are the problems that GANs face while training. Describe them if any.

Mode Collapse: This occurs when the generator starts to produce same output (or a small set of outputs) repeatedly, instead of generating diverse samples from the data distribution.

Vanishing Gradients: This can occur when the discriminator becomes too good at distinguishing real data from fake data. The generator would receive little to no feedback and thus, would not be able to improve.

Difficulty in achieving Nash Equilibrium: In theory, GAN training aims to reach a Nash Equilibrium of the minimax game where neither the generator nor the discriminator can improve unilaterally. In practice, achieving this equilibrium is hard due to the discrete parameter updates of the networks.

Question 6

How can you evaluate GANs? Please provide objective answers.

Evaluating GANs can be challenging due to the lack of an explicit evaluation metric. Some commonly used methods include:

Visual Inspection: For image generation tasks, qualitative evaluation by human is often used. However, this is subjective and does not scale well.

Inception Score (IS): IS measures both the quality and diversity of images generated by a GAN. It uses a pre-trained Inception network and computes the KL-divergence between the conditional class distribution and the marginal class distribution over the generated samples.

Fréchet Inception Distance (FID): FID computes the Wasserstein-2 distance between the distribution of real images and the distribution of generated images in the feature space of a pre-trained Inception network. Lower FID indicates better quality and diversity.

Precision and Recall: These metrics can measure the quality and diversity of generated samples, where precision checks how many generated samples are close to real ones and recall checks how many real samples are close to generated ones.

Using a held-out validation set: This can be used to measure overfitting, a common problem in GANs, by observing the generator's performance on a held-out validation set.

However, these metrics all have their own limitations and none of them perfectly captures the quality and diversity of the generated samples, thus they should be used together and interpreted cautiously.