CS6476: Assignment 5 – Generative Models

The aim of this assignment is to implement and train a simple generative-discriminative network on the MNIST dataset and generate novel images.

Learning objectives:

- Building custom dataset and dataloader.
- Understand the generator-discriminator network.
- Implement training loop for generative adversarial network. (GAN)
- Train and evaluate GAN on MNIST dataset.

Grading schema:

The assignment will be graded out of 16:

- Generator: 4 points
- Discriminator: 2 points
- End-to-end working: 2 points
- Write-up: 8 points

Programming assignment:

Head over to `generative_adversarial_networks.ipynb` for detailed instructions on how to complete this assignment. Return to the write-up when you are done executing the code.

Writeup

Question 1

a. Calculate the number of parameters in the generator for `noise_dim={96, 192, 256}`. Which layer(s) are responsible for the increase in number of parameters?
b. Visualize and attach generated images from the last epoch of training for both vanilla GANs and LS loss GANs. Compare and contrast the results of both methods and provide reasons for the same.

Question 2

a. Comment on the choice of generator and discriminator models. How would you change the model architectures if the dataset would have been more complex, for example the ILSVRC-2012 dataset?
b. Should the generator and discriminator be scaled up equally when the dataset size increases?

Question 3

a. Training GANs has been notoriously difficult. Many follow-up works in this domain focus on this aspect. What do you think is the biggest contributor to training instability in GANs? Potential answers might be the model architecture, loss function, preprocessing, etc. Justify your answer.
b. Briefly discuss how Wasserstein GANs have helped mitigate the shortcomings you have discussed above?

Question 4

a. After training, one can qualitatively evaluate the generated images. However, to concretely compare generator performance, there have been several metrics used in literature so far. Which one do you think makes the most sense? Feel free to critique existing metrics, or propose changes to the current ones.
b. Briefly explain “mode collapse” in GANs and how specific metrics help identify it.