

Homework 7: DCGAN

CS 1470/2470

Optional Assignment

1 Conceptual Questions

1. What is the discriminator loss? What is the generator loss? What are they each measuring? (2-4 sentences)
2. Why do we alternate between training the generator and training the discriminator? (2-4 sentences)
3. One common problem encountered during GAN training is mode collapse. Explain what mode collapse is and why it occurs. (2-4 sentences)
4. Recall that the generator's objective is to minimize the following loss function:

$$L_G = \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))]$$

The authors of the [original GAN paper](#) propose using a slightly modified loss function for training the generator. State the function and describe why it is favorable, in your own words. (2-4 sentences)

5. We use Frechet Inception Distance (FID) to evaluate the performance of our model. What does FID measure? What are some of its limitations? (2-4 sentences)

2 Ethical Implications

1. Question 1 Lawmakers are increasingly interested in regulating deepfakes; for instance, [this](#) is an analysis of the DEEPFAKES bill in the US house. (California and Texas have passed related bills.)
 - (a) A lawmaker has reached out to you to explain to them how deepfakes “work” – they have read some news articles that talk about something called GANs, but want to know more. Explain in laypersons’ terms what a “GAN” is, and how it can be used to produce something like a deepfake. (4-6 sentences)

- (b) Suppose you've just published a paper describing an algorithm where, if given a video, the algorithm can tell whether or not it is a deepfake. You want to use this new development to reduce potential harms of deepfakes; in what settings might this algorithm be used, and by whom? What (potentially non-technical) considerations need to be made when using this algorithm (e.g. free speech), and how will those decisions be made? (3-4 sentences)

3 CS2470-only Questions

1. The model you implemented for this assignment is built upon convolutional layers. A newly proposed model, authored by none other than Ian Goodfellow, is the Self-Attention Generative Adversarial Network (SAGAN), which has achieved state-of-the-art results in image generation tasks for multi-class datasets. You can find the paper [here](#). What is the motivation for using self-attention in the image setting? How do the authors incorporate self-attention into their model?
2. Regular GAN's suffer from the vanishing gradient problem. An attempt to alleviate this issue is using Wasserstein loss. What is Wasserstein loss, and how is it different from the regular GAN loss function? Why does it not suffer as badly from vanishing gradients?
3. Find a GAN paper that interests you. Describe in detail the modifications it makes to regular GAN's, specifically in regards to its loss function and model architecture.