Questions

"Training"

1. List three generative models we discussed in class. How are they similar? How are they different?
2. Why do autoencoders produce blurry images?
3. What does the generator of a GAN optimize? What does the discriminator of a GAN optimize?
4. Why did we need two different `torch.optim` optimizers to optimize the generator and discriminator?
5. What is batch normalization? How is it similar to normalizing the input? How is it different?
6. Why is training a GAN difficult?
7. What is the difference between a targeted and untargeted adversarial attack?
8. What is the difference between a white-box and black-box adversarial attack?

"Generalization"

1. Why is it that a GAN is more likely to suffer from mode collapse than an autoencoder?
2. Would our RNN text generation model also suffer from mode collapse? Why or why not?
3. Why is the balance between generator and discriminator important? Which one is more likely to become too powerful and cause an issue?
4. Suppose we are training a conditional image generation model, conditioned on text. That is, we use a recurrent neural network to embed a sentence description of an image, then generate an image based on the description. Suppose we have millions of training examples in the form of (sentence, image) pairs. After using weeks of GPU processing time, you begin to get fairly good results, but those results are **blurry**. Provide one possible explanation why, based on the loss function that you might have used.