Rensselaer Polytechnic Institute
Department of Electrical, Computer, and Systems Engineering
ECSE 4964/6964: Computational Creativity, Fall 2023

Homework #2: due Thursday, October 12th, at the beginning of class.
Show all work for full credit!

Submit your homework writeup using Gradescope, and your images/video/code to the `HW2\yourteamname` folder on Box that you will share with me. Your team only needs to submit one assignment on Gradescope (there is a way to indicate who is on what team when you upload). Note that these homeworks are meant to be true team efforts; that is, all of your members should be participating in solving each problem, as opposed to assigning one problem per person!

1. (30 points) Collect a set of about 10 images that are from the same class but have varying (to human eyes) degrees of similarity. You could use images from your HW1 dataset, or a totally different set of images. In addition to raw, high quality images, include some noisy, blurry, tilted, or cropped versions of your originals. Get together with your team to agree on a general set of perceptual similarity rankings between images.

   Then use the LPIPS metric to quantitatively evaluate the similarity between pairs of your images. How do these compare to the similarities you get from MSE and SSIM, and to your group's human rankings? Comment on situations where MSE, SSIM, and LPIPS seem to agree well with human similarly, and situations where one or more of these is way off.

2. (40 points) Use the dataset you collected in HW1 to build a GAN to generate realistic images from the same class. There are many choices of GAN architectures; some you may be able to train fully on your own machine, some may be possible to download fixed weights and fine-tune, and some may require some cloud computing. If you don't know where to start, this video from Eryk Salvaggio describes the steps of using RunwayML to fine-tune StyleGAN2 on a user-uploaded dataset. Students in the grad level of the course, I would be be more impressed if you could successfully train the GAN yourself!

   (a) Training a GAN is notoriously finicky so I'd like you to keep track of and report on all the tweaks you had to make to the architecture and hyperparameters to achieve a good result. What kind of GAN was it? What were the dimensions of your latent space? How long did it take? Did you have to lower the resolution of the data? What problems did you encounter and overcome? What did you learn that you would tell another student trying to do the same thing?

   (b) Generate 20–30 examples of new images by sampling from the GAN in latent space, and critically assess the results. What features of your input dataset did the GAN seem to pick up on? Are the images reasonable? How do they compare to your VAE-generated images from HW1?

   (c) Demonstrate some type of cool image manipulation using your GAN; the easiest thing is morphing between images, but if you used an advanced GAN you may be able to do something more interesting (e.g., StyleGAN-based combinations between disparate source images as used in GANBreeder or some sort of image-to-image GAN). Be creative here!

(continued on next page)

3. (30 points) Finally, use the same dataset to build a denoising diffusion model to generate realistic images from the same class. (Remember, we're not at the point of using prompts to generate images; this problem is just the single-class diffusion model we discussed in Lecture 10/Foster Chapter 8.)

   (a) As above, describe your architecture, parameter choices, and tuning process. What were the dimensions of your latent space? How long did it take to train? Did you have to lower the resolution of the data? What problems did you encounter and overcome? What did you learn that you would tell another student trying to do the same thing?

   (b) Generate 20–30 examples of new images by sampling from the diffusion model in latent space, and critically assess the results. What features of your input dataset did the diffusion model seem to pick up on? Are the images reasonable? How do they compare to your GAN-generated images from Problem 2?

   (c) Demonstrate that your diffusion model can generate morphs between images by sampling from a trajectory in the latent space. Can you do anything else interesting with your model?

Rensselaer Polytechnic Institute
Department of Electrical, Computer, and Systems Engineering
ECSE 4964/6964: Computational Creativity, Fall 2023

Homework #2: due Thursday, October 12th, at the beginning of class.
Show all work for full credit!

1. (20 points) For the 6964 level of the class only: Provide one-page reviews of **two** machine learning or computer vision papers published in recent years.

   (a) The first paper should address an innovation related to GANs (i.e., related to Lectures 7–9 of class and Foster Chapter 4), not duplicating any of the papers we mentioned already.

   (b) The second paper should address an innovation related to diffusion models or neural style transfer (i.e., related to Lectures 10–11 of class), not duplicating any of the papers we mentioned already.

   Note that we haven't yet discussed CLIP or language models yet, so please try to avoid papers that are overly language-focused since that will be a separate topic later.

   Claim your paper on this Google sheet. No paper should be covered by more than one person.