ELEC/COMP 447/546
Assignment 6
Due: Apr. 30, 2023, Midnight

Please use this provided Colab notebook for the assignment.

Problem 1: StyleGAN (7 points)

In this problem, you will use StyleGAN2 for controlled image generation. Make sure to run the first 3 code cells of the provided Colab notebook. The first cell installs StyleGAN2 and its dependencies. The second cell loads a pre-trained StyleGAN2 model for faces. The third cell provides you with some useful utility functions. Some preliminary code on how to generate synthetic faces using the utility functions is also provided in cell 4.

StyleGAN2's generator converts a vector $z \in R^{512}$ drawn from the standard Normal distribution into a 'style' vector $w \in R^{512}$. The generator then processes the style vector to produce an image $I \in R^{1024 \times 1024 \times 3}$. In this problem, you will find a direction in the style space corresponding to perceived gender and use that direction to alter the perceived gender of synthetic faces.

a. *Interpolating between images*: Choose two random noise vectors `z0`, and `z1`, such that the two generated faces have different perceived genders based on the `face_is_female` function[1]. This function uses a pre-trained face gender classifier to make its prediction.
   i. Interpolate between the **latent vectors** `z0`, and `z1` with 5 intermediate points. Show a strip of 7 faces along with the classifier predictions in your report.
   ii. Interpolate between the **style vectors** `w0`, and `w1` with 5 intermediate points. Show a strip of 7 faces along with the classifier predictions in your report.
   iii. *Question*: What differences do you notice when interpolating in latent space versus style space? Do the intermediate faces look realistic?

b. *Image manipulation with latent space traversals*
   i. Sample 1000 random `z` vectors, convert them to style vectors `w`, and get their corresponding perceived genders using the trained classifier. This may take a few minutes.

[1] We use binary gender attributes in this assignment for simplicity.

ii. Train a linear classifier (use scikit-learn's linear SVM) that predicts gender from the style vector. The model's coefficients (attribute `coef_`) specify the normal vector to the hyperplane used to separate the perceived genders in style space. Remember to convert your cuda tensors to numpy arrays before sending to scikit-learn's functions.

iii. Sample 2 random `w` vectors. For each `w` vector, display a strip of 5 images. The center image will be the image generated by `w`. The two images to the left will correspond to moving toward the "more male" direction, and the two to the right will correspond to "more female". To generate the latter 4 images, move along the SVM hyperplane's normal vector in both directions using some appropriate step size.

iv. *Question*: Do you notice any facial attributes that seem to commonly change when moving between males and females? Why do you think that occurs?

Problem 2: Using CLIP for Zero-Shot Classification (5 points)

In this problem, you will use Contrastive Language-Image Pre-Training (CLIP) to perform zero-shot classification of images. You can read more about CLIP in this blog post, and check out the example in the official GitHub repository. We will reuse the CIFAR dataset introduced in Assignment 4. Download that dataset as one .npz file here and place it in your Google Drive folder.

a. Perform classification of each test image (last 10,000 images of the dataset) using CLIP. To do so, create 10 different captions (e.g., "An image of a [class]") corresponding to each of the 10 object classes. Then, for each image, store the label that provides the highest probability score. Report overall accuracy.

b. ELEC/COMP 546 ONLY (3 points). Engineer the caption prompts to try to obtain better accuracy. To do so, give a *set* of possible captions per class instead of just one. For example, "A bad photo of a [class]" or "A drawing of a [class]". Report your accuracy.

Submission Instructions

All code must be written using Google Colab (see course website). Every student must submit a zip file for this assignment in Canvas with 2 items:

1. An organized report submitted as a PDF document. The report should contain all image results (intermediate and final), and answer any questions asked in this document. It should also contain any issues (problems encountered, surprises) you may have found as you solved the problems. **Please add a caption for every image specifying what problem number it is addressing and what it is showing**. The heading of the PDF file should contain:
    1. Your name and Net ID.
    2. Names of anyone you collaborated with on this assignment.
    3. A link to your Colab notebook (remember to change permissions on your notebook to allow viewers).
2. A pdf copy of your Colab notebook.

Collaboration Policy

I encourage collaboration both inside and outside class. You may talk to other students for general ideas and concepts, but you should write your own code, answer questions independently, and submit your own work.

Plagiarism

Plagiarism of any form will not be tolerated. You are expected to credit all sources explicitly. If you have any doubts regarding what is and is not plagiarism, talk to me.