Capstone Proposal: GANs for Generative Art

The capabilities of Generative Adversarial Networks (GANs) have improved rapidly since the release of the initial GAN paper in 2014. In the process of learning to generate "fakes" with the same underlying statistics in the training data, a GAN must learn mappings of input noise vectors to generated data. This is the GANs "latent space". My goal for this project is to train NVIDIA's StyleGAN2-ADA on image datasets and explore the creative potential of the latent spaces learned by the trained models. Interpolation videos made with GANs can serve as a novel way to promote businesses on social media platforms or to generate other artwork for commercial purposes (like NFTs).

I have chosen StyleGAN2-ADA for this project because it can produce images up to 1024x1024 resolution, and its "Adaptive Discriminator Augmentation" makes it trainable on datasets much smaller than 30K samples.

malliGAN
The StyleGAN2-ADA architecture will be trained on images of food from Malli, a Los Angeles pop-up restaurant business that opened towards the beginning of the COVID pandemic. All images will be provided by the business owner herself. If the GAN can be trained to produce realistic fake dishes, it could perhaps even give the chef some ideas for new dishes. More importantly, interpolation videos made by morphing from one dish into another can be used to promote her business on social media. Interpolations can be done with the fake dishes or GAN projections of the real dishes (the training images).

grappleGAN
grappleGAN will be trained on images of people grappling/training brazilian jiu jitsu. At least 2K images will be scraped from Instagram and Flickr, where the main subjects are two martial artists actively grappling in the gym or in a competition setting. Since jiu jitsu is a complex sport where bodies can become entangled in a variety of positions, and since StyleGAN2 does not have any sense of human anatomy, I expect the results to be rather abstract. If the quality of the output is not suitable for promoting jiu jitsu brands or tournaments, its use case could include generating abstract art and NFTs.

There will be some prep work involved in building the datasets from the raw images. The width and height must be a power of 2 (256, 512, etc.) and all images must be in the RGB color space. All images will be cropped to a square and resized to 1024x1024.

Training will take place on Google Colab with Colab Pro, a cloud-based IDE that gives access to a dedicated GPU. GAN progress can be measured quantitatively with metrics such as the Frechet Inception Distance (FID) and qualitatively by examining the generator output (samples of fake images). The GAN will be evaluated at regular intervals throughout training. Interpolation videos will be created after the models complete training, or when GAN progress appears to plateau.