**Progress Report**

**Team**: Cem Beyenal, Mason Menser

**Project**: Comparing DDIM and GAN Performance for Image Generation

***Overview***

The goal of our project is to evaluate the effectiveness of Denoising Diffusion Implicit Models (DDIMs) compared to Generative Adversarial Networks (GANs) across multiple datasets. Specifically, we are interested in analyzing differences in training time, sampling speed, and output quality. Our initial proposal focused on running experiments on datasets such as CIFAR-10 and CelebA, with the potential to expand to simpler datasets like MNIST and SVHN.

***Progress So Far***

Since submitting our proposal, we have made significant progress in both implementation setup and data exploration:

Initial: We corresponded with James Halvorsen, the contact point for this project, on the best way to get started. He gave us some great tips and resources to begin training models.

Model Implementation - We have implemented both Wasserstein GAN (WGAN) and DDIM models using Google Colab, ensuring GPU acceleration for efficient training. Our codebase is organized and version-controlled through a GitHub repository, allowing for collaborative development and structured experimentation.

Datasets - Initially, our focus was on CIFAR-10 and MNIST; however, we have found another dataset involving orchard flowers. The dataset, Oxford Flowers 102, contains high-resolution, color-rich images that can help us evaluate generative quality on more visually complex data. Using this additional dataset will allow us to test whether DDIM’s improvements in stability and quality hold beyond simpler datasets.

Preliminary Experiments - We have begun running early training sessions to make sure that both models can properly learn distributions from smaller image sets. These tests helped us verify our DDIM and WGAN codebases and tune hyperparameters such as learning rate, batch size, and diffusion steps. We were able to generate images using a DDIM model and the Oxford Flower dataset through Colab.

***Next Steps***

Our next goals include:

1. Running full experiments on datasets to establish a baseline comparison between DDIM and WGAN in terms of training time, sample generation speed, and image quality.
2. Applying both models to the oxford flower dataset, analyzing differences in how each model handles color and texture generation.
3. Quantitative Evaluation: Implement metrics such as Frechet Inception Distance (FID) and Inception Score (IS) for sample quality comparison.
4. Qualitative Evaluation: Conduct visual side-by-side comparisons of generated vs. real samples to complement quantitative findings.

***Reflections and Adjustments***

We initially planned to focus on CIFAR-10 and CelebA but decided to pivot to the oxford flower dataset. We think this will be easier to compare between models and allows for more diverse insights into model generalization. Further, our current implementation is designed for images with 3 color channels and doesn’t currently support greyscale. We have also improved our workflow by using GitHub for version control and Colab for compute scalability, allowing faster iteration and better collaboration. Another thing to note, is that training time can be significant, so we are emphasizing on consistent work sessions to maintain a high throughput on the project.

***Conclusion***

Our implementation and dataset setup phases are complete, and we are now entering the model evaluation stage. The next phase will focus on rigorous comparison using standardized metrics to quantify performance differences between WGAN and DDIM. We anticipate that these findings will clarify how diffusion-based generative models perform relative to adversarial methods under varying dataset complexity and color richness.