**CS5624 Natural Language Processing: Project Proposal**

**Project Title:** IStyleGAN – Improved Style-based Generative Adversarial Network

**Team Members:** Jiayue Lin (PID: jiayuelin), Ziyi Wei (PID: ziyiw)

**Introduction of the Problem:**

In recent years, significant progress has been made in the field of generative image modeling. Applications like Midjourney and Deep Dream by Google demonstrate the creative potential of image generation techniques. These technologies have opened up new possibilities, but challenges remain.

Our project focuses on improving the quality of generated images using a cutting-edge technology called StyleGAN[1]. StyleGAN is a powerful tool for creating images, but it has its limitations. Our goal is to address these limitations and enhance the quality of generated images.

Why does image quality matter? Well, it is crucial in various fields like computer vision, graphic design, and entertainment. Think about creating realistic avatars, enhancing the clarity of images, or generating lifelike scenes for video games. The quality of the images directly affects how useful and appealing they are.

StyleGAN, while impressive, has some issues that can affect image quality. These issues can make it challenging to create truly realistic and visually pleasing images. They also complicate the process of identifying where an image comes from, which is important in applications like content creation and forensics.

Our project takes a comprehensive approach to tackle these challenges. We plan to make changes to the way StyleGAN works and how it is trained. By doing this, we aim to generate higher-quality images that are more faithful to the input instructions. We also want to make it easier to figure out where a generated image comes from, which is valuable in various real-world scenarios.

In summary, our project aims to enhance the quality of generated images produced by StyleGAN. We will do this by making improvements to the underlying technology, ensuring it produces better results. This will have implications in various domains where high-quality image generation is essential.

**Related Work:**

The StyleGAN is widely applied to many modern scenarios, such as art[2], healthcare[3] and so on. There are many extensions for StyleGAN[4]. Up to now, some papers have shed light on improving the StyleGAN algorithms, especially StyleGAN2[5].

**Datasets:**

When it comes to the datasets we are planning to use, there are a variety of choices that will enable us to effectively address our project's goal of improving image quality using the IStyleGAN architecture. Here is a list of the datasets we intend to use:

- CIFAR-10 (Canadian Institute for Advanced Research, 10 classes)
- LSUN (Large-scale Scene UNderstanding Challenge)
- FFHQ (Flickr-Faces-HQ)
- ArtBench-10 (32x32)

We will continue to explore potential datasets that align with our project's objectives to further enhance image quality using IStyleGAN, ensuring a robust and comprehensive evaluation of our methods.

**Method Development:**

Our project will begin by focusing on StyleGAN, as it provides a solid foundation for us, as newcomers, to delve into the world of generative image modeling. We recognize that StyleGAN2 is available, but we believe starting with StyleGAN will enable us to build a strong understanding of the fundamentals.

Our initial step involves a thorough review of research papers related to StyleGAN and its subsequent improvements, particularly StyleGAN2. This literature review is essential to grasp the algorithm's nuances and potential shortcomings. It will serve as the bedrock upon which we can build our own insights and innovations.

We aim to identify areas where StyleGAN may face challenges or limitations. This critical understanding will guide our efforts to enhance the algorithm from our unique perspective. Our goal is not only to replicate existing improvements but to exploring refinements to the technology.

If time and resources permit, we are also interested in exploring the possibilities of improving StyleGAN2. However, our primary focus will remain on StyleGAN, ensuring that we establish a strong foundation before venturing into more advanced technologies.

**Reference:**

[1] Karras, Tero, Samuli Laine, and Timo Aila. "A style-based generator architecture for generative adversarial networks." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2019.

[2] Dobler, Hübscher, er al. “Art Creation with Multi-Conditional StyleGANs.” Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence AI and Arts. Pages 4936-4942.

[3] Hagag, Amr, et al. "Deep Learning for Cancer Prognosis Prediction Using Portrait Photos by StyleGAN Embedding." *arXiv preprint arXiv:2306.14596* (2023).

[4] Nie, Weili, et al. "Semi-supervised stylegan for Semi-Supervised StyleGAN for Disentanglement Learning." *Proceedings of the 37th International Conference on Machine Learning*. 2020.

[5] Viazovetskyi, Y., Ivashkin, V., & Kashin, E. (2020). Stylegan2 distillation for feed-forward image manipulation. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXII 16* (pp. 170-186). Springer International Publishing.