

Anime Character Generation Using Stylegan and Latent Space Manipulation

¹Dr. Geetha M , ²Arpitha G K

¹Associate Professor, Department of Master of Computer Applications, BIET, Davangere-577004.

bmgeethamallad@gmail.com, Davangere, Karnataka, India

²Student, Department of Master of Computer Applications, BIET, Davangere-577004.

arpithagkbada@gmail.com, Davangere, Karnataka, India

ABSTRACT

Generative Adversarial Networks have revolutionized image synthesis by enabling the generation of highly realistic images across various domains. Among these models, StyleGAN has demonstrated remarkable performance in producing high-resolution and controllable images. This paper presents a framework for anime character generation using StyleGAN combined with latent space manipulation techniques. The proposed system is trained on large-scale anime face datasets and utilizes style-based generative architecture to control attributes such as hair color, facial expression, eye shape, and accessories. Latent space exploration and vector arithmetic techniques are employed to enable semantic editing and interpolation between characters. Experimental results demonstrate that StyleGAN produces high-quality anime characters with improved diversity and controllability compared to traditional GAN models. The study highlights the potential of latent space manipulation for creative content generation and interactive character design.

Keywords: StyleGAN, Generative Adversarial Networks, Anime Character Generation, Latent Space Manipulation, Deep Learning, Image Synthesis

I. INTRODUCTION

Recent advancements in deep learning have significantly improved the quality of synthetic image generation. Generative Adversarial Networks have emerged as a powerful framework for creating realistic images by training two neural networks in competition: a generator and a discriminator. Anime character generation is a specialized image synthesis task that requires capturing stylized facial structures, color patterns, and artistic textures unique to anime art.

StyleGAN, a style-based generative architecture introduced by NVIDIA researchers, has shown superior capability in producing high-resolution and controllable synthetic images. Unlike traditional GANs, StyleGAN introduces style modulation and adaptive instance normalization, enabling better control over visual features. This research focuses on leveraging StyleGAN for anime character generation and applying latent space manipulation to control and modify character attributes.

II. RELATED WORK

Early GAN architectures such as DCGAN demonstrated the feasibility of generating synthetic images but suffered from instability and limited control over features. Progressive GAN improved image resolution by progressively growing network layers during training.

StyleGAN introduced a style-based generator architecture that separates high-level attributes from stochastic variation, allowing improved image quality and disentanglement of features. StyleGAN2 and StyleGAN3 further enhanced training stability and image fidelity.

In the context of anime generation, several studies utilized GAN-based models trained on datasets such as Anime Faces Dataset and Danbooru dataset. Conditional GANs have been applied to control attributes such as hair color and facial expressions. Latent space manipulation techniques, including vector arithmetic and semantic direction discovery, have enabled interpretable editing of generated images.

III. METHODOLOGY

The proposed methodology consists of dataset preparation, model training using StyleGAN, latent space exploration, and semantic attribute manipulation.

The dataset used in this research includes large-scale anime face image collections such as the Anime Faces Dataset and Danbooru dataset. Images are preprocessed by resizing to a fixed resolution, typically 256x256 or 512x512 pixels. Data cleaning ensures removal of corrupted or low-quality images. Data augmentation techniques such as horizontal flipping and color normalization are applied to improve generalization.

StyleGAN architecture is used for image generation. The generator consists of a mapping network and a synthesis network. The mapping network transforms a random latent vector into an intermediate latent space. The synthesis network generates images using style modulation and noise injection at different layers. The discriminator is trained to distinguish real anime images from generated ones.

Training involves adversarial learning where the generator aims to produce realistic anime characters while the discriminator attempts to identify fake images. The objective function includes adversarial loss and optional regularization techniques to improve stability. Training is performed on GPU hardware due to high computational requirements.

Latent space manipulation is performed after training. The intermediate latent space is explored to identify semantic directions corresponding to attributes such as hair color, smile intensity, or presence of accessories. Vector arithmetic techniques are applied to modify specific attributes by adding or subtracting learned attribute vectors from the latent code. Interpolation between two latent vectors is performed to generate smooth transitions between characters.

The overall workflow is described below:

- ✓ Dataset Collection and Preprocessing
- ✓ StyleGAN Model Training
- ✓ Latent Space Sampling
- ✓ Semantic Direction Discovery
- ✓ Attribute Manipulation
- ✓ Image Generation and Visualization

This structured pipeline enables both automated generation and interactive editing of anime characters.

IV. RESULTS AND DISCUSSION

Experimental results demonstrate that StyleGAN generates high-resolution anime characters with sharp facial features, consistent color patterns, and diverse stylistic variations. Compared to DCGAN and basic GAN architectures, StyleGAN provides improved stability and visual fidelity.

Latent space interpolation produces smooth transitions between characters, demonstrating continuity in the learned feature space. Attribute manipulation experiments show that specific features such as hair color and facial expressions can be controlled without significantly affecting other attributes, indicating partial disentanglement of features.

Quantitative evaluation using metrics such as Fréchet Inception Distance confirms improved generation quality. Qualitative assessment highlights high diversity and artistic realism. The results suggest that StyleGAN combined with latent manipulation is suitable for creative content generation and digital art applications.



output images

V. CONCLUSION AND FUTURE WORK

This paper presented an anime character generation framework using StyleGAN and latent space manipulation. The model successfully generates high-quality stylized anime characters and enables semantic editing through vector operations in latent space.

Future work includes improving feature disentanglement using supervised attribute learning, integrating conditional inputs for precise control, exploring StyleGAN3 for enhanced alias-free generation, and developing user-friendly interfaces for interactive character customization. Additional research may also investigate multimodal generation combining text prompts with style-based generative models.

VI. REFERENCES

- [1] I. Goodfellow et al., "Generative Adversarial Nets," NIPS, 2014.
- [2] A. Radford et al., "Unsupervised Representation Learning with Deep Convolutional GANs," ICLR, 2016.
- [3] T. Karras et al., "Progressive Growing of GANs for Improved Quality," ICLR, 2018.
- [4] T. Karras et al., "A Style-Based Generator Architecture for Generative Adversarial Networks," CVPR, 2019.
- [5] T. Karras et al., "Analyzing and Improving the Image Quality of StyleGAN," CVPR, 2020.
- [6] T. Karras et al., "Alias-Free Generative Adversarial Networks," NeurIPS, 2021.
- [7] J. Shen et al., "Interpreting the Latent Space of GANs for Semantic Face Editing," CVPR, 2020.
- [8] E. Härkönen et al., "GANSpace: Discovering Interpretable GAN Controls," NeurIPS, 2020.
- [9] Y. Choi et al., "StarGAN: Unified Generative Adversarial Networks for Multi-Domain Image-to-Image Translation," CVPR, 2018.
- [10] A. Brock et al., "Large Scale GAN Training for High Fidelity Natural Image Synthesis," ICLR, 2019.
- [11] M. Mirza and S. Osindero, "Conditional Generative Adversarial Nets," arXiv, 2014.
- [12] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," ICLR, 2015.