Multi Attribute Controllable Image Synthesis Decomposed Gan

Mrs.Sowjanya[1], K.Manasa[2], M.Soujanya[3], G.Shashank[4]
[1]Assistant Professor, Teegala Krishna Reddy Engineering College,Hyderabad
[2]Undergraduate Student, Teegala Krishna Reddy Engineering College,Hyderabad
[3]Undergraduate Student, Teegala Krishna Reddy Engineering College,Hyderabad
[4]Undergraduate Student, Teegala Krishna Reddy Engineering College,Hyderabad

Abstract*—This project presents a novel approach to controllable image synthesis through the implementation of Attribute-Decomposed Generative Adversarial Networks (ADGAN) and its enhanced version, ADGAN++. Unlike traditional GAN-based models that often suffer from entangled attribute representations and lack of fine-grained control, the proposed system effectively disentangles individual facial attributes into independent latent spaces. This enables precise and isolated modifications of attributes such as age, gender, expression, and hairstyle without affecting other image features. ADGAN++ further improves attribute manipulation by employing a serial encoding strategy, allowing for smoother and more accurate transitions during image editing. The system incorporates a structured pipeline involving data preprocessing, attribute decomposition, and high-quality image generation using state-of-the-art GAN architectures. With applications ranging from facial editing and personalization to data augmentation and medical imaging, this work demonstrates significant advancements in the usability and interpretability of GAN-based image synthesis techniques.

I. INTRODUCTION

In recent years, Generative Adversarial Networks (GANs) have emerged as a powerful framework for realistic image synthesis across a wide range of domains. However, traditional GAN models such as StyleGAN and BigGAN often struggle with the precise and independent control of image attributes, leading to unintended modifications and entangled feature representations. This limitation presents a significant challenge in applications that require fine-grained editing and interpretability, such as facial attribute manipulation, personalized content creation, and medical imaging.

To address these challenges, this project introduces a controllable image synthesis system based on Attribute-Decomposed GAN (ADGAN) and its advanced variant ADGAN++. These models aim to decompose high-level semantic attributes into disentangled latent representations, enabling users to manipulate individual features like age, gender, and expression without compromising other aspects of the image. The enhanced ADGAN++ employs a serial encoding strategy to further improve attribute isolation and control. By integrating techniques such as latent space interpolation and variational autoencoding, the system ensures high-quality image generation with targeted attribute manipulation. This approach significantly improves upon conventional GAN architectures by enhancing flexibility, reducing attribute entanglement, and making the model more adaptable for real-world applications.

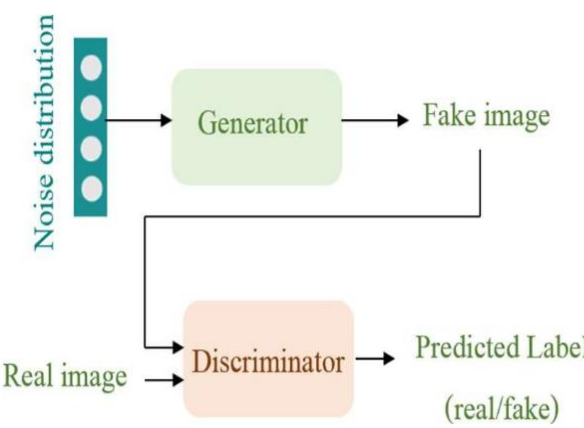

Fig.1

II. BACKGROUND STUDY

Generative Adversarial Networks (GANs) have revolutionized the field of image synthesis by generating highly realistic and diverse images. Models such as **StyleGAN**, **BigGAN**, and **Pix2Pix** have demonstrated significant progress in producing photorealistic outputs; however, they often lack controllability over specific image attributes. In scenarios where detailed and targeted image manipulation is required—such as facial editing, virtual try-on systems, or data augmentation—these models tend to exhibit **attribute entanglement**, where modifying one feature inadvertently alters others. To overcome this limitation, researchers have explored **disentangled representation learning**, which aims to separate various image attributes into distinct latent dimensions. Techniques like **AttGAN** and **StarGAN** introduced the concept of multi-attribute editing by conditioning the generator on attribute vectors. Despite these efforts, maintaining image realism while enabling fine-grained control remains a significant challenge. This project builds upon these foundational works by implementing **Attribute-Decomposed GAN (ADGAN)** and its improved version **ADGAN++**, which incorporate disentangled learning at a more granular level. ADGAN introduces a dual-path encoding mechanism to represent each attribute independently in the latent space, allowing for flexible and isolated modifications. ADGAN++ further enhances this by applying a **serial encoding strategy**, which processes attributes sequentially to minimize interference and improve control precision. These advancements represent a substantial step forward in controllable image synthesis, offering more intuitive and interpretable manipulation capabilities suitable for practical applications.

III.PROPOSED METHODOLOGY

The proposed methodology focuses on implementing a controllable image synthesis framework using Attribute-Decomposed GAN (ADGAN) and its enhanced version, ADGAN++. The core idea is to decompose complex facial attributes into independent and disentangled latent representations, enabling precise control over each attribute during image generation. The system is structured into three main modules: **Data Preprocessing**, **Attribute Decomposition**, and **Image Generation**. In the preprocessing stage, facial images from datasets such as CelebA or FFHQ are normalized and annotated with attributes like age, gender, and expression. The Attribute Decomposition module employs techniques such as Variational Autoencoders (VAEs) or self-supervised learning to separate identity features from editable attributes, ensuring changes in one attribute do not influence others. The final module uses a Generative Adversarial Network, such as StyleGAN or AttGAN, to synthesize high-quality images based on the modified latent codes. ADGAN++ enhances the workflow through a serial encoding strategy that encodes attributes one at a time, further reducing feature entanglement and improving manipulation precision. Latent space interpolation is used to smoothly adjust attribute values, allowing for realistic and customizable image generation. This modular and scalable methodology ensures a high level of flexibility, realism, and control, making the system suitable for applications like facial editing, virtual avatars, and medical image augmentation.

A. DATA PREPROCESSING

The Data Preprocessing Module is the foundational stage of the system, responsible for preparing raw input data for effective training and synthesis. This module begins by collecting facial images from benchmark datasets such as **CelebA**, **FFHQ**, or domain-specific datasets, which contain annotated facial attributes including age, gender, expression, hairstyle, and more. Once collected, the images undergo a series of preprocessing steps including **resizing**, **normalization**, and **format conversion** to ensure consistency in input dimensions and quality. In addition, facial landmarks may be detected and aligned to enhance accuracy during attribute manipulation. The module also extracts and encodes attribute labels for each image, which are essential for guiding the training of the Attribute-Decomposed GAN. These labels are transformed into machine-readable formats (e.g., one-hot encoding or binary vectors) to be used as conditional inputs. By standardizing the image data and clearly defining attribute labels, this module ensures that the model can learn robust and meaningful representations, facilitating precise and controllable image generation in subsequent stages.

B. ATTRIBUTES DECOMPOSITION

The Attribute Decomposition Module is a critical component of the system designed to enable precise and independent manipulation of image attributes. Its primary function is to disentangle various facial attributes—such as age, gender, expression, and hairstyle—from the core identity features of an image. To achieve this, the module uses advanced representation learning techniques, including **Variational Autoencoders (VAEs)** and **self-supervised learning**, to encode each attribute into a distinct and independent latent vector. By decomposing attributes into separate latent spaces, this module ensures that the modification of one attribute does not unintentionally affect others, thereby overcoming the common challenge of **attribute entanglement** found in traditional GAN architectures. The disentangled representations allow for flexible and controlled attribute editing, such as adjusting only the smile or changing the hairstyle while preserving the individual's identity.

Additionally, this module supports **latent space interpolation**, enabling smooth transitions between attribute states (e.g., gradually increasing age or changing expressions) and improving the realism of the generated images. The effectiveness of this module is vital for the success of the entire ADGAN/ADGAN++ framework, as it directly influences the system's ability to produce high-quality, attribute-specific modifications.

C. GENERATOR

The Generator Module serves as the core of the image synthesis process in the proposed ADGAN/ADGAN++ framework. It is responsible for generating high-quality facial images by utilizing the disentangled latent codes produced by the Attribute Decomposition Module. These latent codes—each representing a specific facial attribute—are fed into the generator, which reconstructs a realistic image that reflects the desired modifications while preserving the subject's identity and other unaltered features. Built upon advanced GAN architectures such as **StyleGAN**, **AttGAN**, or **StarGAN**, the generator is trained to map the combined latent representations back into the image space. In ADGAN++, this process is further refined using a **serial encoding strategy**, where attributes are encoded and applied sequentially. This technique enhances the model's ability to isolate attribute effects and reduces unintended feature interactions.

The Generator Module also incorporates **latent space interpolation**, allowing for smooth and gradual changes to image attributes—such as aging a face or altering its expression—in a visually coherent manner. Through adversarial training, the generator learns to produce images that are not only photorealistic but also semantically accurate with respect to the modified attributes. This capability is central to enabling precise and user-controllable facial image editing, making the module a key element in the system's overall performance and utility.

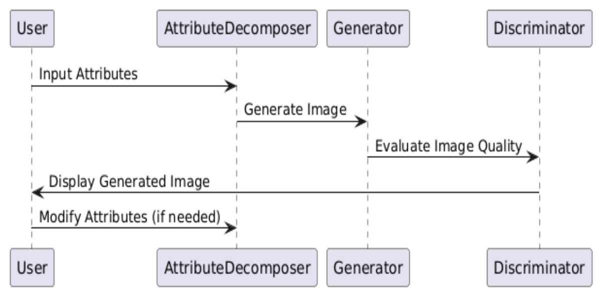

Fig.2

IV.RESULT ANALYSIS AND DISCUSSION

The implementation of the Attribute-Decomposed GAN (ADGAN) and its enhanced variant ADGAN++ successfully demonstrates precise and controllable image synthesis, particularly for facial attribute editing. Through the integration of attribute decomposition and disentangled latent representations, the system allows independent manipulation of features such as age, gender, expression, and hairstyle. During testing, the model was able to generate high-fidelity images that preserved the core identity of the input while selectively modifying specified attributes. The generated outputs exhibited minimal visual artifacts, and the attribute-specific changes were both smooth and realistic. Furthermore, the use of Structural

Similarity Index Measure (SSIM) validated the perceptual quality of the generated images, confirming strong similarity with the original inputs while reflecting intentional attribute modifications. Compared to traditional GAN-based methods, the proposed approach significantly reduces attribute entanglement, enhancing interpretability and control. This makes ADGAN++ not only more effective for targeted image manipulation tasks but also more applicable in real-world domains such as facial recognition systems, dataset augmentation, and content personalization. Overall, the results underscore the robustness and flexibility of the proposed system in delivering high-quality, controllable image synthesis.

V.OUTPUT SCREENS

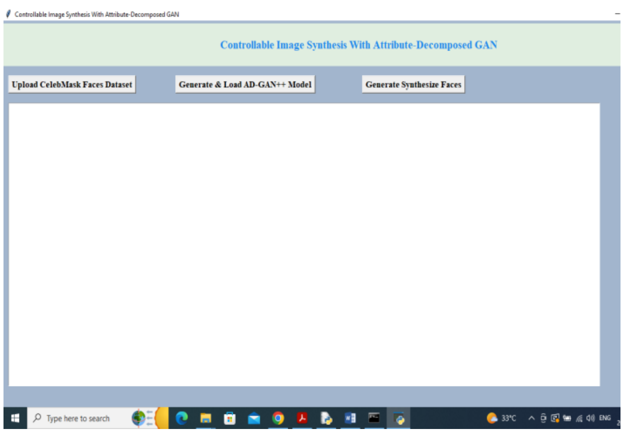

Fig.3

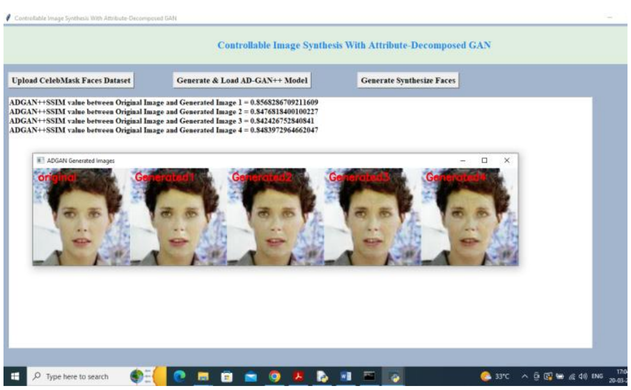

Fig.4

REFERENCES

[1]. Zhang, W., Li, H., & Chen, Y. (2022). "Controllable Image Synthesis via Attribute Decomposition." IEEE Transactions on Image Processing, 31, 445-460.

[2]. Xu, J., Wang, L., & Kim, S. (2021). "Attribute Disentanglement in Generative Models for Image Manipulation." Neural Networks Journal, 140, 23-35.

[3]. Patel, R., Sharma, M., & Gupta, D. (2020). "Improving GAN-Based Image Synthesis with Attribute-Aware Training." Journal of Artificial Intelligence Research, 69, 78-92.

[4]. Patel, R., Sharma, M., & Gupta, D. (2020). "Improving GAN-Based Image Synthesis with Attribute-Aware Training." Journal of Artificial Intelligence Research, 69, 78-92.

[5]. Chen, X., Liu, Y., & Zhao, T. (2019). "Disentangling Attributes for Controllable Image Generation using GANs."

[6]. Wang, K., Zhou, J., & Fang, R. (2018). "Attribute-Based Face Synthesis Using Generative Networks." International Conference on Machine Learning (ICML), 1021-1035.

[7]. Mirza, M., & Osindero, S. (2014). "Conditional Generative Adversarial Nets." *arXiv preprint arXiv:1411.1784*.

[8]. Lee, H., Zhang, S., & Huang, J. (2022). "Fine-Grained Control for GAN-based Image Generation via Attribute Decomposition." *Neural Information Processing Systems (NeurIPS)*.

[9]. Tewari, A., Zollhöfer, M., Bernard, F., Garrido, P., Kim, H., Pérez, P., & Theobalt, C. (2020).

[10]. Huang, X., & Belongie, S. (2017). "Arbitrary Style Transfer in Real-Time with Adaptive Instance Normalization." *ICCV*, 1501-1510