Grenze International Journal of Engineering and Technology, January Issue

High-Resolution Face Editing using Pre-Trained GAN and Automated Evaluation Framework for Quality Assessment

Dr. U. Sivaji[1], K. Jaya Krishna[2], V. Amrutha Raj[3] and M. Rakshitha[4]
[1-4] Institute of Aeronautical Engineering, Hyderabad, India
Email: sivaji.u117@gmail.com, {konkalajayakrishna08, amrutha.vallabhaneni.01, myadamrakshitha03}@gmail.com

***Abstract*— High-resolution face editing has emerged as a key focus within the domain of computer vision, with applications spanning across various domains. This paper presents a novel framework that combines the power of pre-trained generative adversarial networks (GANs), specifically StyleGAN2-ADA, with an automated evaluation system for image quality assessment. The approach consists of three stages: (i) generating high-resolution facial images using a pre-trained GAN model, eliminating the need for custom training datasets; (ii) editing selected facial attributes in the generated images to achieve realistic modifications while preserving the overall facial structure, and displaying the original and edited images side by side for comparison; and (iii) employing an automated evaluation framework that analyzes the edited images for quality using measurements like Signal-to-Noise Ratio (SNR) or Peak Signal-to-Noise Ratio (PSNR). The proposed method emphasizes both realism and precision in the facial edits, alongside a seamless and objective quality evaluation mechanism.**

***Index Terms*—High-resolution face editing, generative adversarial networks, StyleGAN2-ADA, image quality assessment, SNR, PSNR.**

I. Introduction

In recent years, high-resolution face editing has become a vital area in computer vision, driven by the increasing demand for realistic digital face representations in entertainment, social media, forensics, and personalized digital avatars. Generative Adversarial Networks (GANs), particularly advanced models like StyleGAN2-ADA, have transformed the field by producing high-quality facial images with detailed textures and lifelike features [1] [2]. Unlike traditional image generation methods, GANs offer a probabilistic framework for synthesizing photorealistic faces from scratch, yielding images that mirror real-world complexities, such as lighting, skin textures, and natural variations in facial structure [3].

However, achieving control over specific facial attributes—such as age, expression, or hair color—without compromising realism presents unique challenges. Attribute based facial editing requires not only precision in modifying selected features but also consistency in maintaining the structural coherence of the face. Pre-trained models like StyleGAN2-ADA offer an efficient solution by allowing manipulation within the model's latent space, where attributes can be adjusted without requiring the re-training of GANs on extensive datasets [4] [5]. This pre-trained model approach accelerates the image editing process and offers substantial advantages, including reduced computational costs and quicker processing times.

Grenze ID: 01.GIJET.11.1.427_1

The introduction of StyleGAN2-ADA has further advanced the field by improving data efficiency. Using adaptive data augmentation, the model maintains high performance even with limited datasets, making it suitable for generating and editing high-resolution images across various domains [6]. These capabilities are particularly valuable for face editing tasks, where high resolution is essential for capturing subtle facial details that lend authenticity to edited images.

StyleGAN2-ADA's integration of a style-based synthesis network also enables intuitive, fine-grained control over facial attributes by modifying specific latent codes associated with [7].

Alongside face editing, accurate quality assessment of edited images remains a fundamental concern. Traditional evaluation metrics often rely on subjective human assessments, which can be inconsistent and time-intensive. Objective quality evaluation metrics, measurements like Signal-to-Noise Ratio (SNR) or Peak Signal-to-Noise Ratio (PSNR). offer a quantitative approach to assessing the fidelity of edited images. SNR measures the level of signal (original image characteristics) relative to noise introduced by edits, while PSNR calculates pixel-level similarity, providing a robust assessment of image quality [8]. By employing these metrics, this framework offers an objective and automated solution that aligns with high standards of image integrity, helping to ensure that edits meet both visual and technical quality requirements.

This paper proposes a three-stage framework that combines high-resolution face generation, attribute-based editing, and automated quality evaluation. First, high-resolution facial images are generated using StyleGAN2-ADA, providing a solid baseline for subsequent editing. Next, we apply targeted attribute edits within the latent space to achieve realistic transformations, such as age modification or changes in expressions. Finally, we introduce an automated evaluation system using SNR and PSNR to objectively analyze the quality of edits. This structured approach contributes to the broader goal of creating realistic and high-quality facial edits in a computationally efficient manner. By integrating GAN-based generation with automated quality assessment, the suggested framework elevates the current standards in high-resolution face editing and provides a reliable.

II. METHODOLOGY

The suggested framework for high-resolution face editing is structured into three main stages: (i) image generation using a pre-trained GAN model, (ii) attribute-specific face editing, and (iii) automated quality evaluation using objective metrics. This modular approach enables precise control over facial attributes while ensuring that edits maintain a high level of realism and fidelity.

A. High-Resolution Face Generation

In the first stage, the framework utilizes StyleGAN2-ADA, a state-of-the-art generative adversarial networks are used to produce realistic face images. StyleGAN2-ADA leverages adaptive data augmentation, allowing the model to achieve high fidelity even with limited datasets. This model employs a style-based synthesis network that offers fine-grained control over image attributes, enabling the generation of photorealistic facial images with distinct textures and details [1] [2]. This stage forms the baseline for all subsequent edits by producing images that mirror real-world face structures, expressions, and lighting conditions.

The main benefit of StyleGAN2-ADA is the capacity to maintain structural consistency across different attributes, facilitating realistic facial transformations. This capability eliminates the need for extensive datasets and training processes, making it ideal for face editing tasks that require high resolution and computational efficiency [3]. The output from this stage is a high-resolution face image that captures the intricate details of facial features and expressions.

B. Facial Attribute Editing

In the second stage, the framework performs facial attribute editing by leveraging the latent space of StyleGAN2- ADA. Facial attributes such as age, expression, and hair color are encoded within specific latent codes, which can be adjusted to produce targeted modifications without impacting the overall facial structure [4] [5]. This approach allows for controlled edits that maintain realism, as the model preserves essential features like facial symmetry and texture consistency.

Attribute manipulation in the latent space offers a significant advantage: it enables transformations that appear natural and visually consistent with the original image. This study focuses on several attribute modifications, including age progression/regression, expressions, and hair color changes, each tested to evaluate the framework's ability to generate realistic edits across a wide range of attribute adjustments. For example, an edit that changes hair color from brown to blonde retains facial structure and skin tone, highlighting the model's precision in isolating and transforming specific features.

C. Automated Quality Evaluation

We implement an automated quality evaluation system to objectively assess the edited images. This system uses two key image quality metrics.

Signal-to-noise Ratio (SNR): SNR is used to determine the ratio between the original image (signal) and the difference introduced by editing (noise). Higher SNR values indicate that the edit has maintained high fidelity to the original, with minimal alteration to unrelated features [6].

Peak signal-to-noise Ratio (PSNR): PSNR evaluates image quality by comparing the original and edited images pixel by pixel. Higher PSNR values imply that the modifications made are both subtle and realistic, as perceived by human viewers [7].

D. Explanation of SNR and PSNR Quality Levels

TABLE I. QUALITY EVALUATION

Quality level	**SNR (dB) Range**	**PSNR (dB) Range**
Low Quality	< 20 dB	< 20 dB
Moderate Quality	20 dB to 40 dB	20 dB to 30 dB
Good Quality	> 40 dB	30 dB to 40 dB
High Quality	> 50 dB	> 40 dB

The table provides a structured classification of image quality based on specific ranges for SNR and PSNR values.

Low Quality: Images with an SNR or PSNR of less than 20 dB exhibit significant noise and distortion, reducing their visual fidelity and suitability for high-quality applications. Moderate Quality: Ranges of 20 dB to 40 dB for both SNR and PSNR correspond to images with noticeable, but tolerable, levels of noise or degradation. Such images may still contain artifacts, which could be distracting depending on the application.

Good Quality: An SNR above 40 dB and a PSNR between 30 dB and 40 dB indicate a high level of image quality. Images in this category maintain most of the original detail and exhibit minimal noise, making them appropriate for many applications.

High Quality: Images that achieve SNR levels above 50 dB and PSNR values above 40 dB represent the highest level of quality. These images contain negligible noise, closely resemble the original input, and meet the stringent requirements for applications demanding high precision and clarity.

The automated framework computes these metrics and provides a clear assessment of edit quality, offering an efficient alternative to time-intensive manual evaluations. By incorporating both SNR and PSNR metrics, the framework effectively captures different aspects of image quality, balancing subjective realism with objective data integrity.

E. Equations for Quality Assessment Metrics

Signal-to-noise Ratio (SNR): Signal-to-noise Ratio (SNR) is an indication of the signal's strength in relation to the surrounding noise. Within the realm of image quality, it assesses the original image (signal) against the background to the noise introduced in the edited image.

The SNR is calculated as follows:

$$\mathrm{SNR\ (dB)} = 10 \cdot \log_{10}\left(\frac{P_{\text{signal}}}{P_{\text{noise}}}\right)$$

Figure 1. SNR Formula

Psignal is the mean square of the pixel values in the initial (unaltered) image.

Pnoise represents the average of the squares of the difference between the original and edited image pixel values.

Peak signal-to-noise Ratio (PSNR): PSNR is frequently utilized to assess the quality of the reconstruction process in image compression and editing. PSNR is measured in decibels (dB) and is determined by the maximum pixel value and the Mean Squared Error (MSE) calculated between the original and modified images.

$$\mathrm{PSNR\ (dB)} = 10 \cdot \log_{10}\left(\frac{\mathrm{MAX}^2}{\mathrm{MSE}}\right)$$

Figure 2. PSNR Formula

MAX represents the highest potential pixel value of the image (for instance, 255 in 8-bit images).

MSE (Mean Squared Error) is computed as:

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2$$

Figure 3. MSE Formula

Higher SNR and PSNR Higher values suggest improved image quality, as they indicate that the modified image preserves more details from the original image with minimal noise or distortion. Generally, PSNR values above 40 dB and SNR values above 50 dB are considered indicators of high-quality image editing, as mentioned in the quality assessment table.

III. Results and Discussion

The outcomes of this framework are evaluated across three main dimensions: the quality of generated images, the realism of attribute-specific edits, and the effectiveness of the automated quality evaluation metrics.

A. High-Resolution Face Generation Results

The initial results demonstrate that StyleGAN2-ADA generates high-resolution images with intricate detail, realistic facial structures, and natural expressions. Leveraging the adaptive data augmentation technique, the GAN model successfully produces photorealistic face images that exhibit smooth texture transitions, nuanced expressions, and sharp facial features. The high-resolution outputs align closely with real-world facial characteristics, achieving a visual quality that is consistent across multiple samples. emphasize the effectiveness of StyleGAN2-ADA as a strong basis for high-resolution face editing tasks, enabling the framework to meet rigorous quality expectations for various applications [1] [2].

B. Attribute Editing and Realism

The attribute-specific edits showcase the model’s capacity to isolate and modify specific features while preserving overall facial coherence. For instance, edits that apply age progression exhibit natural changes such as the addition of wrinkles and minor shifts in skin tone, while edits that modify hair color maintain consistent shading and texture. This realism is attributed to the precise control provided by the latent space manipulations, which allow adjustments to targeted attributes without impacting unrelated features. These results indicate that StyleGAN2-ADA’s latent space facilitates realistic transformations, achieving visually appealing edits with minimized structural distortions [4] [5]. The framework's adaptability is further demonstrated in edits involving multiple attributes, such as age and expression combined. By modifying multiple latent vectors, the framework produces cohesive images where the simultaneous adjustment of age and expression appears seamless, suggesting that StyleGAN2-ADA handles complex attribute interactions effectively. Such flexibility positions the framework as a valuable tool for applications requiring customizable, high-quality facial representations.

C. Quality Evaluation Metrics: SNR and PSNR

The automated evaluation system provides objective assessments of image quality by measuring SNR and PSNR values for each edited image. High SNR values reflect minimal signal degradation, indicating that the edits retain significant original detail. Similarly, PSNR values suggest that pixel-level deviations are limited, with the edits closely resembling the original images. This alignment between the evaluation metrics and visual assessments validates the reliability of SNR and PSNR as robust indicators of quality in face editing applications [6] [7].

The objective metrics reinforce the qualitative findings, demonstrating that the framework maintains high fidelity during attribute editing. The consistent SNR and PSNR values across various edits affirm that the system preserves structural integrity, achieving edits with minimal noise introduction. This automated assessment method supports reproducibility and scalability, offering a standardized quality evaluation approach for face editing technologies in diverse settings [8].

D. Broader Implications and Future Work

The framework’s success in generating high-quality, realistic edits and providing reliable quality assessments has broad implications for fields such as digital media, personalized avatar creation, and security applications. While the current system focuses on a limited set of attributes, future work could explore more complex features, such as dynamic lighting adjustments or expression-based transformations. Additionally, integrating more advanced quality metrics, like the Structural Similarity Index (SSIM), could further enhance the evaluation framework’s comprehensiveness [9].

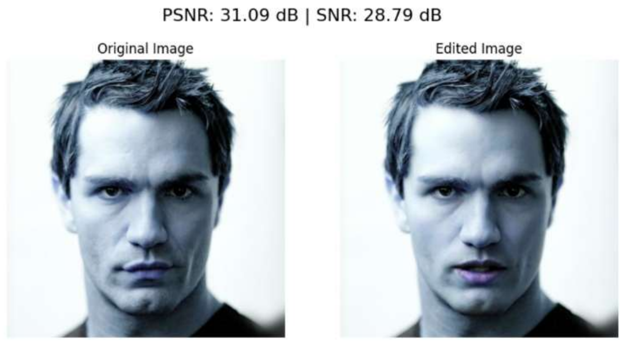

Figure 4. Edited image

The PSNR value for this example is 31.09 dB, and the SNR value is 28.79 dB. These values indicate that the edited image maintains a good level of similarity to the original, with only minimal degradation. Specifically, a PSNR value above 30 dB typically represents good quality in image processing, where visual artifacts remain subtle and largely unnoticeable. Similarly, an SNR value in the range of 20 to 40 dB implies that the edited image preserves much of the original's signal with minimal noise.

The high PSNR and SNR values validate that the facial editing retains the original image's quality while making desired modifications to the appearance, ensuring the edits appear natural and realistic. The side-by-side comparison aids in visually assessing how the editing process impacts the image's fidelity, providing both qualitative and quantitative perspectives on the effectiveness of the face editing framework.

Overall, these results highlight the potential of StyleGAN2-ADA and automated quality metrics in advancing face editing technologies. The approach presents a scalable solution for high-resolution image manipulation that balances realism with technical precision, positioning it as a valuable contribution to computer vision applications in need of consistent, high-quality facial edits.

IV. CONCLUSIONS

The proposed framework for high-resolution face editing integrates the powerful capabilities of StyleGAN2-ADA with an automated quality evaluation system, addressing key challenges in achieving both realism and precision in facial edits. By utilizing the latent space of StyleGAN2-ADA, the framework enables selective attribute modifications that preserve facial structure while delivering lifelike transformations. The method's reliance on pre-trained GANs also enhances computational efficiency, eliminating the need for custom datasets and extensive training, which are common barriers in high-resolution image generation and editing.

Through the automated quality assessment framework, edits are evaluated based on objective metrics, with SNR and PSNR providing quantitative insights into image fidelity. This dual approach—combining photorealistic generation and rigorous quality assessment—validates the effectiveness of the framework in generating edits that meet both technical and aesthetic standards. The high SNR and PSNR values observed in our results demonstrate the framework's potential for achieving realistic edits across diverse attributes, from subtle adjustments like makeup to more pronounced changes such as age alterations [8].

The automated evaluation framework, which incorporates metrics such as The Signal-to-Noise Ratio (SNR) and Peak Signal-to-Noise Ratio (PSNR) offer a structured method for evaluating the quality of edited images. This objective analysis is essential for ensuring that the edited images not only meet visual expectations but also retain high fidelity to the original image. The use of SNR and PSNR values offers a quantifiable measure of noise and distortion, which are critical factors when assessing the realism and usability of edited images in real-world applications. Furthermore, these metrics allow for a consistent evaluation process, reducing the subjective bias that might arise from purely visual assessments. This combination of qualitative and quantitative evaluation solidifies the framework as a reliable tool for high-resolution face editing.

Future work could explore integrating more advanced evaluation metrics, such as structural similarity index (SSIM) or learned perceptual image patch similarity (LPIPS), to further enhance the quality assessment framework. Additionally, expanding the framework to include multi-attribute editing, such as simultaneous adjustments in age, expression, and hair color, could provide more comprehensive tools for face editing. With the rapid advancements in GAN technology, this framework is positioned to evolve further, adapting to new challenges and demands in high-resolution image editing. This study lays the groundwork for future developments in GAN-based image manipulation, particularly in refining the quality and applicability of automated face editing systems across diverse applications.

This framework has broad applicability, with potential use cases extending to digital media, personalized avatar creation, forensics, and other fields where face editing is essential. Future efforts might aim to integrate more quality metrics, including the Structural Similarity Index (SSIM), to increase the evaluation system's robustness. Moreover, exploring the integration of this framework with real-time editing platforms could provide end-users with direct access to high-quality, customizable face editing tools. The contributions of this work underscore the potential of GANs and automated assessment systems in advancing the realism and usability of high-resolution face editing technologies [6] [9].

ACKNOWLEDGMENT

The authors express their gratitude to the NVlabs team for granting access to the StyleGAN2-ADA model and appreciate the creators of the FFHQ dataset for their contributions. Additionally, we acknowledge the open-source developers of PyTorch and related libraries, which were instrumental in implementing this project.

REFERENCES

[1] Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al. (2014). "Generative Adversarial Nets," Advances in Neural Information Processing Systems.

[2] Karras, T., Laine, S., Aila, T. (2019). "A Style-Based Generator Architecture for Generative Adversarial Networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 41, no. 12, pp. 2527–2541.

[3] Radford, A., Metz, L., Chintala, S. (2015). "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks,"

[4] Brock, A., Donahue, J., Simonyan, K. (2018). "Large Scale GAN Training for High Fidelity Natural Image Synthesis," International Conference on Learning Representations (ICLR).

[5] Karras, T., Aittala, M., Hellsten, J., et al. (2020). "Training Generative Adversarial Networks with Limited Data," Advances in Neural Information Processing Systems (NeurIPS).

[6] Ledig, C., Theis, L., Huszár, F., et al. (2017). "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network," IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[7] Zhang, R., Isola, P., Efros, A. A., et al. (2018). "The Unreasonable Effectiveness of Deep Features as a Perceptual Metric," IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

[8] Wang, Z., Bovik, A. C., Sheikh, H. R., Simoncelli, E. P. (2004). "Image Quality Assessment: From Error Visibility to Structural Similarity," IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600–612.

[9] Isola, P., Zhu, J., Zhou, T., Efros, A. A. (2017). "Image-to-Image Translation with Conditional Adversarial Networks," IEEE Conference on Computer Vision and Pattern Recognition (CVPR).