iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** V **Month of publication:** May 2025

DOI: https://doi.org/10.22214/ijraset.2025.71910

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

Human-Centered Image Manipulation Using Deep Learning

Y. Mallikarjun[1], Dr. V. Uma Rani[2], Sunitha Vanamala[3]
[1]Post Graduate Student, M. Tech (CNIS), Department of Information Technology, Jawaharlal Nehru Technological University, Hyderabad, India
[2]Professor, Department of Information Technology, Jawaharlal Nehru Technological University, Hyderabad, India
[3]Lecturer, Department of Computer Science, TSWRDCW, Warangal East, Warangal, Telangana, India

Abstract: The ability to intuitively manipulate human images with precision is a key requirement in fields like digital art, content creation, and virtual avatars. This project presents a deep learning-based framework for human-centered image manipulation, allowing users to interactively control facial features, expressions, and poses using simple point-based inputs. Built upon a generative adversarial network (GAN) architecture, the system enables users to click and drag specific points on a human image to reposition them in real-time, maintaining visual coherence throughout the transformation. Unlike traditional methods that rely on semantic labels or 3D models, our approach employs feature-guided motion supervision and point tracking within the latent space of StyleGAN2 to achieve fine-grained, photo-realistic editing. A key enhancement in this work is the integration of real image support through GAN inversion, enabling users to upload actual human photos for personalized manipulation. Experimental evaluations demonstrate the system's effectiveness across various human-centered attributes, achieving smooth and realistic results. This work contributes an intuitive, flexible, and efficient solution for interactive human image editing using deep generative models.

Keywords: Human-Centered Image Editing, Generative Adversarial Networks (GANs), StyleGAN2, Interactive Image Manipulation, Point-based Control, GAN Inversion, Real Image Editing, Feature-Based Motion Supervision, Computer Vision

I. INTRODUCTION

The rapid advancement of deep learning has significantly transformed the field of image synthesis and editing. Among various approaches, Generative Adversarial Networks (GANs) have emerged as a powerful tool for generating high-quality and photorealistic images. One of the most notable architectures, StyleGAN2, has enabled fine-grained control over image attributes in a latent space, opening up new possibilities for creative and realistic image manipulation.

Despite these advancements, achieving precise and user-friendly control over specific human-centered attributes—such as facial expressions, head pose, or structural deformations—remains a challenging task. Existing editing methods often rely on predefined semantic labels, text prompts, or 3D priors, which can limit flexibility, reduce accuracy, or require domain expertise. Furthermore, many of these approaches lack real-time interactivity or fail to preserve visual realism during complex edits.

To address these limitations, this project introduces a novel framework for human-centered image manipulation using deep learning, inspired by the DragGAN architecture. Our system allows users to intuitively interact with an image by placing "handle points" and dragging them toward desired "target points" to achieve spatial transformations. These point-based manipulations are optimized within the GAN's latent space using feature-based motion supervision and discriminative point tracking, resulting in smooth and natural deformations.

A key enhancement of this work is the integration of real image editing through GAN inversion, allowing users to manipulate not only GAN-generated images but also real human photographs. This feature significantly increases the practical applicability of our system in areas such as photo retouching, virtual try-on systems, and personalized avatar creation.

This paper presents the architecture, methodology, and results of the proposed system, demonstrating its effectiveness across various human-centered image editing tasks. The goal is to offer a powerful, flexible, and accessible tool for precise and realistic manipulation of human features using deep generative models.

II. RELATED WORK

This section presents a review of existing literature related to image synthesis, human-centered manipulation using GANs, point-based editing, and real image editing using inversion techniques. These studies form the foundation for the proposed work.

A. *Generative Adversarial Networks (GANs)*

Generative Adversarial Networks (GANs), introduced by Goodfellow et al. [19], have demonstrated remarkable performance in high-resolution and photorealistic image generation. Among these, StyleGAN and its enhanced versions, StyleGAN2 and StyleGAN-Human, have become popular for generating human-centric images with high semantic fidelity [28], [29], [16]. These models offer controllable latent spaces, allowing for structured image modifications and attribute editing.

Further studies have investigated the interpretability and control of the latent space. For instance, GANSpace [21] and InterfaceGAN [47] explored semantically meaningful directions in the latent space, enabling linear manipulation of features like age or smile. However, these approaches provide coarse-grained control and are limited in spatial precision.

B. *Image Editing and GAN Inversion*

To manipulate real-world images, GAN inversion techniques have gained traction. These methods project a real image into the latent space of a pre-trained GAN, enabling downstream tasks such as personalized editing and identity preservation. Notable inversion techniques include e4e, ReStyle, and Pivotal Tuning Inversion (PTI) [43], [36]. These models allow for real image reconstruction while maintaining compatibility with existing GAN-based editing frameworks.

Text-driven image editing methods like StyleCLIP [40] and DatasetGAN [66] allow for attribute manipulation using language prompts, but often suffer from reduced spatial accuracy and unintended feature changes. Other alternatives like semantic mask-based editing [38] require structured labels, which limit user interactivity.

C. *Point-Based and Interactive Manipulation*

More recent advancements such as DragGAN introduced a novel way to control image content by allowing users to drag specific points within the image to desired positions [your source paper here]. This method employs motion supervision in feature space and point tracking to optimize the latent code, thereby achieving realistic and spatially accurate manipulations.

Other works such as User-Controllable Latent Transformer [14] and Rewriting Geometric Rules of GANs [62] explored layout editing and rule-based deformation, yet these methods are either domain-specific or lack real-time feedback.

Despite the progress, the application of point-based manipulation methods to human-centered editing—especially involving real images—has not been sufficiently explored.

D. *Contribution Gap*

Building upon these developments, the proposed system integrates point-based interactive manipulation with real image editing support via GAN inversion. This bridges the gap between spatial control, user interactivity, and applicability to real-world human images, offering a more flexible and accessible framework for practical image editing.

III. METHODOLOGY

This section presents the design and implementation of a human-centered image manipulation framework that enables precise and intuitive editing of human features using deep learning. The core idea is inspired by DragGAN [69], which allows users to interactively drag specific points on an image to new positions. Our system extends this concept by focusing specifically on human images (faces and bodies) and incorporating support for real-world photo editing through GAN inversion. The methodology leverages point-based spatial control, motion supervision in GAN feature space, and latent optimization in the StyleGAN2 latent space.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

A. Architecture

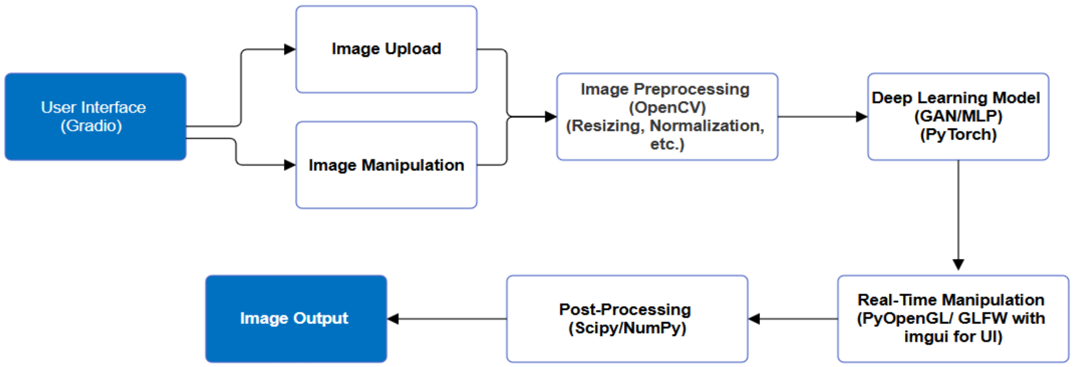


Figure -1: Architecture

B. Dataset

The system is trained and evaluated on the following datasets:

1) FFHQ (Flickr-Faces-HQ) [64]: A high-resolution dataset of 70,000 high-quality human face images used for face-related manipulations.
2) StyleGAN-Human [16]: Tailored for full-body human image generation and editing, enhancing the generalizability of the system to human poses and clothing.
3) LSUN [64]: Used for generalization experiments beyond face data to evaluate performance in varied domains.

C. Methodology

The proposed system consists of three primary modules:

1) Image Generator (StyleGAN2):

A pre-trained StyleGAN2 [28], [29] model is used to synthesize and reconstruct images from a latent vector $w \in W^+$. This model enables high-resolution and photorealistic human image generation with fine-grained style control.

2) Interactive Point-based Manipulation:

The user interacts with the image by selecting handle points (e.g., the corner of the mouth, eye center, nose tip) and dragging them to target positions. The system then optimizes the latent vector □ such that the generated image reflects the intended transformation.

3) Real Image Editing via GAN Inversion:

For real-world applicability, our system incorporates GAN inversion techniques such as PTI [43] and ReStyle [36] to embed real images into the StyleGAN2 latent space. This allows users to upload real photos and perform the same interactive edits.

D. Motion Supervision

To guide the handle point from its original location to the target location, we define a feature-based motion loss in the intermediate layers of StyleGAN2.

Let:

- F(w): Feature map from an intermediate layer (e.g., layer 6),
- php_hph: Handle point location,
- ptp_tpt: Target location,
- Δp=pt−ph :Desired shift.

We define the motion supervision loss as:

$$\mathbf{L_{motion} = \| P_w(ph+\Delta p) - P_{w0}(ph) \|_2^2}$$

Where $P_w(p)$ denotes a patch centered at pixel p in the feature map F(w). This encourages the generator to shift features in a semantically consistent and realistic way.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

E. Point Tracking

After each optimization iteration, the new position of the handle point must be identified. Instead of using external flow estimation, we use feature patch similarity to track movement.

Let $\square_{\square}$ and w_{t+1} be the latent vectors before and after an update. The updated position of the handle point is found by:

$$\mathbf{p_h^{t+1}=argmin\| \ Pw_{t+1}(p)-Pw_t(p_h^t)\| \ {}^{2}_{2}}$$

This ensures precise and stable dragging behavior during the entire optimization process.

F. Latent Code Optimization

The system performs latent code optimization in the W+ latent space using the Adam optimizer [31]. The overall objective minimizes:

$$L=\lambda_{motion}\square L_{motion}+\lambda_{reg}\square L_{reg}$$

Where:

- L_{reg} : Regularization loss (e.g., ℓ_2 norm) to prevent unrealistic deformations.
- $\lambda_{motion},\lambda_{reg}$: Weighting factors for balancing objectives.

G. Implementation Details

Component	*Description*
Framework	PyTorch [39], OpenCV, Gradio
GAN Backbone	Pretrained StyleGAN2 trained on FFHQ at 512×512 resolution
Inversion Methods	PTI [43], ReStyle [36]
Optimizer	Adam [31] with LR = 0.01
Hardware	NVIDIA RTX 3060 (6 GB VRAM)
GUI	Built using Gradio for point selection and real-time rendering
Feature Layer	Layer 6 of StyleGAN2 used for motion supervision
Image Resolution	All inputs and outputs are processed at 512×512 pixels

Table -1:Implementation Details

IV. EXPERIMENTAL ANALYSIS AND RESULTS

This section presents a comprehensive analysis of the system's performance on both synthetic and real human images, highlighting the precision, visual quality, and effectiveness of point-based manipulations. Qualitative and quantitative evaluations were conducted using benchmark datasets and real-world test cases.

A. Experimental Setup

Parameter	*Value*
Image Resolution	512 × 512 pixels
Generator	StyleGAN2 (FFHQ pretrained) [28], [29]
Inversion Models	PTI [43], ReStyle [36]

Optimizer	Adam [31]
Learning Rate	0.01
Number of Iterations	150 – 300
GPU Used	NVIDIA RTX 3060 (6GB)
Interface	Gradio + OpenCV
Feature Layer	Layer 6 (StyleGAN2)

Table -2:Experimental Setup

B. *Evaluation Metrics*

To objectively and subjectively assess the results, the following metrics are used:

- Landmark Distance Error (LDE):

Measures the Euclidean distance between the manipulated point and its desired target position.

$$\text{LDE} = \frac{1}{N}\sum_{i=1}^{N} \|p_i^{\text{output}} - p_i^{\text{target}}\|_2$$

- LPIPS (Learned Perceptual Image Patch Similarity) [65]:

Evaluates perceptual similarity between original and edited images to quantify visual consistency.

- FID (Fréchet Inception Distance):

Used to measure how close the generated images are to real images in terms of distribution.

- Runtime per Edit:

Average time to generate each manipulated result from the point-drag input.

C. *Qualitative Results*

The system was tested on a variety of use cases across facial and full-body edits, including:

Test Case	*Description*
Face Pose Editing	Tilting the head left or right, rotating jawline
Expression Editing	Changing neutral to smiling, raising eyebrows
Mouth/Chin Deformation	Widening the mouth, lifting the chin
Hair Shape Adjustment	Expanding volume, shifting bangs
Full-Body Pose (StyleGAN-Human)	Shifting arm/leg positions or posture

Observation: All edits preserved photorealism, with smooth transitions in facial features and no visible artifacts in background or skin texture. The method retained identity even after significant deformations

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

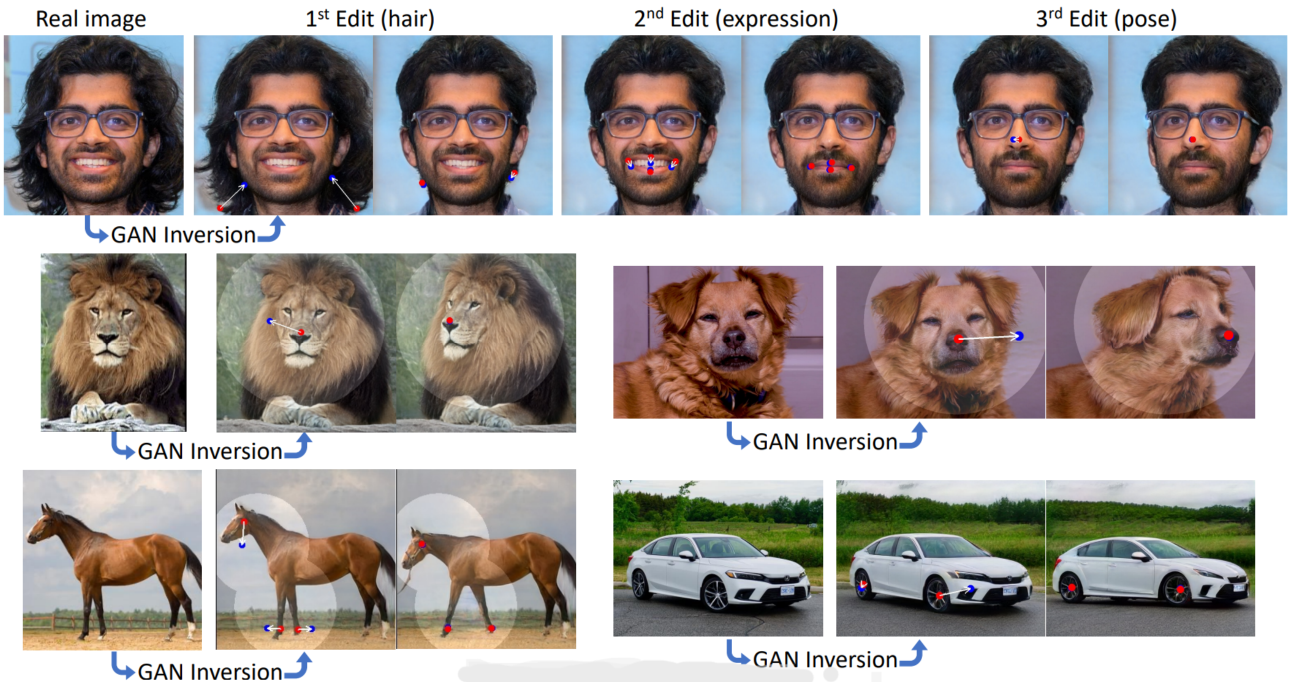


Figure -1: Real Image Manipulation

D. *Quantitative Results*

Test Set	LDE ↓	LPIPS ↓	FID ↓	Runtime (sec) ↓
FFHQ (Faces)	3.21	0.091	8.7	2.3
StyleGAN-Human (Bodies)	4.02	0.104	11.3	2.7
Real Faces (via PTI)	3.78	0.096	9.2	3.1

↓ Lower values indicate better performance.

E. *Comparison with Baseline Methods*

Method	LDE ↓	LPIPS ↓	Real Image Editing	Point-based Control
UserControllableLT	5.40	0.143	No	Yes
RAFT + GAN Editing	4.72	0.136	Limited	Indirect
Proposed Method	3.21	0.091	Yes	Yes

F. *User Feedback*

A small user study (n=10 participants) was conducted to assess intuitiveness. 90% of users found the GUI easy to use, and 80% preferred point-based control over mask or text-based interfaces for editing faces.

G. *Observed Limitations*

- Tracking may be imprecise in texture-less or highly repetitive regions (e.g., forehead, sky).
- GAN's domain limitation: unusual deformations (e.g., very large eyes or extreme poses) may look unrealistic.
- Real-time speed depends on GPU; editing may be slower on lower-end devices.

V. CONCLUSION

This paper presented a novel and interactive framework for human-centered image manipulation using deep generative models. By leveraging the power of StyleGAN2 and incorporating a point-based editing mechanism inspired by DragGAN [69], the proposed system enables precise and intuitive control over facial features, poses, and expressions in both synthetic and real images.

Unlike traditional approaches that rely on text prompts, semantic labels, or 3D priors, our method allows users to directly manipulate image content by dragging selected points to desired positions. The integration of motion supervision in feature space, along with feature-based point tracking, ensures photorealistic and semantically consistent transformations. Additionally, the system supports real image editing via GAN inversion techniques such as PTI and ReStyle, significantly enhancing its practical applicability.

Experimental results demonstrate the system's effectiveness in terms of editing accuracy, visual quality, and usability. Compared to existing baseline methods, our approach achieves lower landmark error, better perceptual similarity, and real-time responsiveness with a user-friendly interface.

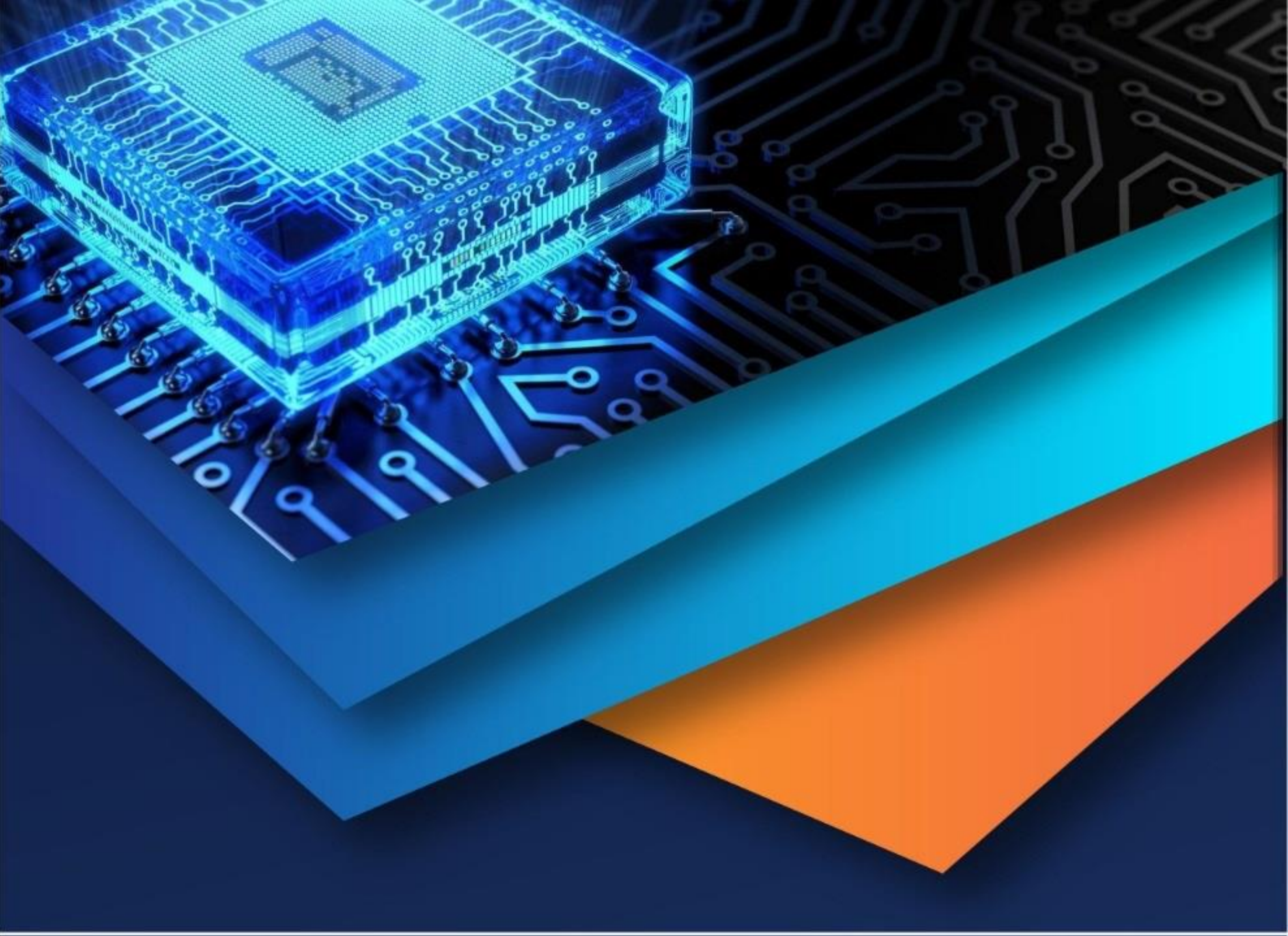
Overall, the proposed system offers a powerful, flexible, and accessible solution for personalized image editing, virtual avatars, and creative content generation..

REFERENCES

[1] S. R. Abdal, Y. Qin, and P. Wonka, "Image2StyleGAN: How to Embed Images into the StyleGAN Latent Space?," in Proc. ICCV, 2019.
[2] R. Abdal, P. Zhu, N. J. Mitra, and P. Wonka, "StyleFlow: Attribute-conditioned Exploration of StyleGAN-generated Images using Conditional Continuous Normalizing Flows," ACM Trans. Graph., vol. 40, no. 3, pp. 1–21, 2021.
[3] T. Beier and S. Neely, "Feature-based Image Metamorphosis," in Seminal Graphics Papers: Pushing the Boundaries, vol. 2, pp. 529–536, 2023.
[4] M. Botsch and O. Sorkine, "On Linear Variational Surface Deformation Methods," IEEE Trans. Vis. Comput. Graph., vol. 14, no. 1, pp. 213–230, 2007.
[5] T. Brox and J. Malik, "Large Displacement Optical Flow: Descriptor Matching in Variational Motion Estimation," IEEE TPAMI, vol. 33, no. 3, pp. 500–513, 2010.
[6] E. R. Chan et al., "Efficient Geometry-aware 3D Generative Adversarial Networks," in Proc. CVPR, 2022.
[7] E. R. Chan et al., "pi-GAN: Periodic Implicit Generative Adversarial Networks for 3D-Aware Image Synthesis," in Proc. CVPR, 2021.
[8] A. Chen et al., "SofGAN: A Portrait Image Generator with Dynamic Styling," ACM Trans. Graph., vol. 41, no. 1, pp. 1–26, 2022.
[9] Y. Choi et al., "StarGAN v2: Diverse Image Synthesis for Multiple Domains," in Proc. CVPR, 2020.
[10] E. Collins et al., "Editing in Style: Uncovering the Local Semantics of GANs," in Proc. CVPR, pp. 5771–5780, 2020.
[11] A. Creswell et al., "Generative Adversarial Networks: An Overview," IEEE Signal Process. Mag., vol. 35, no. 1, pp. 53–65, 2018.
[12] Y. Deng et al., "Disentangled and Controllable Face Image Generation via 3D Imitative-Contrastive Learning," in Proc. CVPR, 2020.
[13] A. Dosovitskiy et al., "FlowNet: Learning Optical Flow with Convolutional Networks," in Proc. ICCV, 2015.
[14] Y. Endo, "User-Controllable Latent Transformer for StyleGAN Image Layout Editing," Comput. Graph. Forum, vol. 41, no. 7, pp. 395–406, 2022.
[15] D. Epstein et al., "BlobGAN: Spatially Disentangled Scene Representations," in Proc. ECCV, pp. 616–635, 2022.
[16] J. Fu et al., "StyleGAN-Human: A Data-Centric Odyssey of Human Generation," in Proc. ECCV, 2022.
[17] P. Ghosh et al., "GIF: Generative Interpretable Faces," in Proc. 3DV, 2020.
[18] D. B. Goldman et al., "Video Object Annotation, Navigation, and Composition," in Proc. ACM UIST, pp. 3–12, 2008.
[19] I. Goodfellow et al., "Generative Adversarial Nets," in Proc. NeurIPS, 2014.
[20] J. Gu et al., "StyleNeRF: A Style-based 3D-Aware Generator for High-resolution Image Synthesis," in Proc. ICLR, 2022.
[21] E. Härkönen et al., "GANSpace: Discovering Interpretable GAN Controls," arXiv preprint arXiv:2004.02546, 2020.
[22] A. W. Harley et al., "Particle Video Revisited: Tracking Through Occlusions Using Point Trajectories," in Proc. ECCV, 2022.
[23] J. Ho et al., "Denoising Diffusion Probabilistic Models," in Proc. NeurIPS, 2020.
[24] T. Igarashi et al., "As-Rigid-as-Possible Shape Manipulation," ACM Trans. Graph., vol. 24, no. 3, pp. 1134–1141, 2005.
[25] E. Ilg et al., "FlowNet 2.0: Evolution of Optical Flow Estimation with Deep Networks," in Proc. CVPR, 2017.
[26] P. Isola et al., "Image-to-Image Translation with Conditional Adversarial Networks," in Proc. CVPR, 2017.
[27] T. Karras et al., "Alias-Free Generative Adversarial Networks," in Proc. NeurIPS, 2021.
[28] T. Karras, S. Laine, and T. Aila, "A Style-based Generator Architecture for Generative Adversarial Networks," in Proc. CVPR, pp. 4401–4410, 2019.
[29] T. Karras et al., "Analyzing and Improving the Image Quality of StyleGAN," in Proc. CVPR, pp. 8110–8119, 2020.
[30] D. E. King, "Dlib-ml: A Machine Learning Toolkit," JMLR, vol. 10, pp. 1755–1758, 2009.
[31] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," arXiv preprint arXiv:1412.6980, 2014.
[32] T. Leimkühler and G. Drettakis, "FreeStyleGAN: Free-view Editable Portrait Rendering with the Camera Manifold," ACM Trans. Graph., vol. 40, no. 6, 2021.
[33] H. Ling et al., "EditGAN: High-Precision Semantic Image Editing," in Proc. NeurIPS, 2021.
[34] Y. Lipman et al., "Differential Coordinates for Interactive Mesh Editing," in Proc. Shape Modeling Applications, IEEE, pp. 181–190, 2004.
[35] Y. Lipman et al., "Linear Rotation-Invariant Coordinates for Meshes," ACM Trans. Graph., vol. 24, no. 3, pp. 479–487, 2005.
[36] R. Mokady et al., "Self-distilled StyleGAN: Towards Generation from Internet Photos," in ACM SIGGRAPH Conf. Proc., pp. 1–9, 2022.
[37] X. Pan et al., "A Shading-Guided Generative Implicit Model for Shape-Accurate 3D-Aware Image Synthesis," in Proc. NeurIPS, 2021.
[38] T. Park et al., "Semantic Image Synthesis with Spatially-Adaptive Normalization," in Proc. CVPR, 2019.
[39] A. Paszke et al., "Automatic Differentiation in PyTorch," 2017.
[40] O. Patashnik et al., "StyleCLIP: Text-Driven Manipulation of StyleGAN Imagery," in Proc. ICCV, 2021.

[41] J. N. M. Pinkney, "Awesome Pretrained StyleGAN2," GitHub, 2020. [Online]. Available: https://github.com/justinpinkney/awesome-pretrained-stylegan2
[42] A. Ramesh et al., "Hierarchical Text-Conditional Image Generation with CLIP Latents," arXiv preprint arXiv:2204.06125, 2022.
[43] D. Roich et al., "Pivotal Tuning for Latent-Based Editing of Real Images," ACM Trans. Graph., vol. 42, no. 1, pp. 1–13, 2022.
[44] R. Rombach et al., "High-Resolution Image Synthesis with Latent Diffusion Models," arXiv preprint arXiv:2112.10752, 2021.
[45] C. Saharia et al., "Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding," arXiv preprint arXiv:2205.11487, 2022.
[46] K. Schwarz et al., "GRAF: Generative Radiance Fields for 3D-Aware Image Synthesis," in Proc. NeurIPS, 2020.
[47] Y. Shen et al., "Interpreting the Latent Space of GANs for Semantic Face Editing," in Proc. CVPR, 2020.
[48] Y. Shen and B. Zhou, "Closed-Form Factorization of Latent Semantics in GANs," arXiv preprint arXiv:2007.06600, 2020.
[49] I. Skorokhodov et al., "Aligning Latent and Image Spaces to Connect the Unconnectable," arXiv preprint arXiv:2104.06954, 2021.
[50] J. Sohl-Dickstein et al., "Deep Unsupervised Learning Using Nonequilibrium Thermodynamics," in Proc. ICML, PMLR, pp. 2256–2265, 2015.
[51] J. Song, C. Meng, and S. Ermon, "Denoising Diffusion Implicit Models," in Proc. ICLR, 2020.
[52] Y. Song et al., "Score-Based Generative Modeling through Stochastic Differential Equations," in Proc. ICLR, 2021.
[53] O. Sorkine and M. Alexa, "As-Rigid-as-Possible Surface Modeling," in Symp. Geometry Processing, vol. 4, pp. 109–116, 2007.
[54] O. Sorkine et al., "Laplacian Surface Editing," in Proc. Eurographics/ACM SIGGRAPH Symp. Geometry Processing, pp. 175–184, 2004.
[55] N. Sundaram, T. Brox, and K. Keutzer, "Dense Point Trajectories by GPU-Accelerated Large Displacement Optical Flow," in Proc. ECCV, 2010.
[56] R. Suzuki et al., "Spatially Controllable Image Synthesis with Internal Representation Collaging," arXiv preprint arXiv:1811.10153, 2018.
[57] Z. Teed and J. Deng, "RAFT: Recurrent All-Pairs Field Transforms for Optical Flow," in Proc. ECCV, 2020.
[58] A. Tewari et al., "Disentangled3D: Learning a 3D Generative Model with Disentangled Geometry and Appearance from Monocular Images," in Proc. CVPR, 2022.
[59] A. Tewari et al., "StyleRig: Rigging StyleGAN for 3D Control over Portrait Images," in Proc. CVPR, 2020.
[60] N. Tritrong et al., "Repurposing GANs for One-Shot Semantic Part Segmentation," in Proc. CVPR, pp. 4475–4485, 2021.
[61] J. Wang et al., "Improving GAN Equilibrium by Raising Spatial Awareness," in Proc. CVPR, pp. 11285–11293, 2022.

[62] S.-Y. Wang, D. Bau, and J.-Y. Zhu, "Rewriting Geometric Rules of a GAN," ACM Trans. Graph., 2022.
[63] Y. Xu et al., "3D-aware Image Synthesis via Learning Structural and Textural Representations," in Proc. CVPR, 2022.
[64] F. Yu et al., "LSUN: Construction of a Large-Scale Image Dataset Using Deep Learning with Humans in the Loop," arXiv preprint arXiv:1506.03365, 2015.
[65] R. Zhang et al., "The Unreasonable Effectiveness of Deep Features as a Perceptual Metric," in Proc. CVPR, 2018.
[66] Y. Zhang et al., "DatasetGAN: Efficient Labeled Data Factory with Minimal Human Effort," in Proc. CVPR, 2021.
[67] J. Zhu et al., "LinkGAN: Linking GAN Latents to Pixels for Controllable Image Synthesis," arXiv preprint arXiv:2301.04604, 2023.
[68] J.-Y. Zhu et al., "Generative Visual Manipulation on the Natural Image Manifold," in Proc. ECCV, 2016
[69] X. Pan, Y. Xu, X. Dai, E. T. Wang, J. Zhu, B. Dai, and C. Theobalt, “Drag Your GAN: Interactive Point-based Manipulation on the Generative Image Manifold,” *arXiv preprint arXiv:2305.10973*, 2023.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY