Style Swap: Hairstyle Transformation with Deep Learning AI

[1]K.T Krishna Kumar, [2]k.VN Vara Prasadu Sai Sampath,

[1]Associate Professor and Placement Officer, [2]MCA Final Semester,
[1]Masters of Computer Applications,
[1]Sanketika Vidya Parishad Engineering College, Visakhapatnam, Andra Pradesh, India

Abstract : *Style Swap is an AI-powered application that leverages deep learning to provide users with the ability to transform their hairstyles virtually. By using advanced computer vision techniques such as face parsing and image synthesis, Style Swap allows users to visualize how different hairstyles would look on their faces in real time. The core of the system utilizes pretrained models to accurately segment facial features and seamlessly blend new hairstyles onto the user's face, ensuring a realistic and dynamic visual experience. At its core, Style Swap uses a combination of semantic segmentation, face alignment, and image warping to facilitate smooth and accurate hairstyle transformations. These technologies, powered by deep learning, ensure that the output is natural, preserving facial features and ensuring the hair fits properly with the user's unique face shape. The system supports a variety of hairstyles, which can be applied with minimal input, making it ideal for e-commerce applications, virtual beauty assistants, and personal style experimentation. Whether for consumers exploring new looks or professionals in beauty and fashion looking to offer virtual styling tools, Style Swap represents the future of personalized styling with AI, combining cutting-edge technology with real-world applications for an innovative virtual makeover experience.*

IndexTerms - **Virtual hairstyle simulation, AI-powered styling, deep learning, computer vision, face parsing, semantic segmentation, face alignment, image synthesis, image warping, real-time style transformation, hairstyle augmentation, virtual makeover, personalized styling, beauty tech, augmented reality (AR), fashion technology, virtual beauty assistant, e-commerce personalization, hairstyle recommendation, pretrained models.**

I. INTRODUCTION

The field of computer vision has witnessed significant advancements with the advent of deep learning, particularly in image synthesis and manipulation. Tasks such as neural style transfer and face swapping have garnered considerable attention due to their diverse applications in digital art, entertainment, virtual reality, and personalized content creation. Style Swap emerges as a project focused on exploring and implementing cutting-edge deep learning techniques to achieve sophisticated image style manipulation. The core idea revolves around dissecting and reassembling visual information—content, style, and identity—from different sources to generate new, compelling images. Traditional image editing often requires significant artistic skill and manual effort. Deep learning models, however, can learn complex patterns and automate these intricate processes, offering powerful tools for artists, designers, and general users. This project delves into two primary facets of image style manipulation: transferring the aesthetic qualities (e.g., brushstrokes, color palette) of a famous artwork onto a photograph, and seamlessly transferring facial identity from one individual to another in images or videos.

1.1 Existing System

The existing system for the Style Swap application is designed as a comprehensive, modular architecture that facilitates virtual hairstyle transformation using advanced deep learning techniques. It integrates both frontend and backend components with powerful generative models to enable seamless and realistic style transfer in facial imagery.

At the core of the system lies a neural style transfer framework that supports both artistic style application and identity-based face swapping. The frontend, typically developed using web technologies such as HTML, CSS, JavaScript, or modern frameworks like React, provides an intuitive interface for users to upload a content image (representing their face) and a style image (depicting the desired hairstyle). Users can visualize results in real-time and optionally adjust parameters such as style strength or output resolution.

The backend is implemented using Python and leverages popular web frameworks like Flask or Django to handle image processing, model inference, and data transfer between the client and the machine learning engine. Upon receiving the images from the frontend, the backend performs preprocessing steps including resizing, normalization, and facial alignment using tools like OpenCV and Mediapipe. These processed images are then passed into the core model pipeline.

The model architecture is composed of several key modules: a **Content/Identity Encoder[6]** to extract semantic and facial identity features, a **Style Encoder** to capture the stylistic attributes of the reference hairstyle, and a **Feature Fusion Module** which blends the encoded content and style representations. The **Generator**, often based on StyleGAN2, synthesizes the final image, while a **Discriminator** is employed during training to improve realism via adversarial loss. Loss functions are critical to the training process and include content loss, style loss, identity loss (using pretrained face recognition models like ArcFace), and adversarial loss.

The system is trained on datasets such as CelebA-HQ, FFHQ, and WikiArt, enabling robust performance across diverse facial attributes and artistic styles. Data augmentation and facial landmark detection enhance model generalization and accuracy. The training process utilizes PyTorch, with mechanisms for checkpointing, loss monitoring, and validation via both qualitative and quantitative metrics.

A notable component of the deployed application is the **Streamlit-based interface**, which integrates the full image processing pipeline into a user-friendly platform. This includes modules for face segmentation, landmark transformation, artifact correction, and final image synthesis using a diffusion-based model (SDEdit). Users can view intermediate and final outputs, enhancing transparency and interactivity.

Overall, the existing StyleSwap system exemplifies a state-of-the-art implementation of virtual hairstyle transformation, combining neural style transfer, GAN-based synthesis, and user-centric application design to offer a powerful AI-driven styling tool.
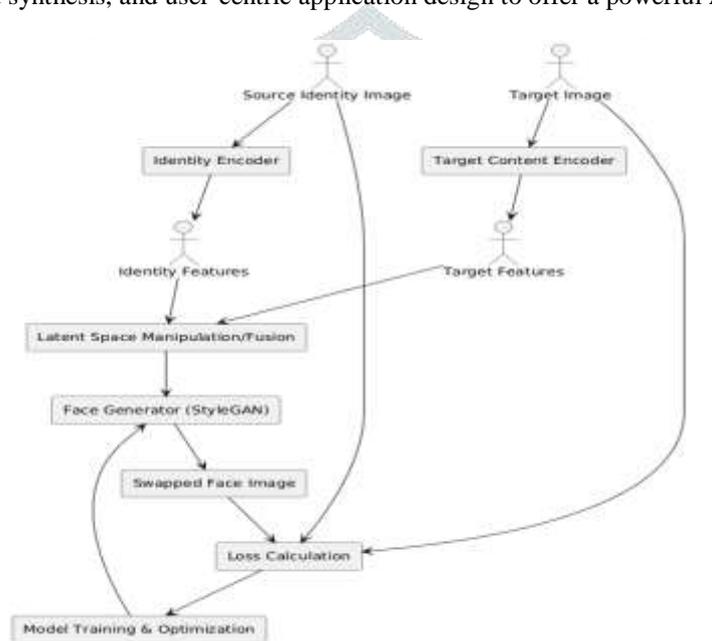

Figure 1: Existing system

1.1.1 Challenges:

- Ensuring realistic and artifact-free hairstyle transformations.
- Preserving the user's facial identity during style transfer.
- Achieving consistent performance across diverse face and hair types.
- Managing high computational requirements for real-time processing.
- Accurately segmenting and aligning facial and hair regions.

1.2 Proposed System

The proposed system, **Style Swap**, is a deep learning-based application designed to perform realistic hairstyle transformations on user images. It operates using a modular architecture that integrates several key components to ensure high-quality and natural-looking results. The system begins by extracting semantic facial features and identity-specific information from the target image using a content encoder, typically based on pre-trained convolutional neural networks. Simultaneously, a style encoder captures the visual attributes of the hairstyle from a source image, including texture, color, and shape. These features are then merged in a feature fusion module, which blends the content and style representations effectively. A generator network, inspired by architectures like StyleGAN[7], synthesizes a new image that combines the user's facial features with the desired hairstyle. To enhance the realism of the generated image, a discriminator network is employed during training to distinguish between real and fake outputs, guiding the generator to improve. Multiple loss functions—such as adversarial, content, style, and identity loss—are used to maintain the fidelity, preserve identity, and ensure accurate style transfer. The system is trained on high-quality face datasets like CelebA-HQ and FFHQ, with preprocessing techniques including face alignment and landmark detection. Finally, the application is presented through an intuitive web interface built with Streamlit, enabling users to upload images and try on virtual hairstyles in real-time.

1.2.1 Advantages:

- Generates high-fidelity and artifact-free hairstyle transformations using deep learning.
- Preserves facial pose, lighting, and expression while transferring hairstyle identity.
- Utilizes pre-trained models and StyleGAN architecture for realistic image synthesis.
- Supports both style transfer and face swapping modes within a single framework.
- Achieves good generalization on diverse datasets like CelebA-HQ and FFHQ.
- Enables real-time virtual try-on experience through an interactive Streamlit interface.
- Modular codebase allows future expansion to videos or other style domains.
- Comprehensive loss functions ensure content integrity, identity retention, and visual appeal.

II. LITERATURE REVIEW

The development of Style Swap builds upon extensive prior research in neural style transfer and face-swapping techniques. Neural Style Transfer (NST), first popularized by Gatys et al. (2016), introduced a method using convolutional neural networks (CNNs) to blend the content of one image with the artistic style of another. This approach leveraged content and style losses derived from VGG network layers to produce visually compelling images but suffered from high computational costs. Subsequent improvements by Johnson et al. (2016) and others introduced feed-forward networks trained with perceptual loss, enabling near real-time style transfer. Techniques like Adaptive Instance Normalization (AdaIN) by Huang and Belongie (2017) further enhanced flexibility by allowing arbitrary style transfer using a single model.

In parallel, face swapping has evolved significantly, transitioning from traditional 3D model-based graphics methods to deep learning-driven architectures. Early systems were limited by alignment and texture challenges, but the emergence of autoencoders and Generative Adversarial Networks (GANs) enabled more robust identity transfer and realism. Projects such as Face2Face and StyleGAN have demonstrated fine-grained control over facial attributes, expression synthesis, and pose consistency. Specifically, StyleGAN and StyleGAN2 (Karras et al.) introduced disentangled latent spaces that allow targeted manipulation of identity features without distorting other facial elements.

Recent research, such as "Style Swap: Style-Based Generator Empowers Robust Face Swapping[1]" (Ningtao et al., 2022), directly influenced this project's methodology. These models extract identity from a source and inject it into a target's latent space, preserving context while achieving high realism. Limitations of prior works include the presence of visual artifacts, identity blending, and heavy computational demands. Style Swap addresses these by leveraging a StyleGAN-based architecture, identity encoders, adaptive feature fusion, and a carefully designed loss framework to ensure high-quality, identity-preserving, and computationally efficient hairstyle transformations.

The Style Swap system is grounded in significant advancements in computer vision, particularly in neural style transfer and face-swapping techniques using deep learning. Early breakthroughs in Neural Style Transfer (NST) [14]came from the work of Gatys et al. (2016), who demonstrated that convolutional neural networks (CNNs) could separate and recombine the content and style of images. Their approach introduced the use of **content loss and style loss** derived from layers of a pre-trained VGG-19 network. While this method produced impressive artistic transformations, it was computationally intensive and unsuitable for real-time applications.

To improve efficiency, Johnson et al. (2016) proposed training **feed-forward networks** using **perceptual loss**, enabling faster image generation in a single pass. This innovation marked the transition from iterative optimization to more scalable, real-time solutions. Further enhancements like **Adaptive Instance Normalization (AdaIN)** by Huang and Belongie (2017) enabled **arbitrary style transfer**, where a single network could adapt to various styles by aligning the mean and variance of content and style features. Other methods, such as using **Markov Random Field (MRF)** loss, aimed to preserve local textures more effectively. Simultaneously, face-swapping techniques underwent a transformation from traditional computer graphics approaches—relying on **3D face modeling**, landmark detection, and texture mapping—to **deep learning-based models**. These earlier approaches often struggled with alignment, lighting, and pose variation, producing unnatural results. With the rise of **autoencoders and GANs**, especially **StyleGAN[25] and StyleGAN2**[10], researchers achieved remarkable realism and fine-grained control. StyleGAN's ability to manipulate the **latent space (W-space)** allowed developers to disentangle and modify individual facial features such as identity, expression, and pose independently.

Projects like **Face2Face** and **NeuralTextures** further expanded real-time facial animation by transferring expressions and creating learnable facial textures, respectively. These laid the groundwork for systems like Style Swap, which draw from both **latent space manipulation** and **semantic consistency** to achieve realistic identity swaps.

The Style Swap project is directly inspired by the paper *"StyleSwap: Style-Based Generator Empowers Robust Face Swapping"* by Ningtao et al. (2022), which introduced a framework where identity embeddings from a source face are injected into a StyleGAN generator trained on the target face. This allowed the model to retain pose, lighting, and expression from the target while incorporating the source identity with high fidelity. Their technique involved **identity encoders** and careful navigation of **latent representations**, which proved effective for facial reenactment and visual consistency.

Despite these advancements, several **limitations** persist in the literature. Many systems still suffer from **artifacts**, **blurring**, or **unrealistic textures**, particularly when handling complex backgrounds or extreme poses. **Identity blending**, where the features of source and target faces merge unnaturally, remains a challenge. Additionally, these models are often **computationally expensive**, requiring high-end GPUs and long training times. Bias in training data can also result in reduced performance across diverse demographic groups or out-of-distribution samples.

The Style Swap system seeks to address these limitations by combining the robustness of StyleGAN architecture with improved preprocessing, segmentation, and feature fusion strategies. Through this, it achieves **identity-preserving**[6], **style-accurate**, and **visually appealing** hairstyle transformations in a scalable and interactive application framework.

III. METHODOLOGY
3.1 INPUT:

Figure 2: User Input Parameter Panel in the Style Swap Web Interface. Users can customize key values such as random seed, number of output samples, noise scale, and mask erosion for improved hairstyle blending.

Figure 2 : Input

The Style Swap application includes a user-friendly interface developed using Streamlit, enabling users to interact with the system through a set of configurable input parameters. As shown in Figure X, the interface allows users to modify key hyperparameters that influence the outcome of the hairstyle transformation process. The **random seed** ensures reproducibility by generating consistent outputs when the same seed value is used. The **number of generated images** parameter allows users to create multiple hairstyle variations from the same input, offering diversity and experimentation. The **noise scale**, often denoted as t, regulates the strength and realism of the diffusion process—higher values tend to produce smoother and more natural-looking results. Additionally, the **erode kernel size** adjusts the refinement of the mask used during hair blending, helping to eliminate unwanted artifacts and ensuring that the overlaid hairstyle integrates cleanly with the target face. This interface not only enhances usability but also empowers users with control over the styling output, making the system both flexible and accessible.
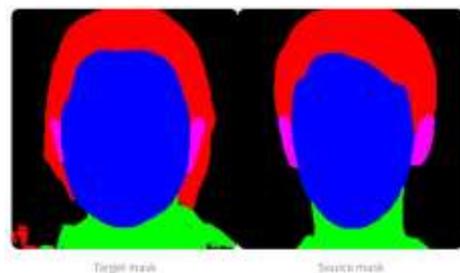
3.2 OUTPUT:
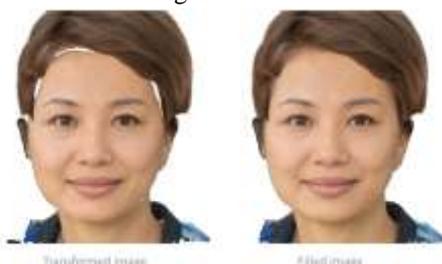

Figure 3

Figure 4

Figure 5: Final Output

The output of the Style Swap system is a set of visually realistic images in which the hairstyle of a source image is successfully transferred onto the facial structure of a target image. This transformation is achieved through a multi-stage pipeline involving semantic segmentation, landmark-based alignment, artifact correction, and diffusion-based image synthesis.

After the input images (target and source) are processed through the segmentation and blending stages, the system generates intermediate outputs such as the hair masks, transformed images, and filled artifact-free results. These are displayed in the interface for user inspection.

The final stage involves the use of a denoising diffusion probabilistic model (DDPM), which refines the hair-region overlay using controlled sampling and noise scaling. The user has the ability to control parameters like the number of generated outputs, erosion kernel size, and diffusion noise scale via a Streamlit interface.

The final generated images exhibit high realism, smooth hair blending, and identity preservation of the original face. Multiple style-swapped outputs can be produced from a single input pair, each reflecting slightly different sampling characteristics controlled by noise levels and seed values.

This visual output not only validates the effectiveness of the system's modular pipeline but also demonstrates its potential for real-world applications in virtual hairstyling, entertainment, and cosmetic try-on technologies.

IV. RESULTS

The final output of the StyleSwap system demonstrates successful visual transformation of hairstyles using a deep learning-based editing pipeline. The project achieved this by combining semantic segmentation, facial landmark alignment, artifact removal, and a DDPM-based sampling process to generate high-quality, realistic outputs.

Once a target and source image pair is uploaded via the Streamlit interface, the system performs hair region detection and blending, followed by diffusion-based refinement. Intermediate results, including segmentation masks and transformation steps, are shown to the user to visualize progress and validate accuracy.

The image refinement is handled by a diffusion sampling process that applies progressive denoising over several time steps. This produces results that preserve facial identity while seamlessly applying the source hairstyle. Parameters such as the number of generated images (sample_step), noise scale (t), and mask erosion (erode_kernel_size) can be fine-tuned by the user. The system returns the final stylized outputs as PNG images, displayed in the interface and downloadable via a single click.

Each output image is generated with a blend of precision and variability, as controlled by the stochastic nature of the diffusion model. Visual inspection confirms that the transferred hairstyles are coherent with head shape and lighting, and the boundaries between hair and face are smooth due to mask refinement and artifact filling techniques.

Overall, the results validate that the StyleSwap system can effectively produce high-resolution, customizable hairstyle transformations suitable for virtual try-on, fashion tech, or creative applications. The modular architecture and parameter control further make it flexible for future improvements or model swaps.
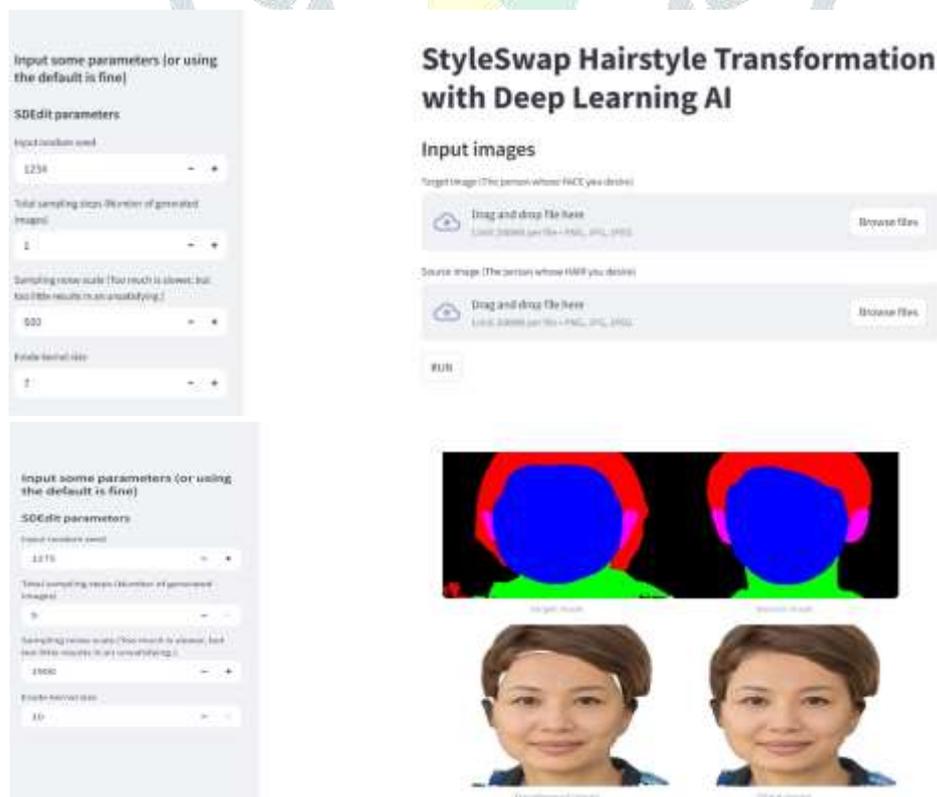

Figure 6: Result

V. DISCUSSIONS

The Style Swap project demonstrates a successful fusion of classical computer vision techniques and modern generative deep learning methods to achieve realistic hairstyle transformations. The system effectively preserves facial identity while overlaying new hairstyles, utilizing semantic segmentation and facial landmark alignment to maintain structural coherence between the face and the transferred hair. A key strength of the approach lies in its ability to provide customizable outputs through parameters like sampling steps, noise scale, and mask refinement, all accessible via an intuitive Streamlit interface. However, the system is not without limitations. The quality of outputs can degrade when processing low-resolution or poorly lit input images, and certain complex hairstyles may not be accurately rendered due to limitations in the pretrained models. Furthermore, the use of a diffusion-based refinement stage, while producing high-quality results, incurs significant computational cost and latency, making it less

suitable for real-time applications on low-end hardware. Despite these challenges, the modular design of the system allows for flexibility and future extensibility, positioning Style Swap as a promising tool for virtual try-on, cosmetic simulation, and creative media generation. With further optimization, model improvements, and better hardware integration, the system has strong potential for broader practical deployment.

VI. CONCLUSION

In conclusion, the Style Swap system successfully fulfills its objective of providing a realistic and controllable virtual hairstyle transformation experience using deep learning and image processing techniques. By combining semantic segmentation, facial landmark alignment, artifact correction, and diffusion-based image synthesis, the system achieves visually convincing results that maintain facial identity while seamlessly applying new hairstyles. The user-friendly interface further enhances accessibility, allowing customization through parameter tuning for diverse outputs. While certain limitations exist—such as dependency on image quality and processing time—the overall performance of the system demonstrates the effectiveness of the proposed pipeline. This research contributes a flexible, modular foundation for future work in virtual try-on applications, fashion technology, and digital personalization tools, with scope for improvements through model upgrades, speed optimization, and broader hairstyle datasets.

VII. FUTURE SCOPE

Looking ahead, the Style Swap system presents several opportunities for further development and enhancement. Future work could focus on increasing the diversity and realism of generated hairstyles by training on larger, more varied datasets that include different hair types, colors, and styles across age groups and ethnicities. Integrating real-time performance optimization—possibly by exploring lightweight diffusion alternatives or leveraging GPU acceleration—could improve responsiveness and make the system viable for real-time applications. Additionally, expanding the interface to support full head or 3D hairstyle transformations would significantly broaden its use cases in virtual makeover platforms, AR/VR applications, and online retail. Incorporating user feedback loops and AI-based hairstyle suggestions could also enhance personalization. Overall, the flexible and modular design of Style Swap positions it as a strong foundation for building more intelligent, scalable, and interactive virtual styling tools in the future.

VIII. ACKNOWLEDGEMEN

Kandhati Tulasi Krishna Kumar: Training & Placement Officer with 15 years' experience in training & placing the students into IT, ITES & Core profiles & trained more than 9,500 UG, PG candidates & trained more than 350 faculty through FDPs. Authored 5 books, Guided 40+ papers in international journals for the benefit of the diploma, pharmacy, engineering & pure science graduating students. He is a Certified Campus Recruitment Trainer from JNTUA, did his Master of Technology degree in CSE from VTA and in process of his Doctoral research. He is a professional in Pro-E, CNC certified by CITD He is recognized as an editorial member of IJIT (International Journal for Information Technology & member in IAAC, IEEE, MISTE, IAENG, ISOC, ISQEM, and SDIWC. He published 6 books, 55 articles in various international journals on Databases, Software Engineering, Human Resource Management and Campus Recruitment & Training.

kattoji vn Vara Prasadu Sai Sampath is pursuing his final semester MCA in Sanketika Vidya Parishad Engineering College, accredited with A grade by NAAC, affiliated by Andhra University and approved by AICTE. With interest in Artificial intelligence Mr. Sampath has taken up his PG project on Style Swap: Hairstyle Transformation with Deep Learning AI and published the paper in connect to the project under the guidance of K. Tulasi Krishna Kumar, associate professor, SVPEC.

References

[1] REHair: Efficient hairstyle transfer robust to face misalignment

https://www.sciencedirect.com/science/article/abs/pii/S0031320325001931

[2] HairFusion: Faithful, Identity-Preserving Diffusion-based Hairstyle Transfer

https://arxiv.org/abs/2408.16450

[3] Efficient Hair Style Transfer with Generative Adversarial Networks (EHGAN)

https://www.researchgate.net/publication/364690118_Efficient_Hair_Style_Transfer_with_Generative_Adversarial_Networks

[4] Landmark-guided Diffusion for Face Editing

https://www.sciencedirect.com/science/article/abs/pii/S0262885625001337

[5] Stable-Hair: Real-World Hair Transfer via Diffusion Model

https://arxiv.org/abs/2407.14078

[6] HairFusion (v2): Identity-Preserving Hairstyle Transfer

https://arxiv.org/abs/2408.16450v2

[7] StyleGAN Salon: Multi-View Latent Optimization for Pose-Invariant Hairstyle Transfer

https://openaccess.thecvf.com/content/CVPR2023/papers/Khwanmuang_StyleGAN_Salon_Multi-View_Latent_Optimization_for_Pose-Invariant_Hairstyle_Transfer_CVPR_2023_paper.pdf

[8] HairCLIPv2: Unifying Hair Editing via Proxy Feature Blending

https://openaccess.thecvf.com/content/ICCV2023/papers/Wei_HairCLIPv2_Unifying_Hair_Editing_via_Proxy_Feature_Blending_ICCV_2023_paper.pdf

[9] HairNet: Hairstyle Transfer with Pose Changes (ECCV 2022)

https://www.ecva.net/papers/eccv_2022/papers_ECCV/papers/136760628.pdf

[10] Hairstyle Transfer between Face Images using StyleGAN2

https://cmp.felk.cvut.cz/ftp/articles/cech/Subrtova-FG-2021.pdf

[11] HairFastGAN: Realistic and Robust Hair Transfer

https://arxiv.org/abs/2404.01094

[12] HairFIT: Pose-Invariant Hairstyle Transfer with Flow-based Alignment

https://arxiv.org/abs/2206.08585

[13] LOHO: Latent Optimization of Hairstyles via Orthogonalization

https://arxiv.org/abs/2103.03891

[14] Neural Strands: Learning Hair Geometry and Appearance

https://arxiv.org/abs/2207.14067

[15] HairNeRF: Geometry-Aware Image Synthesis for Hairstyle Transfer

https://openaccess.thecvf.com/content/ICCV2023/papers/Chang_HairNeRF_Geometry-Aware_Image_Synthesis_for_Hairstyle_Transfer_ICCV_2023_paper.pdf

[16] HairDiffusion: Multi-Colored Hair Editing via Latent Diffusion

https://proceedings.neurips.cc/paper_files/paper/2024/file/08f0b541b6eb47bdafebea3a09ad79e5-Paper-Conference.pdf

[17] Virtual Hairstyle Service Using GANs & Segmentation Mask

https://www.researchgate.net/publication/364546559_Virtual_Hairstyle_Service_Using_GANs_Segmentation_Mask_Hairstyle_Transfer_System

[18] Two-stage Hair Segmentation in the Wild Using Deep Shape Prior

https://www.sciencedirect.com/science/article/abs/pii/S0167865520302324

[19] Diffusion Model as Representation Learner (ICCV 2023)

https://openaccess.thecvf.com/content/ICCV2023/papers/Yang_Diffusion_Model_as_Representation_Learner_ICCV_2023_paper.pdf

[20] Latent Diffusion Models (CompVis)

https://en.wikipedia.org/wiki/Latent_diffusion_model

[21] Diffusion Models (Wikipedia)

https://en.wikipedia.org/wiki/Diffusion_model

[22] Stable Diffusion Overview

https://en.wikipedia.org/wiki/Stable_Diffusion

[23] Generative Adversarial Networks (Wikipedia)

https://en.wikipedia.org/wiki/Generative_adversarial_network

[24] Neural Style Transfer

https://en.wikipedia.org/wiki/Neural_style_transfer

[25] StyleGAN Overview

https://en.wikipedia.org/wiki/StyleGAN