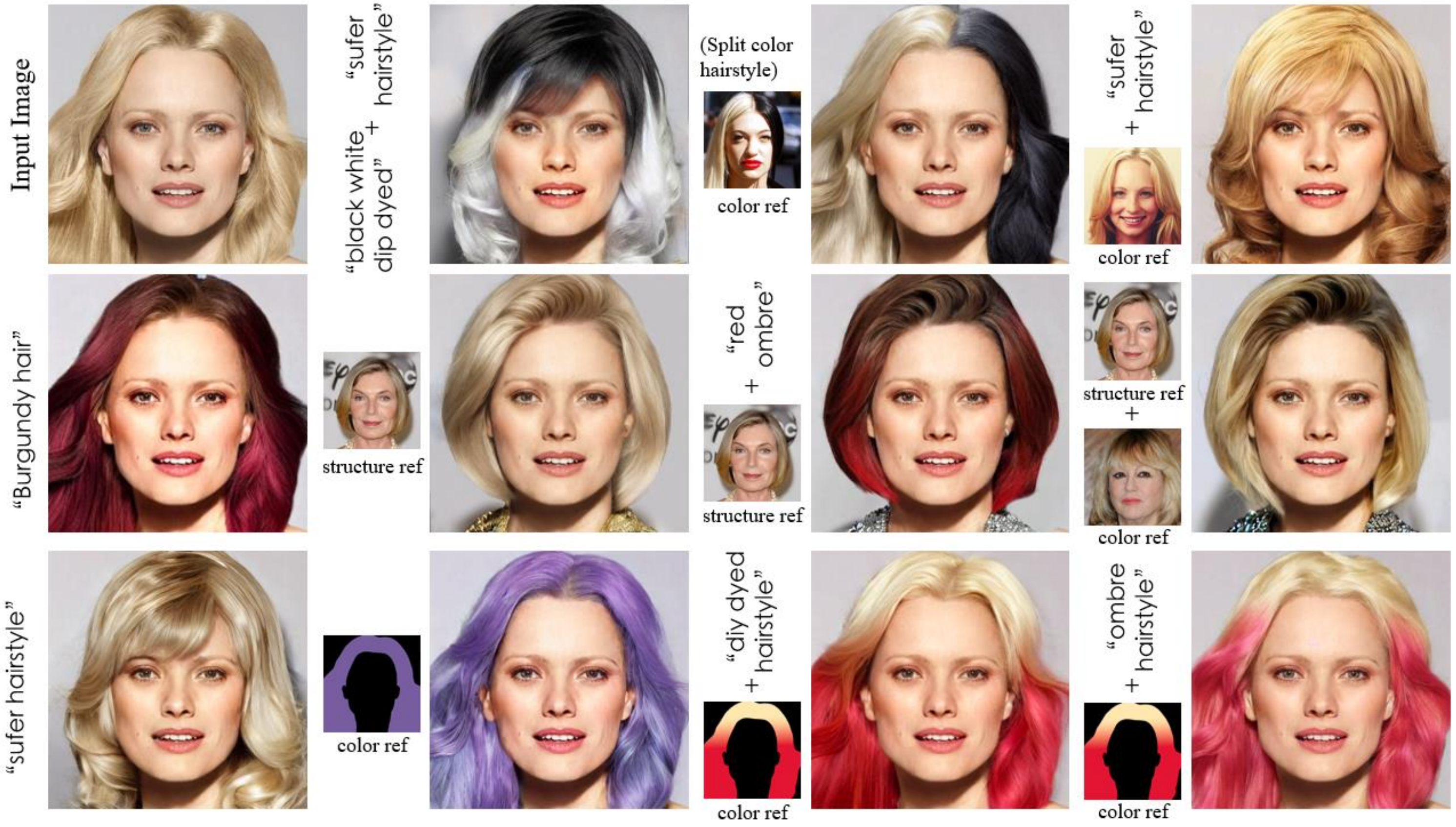
HairDiffusion

Vivid Multi-Colored Hair Editing via Latent Diffusion

Yu Zeng[1], **Yang Zhang**[1*], **Jiachen Liu**[1], **Linlin Shen**[1,2,3]
Kaijun Deng[1], **Weizhao He**[1], **Jinbao Wang**[3,4]
[1]Computer Vision Institute, School of Computer Science & Software Engineering, Shenzhen University
[2]Shenzhen Institute of Artificial Intelligence and Robotics for Society
[3]National Engineering Laboratory for Big Data System Computing Technology, Shenzhen University
[4]Guangdong Provincial Key Laboratory of Intelligent Information Processing

Background



- Previous methods have overlooked the **hair color structure**.

"ombre hair" means a hair color structure with a gradient transition from top-to-bottom,
"split color hair" exhibits a hair color structure with a left-to-right transition···

Background

GAN-based models

1) Difficulty in generating intricate hair color and hairstyle due to the insufficient diversity in the training data`s multi-color hair distributions;
2) Challenges in preserving the original facial information when editing the latent code after mapping images into latent space, leading to difficulties in editing images while preserving irrelevant attributes.

Diffusion-based models

1) Lack of tailored masks for hairstyle inpainting, necessitating consideration of hairstyle regions while preserving irrelevant attributes
2) Difficulty in providing sufficient control for the hair editing task, which requires faithful transfer of hair color from another image or retaining the original hair color of the image
3) Limitations in text and semantic understanding related to hair color and hairstyle, hinder the precision of CLIP-guided diffusion processes

Contribution

1) We present a **warping module** designed for hair warping, allowing the alignment of the target

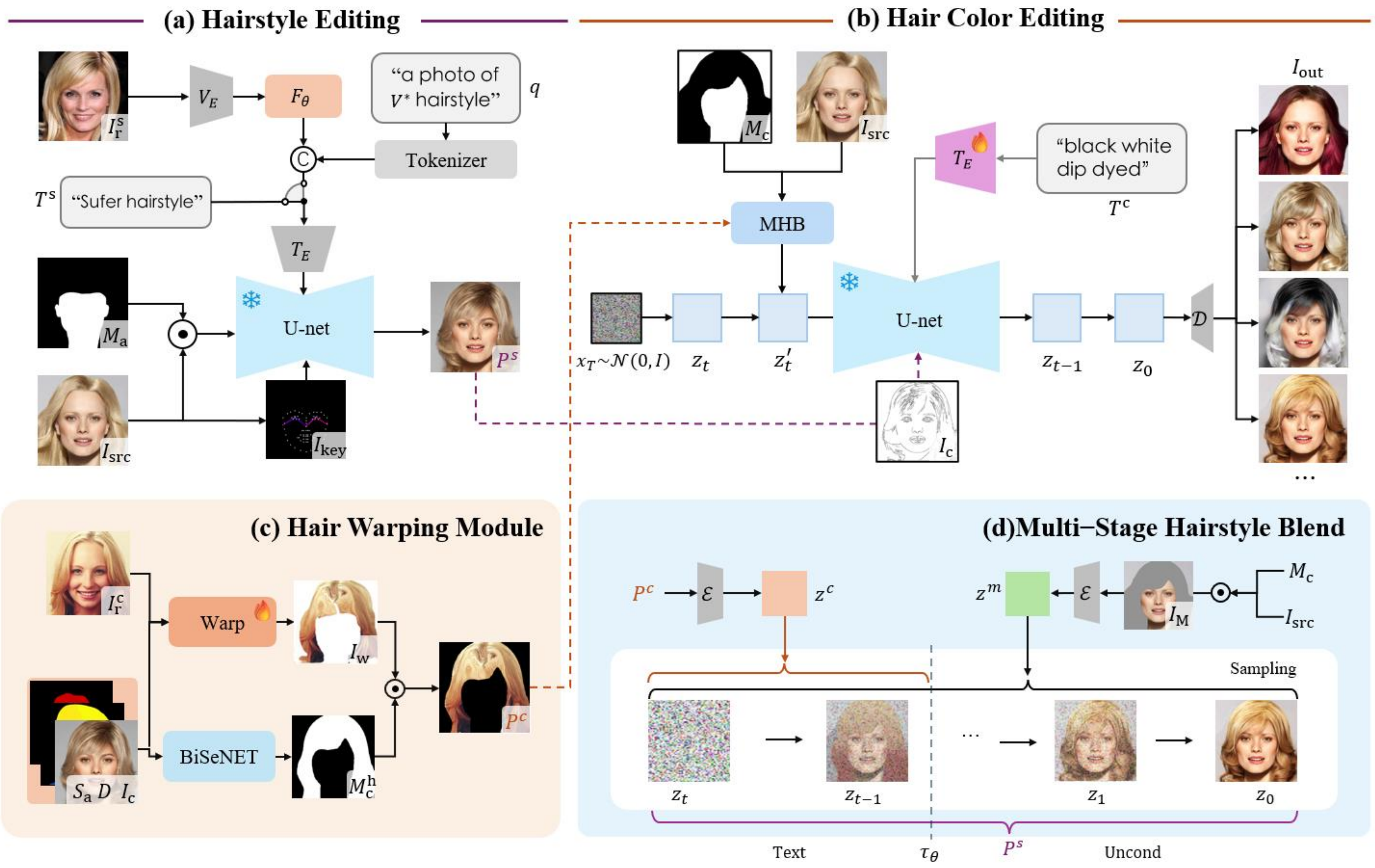
hair mask with precision and enabling comprehensive hair color structure editing through reference images.

2) The **MHB method** is proposed within LDMs, which enables the decoupling of hair color and hairstyle, thereby effectively achieving high-quality hair color and hairstyle editing.

3) Through extensive qualitative and quantitative evaluations, we showcase the superior performance of our method in text-based hairstyle editing, **reference image-based hair color editing**, and **preservation of facial attributes**

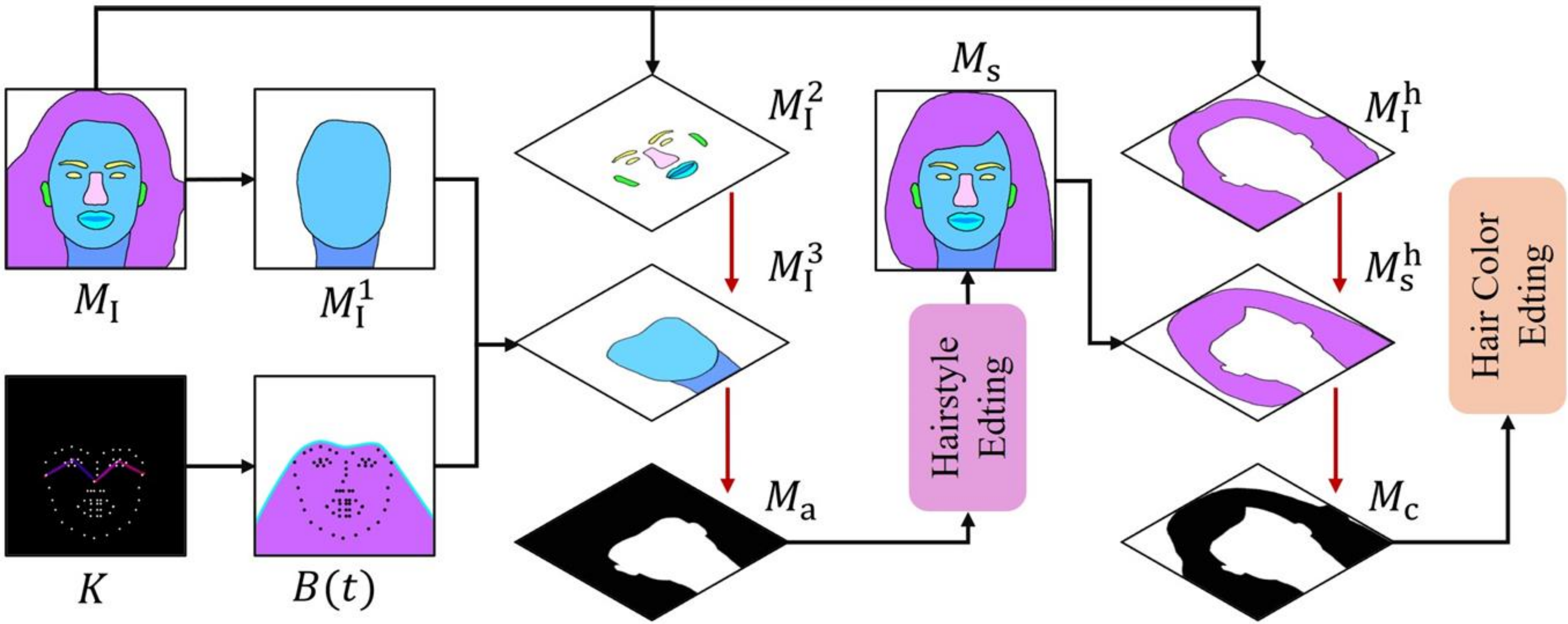
4) The application of LDMs to address the challenge of text and image-based hair editing is pioneered through the introduction of **hair-agnostic facial representation masks**, reframing hair editing as an inpainting task and representing a novel approach. To the best of our knowledge, this method has not been previously explored in this domain.

Quick Preview

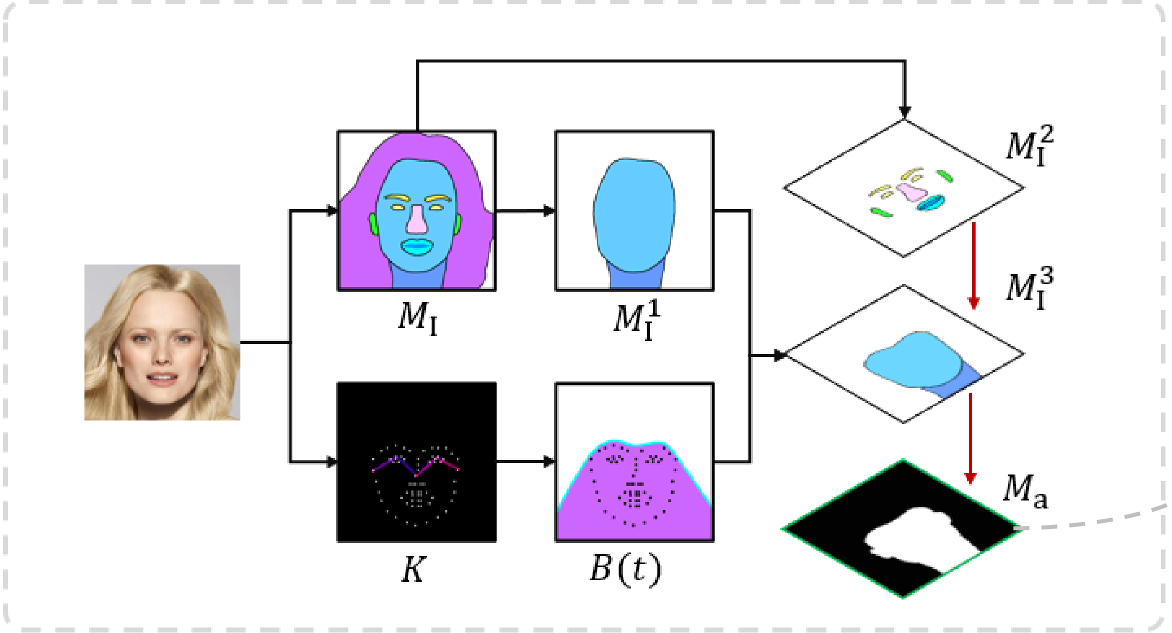


Overview of HairDiffusion

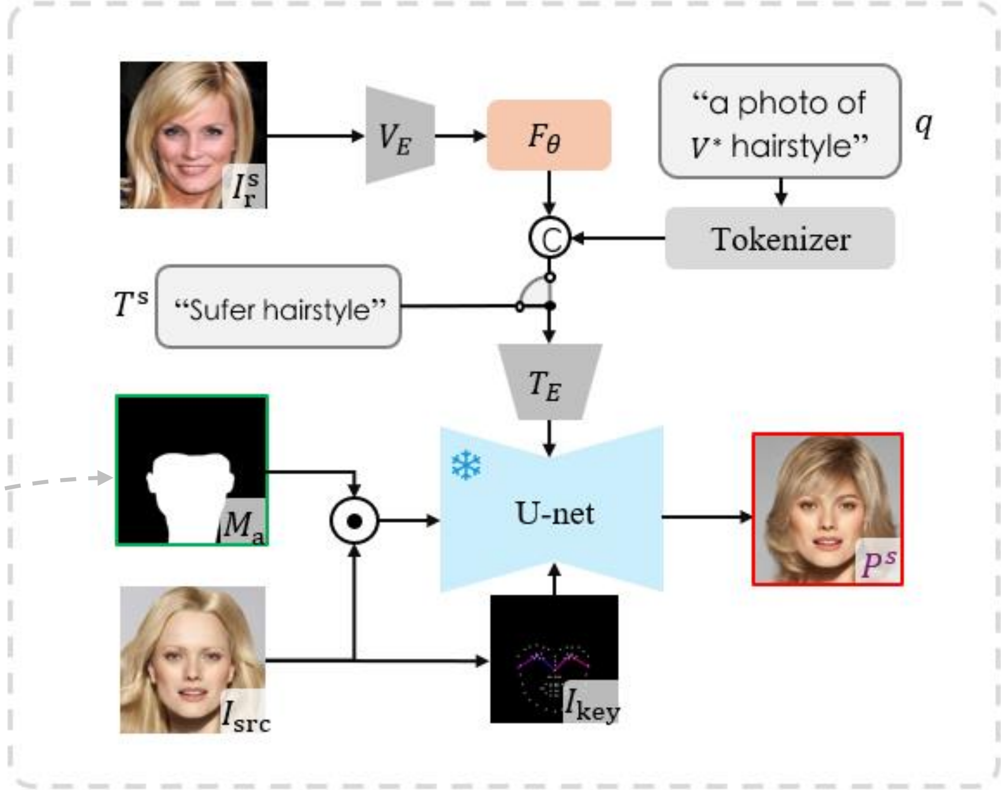
Design of Hair-Agnostic Masks



Hairstyle Editing

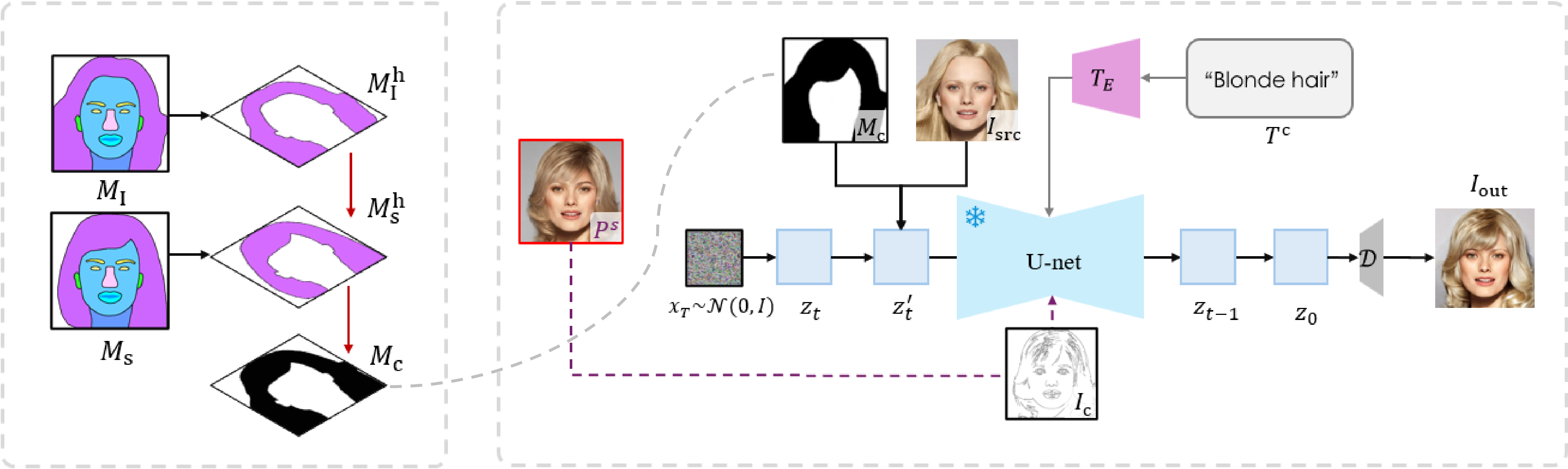


(a) Agnostic Mask in Hairstyle editing Stage



(b) Hairstyle Editing

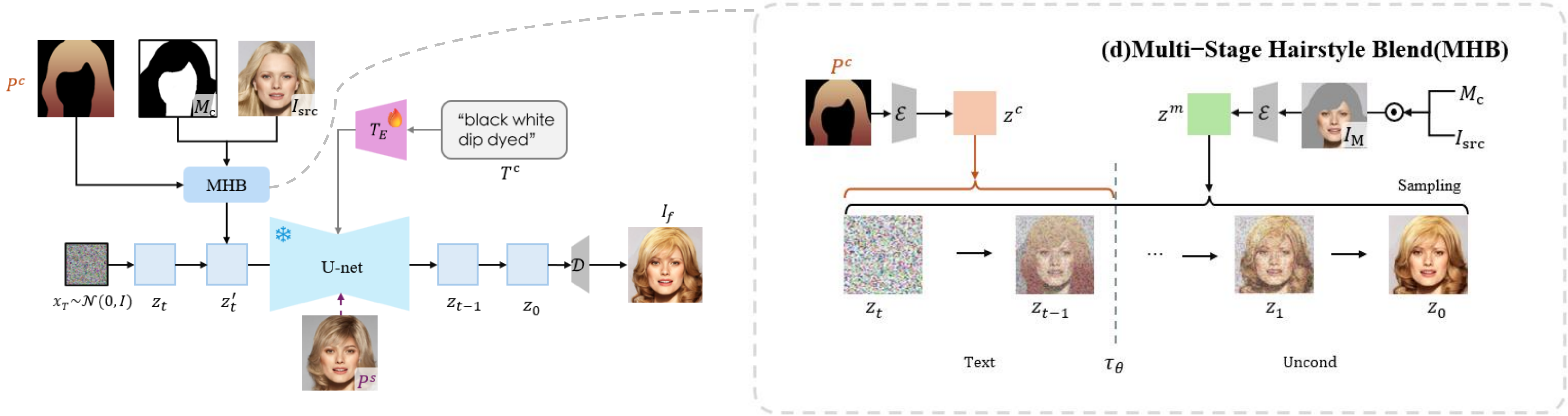
Hair Color Editing



(a) Agnostic Mask in Hair color editing Stage

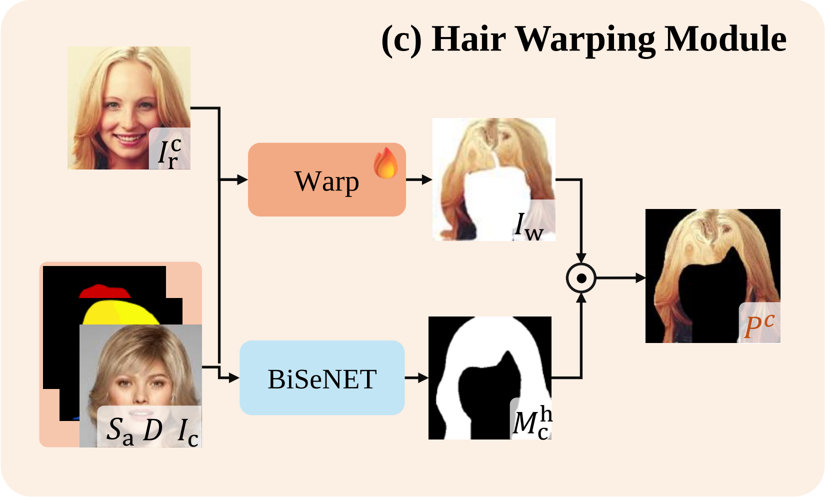
(b) StyleProxy control the hairstyle

Multi-stage Hairstyle Blend



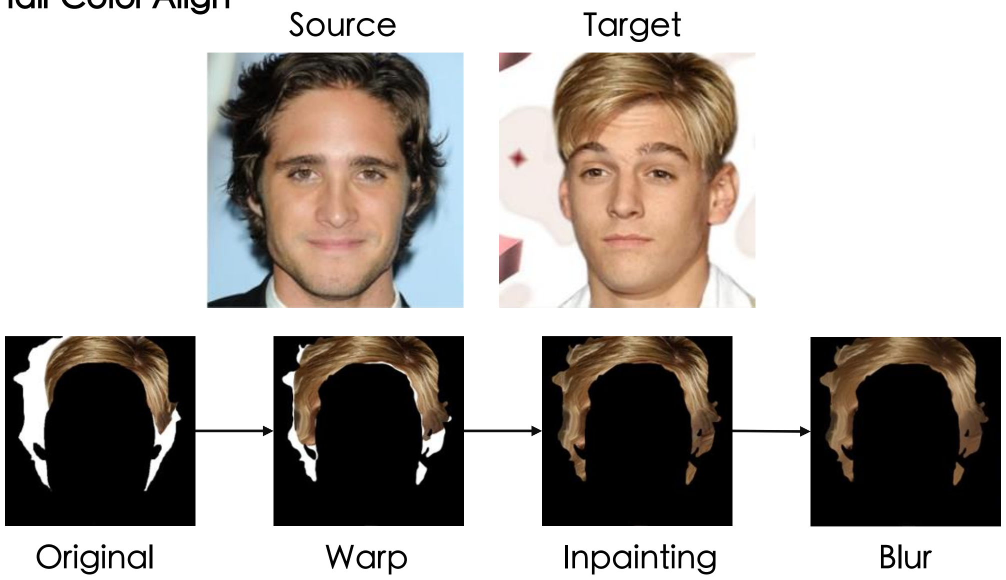
$$z_t' = B(z_t^{\mathrm{c}}, z_t^{\mathrm{m}}, t) = \begin{cases} z_t^{\mathrm{c}} \odot m_{\mathrm{c}} + z_t^{\mathrm{m}} \odot (1 - m_{\mathrm{c}}), & \text{if } t = \tau \\ z_t^{\mathrm{m}}, & \text{otherwise,} \end{cases}$$

Warping Module

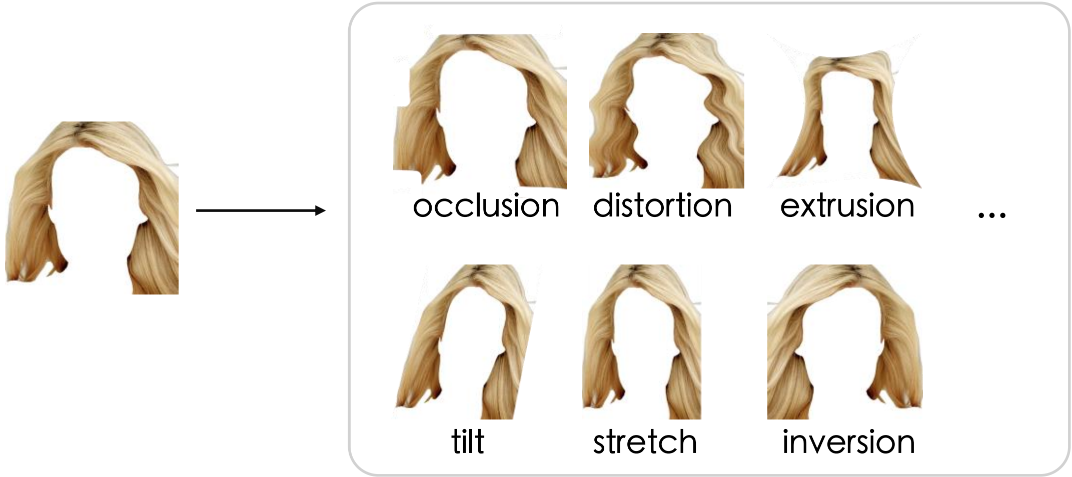


Inspired by VITON

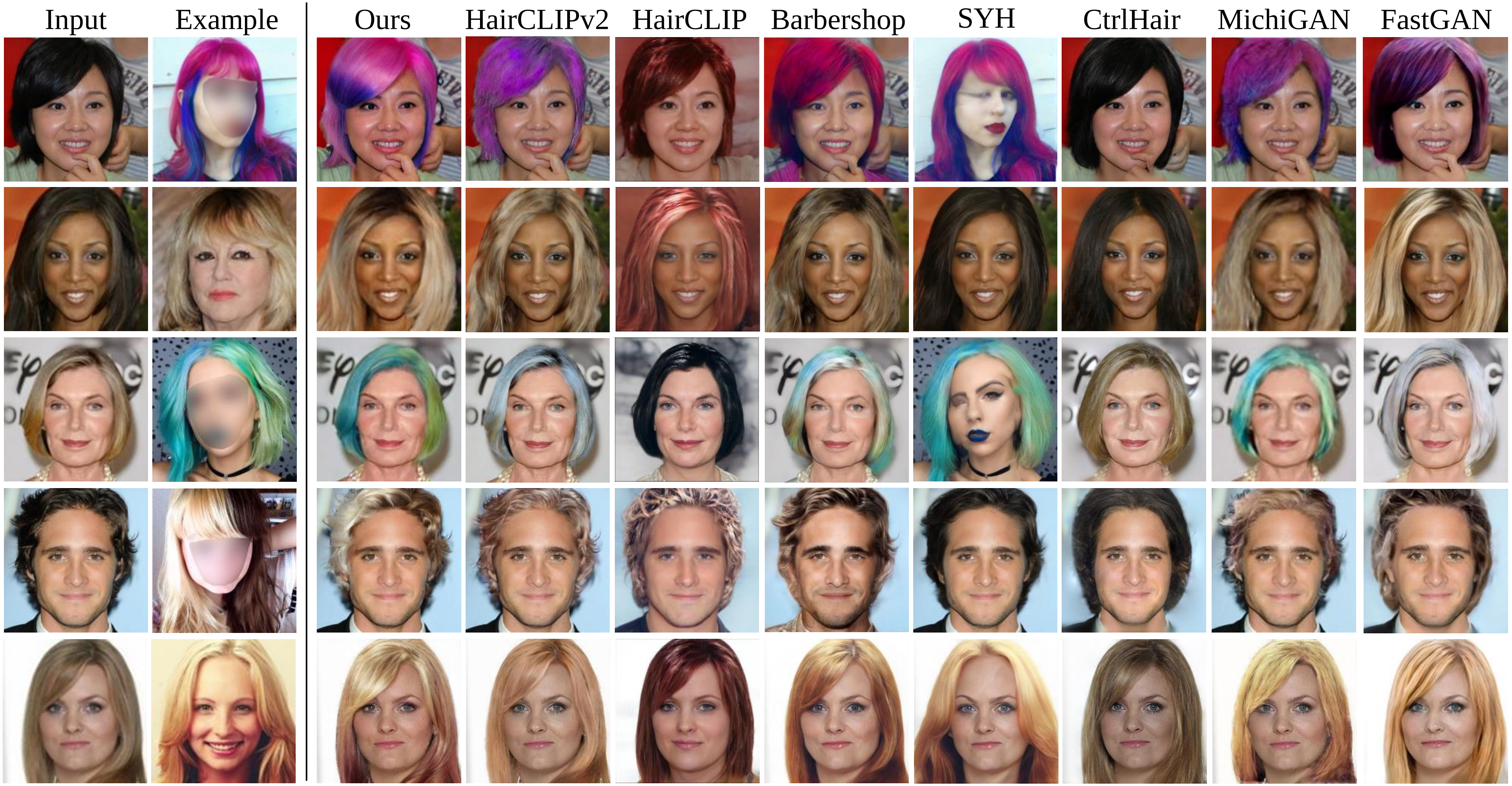
Hair Color Align



Data augmentation



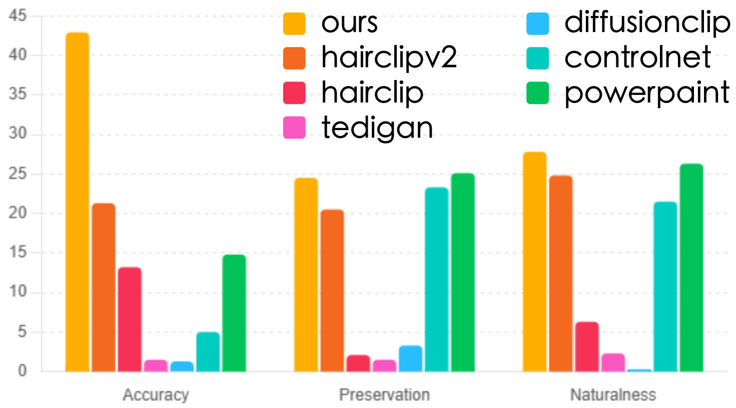
Reference Image-Based Hair Color Transfer



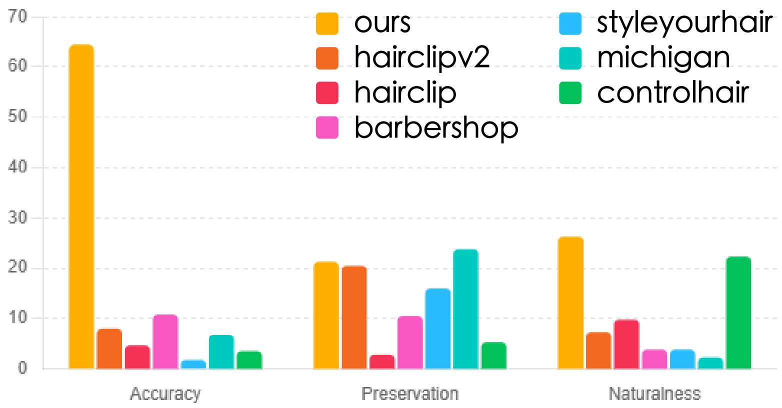
Text-based Hair Color Transfer

User Study

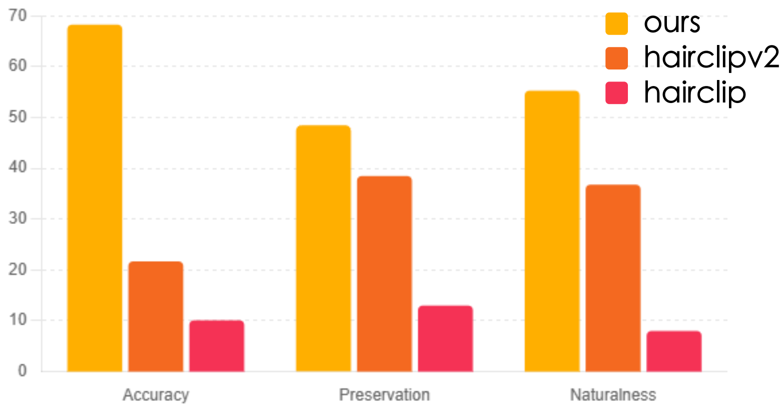
	Text-Driven							Color Transfer							Cross-Model		
Metrics	Ours	[34]	[33]	[36]	[17]	[40]	[44]	Ours	[34]	[33]	[42]	[18]	[31]	[10]	Ours	[34]	[33]
Accuracy	42.9	21.3	13.2	1.5	1.3	5.0	14.8	64.5	8.0	4.7	10.8	1.8	6.8	3.5	68.3	21.7	10.1
Preservation	24.5	20.5	2.1	1.5	3.3	23.3	25.1	21.3	20.5	2.8	10.5	16.0	23.8	5.3	48.5	38.5	13.0
Naturalness	27.8	24.8	6.3	2.3	0.3	21.5	26.3	26.3	7.3	9.8	3.8	28.3	2.3	22.3	55.3	36.8	8.0



(a)Text-Driven

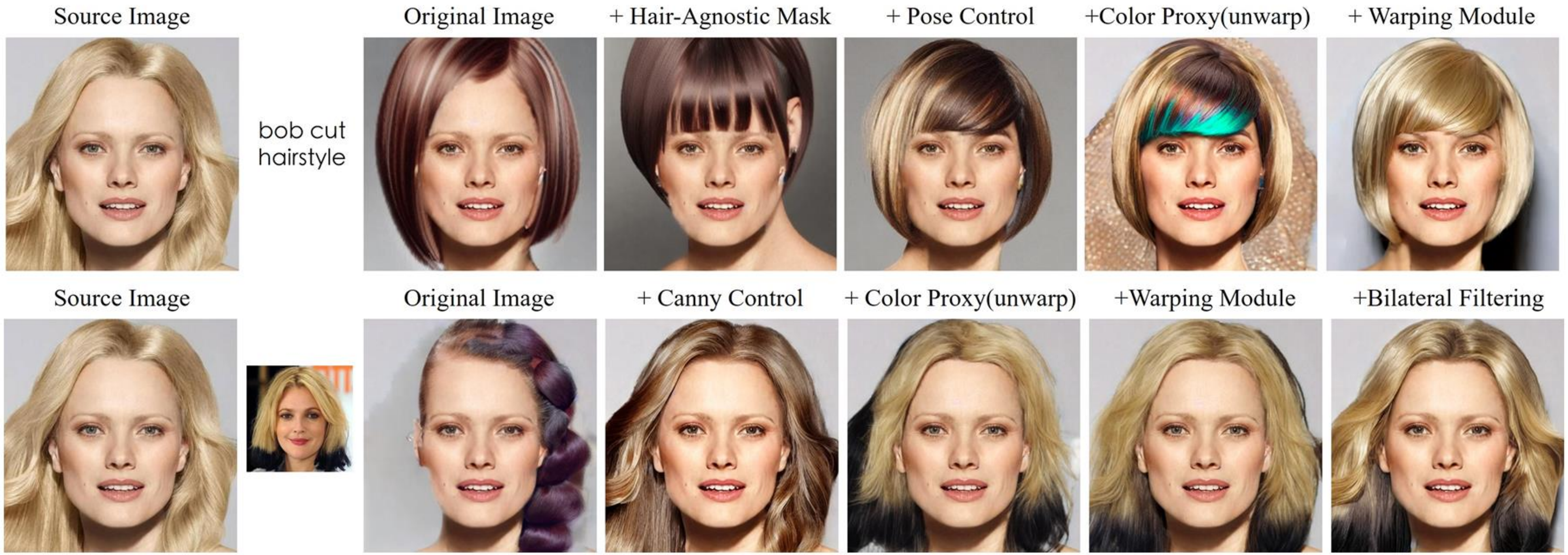


(b)Color Transfer

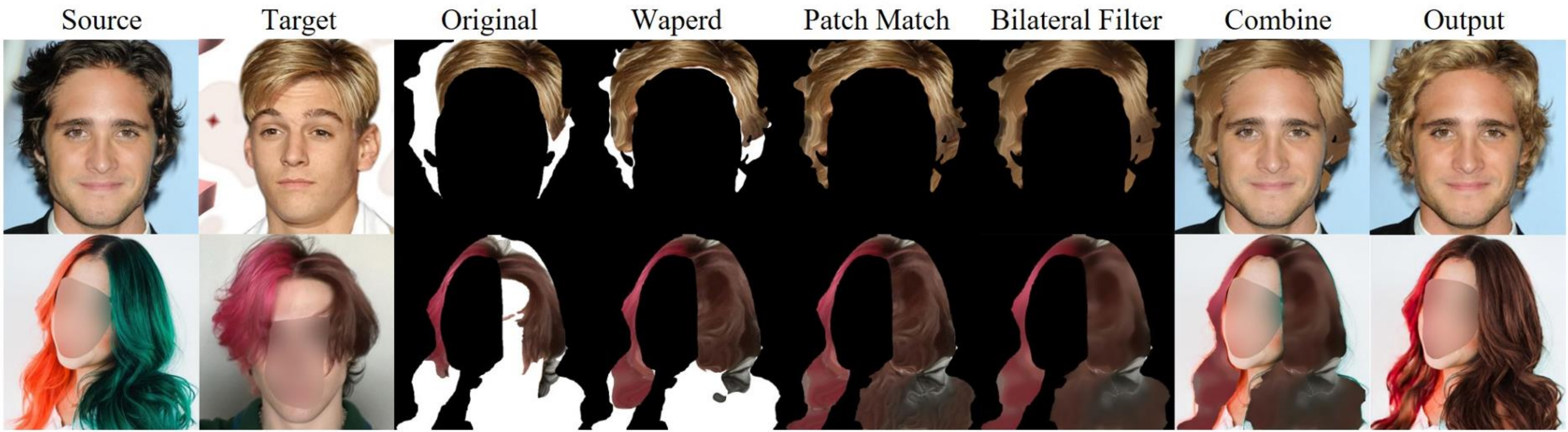


(c)Cross-Model

Ablation Study



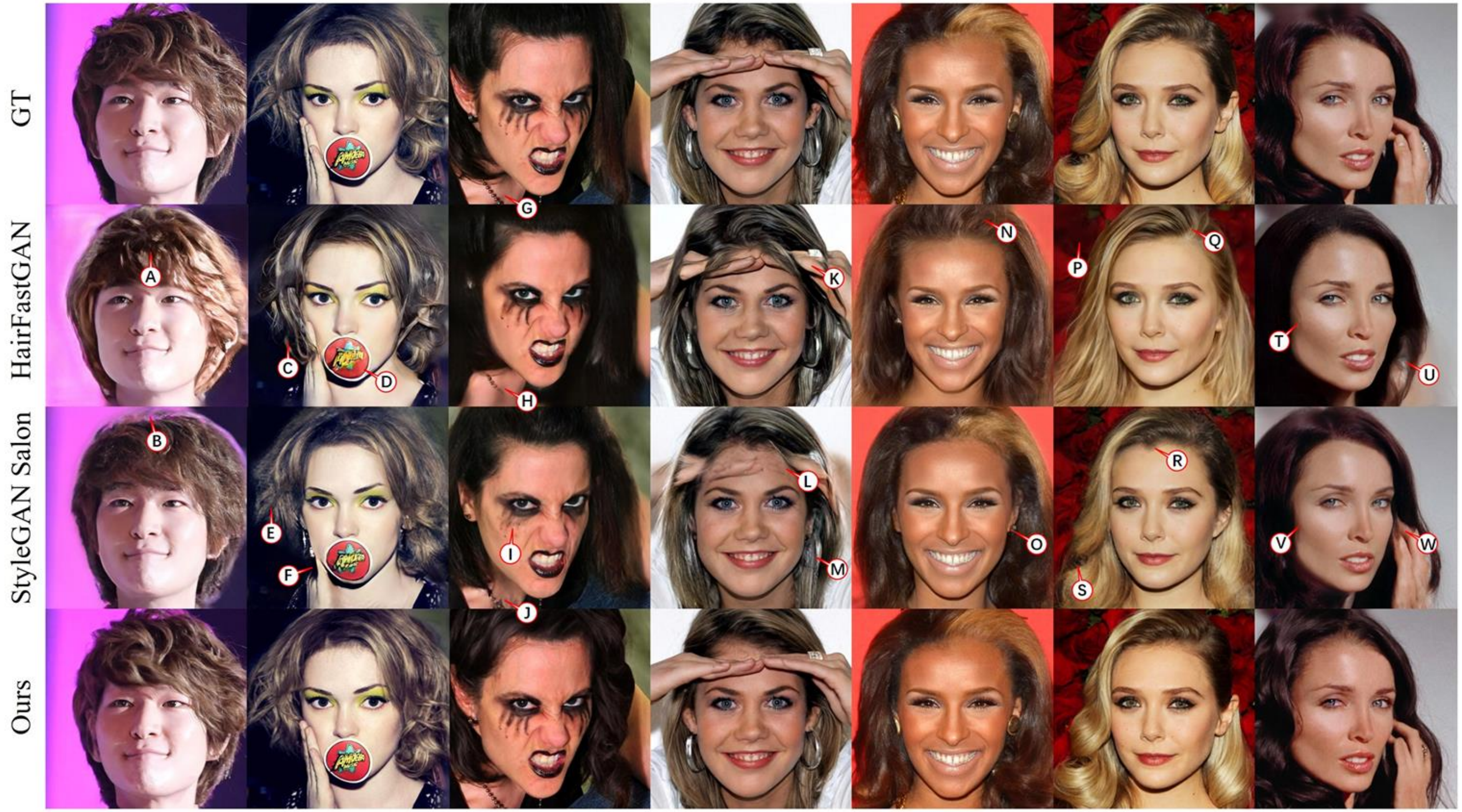
Ablation Study



Model	FID↓	FID_{CLIP}↓	SSIM↑
w/o Warping module.	33.17	12.53	0.62
w/o Bilateral filtering.	20.85	6.02	0.74
w/o PatchMatch.	27.74	8.51	0.70
HairDiffusion	**20.83**	**5.96**	**0.76**

Reconstruction

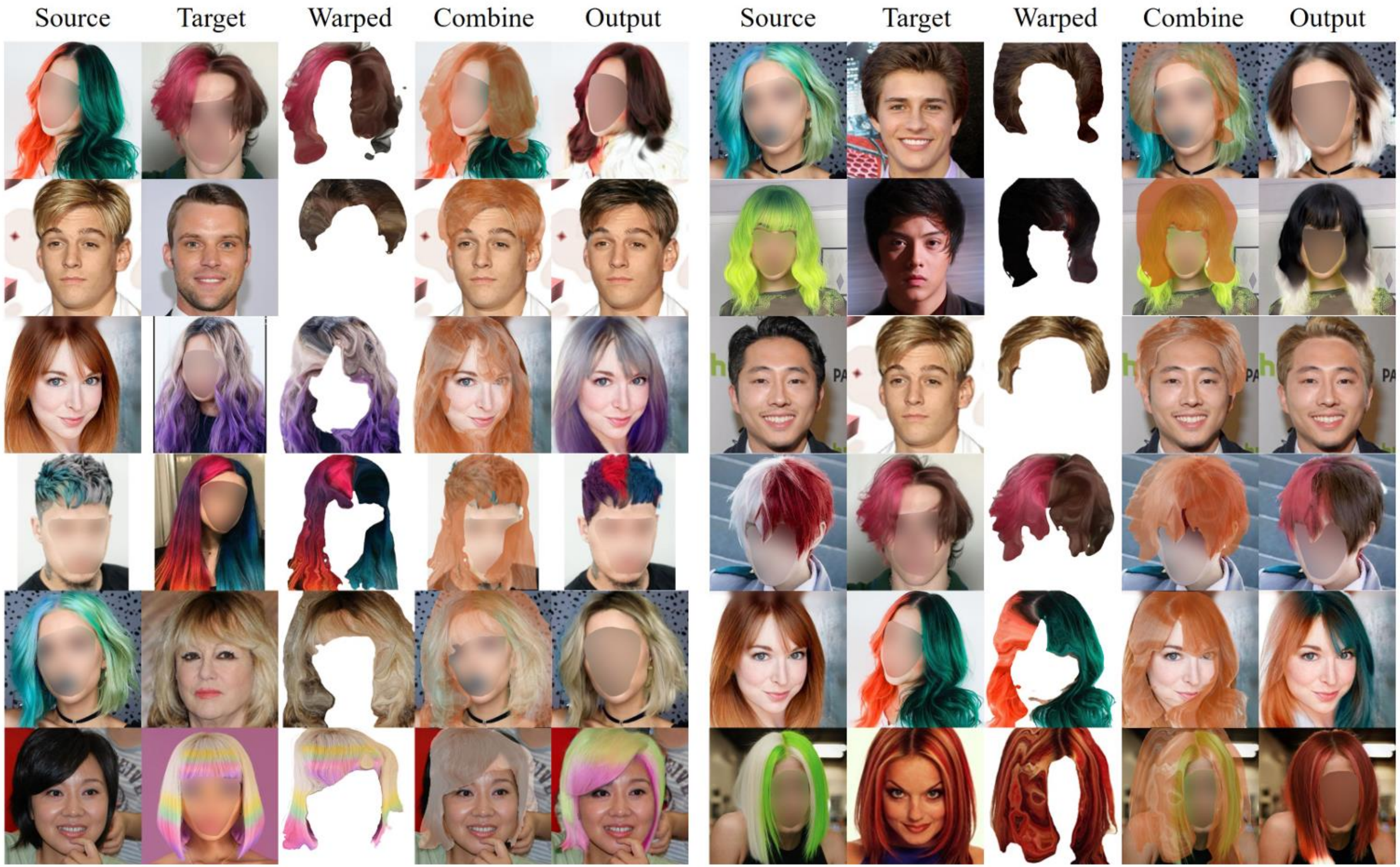
Model	Reconstruction			
	LPIPS↓	PSNR↑	FID↓	FID_{CLIP}↓
HairCLIP	0.36	14.08	35.49	10.48
HairCLIPv2	0.16	19.71	10.09	4.08
CtrlHair	0.15	19.96	<u>8.03</u>	1.25
StyleYourHair	0.14	21.74	10.69	2.73
Barbershop	0.11	21.18	13.37	2.61
HairFast	<u>0.08</u>	<u>23.45</u>	9.72	<u>0.97</u>
HairDiffusion	**0.07**	**31.66**	**5.41**	**0.68**



Preservation of facial attributes

Methods	IDS↑	PSNR↑	SSIM↑
Ours	**0.94**	**33.1**	**0.95**
HairCLIPv2 [34]	0.84	29.5	0.91
HairCLIP [33]	0.45	21.6	0.74
TediGAN [36]	0.16	22.5	0.74
DiffCLIP [17]	0.71	26.8	0.86

Limination


Source
Target
Warped
Combine
Output
Source
Target
Warped
Combine
Output

Limination

Model	Single Color transfer	
	FID↓	FID_{CLIP}↓
HairCLIP	40.08	10.94
HairCLIPv2	20.21	6.55
CtrlHair	**19.65**	3.62
StyleYourHair	-	-
Barbershop	20.54	3.89
HairFast	20.17	**3.00**
HairDiffusion	20.83	5.96

Conclusion

- We first propose the latent diffusion-based approach for hair editing.
- We introduce the MHB module and hair-agnostic masks, which enable the diffusion model to effectively control hairstyle and hair color independently while preserving unrelated attributes.
- We employ a warping module for the first time in this task to ensure alignment of hair color, demonstrating its capability in hair color manipulation and preservation.
- By collecting image-text pairs focused on hair color structure, we further enhance our model' s ability to finely control hair color using both text and reference images.