Lipstick ain't enough: Beyond Color Matching for In-the-Wild Makeup Transfer
— Supplementary material —

Thao Nguyen[1] Anh Tuan Tran[1] Minh Hoai[1,2]
[1]VinAI Research, Hanoi, Vietnam, [2]Stony Brook University, Stony Brook, NY 11790, USA
{v.thaontp79,v.anhtt152,v.hoainm}@vinai.io

Abstract

This document provides additional quantitative and qualitative results that could not be included in the main paper due to the page limit. In the first section, we give further details for the Pattern Transfer Branch, including building synthesis datasets and quantitative results of different backbones. In the second section, we provide additional qualitative results for each used dataset, together with multiple applications like partial makeup transfer and makeup interpolation. These examples demonstrate that our model can handle a variety of makeup styles, from simple to complicated, among variances of races and ages. In the last section, we show some difficult and failure cases.

1. Pattern Transfer Branch

As stated in the main paper, we train Pattern Branch on CPM-Synt-1 and compare results on both CPM-Synt-1 and CPM-Synt-2 dataset. In this section, we further describe the creation of these two datasets. We also report the results of different backbones.

1.1. Building CPM-Synt-1 & CPM-Synt-2 Dataset

Together with other datasets, a collection of patterns, called Sticker Dataset, will also be published upon the acceptance of the paper. From raw images crawled from Google Image Search, we discarded all images smaller than 64×64 or the ones without the alpha-channel. The final Sticker Dataset contains 577 high-quality PNG images. Some examples are shown in Figure 1.

Figure 1: Sticker Dataset.

While CPM-Synt-1 Dataset is straight forward as described in the main paper, CPM-Synt-2 Dataset might need more explanation. As the intention that CPM-Synt-2 is built for Pattern-Transfer Evaluation only, we need to minimize the difference between the source and reference in terms of color makeup styles. It is also fairer for other methods since they didn't distinguish between color and pattern-makeup.

Yet, when transferring the makeup style from reference to the source image, it is non-trivial to separate the color and pattern. Since the color style of the source image and reference image can be naturally different e.g., skin color, lip color. As a result, other methods can transfer these natural colors as makeup-color. To deal with that, we first push two non-makeup images into the same color-style. We use BeautyGAN [3] to generate the color-makeup images, using the same style-image. Then we add patterns to these color-makeup images and create the triplet (source, reference, ground-truth). Some examples are visualized in Figure 2.

1.2. Pattern Transfer Branch

For the Pattern Transfer Branch, besides UNet structure with Resnet-50 as the pre-trained encoder, we also conduct experiments with several different backbones. Quantitative results are shown in Table 1, conducted on CPM-Synt-1 test

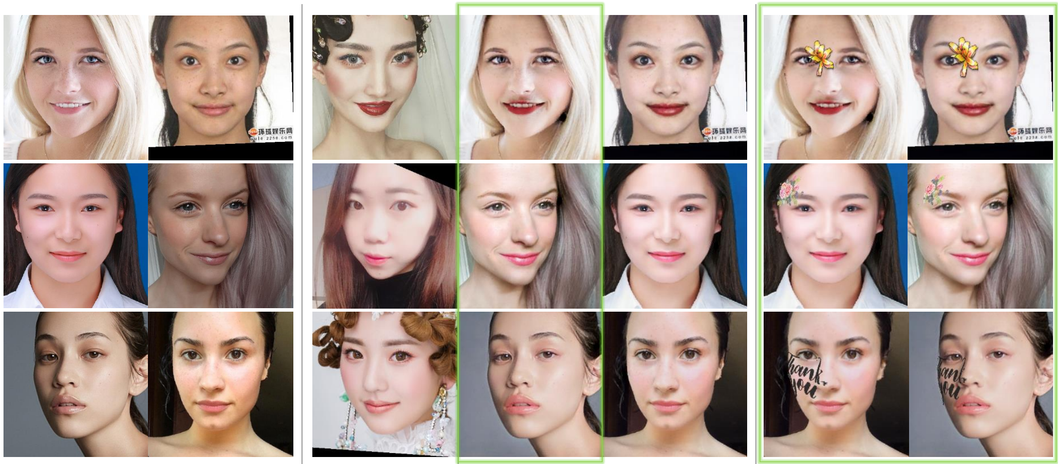

Figure 2: CPM-Synt-2 Dataset. 'GT' denotes ground-truth. Green bounding boxes indicates the triplet used to form CPM-Synt-2 Dataset.

set. The accuracy and mIOU vary among different backbones. Overall, Resnet-50-Unet provides the best segmentation results, particularly on mIOU, and we use it as our final model. Qualitative results are shown in Fig. 3.

Decoder	Encoder	acc	mIOU
UNet [5]	Resnet50	0.902	**0.788**
	Resnet101	**0.904**	0.783
	Vgg16	0.865	0.737
	Vgg19	0.870	0.732
FPN [4]	Resnet50	0.881	0.698
	Resnet101	0.874	0.702
	Vgg16	0.872	0.694
	Vgg19	0.857	0.692

Table 1: Comparison between different backbones used for the Pattern Segmentation model.

Thanks to the triplets from CPM-Synt-2, we can conduct quantitative results (MS-SSIM) to evaluate pattern-transfer among different methods. Along with the quantitative table shown in the main paper, we here provide more qualitative results on CPM-Synt-2 Dataset (Fig. 5).

As the color-style difference between the source and reference image is minimized, color-driven methods like BeautyGAN [3], DMT [6], and PSGAN [2] yield after-makeup images nearly the same as the source image. LADN can reproduce the pattern in the reference image, yet the results are imperfect in terms of textures, colors, and quality. Our method is the only one that can successfully transfer the pattern while keeping all patterns' characteristics.

2. Additional Qualitative Results

In this section, we provide additional qualitative results on MT-Dataset [3], CPM-Synt-2 and CPM-Real Dataset.

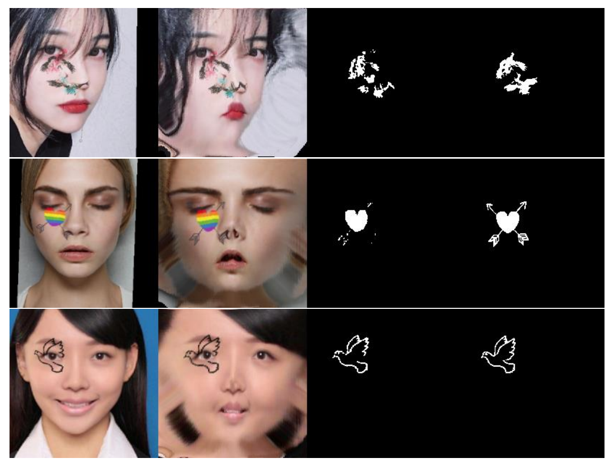

Figure 3: Makeup pattern segmentation. Images are from CPM-Synt-1 Dataset. From left to right: Original image, texture map, ground truth (GT), and prediction.

Experiments in Section 2.1 are intended to compare complete makeup transfer (both pattern and color transfer). Later, we illustrate extra examples on partial makeup transfer and style-interpolation.

2.1. Additional Qualitative Results on MT and CPM-Synt-1 Dataset

In this section, we provide additional qualitative results on MT Dataset [3], CPM-Synt-1, and CPM-Real in Fig. 6, 7, and 8 respectively. Same as in the main paper, we compare our results against DMT [6], BeautyGAN [3], LADN [1], and [2].

While the general color makeup style can be hard to compare among methods, pattern-based styles are easy to evaluate. Our method is the only one that can replicate the

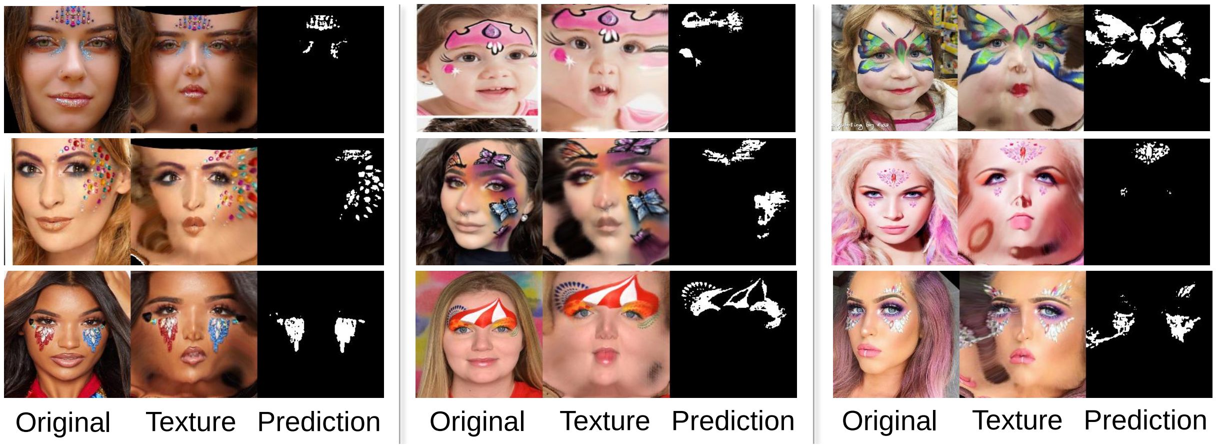

Figure 4: Makeup pattern segmentation. Images are from CPM-Real Dataset

pattern of the reference image.

2.2. Interpolation and Partial Makeup Transfer

We provide examples of mixed single-styles to quadruple-styles, in Fig. 10 and Fig. 11. The interpolated results are smooth and natural.

Partial makeup styles transfer is shown in Figure 12, 13. Besides the partial facial region makeup (lip, eye-shadow, skin), we add the pattern-only makeup. As stated in the main paper, Pattern and Color Branch can be used separately. Thus, it empowers the combination of makeup styles.

3. Limitation

Although showing promising results on general cases, our Pattern Segmentation still suffers to detecting all the pattern's regions on some difficult cases. In Fig. 4, we show some pattern segmentation prediction on complicated real-life makeup styles. Some facial jewelers are too tiny, while some face paintings are hard to be covered entirely.

In Fig. 9, we demonstrate some under-performing after-makeup results of difficult cases. In some cases, the facial pieces of jewelry are too tiny (Row 1–3). In another case, Pattern Segmentation can not detect the whole face-painting (Row 4–7). The reference pose can also lead to wrong transferred styles. For example, closed eye-lid in reference leads to a wrong makeup position (Row 8).

References

[1] Qiao Gu, Guanzhi Wang, Mang Tik Chiu, Yu-Wing Tai, and Chi-Keung Tang. Ladn: Local adversarial disentangling network for facial makeup and de-makeup. In *ICCV*, 2019. 4322, 4324, 4325, 4326, 4327, 4328

[2] Wentao Jiang, Si Liu, Chen Gao, Jie Cao, Ran He, Jiashi Feng, and Shuicheng Yan. Psgan: Pose and expression robust spatial-aware gan for customizable makeup transfer. In *CVPR*, 2019. 4322, 4324, 4325, 4326, 4327, 4328

[3] Tingting Li, Ruihe Qian, Chao Dong, Si Liu, Qiong Yan, Wenwu Zhu, and Liang Lin. Beautygan: Instance-level facial makeup transfer with deep generative adversarial network. In *Proceedings of the 26th ACM international conference on Multimedia*, 2018. 4321, 4322, 4324, 4325, 4326, 4327, 4328

[4] T. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie. In *CVPR*. 4322

[5] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *MICCAI*, 2015. 4322

[6] Honglun Zhang, Wenqing Chen, Hao He, and Yaohui Jin. Disentangled makeup transfer with generative adversarial network. 2019. 4322, 4324, 4325, 4326, 4327, 4328

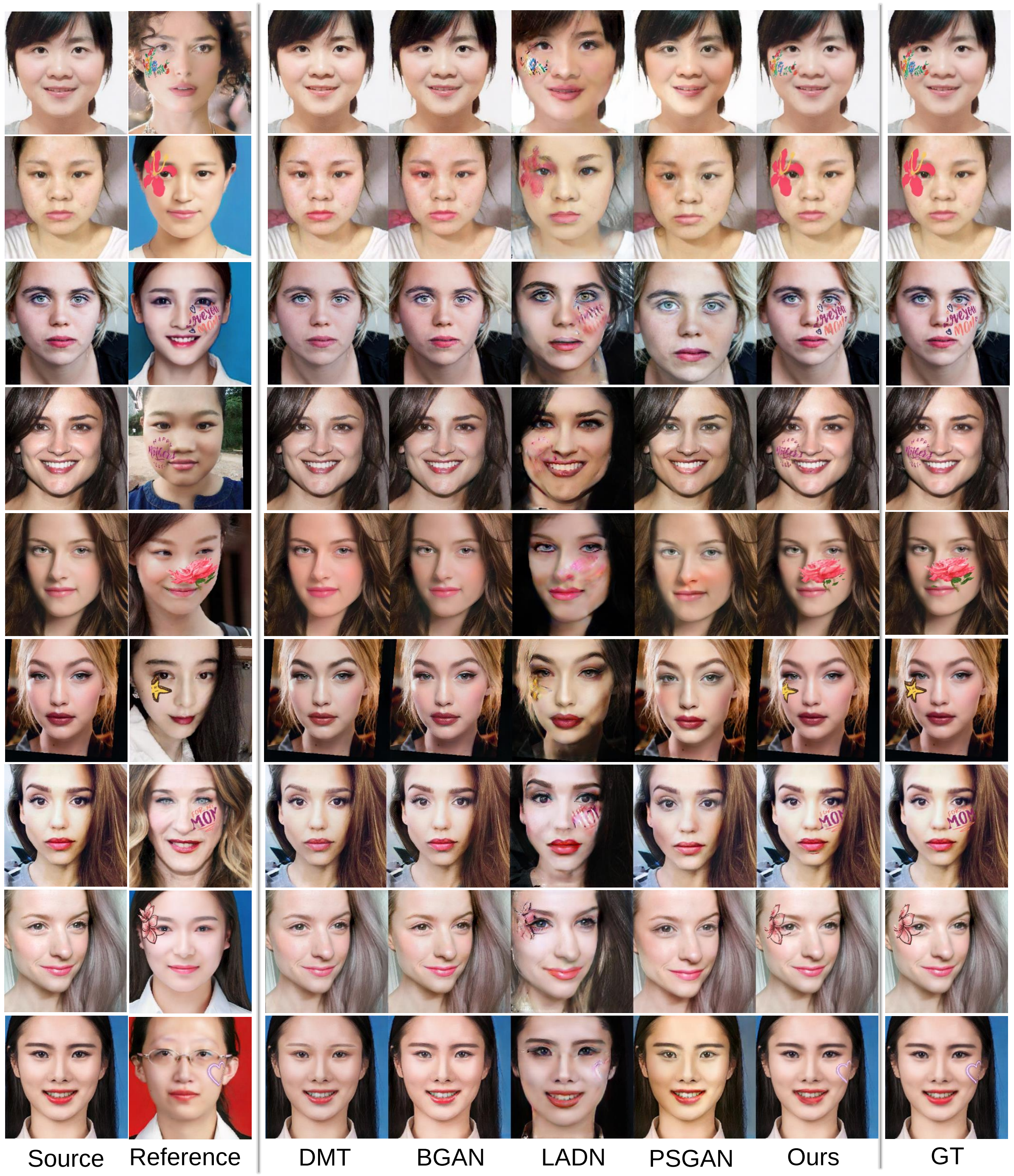

Figure 5: Qualitative results on CPM-Synt-2. From left to right: source image, ground Truth (GT), DMT [6], BeautyGAN [3], LADN [1], PSGAN [2], ours, and reference image.

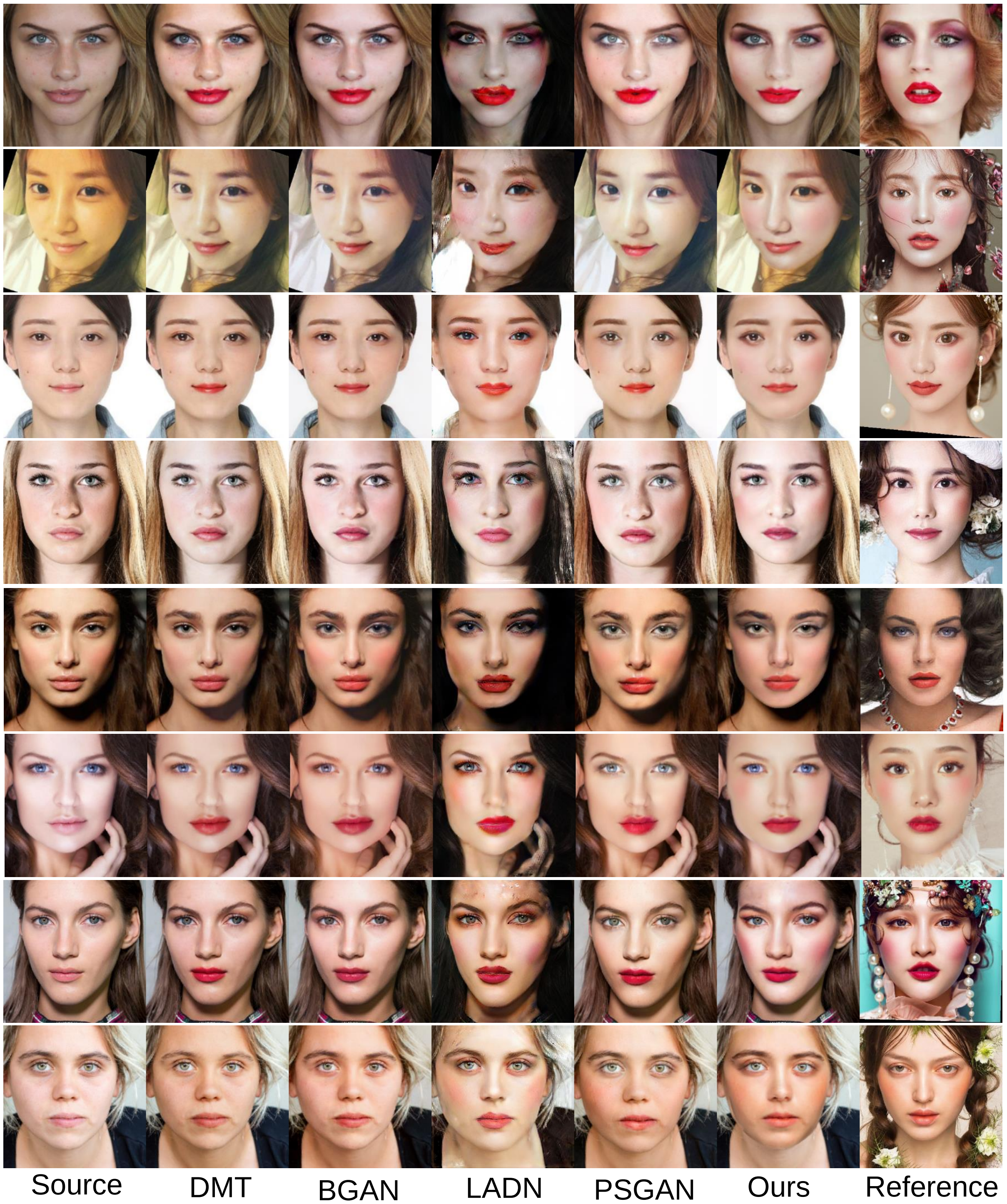

Figure 6: Additional qualitative results on MT [3]. From left to right: source image, DMT [6], BeautyGAN [3], LADN [1], PSGAN [2], ours and reference image.

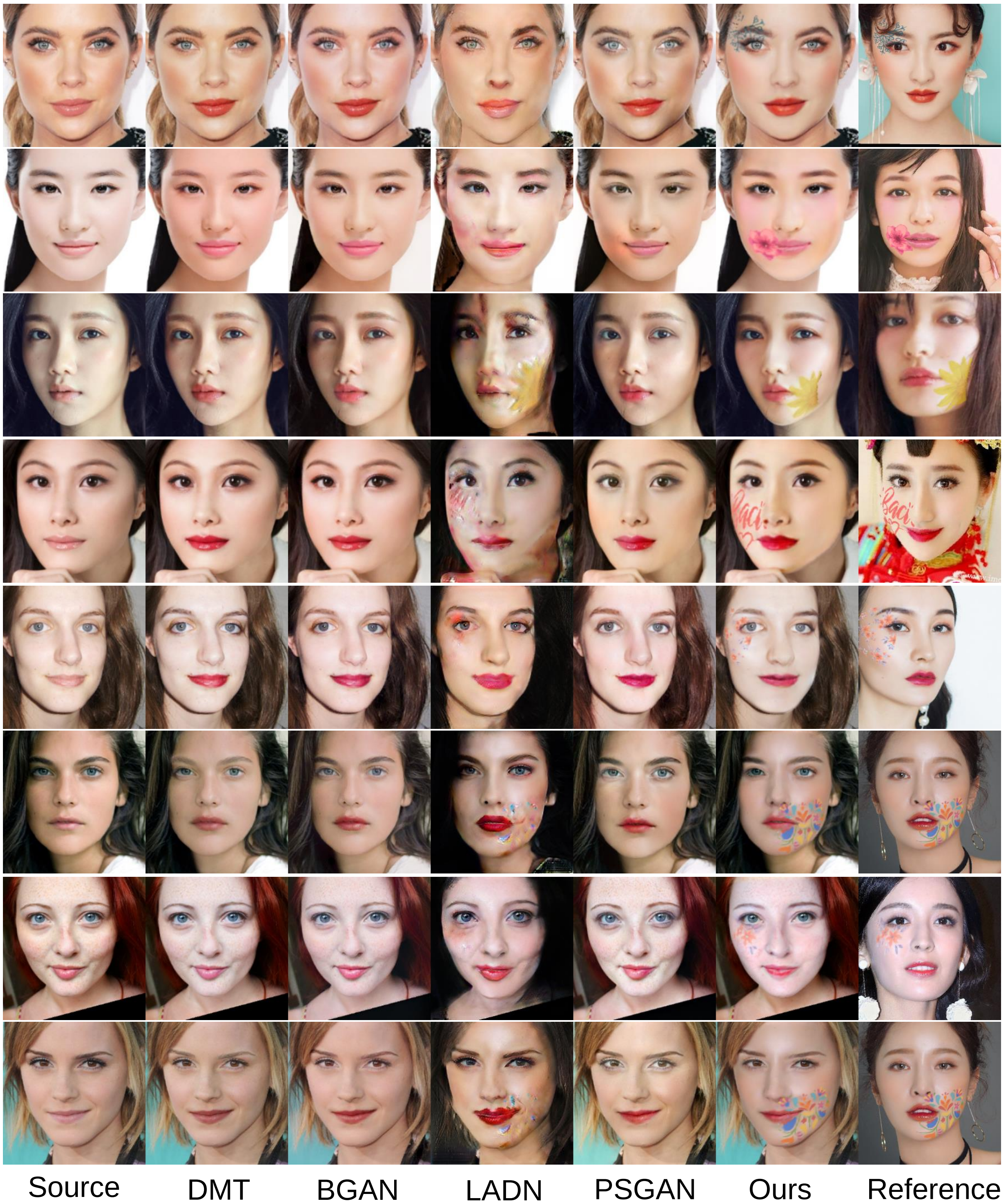

Figure 7: Additional qualitative results on CPM-Synt-1 Dataset. From left to right: Source image, DMT [6], BeautyGAN [3], LADN [1], PSGAN [2], Ours and Reference image.

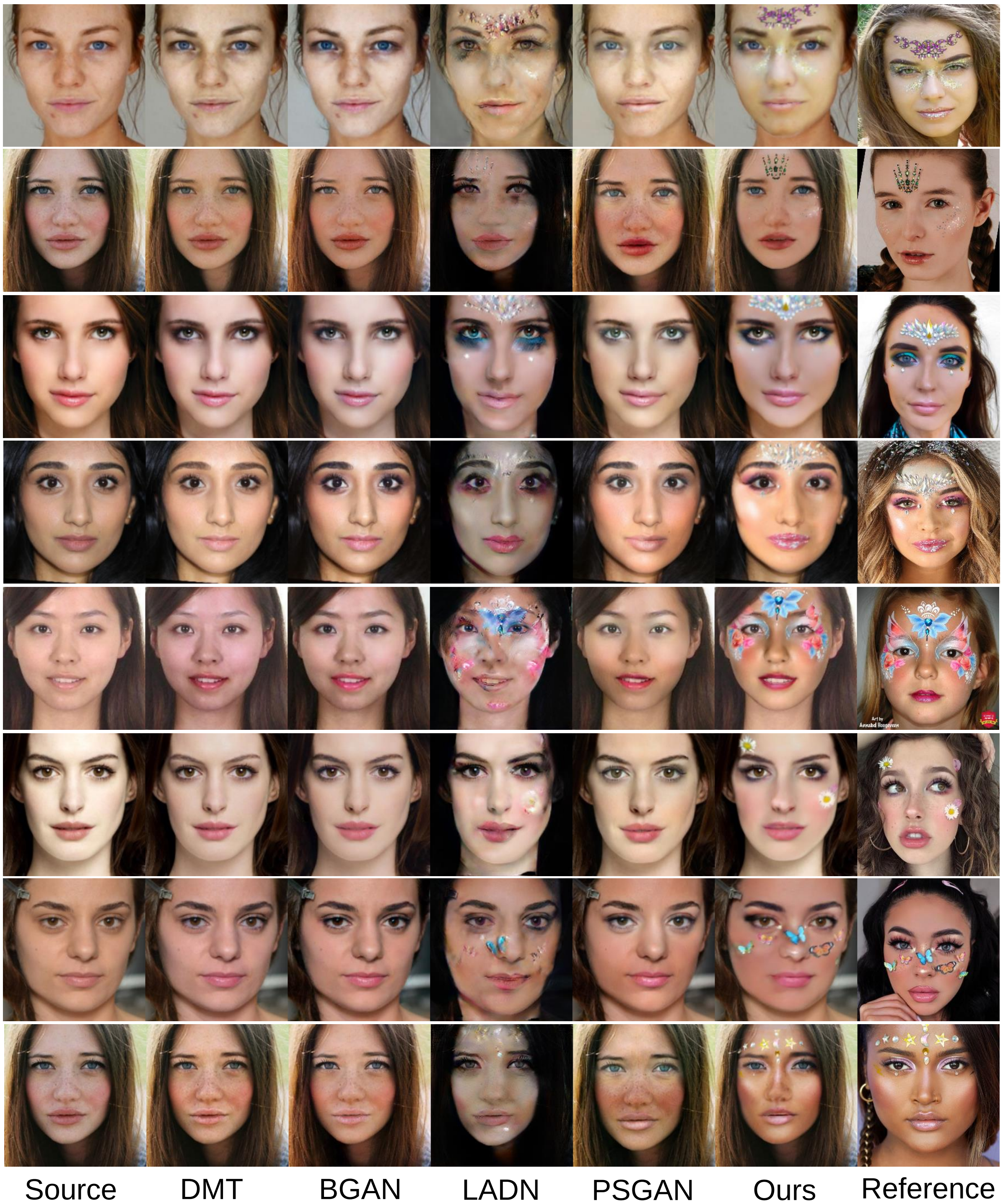

Figure 8: Additional qualitative results on CPM-Real. From left to right: Source image, DMT [6], BeautyGAN [3], LADN [1], PSGAN [2], Ours and Reference image.

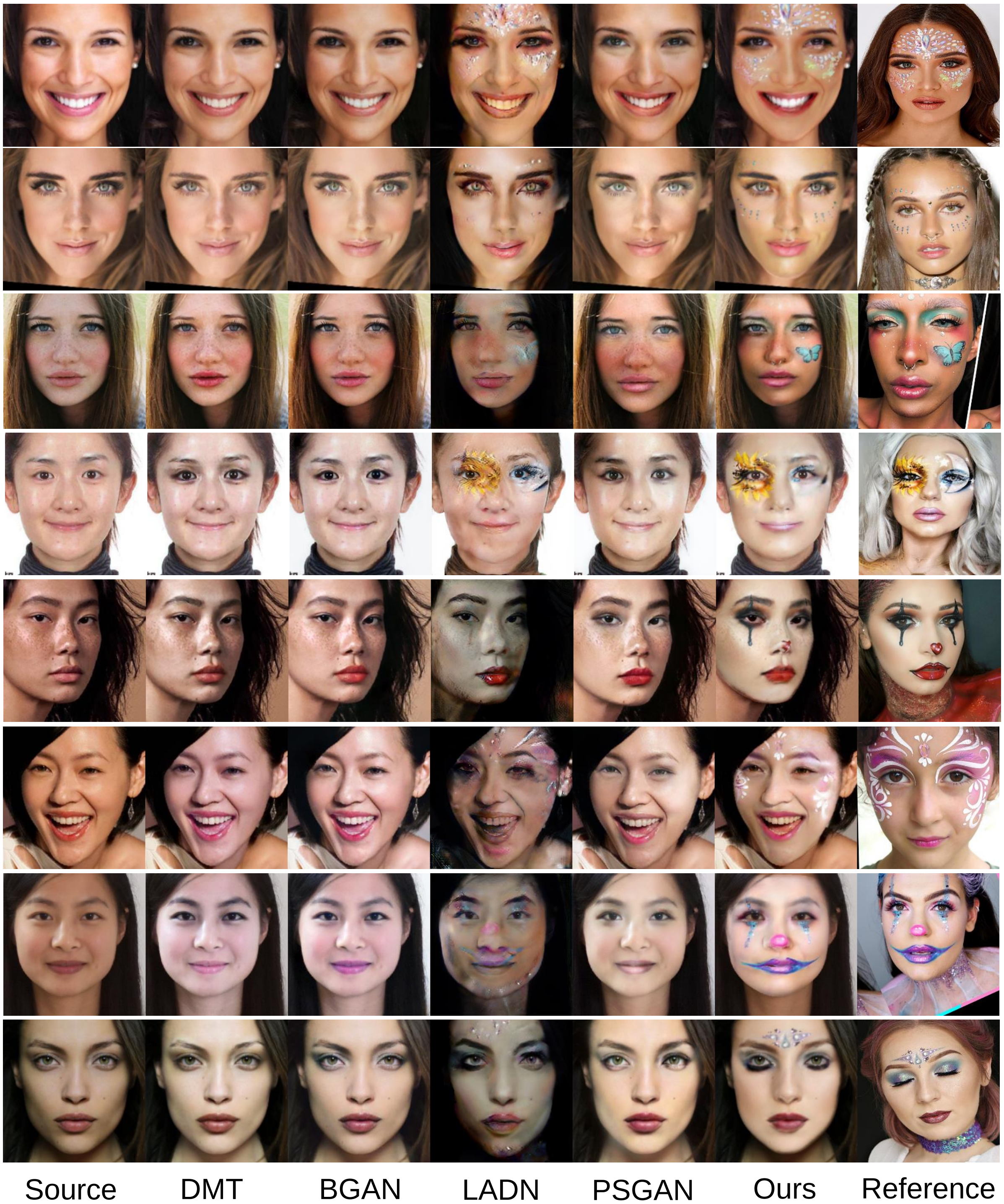

Figure 9: Some difficult cases. From left to right: Source image, DMT [6], BeautyGAN [3], LADN [1], PSGAN [2], Ours and Reference image.

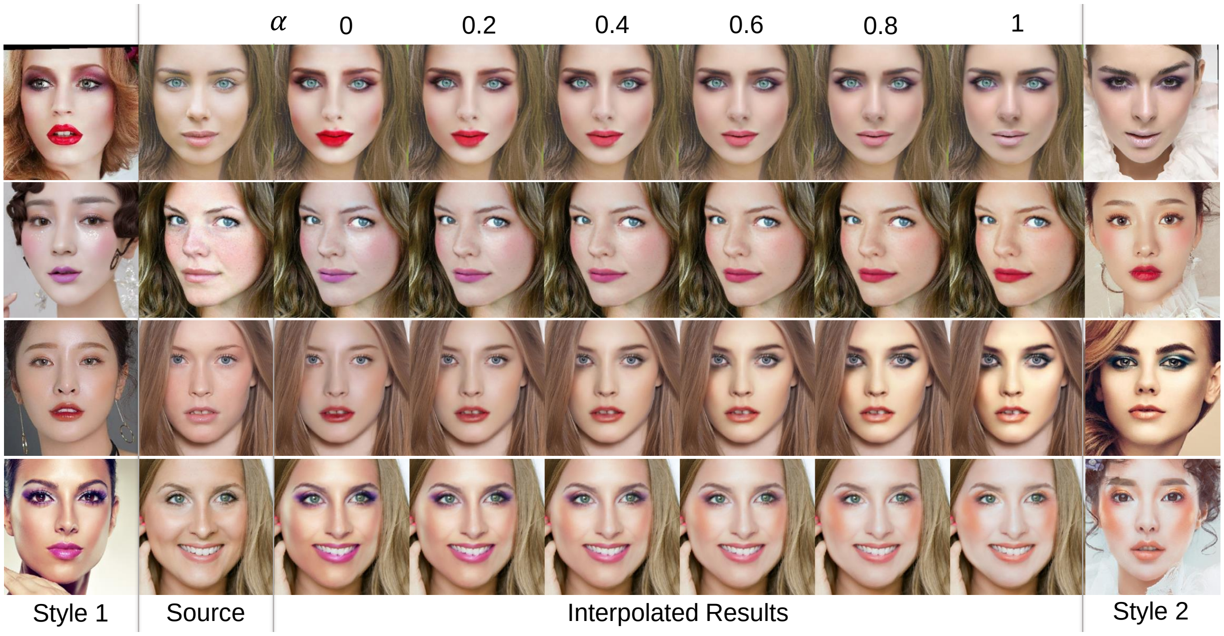

Figure 10: Additional qualitative results for interpolation

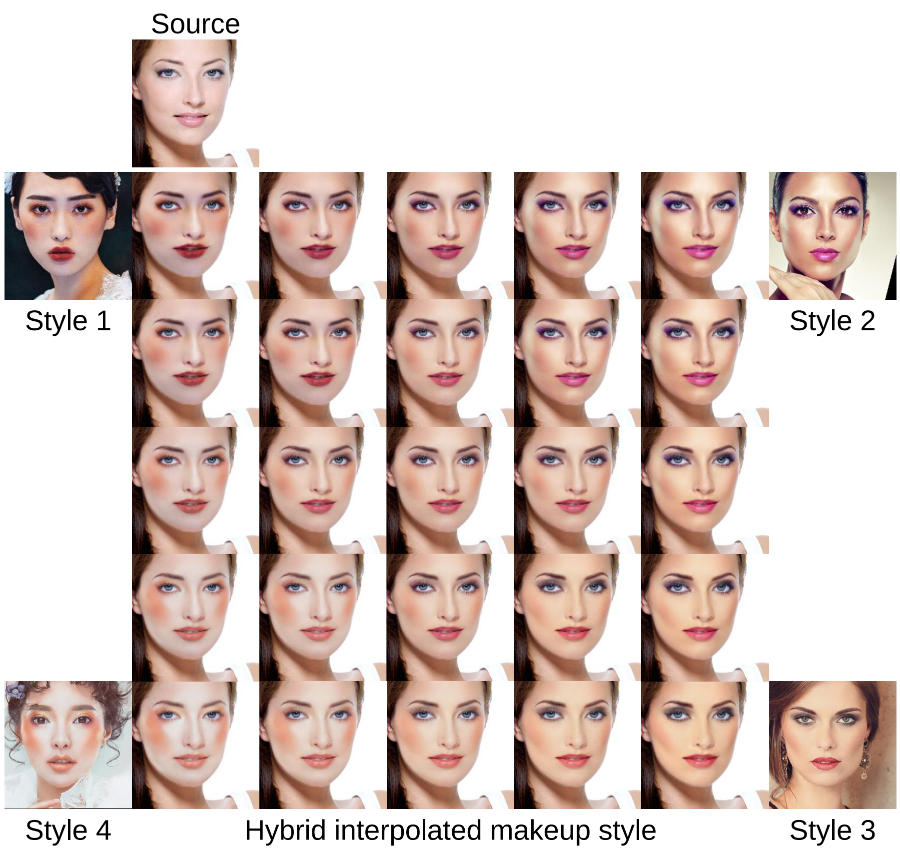

Figure 11: Mixed multiple makeup styles

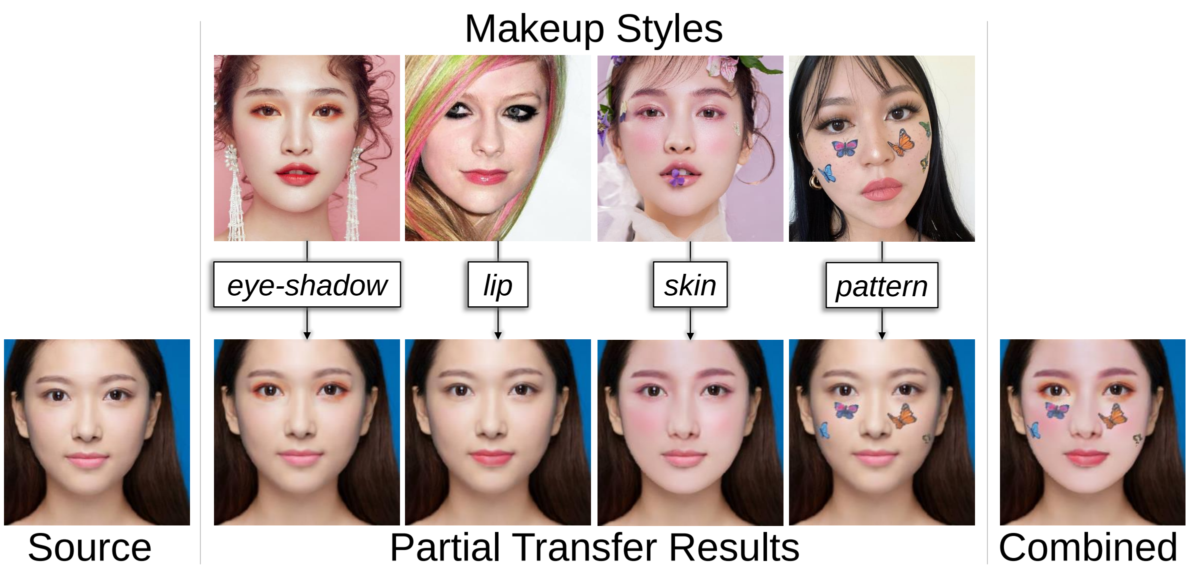

Figure 12: Additional qualitative results for Partial Makeup Transfer

Figure 13: Additional qualitative results for Partial Makeup Transfer