

Supplementary Material

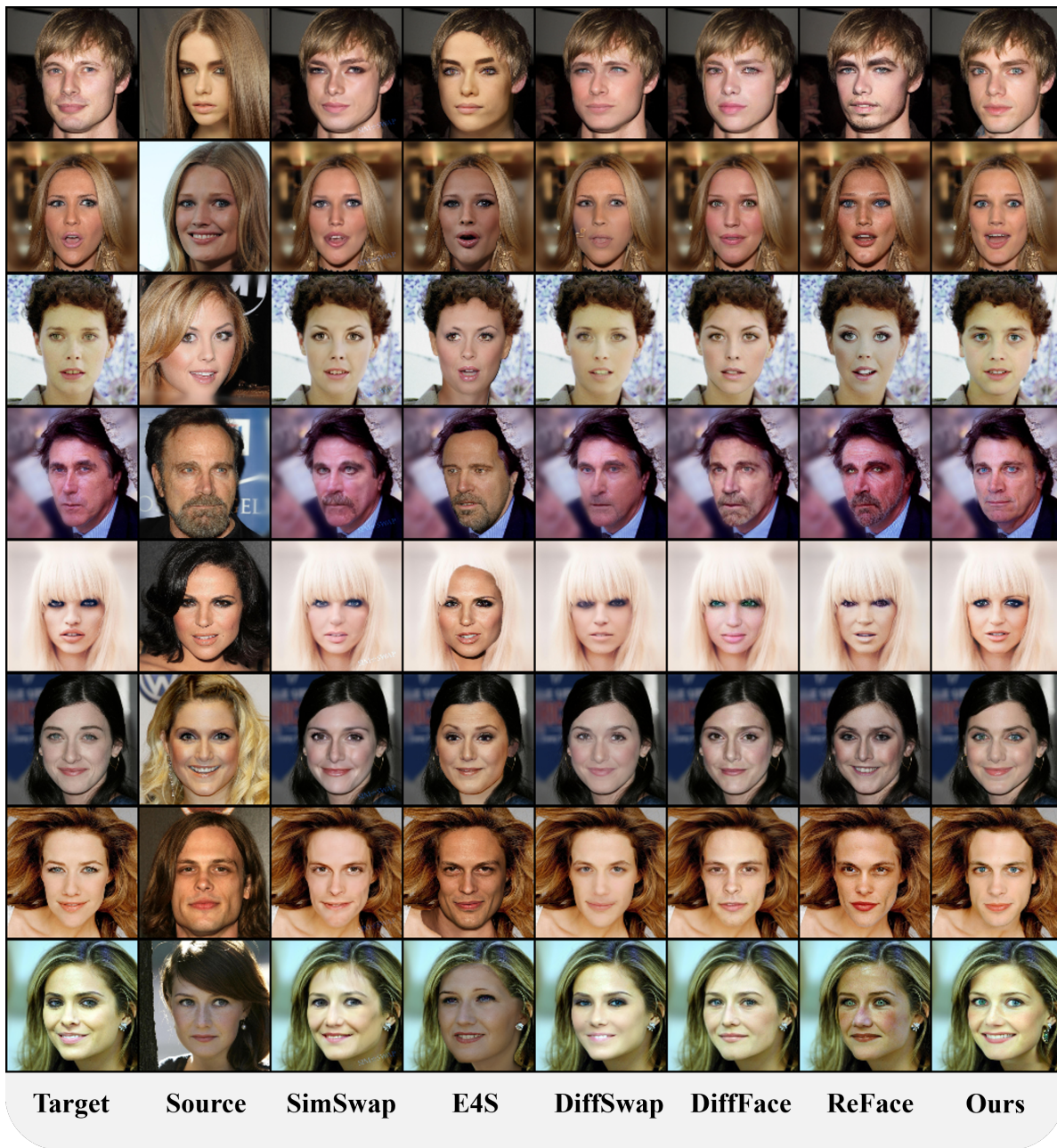


Figure 1. CelebA-HQ results with various approaches.

1. Additional results on CelebA-HQ

To further demonstrate our model, we evaluate it in the CelebA-HQ dataset and post the samples in Figure 1 without cherry pick. Our results well balance the image quality, identity preservation and attribute consistency, as discussed in the main body.

2. Training details

All the experiments are conducted on NVIDIA A100. We use the AdamW optimizer to perform gradient descent. In stage-1 and stage-2 training, we use a fixed learning rate of 10^{-5} , effective batch-size of 64 and train 200000 steps for each. In stage-3 training (refinement post-tuning), we use the same hyperparameters except for a reduced learning rate of 10^{-6} and training 50000 steps.