Supplementary Material for 3D-Aware Face Swapping

Yixuan Li Chao Ma* Yichao Yan* Wenhan Zhu Xiaokang Yang
MoE Key Lab of Artificial Intelligence, AI Institute, Shanghai Jiao Tong University, China
{lyx0208, chaoma, yanyichao, zhuwenhan823, xkyang}@sjtu.edu.cn

This supplementary material contains additional details of the main manuscript. Sec. A provides extra experimental results of the proposed 3dSwap while Sec. B further discusses the role our mode plays in preventing negative consequences caused by DeepFake techniques.

A. Additional Experimental Results

In this section, supplementary experimental results for the main manuscript are provided. In Sec. A.1, we exhibit more visualized face swapping results on celebA-HQ [2] to show the generalization of the proposed 3dSwap under several challenging settings. In Sec. A.2, we extend Sec. 4.4 of the main manuscript by adding more visualization. In Sec. A.3, we provide extra multi-view reconstruction results synthesized by our 3D GAN inversion module from one-shot input images as a supplement of the ablation studies.

A.1. More Qualitative Results

As shown in Fig. 2, we provide more qualitative results of 3dSwap on CelebA-HQ. Specifically, face swapping is performed on several challenging settings, such as different genders (Fig. 2(a)), large age gaps (Fig. 2(b)), different skin colors (Fig. 2(c)), distinguishing lighting conditions (Fig. 2(d)) and large pose variations (Fig. 2(e)). High-fidelity swapped faces under all these challenging cases validate the effectiveness and robustness of the proposed method.

A.2. Analysis on 3D-Aware Face Swapping

As indicated in Sec. 4.4 of the main manuscript, we employ nine fixed poses to measure the identity consistency of 3D-aware face swapping. In Fig. 1, the nine poses are visualized. Moreover, we show several multi-view face swapping results in Fig. 3, where the swapped faces are synthesized in target views (*i.e.*, in the viewing direction of target images), source views, and three fixed poses (*i.e.*, left, frontal and right). *Please refer to the attached video for better visualization.*

* Corresponding authors.

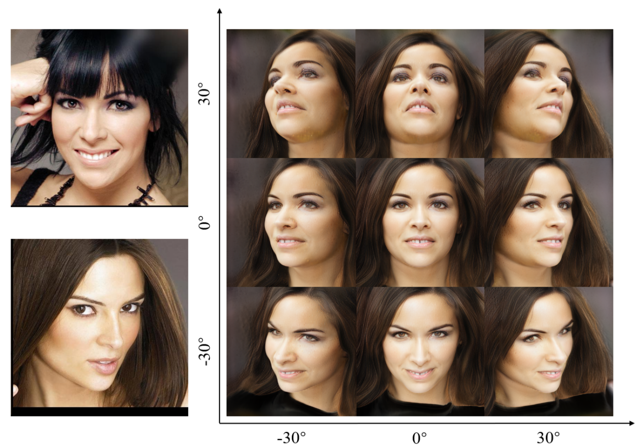


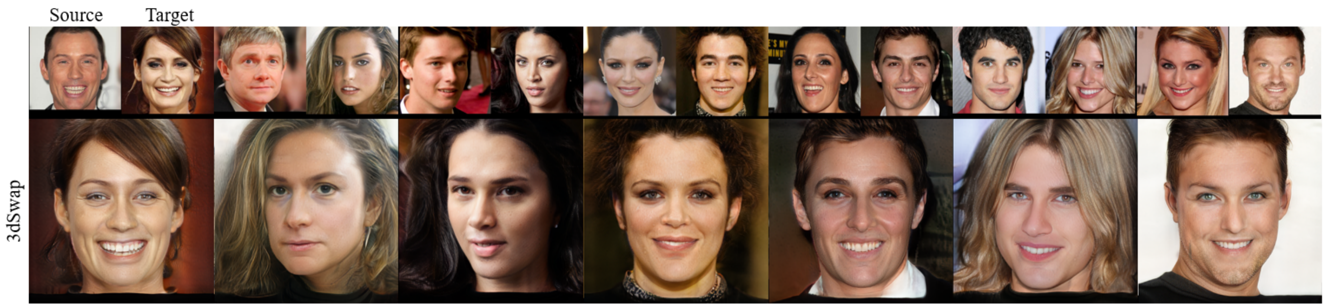
Figure 1. Nine different fixed poses were used to calculate the "Average Identity Similarity" in Section 4.4 of the main manuscript.

A.3. More Results on 3D Face Reconstruction

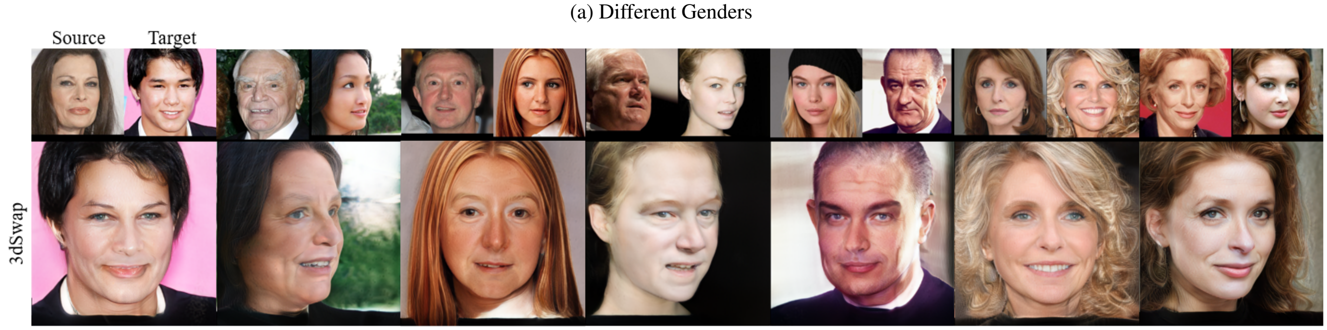
We provide extra results of our 3D GAN inversion in Fig. 4 and follow the setting in ablation studies that the pre-trained generator [1] is tuned for 500 steps to have a better visualization. As observed from Fig. 4, our 3D GAN inversion module can perform accurate reconstruction despite the variation of gender, age, pose, make-up, *etc*.

B. Broader Impact

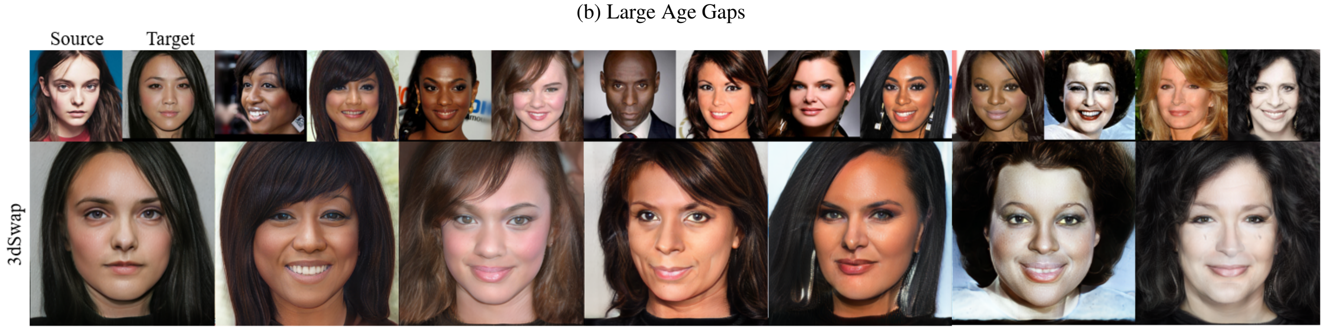
As mentioned in the main manuscript, face swapping techniques may have negative effects on the society. In order to build an effective face forgery detection model, it is necessary to first understand the working principle of Deepfake techniques. The first 3D-aware face swapping approach presented in our research directly synthesize the swapped face by leveraging a pre-trained generator without forging the original target images, and might challenge current forgery detection approaches. Thus the fake faces generated by our method can be utilized to help improving the data-driven forgery detectors. Moreover, a brand new dataset can be built on our multi-view results for any future research on 3D forgery detection.



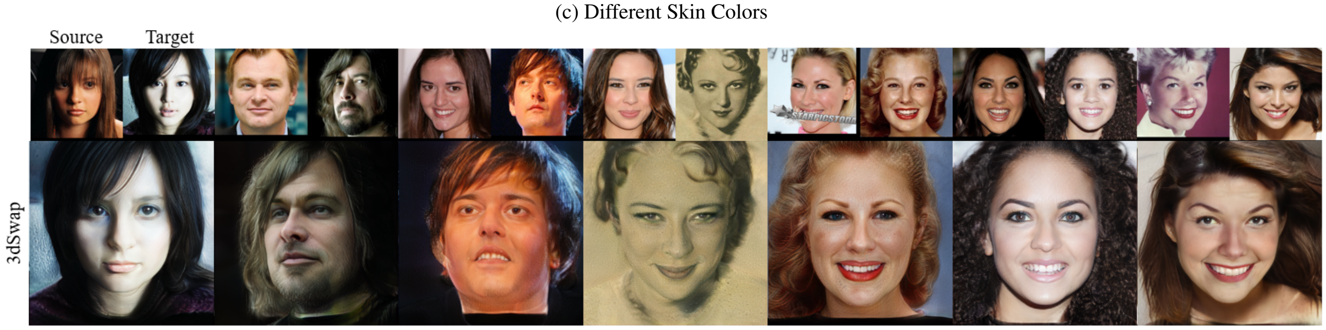
(a) Different Genders



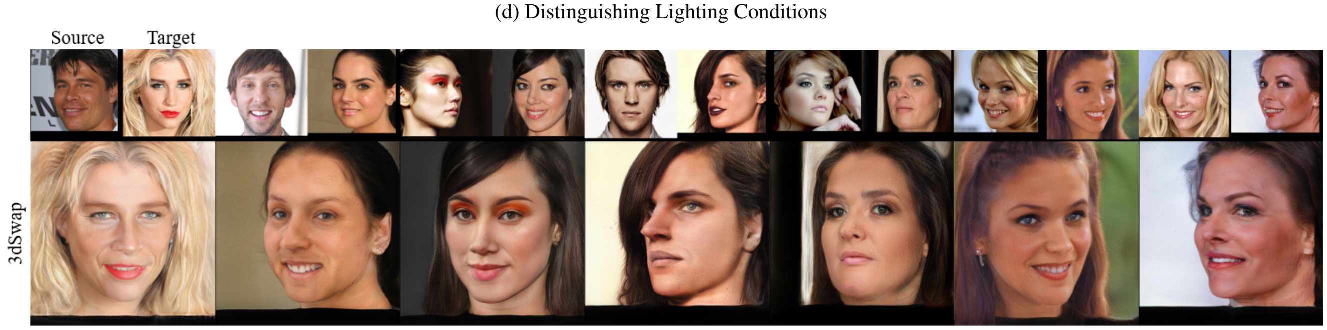
(b) Large Age Gaps



(c) Different Skin Colors



(d) Distinguishing Lighting Conditions



(e) Large Pose Variations

Figure 2. **More qualitative results on CelebA-HQ.** All the results are under a certain challenging setting: (a) different genders, (b) large age gaps, (c) different skin colors, (d) distinguishing lighting conditions, (e) large pose variations.

Figure 3. **Multi-view face swapping results on CelebA-HQ.** The swapped faces are in the viewing direction of target images, source images, and three fixed poses (*i.e.*, left, frontal and right). With the proposed 3dSwap, we can not only generate multi-view consistent swapped faces, but the identities of source images are also recognizable from any viewing direction.

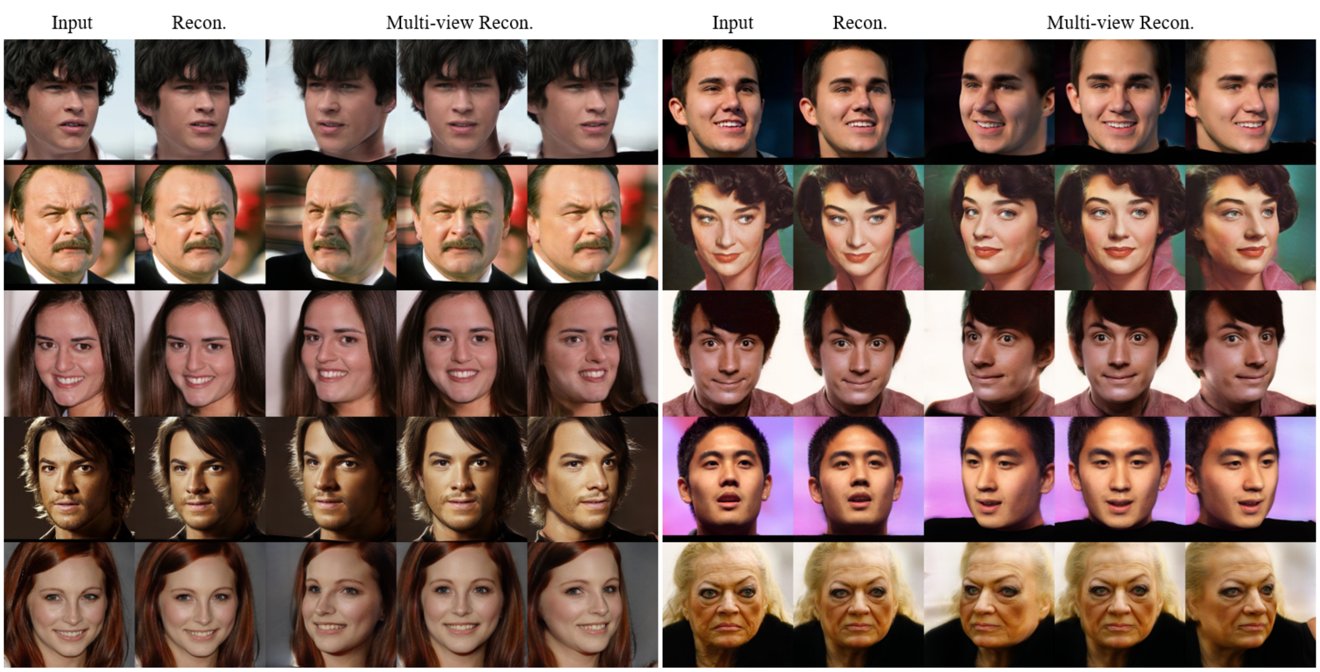


Figure 4. **More 3D GAN inversion results on CelebA-HQ.** With 3D GAN inversion, we are able to perform an accurate reconstruction of arbitrary input images. We also show the reconstruction results in right, frontal and left views to demonstrate their multi-view consistency.

References

[1] Eric R. Chan, Connor Z. Lin, Matthew A. Chan, Koki Nagano, Boxiao Pan, Shalini De Mello, Orazio Gallo, Leonidas J. Guibas, Jonathan Tremblay, Sameh Khamis, Tero Karras, and Gordon Wetzstein. Efficient geometry-aware 3d generative adversarial networks. In *CVPR*, pages 16123–16133, 2022.

[2] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. Progressive growing of gans for improved quality, stability, and variation. In *ICLR*, 2018.