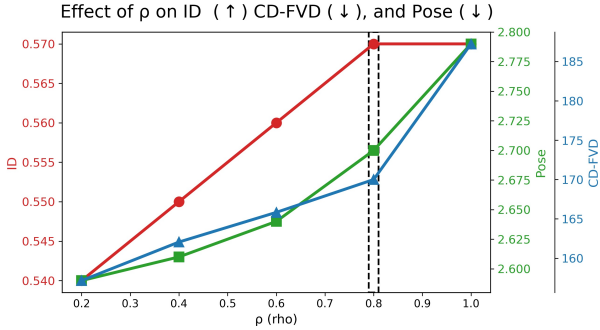


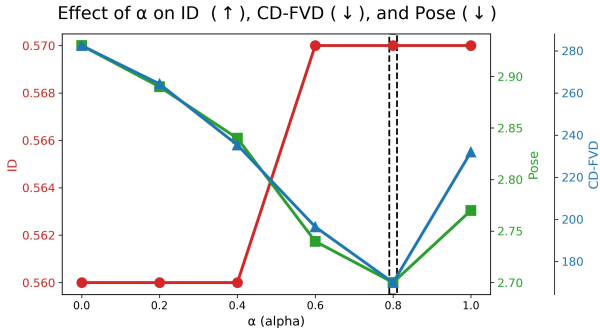
Supplementary material for VFace: A Training-Free Approach for Diffusion-Based Video Face Swapping

1. Hyper parameter analysis

We conduct hyper parameter analysis on ρ the FSAI split ratio, α the FATS blending ratio which is shown in Fig. 1a and Fig. 1b respectively. We find that values around 0.8 for both parameters yield the highest ID while maintaining considerably lower CD-FVD and Pose errors. Further, we observe that the FVD values for FATS time steps $T_1 = 10$ and $T_1 = 20$ are 230 and 639, respectively. These results indicate that, while FATS can improve video fidelity performance, excessive usage can lead to severe degradation.



(a) Effect of ρ (frequency split ratio).



(b) Effect of α (blending ratio).

Figure 1. Parameter ablations on frequency split ratio ρ and blending ratio α .

2. Ablations on Image Datasets and Hyperparameters

We provide additional ablations on **FFHQ** and **CelebA** to isolate the effects of each module. As shown in Table 2 and Table 1, these results quantify the impact of TSG and FSAI, confirming that both components contribute positively to frame-level fidelity and identity preservation across datasets. While TSG improves pose accuracy, it slightly degrades ID transferability. FSAI compensates for this degradation, restoring ID performance to baseline levels while preserving the pose improvements introduced by TSG.

Method	FID \downarrow	ID retrieval \uparrow		Pose \downarrow	Expr. \downarrow
		Top-1	Top-5		
REFace [1]	5.53	95.4%	98.7%	3.74	1.04
[1] + TSG	4.69	94.1%	98.2%	2.82	0.86
[1] + TSG + FSAI	4.29	96.2%	98.4%	2.90	0.86

Table 1. Ablation study on the FFHQ dataset.

Method	FID \downarrow	ID retrieval \uparrow		Pose \downarrow	Expr. \downarrow
		Top-1	Top-5		
REFace [1]	6.09	98.8%	99.6%	3.51	0.96
[1] + TSG	5.48	94.8%	97.8%	2.52	0.78
[1] + TSG + FSAI	5.42	96.2%	98.4%	2.63	0.77

Table 2. Ablation study on the CelebA dataset.

3. Limitation

We observe some flickering remains in our VFace outcomes. However, our method substantially improves temporal consistency over the base image-based models. Failure cases, such as severe occlusions, large appearance gaps, and identity leakage, mainly arise from limitations of the underlying models (see Fig. 2).

Figure 2. Some failure cases with occlusion and high pose variations. Play the video with adobe reader.

4. Ethics and Societal Impact

Our goal is to make face swapping and generation more accessible and creative for all users. However, we acknowledge the risk of misuse, including the creation of misleading or harmful media. Such techniques are required to be able to create synthetic data responsibly for learning deepfake detectors under varied scenarios. To promote responsible use, we encourage the development of safeguards such as bias detection, misuse prevention, and transparency mechanisms.

5. Adaptability to Other Image Based Face Swapping Methods

We include results on **FaceAdapter** [2], a recent diffusion-based image face swapping approach, which benefits from our modules, demonstrating the broader applicability of our method. We used the same hyperparameter settings as in ReFace in Table 3; however, due to differences in model implementations and underlying diffusion architectures, further hyperparameter tuning may improve performance.

Method	CD-FVD↓	FVD↓	ID Sim.↑	ID retrieval ↑		Pose↓	Expr.↓
				Top-1	Top-5		
Face Adapter	426.51	435.35	0.43	92.48%	98.33%	4.58	1.38
[14]+VFace	358.34	358.77	0.43	91.67%	98.12%	4.30	1.34

Table 3. VFace Implementation on FaceAdapter.

References

- [1] Sanoojan Baliah, Qinliang Lin, Shengcai Liao, Xiaodan Liang, and Muhammad Haris Khan. Realistic and efficient face swapping: A unified approach with diffusion models. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2025. Oral. [1](#)
- [2] Yue Han, Junwei Zhu, Keke He, Xu Chen, Yanhao Ge, Wei Li, Xiangtai Li, Jiangning Zhang, Chengjie Wang, and Yong Liu. Face adapter for pre-trained diffusion models with fine-grained id and attribute control, 2024. [2](#)