Controllable 3D Face Synthesis with Conditional Generative Occupancy Fields

Keqiang Sun*
CUHK MMLab

Shangzhe Wu*
Oxford VGG

Zhaoyang Huang
CUHK MMLab

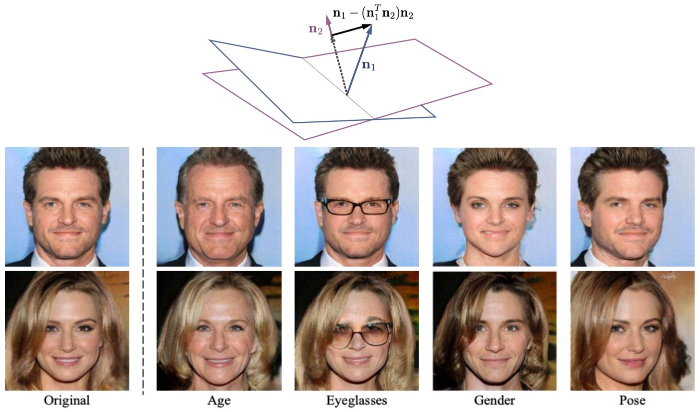
Ning Zhang
SenseTime

Quan Wang
SenseTime

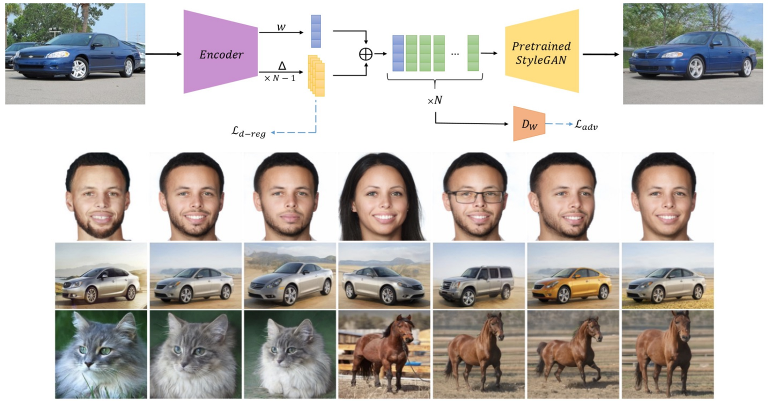
Hongsheng Li
CUHK MMLab

* Denotes Equal Contribution

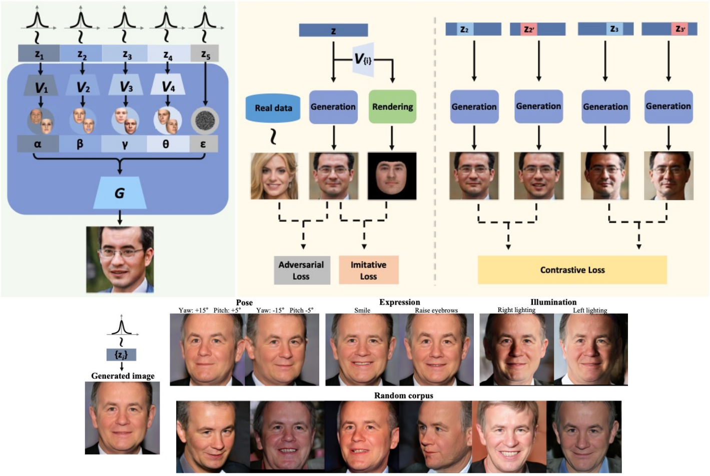
2D Generator Based



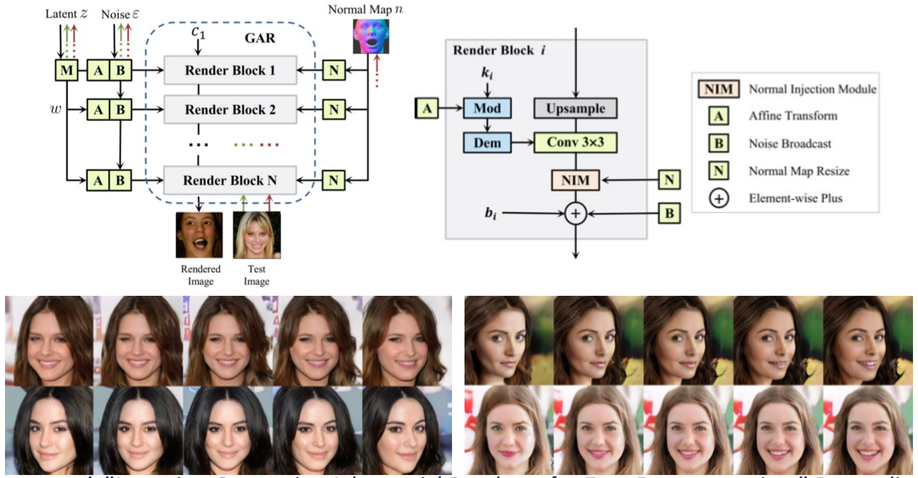
Shen, Yujun, et al. "Interpreting the latent space of gans for semantic face editing." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020.



Omer Tov, Yuval Alaluf, Yotam Nitzan, Or Patashnik, and Daniel Cohen-Or. Designing an encoder for stylegan image manipulation. SIGGRAPH, 2021.

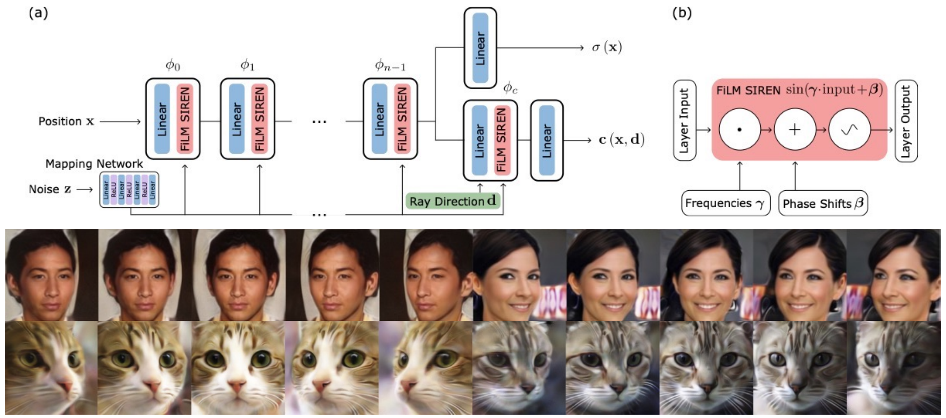


Deng, Yu, et al. "Disentangled and controllable face image generation via 3d imitative-contrastive learning." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2020.

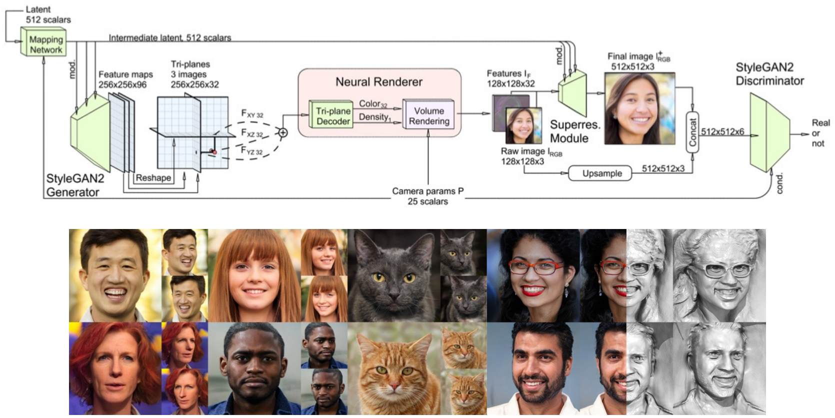


Piao, Jingtan, et al. "Inverting Generative Adversarial Renderer for Face Reconstruction." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021.

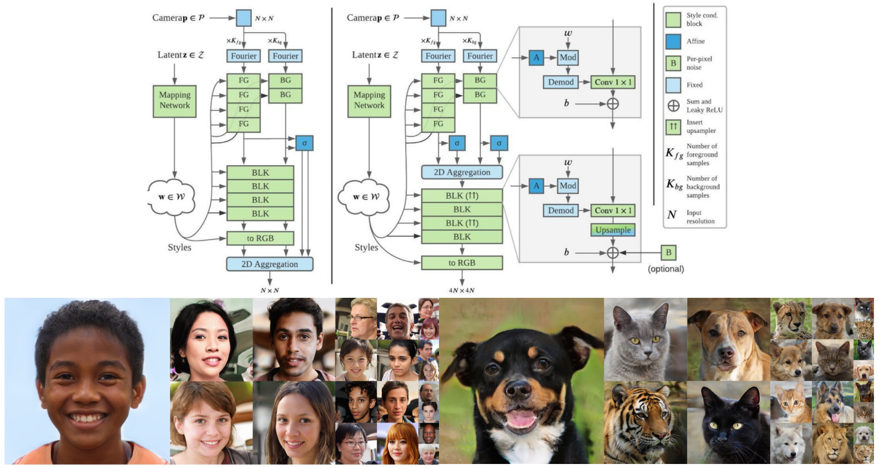
3D Generator



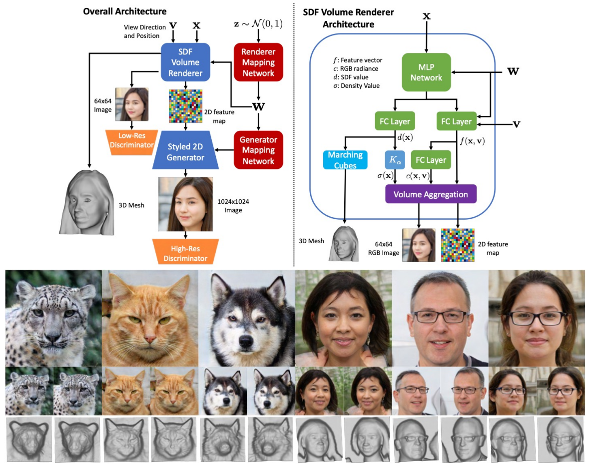
Chan, Eric R., et al. "pi-gan: Periodic implicit generative adversarial networks for 3d-aware image synthesis." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2021.



Chan, Eric R., et al. "Efficient geometry-aware 3D generative adversarial networks." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022.

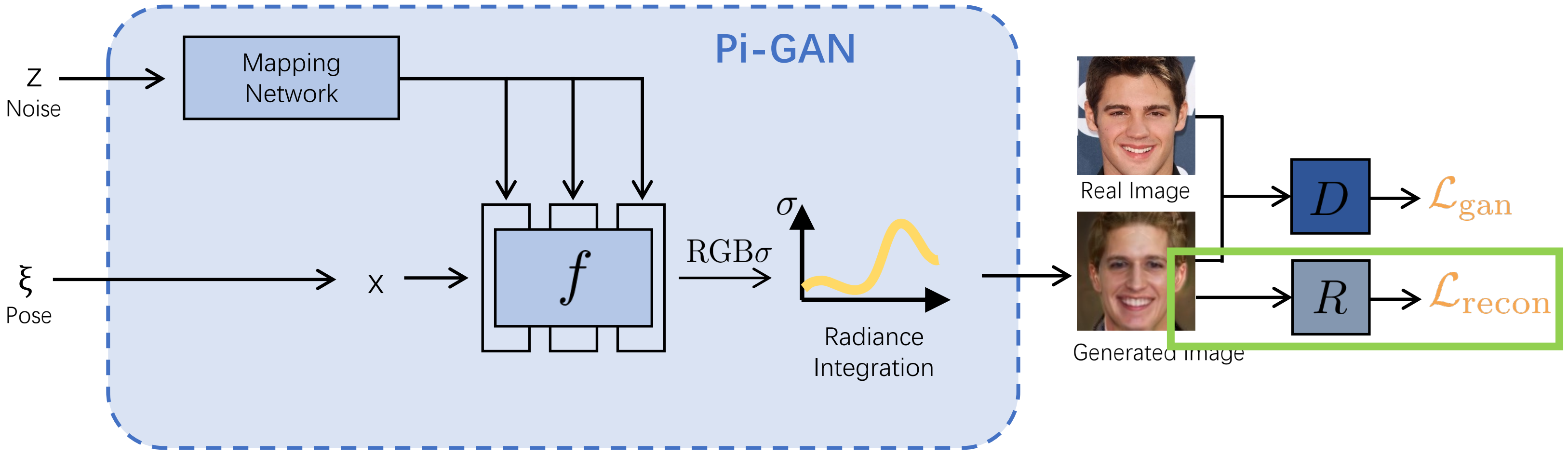


Gu, Jiatao, et al. "Stylenerf: A style-based 3d-aware generator for high-resolution image synthesis." *arXiv preprint arXiv:2110.08985* (2021).



Or-El, Roy, et al. "Stylesdf: High-resolution 3d-consistent image and geometry generation." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022.

Conditional Generative Occupancy Fields

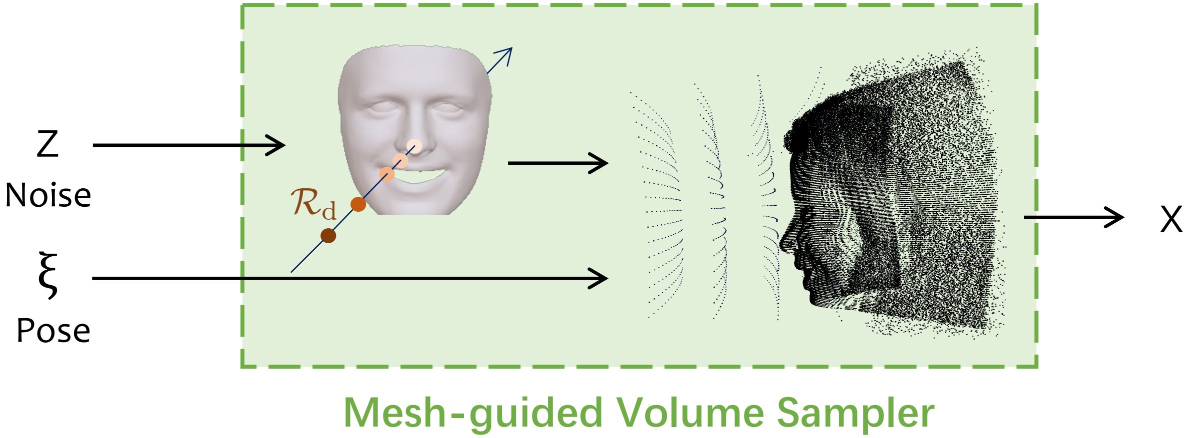


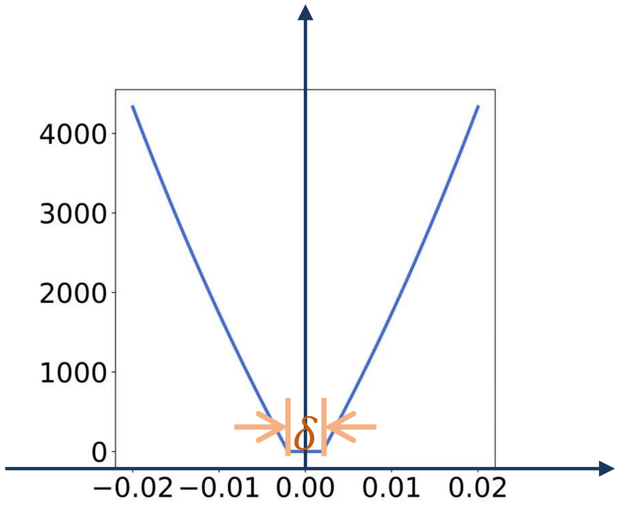
$$\mathcal{L}_{\text{recon}} = \|\hat{\mathbf{z}} - \mathbf{z}\|_1, \quad \text{where} \quad \hat{\mathbf{z}} = \tau(R(G(\mathbf{z}, \xi))).$$

Index	Loss	CD ↓	LD ↓	LC ↑	DS_s ↑	DS_e ↑	DS_p ↑	FID (128) ↓
1	$\mathcal{L}_{\text{gan}}$	1.09	5.04	2.04	2.13	2.54	7.16	**18.76**
2	$+\mathcal{L}_{\text{recon}}$	0.87	3.85	26.15	3.56	5.03	11.00	21.97

(baseline)

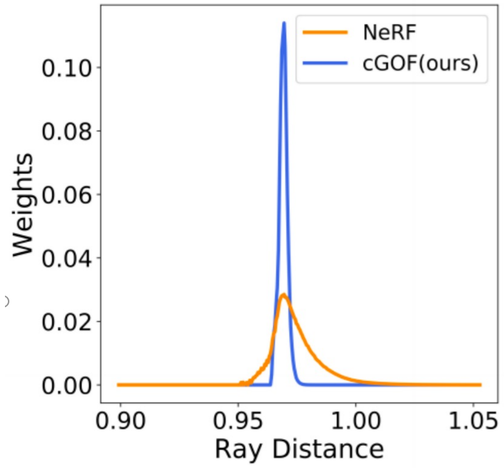
Conditional Generative Occupancy Fields



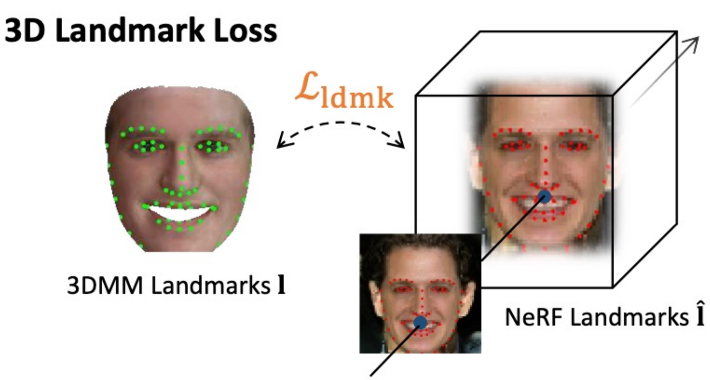


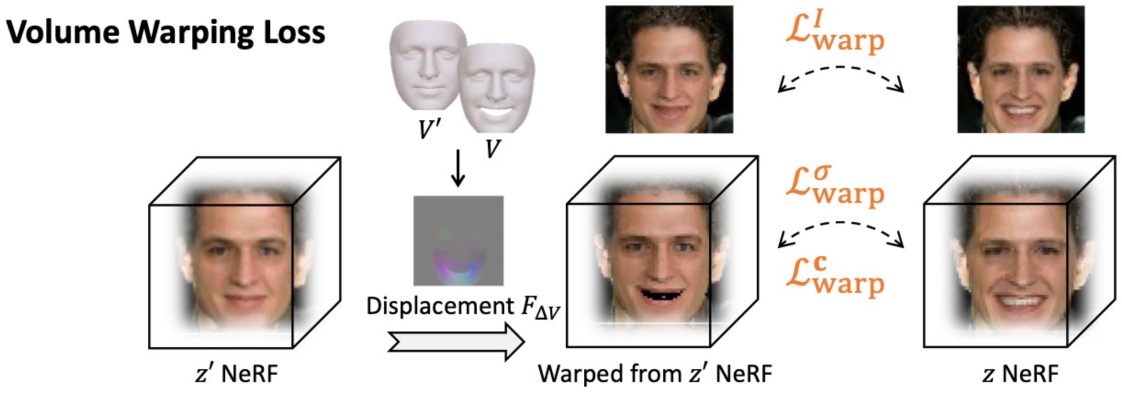
$$R_\mathrm{d} = \sum_{i=0}^{N_\mathrm{vol}} \sigma_i \cdot \underbrace{\left[\exp\left(\alpha \cdot \max(d_i - \delta/2, 0)\right) - 1\right]}_{\text{Weight Term}}$$

Index	Loss	CD ↓	LD ↓	LC ↑	DS_s ↑	DS_e ↑	DS_p ↑	FID (128) ↓
1	$\mathcal{L}_{\text{gan}}$	1.09	5.04	2.04	2.13	2.54	7.16	**18.76**
2	+ $\mathcal{L}_{\text{recon}}$	0.87	3.85	26.15	3.56	5.03	11.00	21.97
3	(+ $\mathcal{L}_{depth}$)*	1.65	6.16	0.06	0.93	1.34	1.09	71.67
4	+ MgS	0.29	3.98	27.45	3.55	4.90	9.48	38.91
5	+ $\mathcal{R}_\mathrm{d}$	0.31	3.51	51.74	3.56	5.31	15.94	29.62



Conditional Generative Occupancy Fields



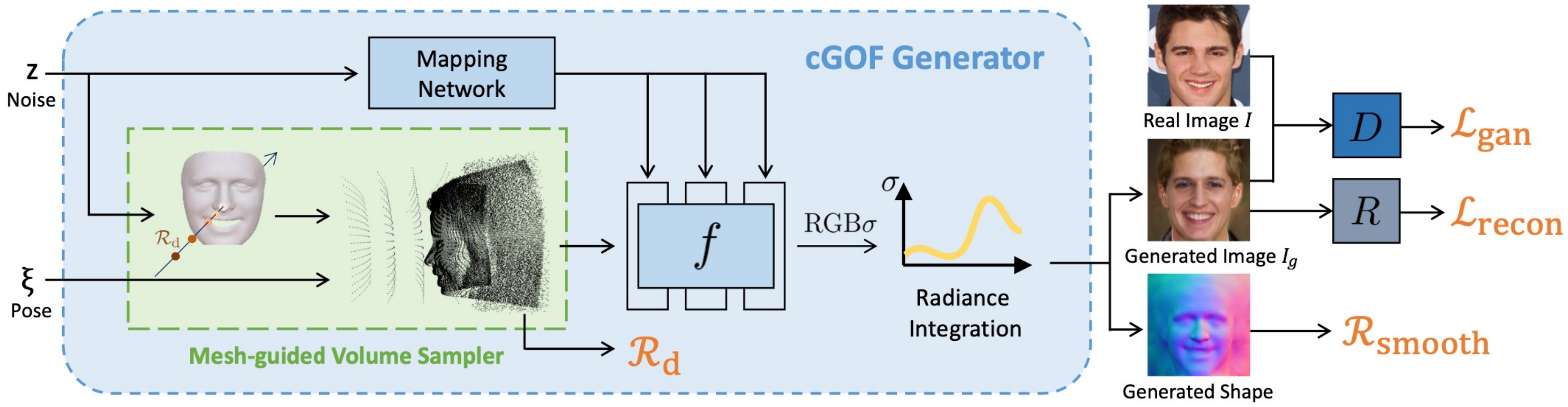


$$\mathcal{L}_{\text{ldmk}} = \sum_{k=1}^{N_{\text{k}}} \|\hat{\mathbf{l}}_k - \mathbf{l}_k\|_1 + \sum_{k=18}^{N_{\text{k}}} \|\hat{\mathbf{l}}'_k - \mathbf{l}_k\|_1 .$$

$$\mathcal{L}_{\text{warp}} = \beta_{\text{d}} \cdot \sum_{i}^{N_{\text{surf}}} \|\sigma'_i - \sigma_i\|_1 + \beta_{\text{c}} \cdot \sum_{i}^{N_{\text{surf}}} \|\mathbf{c}'_i - \mathbf{c}_i\|_1 + \beta_{\text{I}} \cdot \|\hat{I}_{\text{g}} - I_{\text{g}}\|_1 ,$$

Index	Loss	CD ↓	LD ↓	LC ↑	DS_{s} ↑	DS_{e} ↑	DS_{p} ↑	FID (128) ↓
1	$\mathcal{L}_{\text{gan}}$	1.09	5.04	2.04	2.13	2.54	7.16	**18.76**
2	+ $\mathcal{L}_{\text{recon}}$	0.87	3.85	26.15	3.56	5.03	11.00	21.97
3	(+ $\mathcal{L}_{depth}$)*	1.65	6.16	0.06	0.93	1.34	1.09	71.67
4	+ MgS	0.29	3.98	27.45	3.55	4.90	9.48	38.91
5	+ $\mathcal{R}_{\text{d}}$	0.31	3.51	51.74	3.56	5.31	15.94	29.62
6	+ $\mathcal{R}^{\text{norm}}_{\text{smooth}}$	0.27	4.72	30.25	3.18	4.43	16.26	31.63
7	+ $\mathcal{L}_{\text{ldmk}}$	0.39	1.86	84.43	16.54	16.65	21.24	56.90
8	+ $\mathcal{L}_{\text{warp}}$	**0.26**	1.44	89.91	20.47	22.04	22.91	47.18

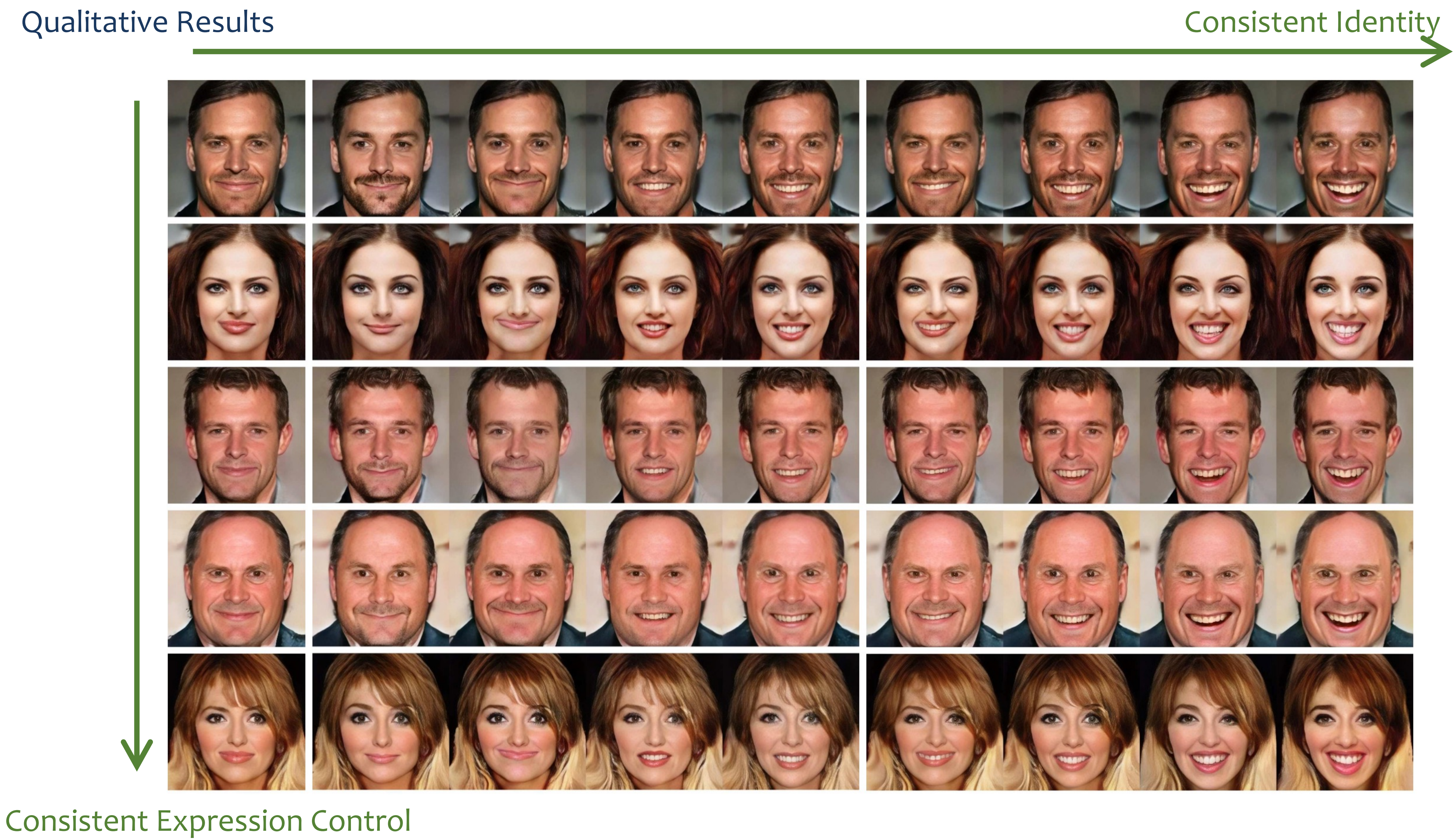
Conditional Generative Occupancy Fields



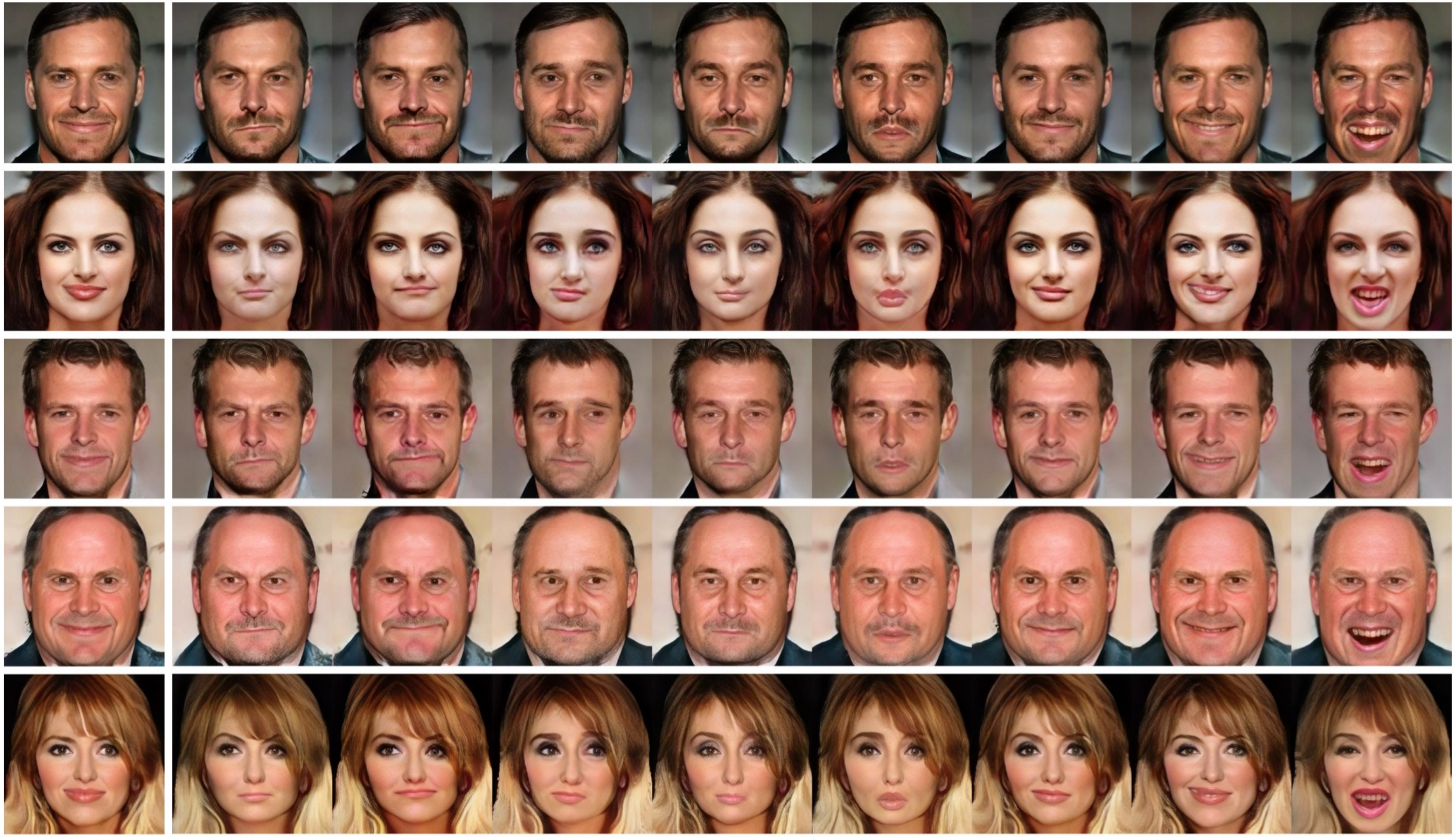
$$\mathcal{L} = \lambda_{\text{gan}}\mathcal{L}_{\text{gan}} + \lambda_{\text{recon}}\mathcal{L}_{\text{recon}} + \lambda_{\text{d}}R_{\text{d}} + \lambda_{\text{ldmk}}\mathcal{L}_{\text{ldmk}} + \lambda_{\text{warp}}\mathcal{L}_{\text{warp}} + \lambda_{\text{smooth}}R_{\text{smooth}}$$

Index	Loss	CD ↓	LD ↓	LC ↑	DS_s ↑	DS_e ↑	DS_p ↑	FID (128) ↓
1	$\mathcal{L}_{\text{gan}}$	1.09	5.04	2.04	2.13	2.54	7.16	**18.76**
2	+ $\mathcal{L}_{\text{recon}}$	0.87	3.85	26.15	3.56	5.03	11.00	21.97
3	(+ $\mathcal{L}_{depth}$)*	1.65	6.16	0.06	0.93	1.34	1.09	71.67
4	+ MgS	0.29	3.98	27.45	3.55	4.90	9.48	38.91
5	+ $\mathcal{R}_{\text{d}}$	0.31	3.51	51.74	3.56	5.31	15.94	29.62
6	+ $\mathcal{R}_{\text{smooth}}^{\text{norm}}$	0.27	4.72	30.25	3.18	4.43	16.26	31.63
7	+ $\mathcal{L}_{\text{ldmk}}$	0.39	1.86	84.43	16.54	16.65	21.24	56.90
8	+ $\mathcal{L}_{\text{warp}}$	**0.26**	1.44	89.91	20.47	22.04	22.91	47.18
9	+ $\mathcal{R}_{\text{smooth}}^{\text{depth}}$	0.27	**1.26**	**92.88**	**23.24**	**29.13**	**23.45**	26.64

Qualitative Results



Qualitative Results



Out-of-distribution Expression Control Results

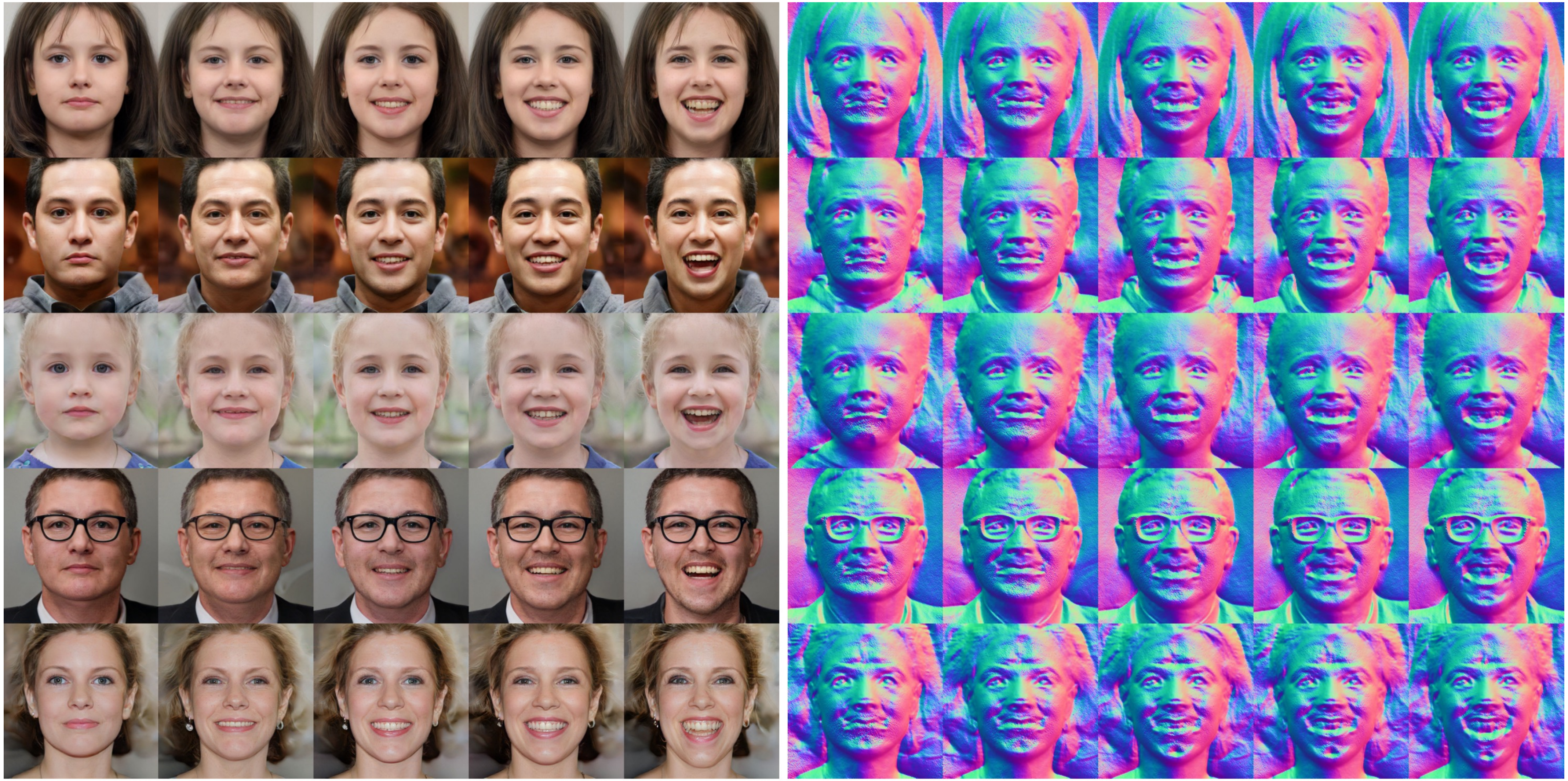
Qualitative Results

Head Pose Control Results

Qualitative Results

Head Pose Control Results

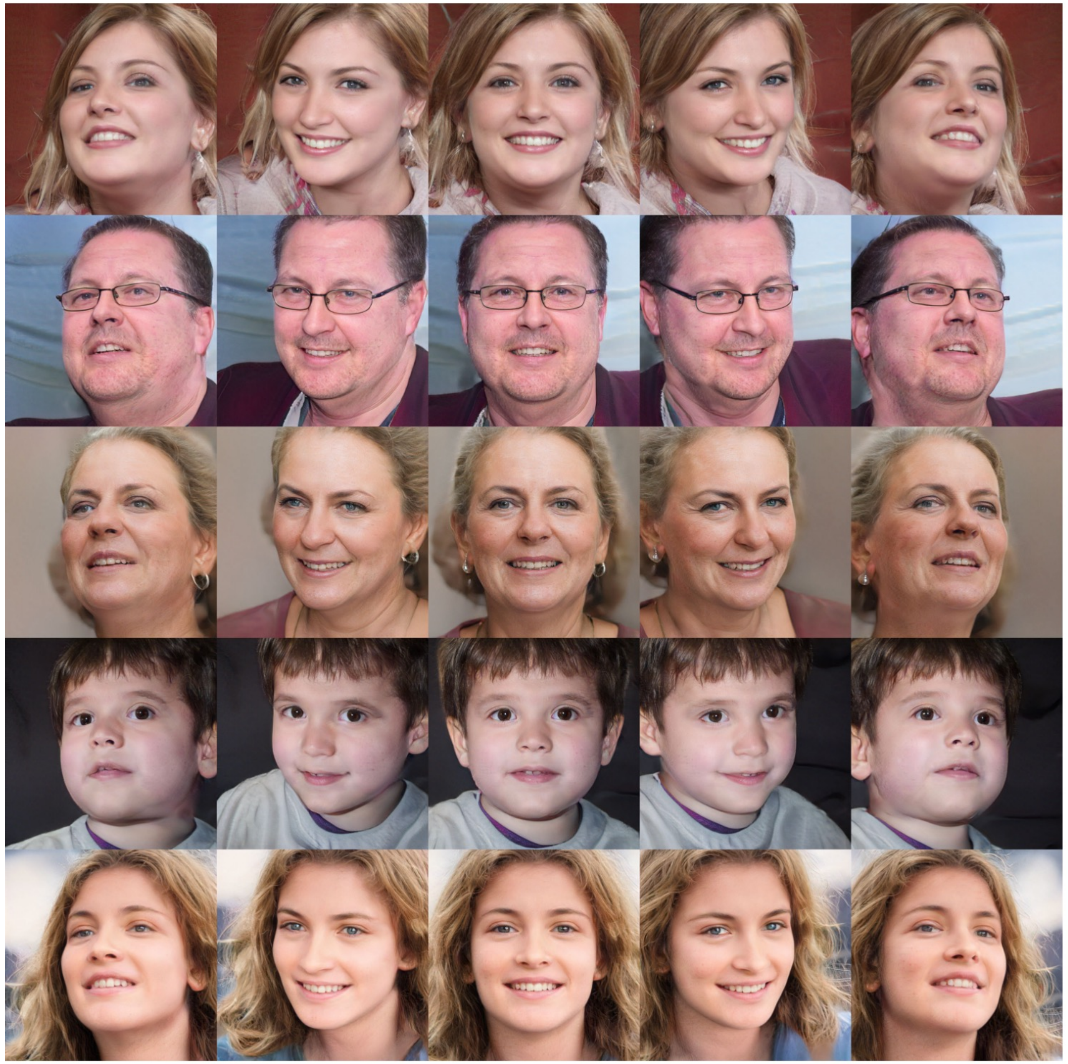
Extend to EG3D[1]



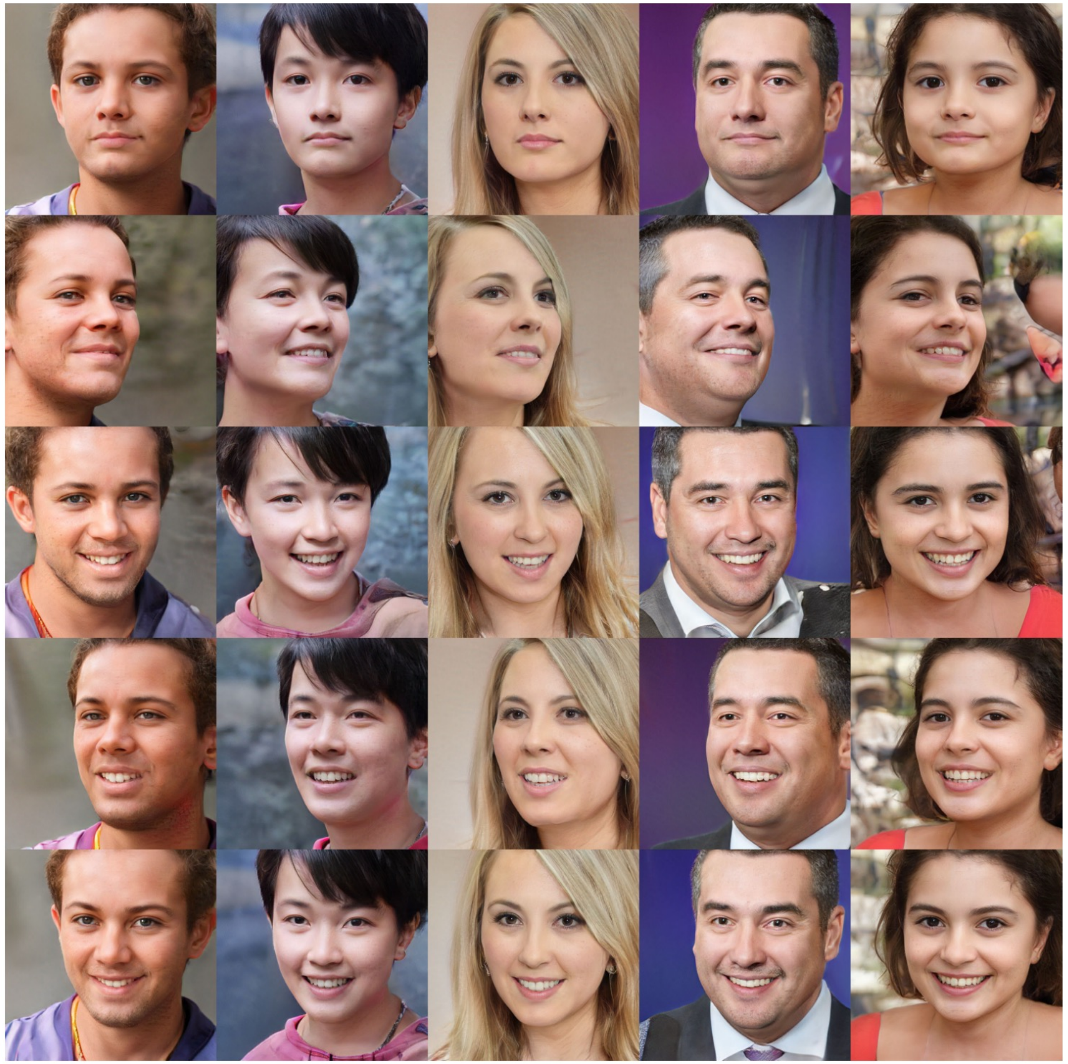
(a) Varying Expressions (b) Corresponding Normal Maps for (a)

[1] Chan, Eric R., et al. "Efficient geometry-aware 3D generative adversarial networks." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022.

Extend to EG3D[1]



(c) Varying Poses



(d) Varying Identities

[1] Chan, Eric R., et al. "Efficient geometry-aware 3D generative adversarial networks." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022.

Extend to EG3D[1]

Input Inversion Varying Expression

[1] Chan, Eric R., et al. "Efficient geometry-aware 3D generative adversarial networks." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022.

Thank You