

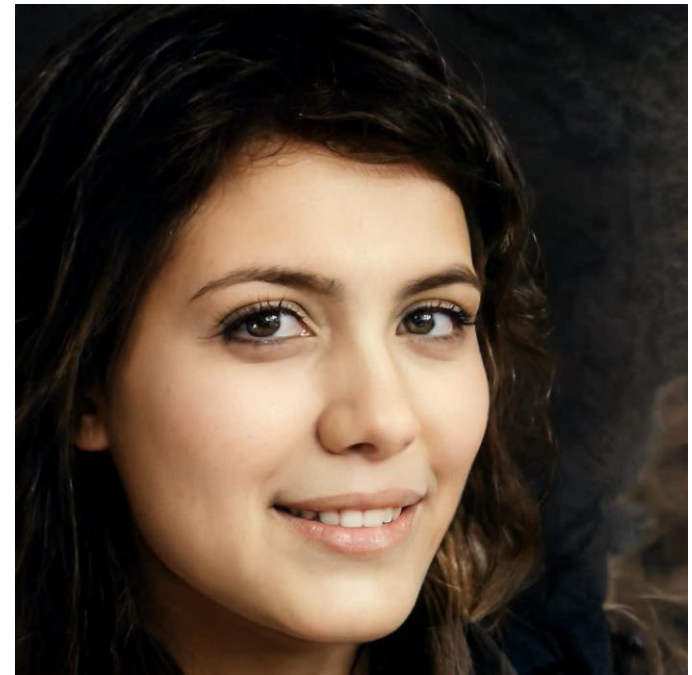
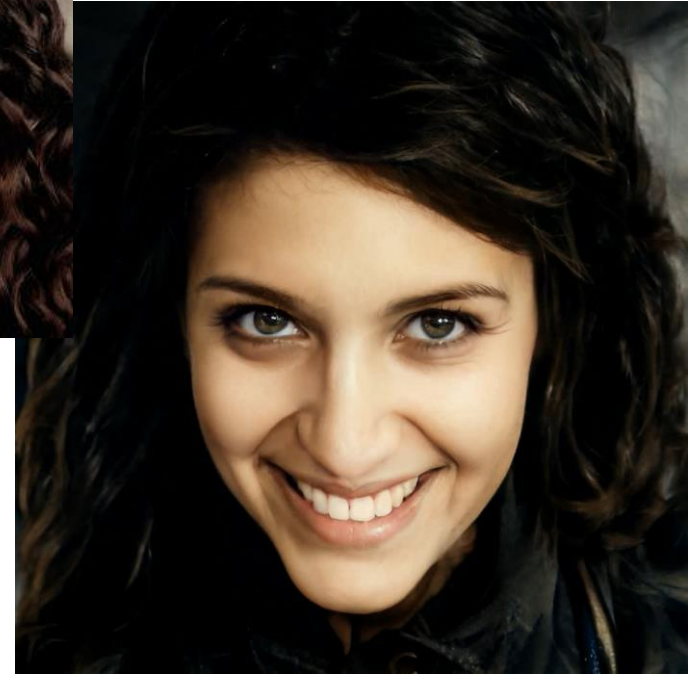
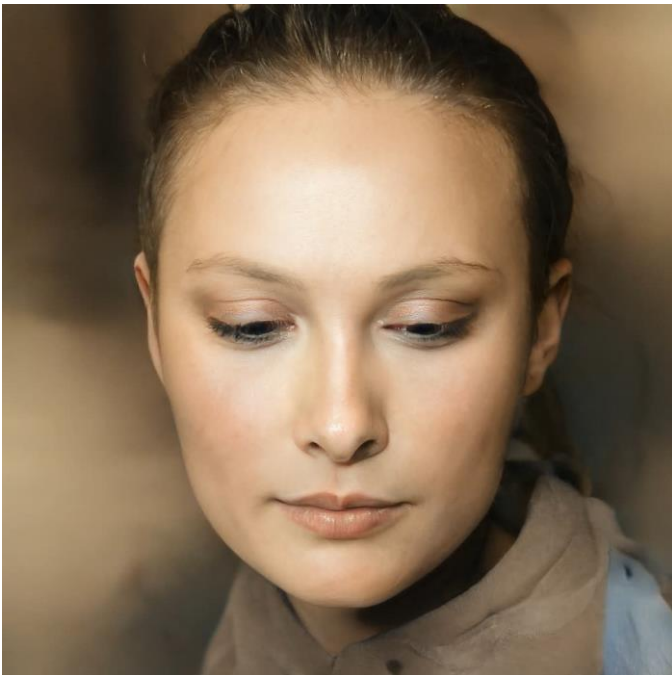
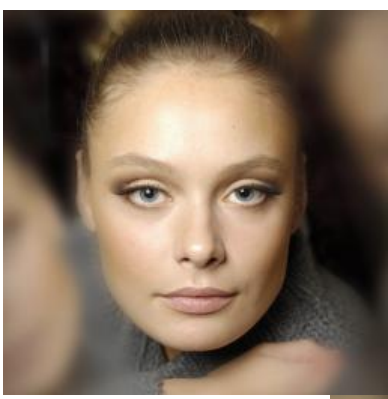


ReDirTrans: Latent-to-Latent Translation for Gaze and Head Redirection

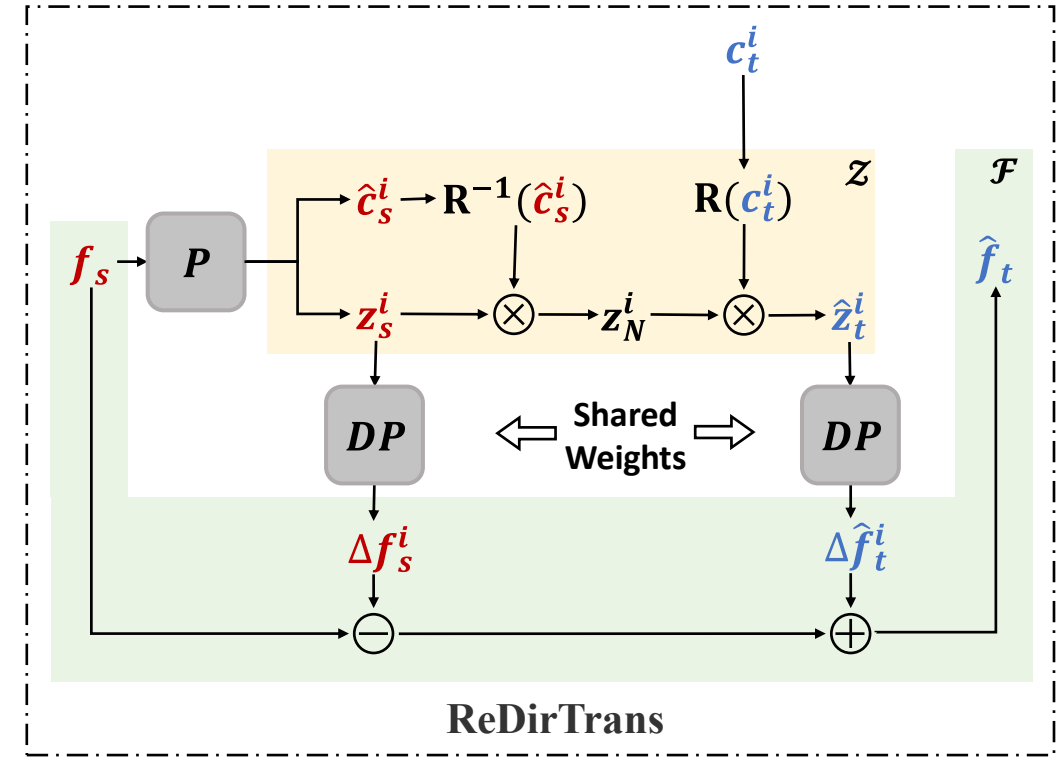
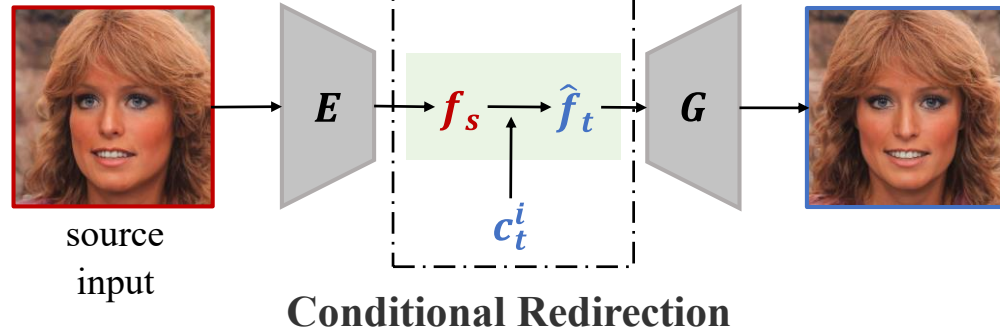
Shiwei Jin ¹, Zhen Wang ², Lei Wang ², Ning Bi ², Truong Nguyen ¹

¹ ECE Dept. UC San Diego, ² Qualcomm Technologies, Inc.





ReDirTrans



[4] Y. Zheng, S. Park, X. Zhang, S. De Mello, and O. Hilliges, "Self-learning transformations for improving gaze and head redirection," Advances in Neural Information Processing Systems, vol. 33, pp. 13127–13138, 2020.

[5] Y. Dalva, S. F. Altındaş, and A. Dundar, "Vecgan: Image-to-image translation with interpretable latent directions," in *Computer Vision--ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23--27, 2022, Proceedings, Part XVI*, pp. 153-169, 2022.

Motivation

Task	Method	Category	Image	Resolution	DoF	Condition
Gaze Redirection	DeepWarp [1]	Warping-	Eye	51×41	2	Pitch & Yaw
	Yu <i>et al.</i> [2]	Warping-	Eye	36×60 (60×75)	2	
	FAZE [3]	Generator-	Eye	256×64	2	
	ST-ED [4]	Generator-	Face (Restricted)	128×128	2	
Face Editing	StyleSapce [5]	Generator-	Face	1024×1024	1	No physical meanings

[1] Y. Zheng, S. Park, X. Zhang, S. De Mello, and O. Hilliges, “Self-learning transformations for improving gaze and head redirection,” *Advances in Neural Information Processing Systems*, vol. 33, pp. 13127–13138, 2020.

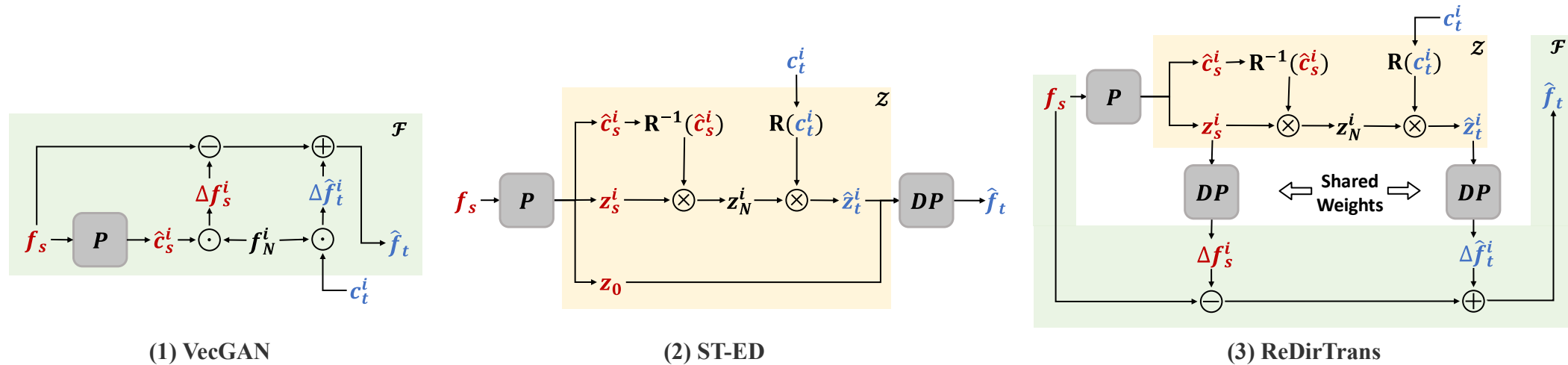
[2] Y. Ganin, D. Kononenko, D. Sungatullina, and V. Lempitsky, “Deepwarp: Photorealistic image resynthesis for gaze manipulation,” in *European conference on computer vision*, pp. 311–326, Springer, 2016.

[3] Y. Yu and J.-M. Odobez, “Unsupervised representation learning for gaze estimation,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7314–7324, 2020.

[4] S. Park, S. D. Mello, P. Molchanov, U. Iqbal, O. Hilliges, and J. Kautz, “Few-shot adaptive gaze estimation,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 9368–9377, 2019.

[5] Z. Wu, D. Lischinski, and E. Shechtman, “StyleSpace analysis: Disentangled controls for stylegan image generation,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 12863–12872, 2021.

Latent Vectors Editing in cGAN



Method	Latent Vector Compression	Initial Condition Estimation	Interpretability*	Portability	Editability	Physical Meaning of Conditions
VecGAN [5]	No	Yes	No	Yes	No	No
ST-ED [4]	Yes	Yes	Yes	No	-	Yes
ReDirTrans	No	Yes	Yes	Yes	Yes	Yes

* Transformation equivariant mappings between the embedding space and image space.

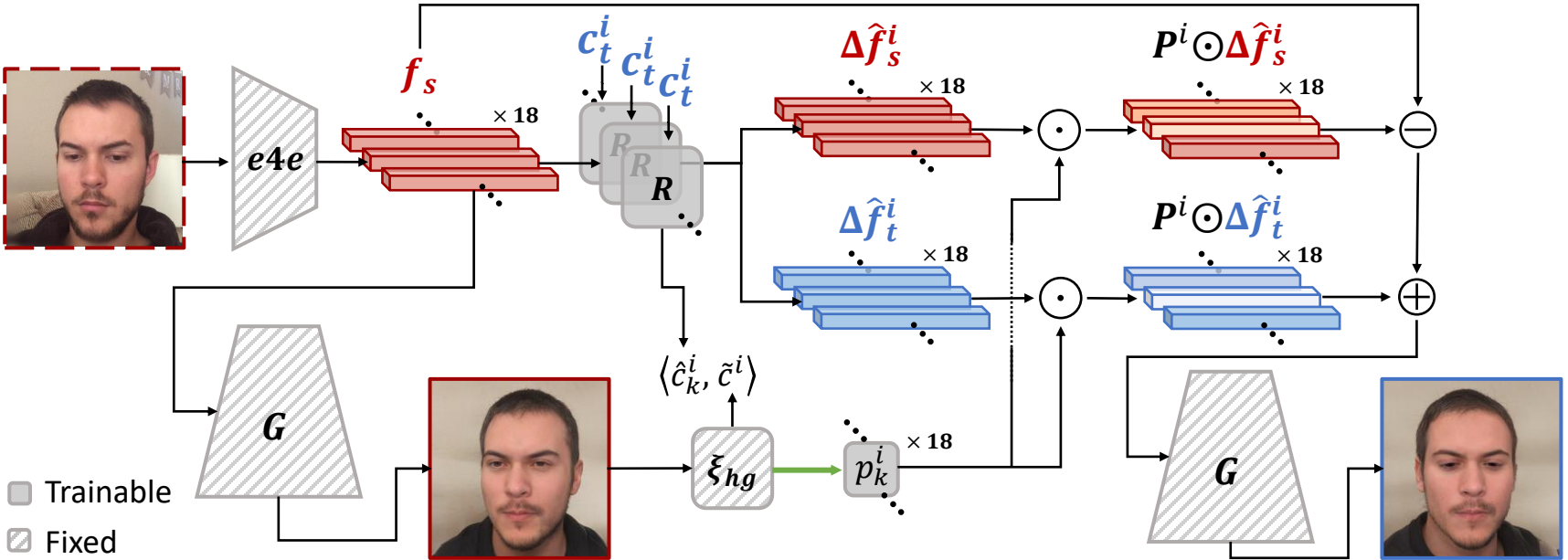
[4] Y. Zheng, S. Park, X. Zhang, S. De Mello, and O. Hilliges, "Self-learning transformations for improving gaze and head redirection," Advances in Neural Information Processing Systems, vol. 33, pp. 13127–13138, 2020.

[5] Y. Dalva, S. F. Altındaş, and A. Dundar, "Vecgan: Image-to-image translation with interpretable latent directions," in *Computer Vision--ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23--27, 2022, Proceedings, Part XVI*, pp. 153-169, 2022.

ReDirTrans-GAN

$$\hat{f}_t^i = f_s + \sum_{i=1}^M P^i \odot (-\Delta f_s^i + \Delta \hat{f}_t^i), i \in [1, M]$$

- ReDirTrans works with GAN inversion



[6] O. Tov, Y. Alaluf, Y. Nitzan, O. Patashnik, and D. Cohen-Or, "Designing an encoder for stylegan image manipulation," ACM Transactions on Graphics (TOG), vol. 40, no. 4, pp. 1–14, 2021.
 [7] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, "Analyzing and improving the image quality of stylegan," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 8110–8119, 2020.

Results

- Quantitative Comparison

Within-dataset: GazeCapture Test Subset [8]

	Gaze Redir	Head Redir	Gaze Induce	Head Induce	LPIPS
StarGAN	4.602	3.989	0.755	3.067	0.257
He <i>et al.</i>	4.617	1.392	0.560	3.925	0.223
VecGAN	2.282	0.824	0.401	2.205	0.197
ST-ED	2.385	0.800	0.384	2.187	0.208
ReDirTrans	2.163	0.753	0.429	2.155	0.197

Cross-dataset: MPIIFaceGaze [9]

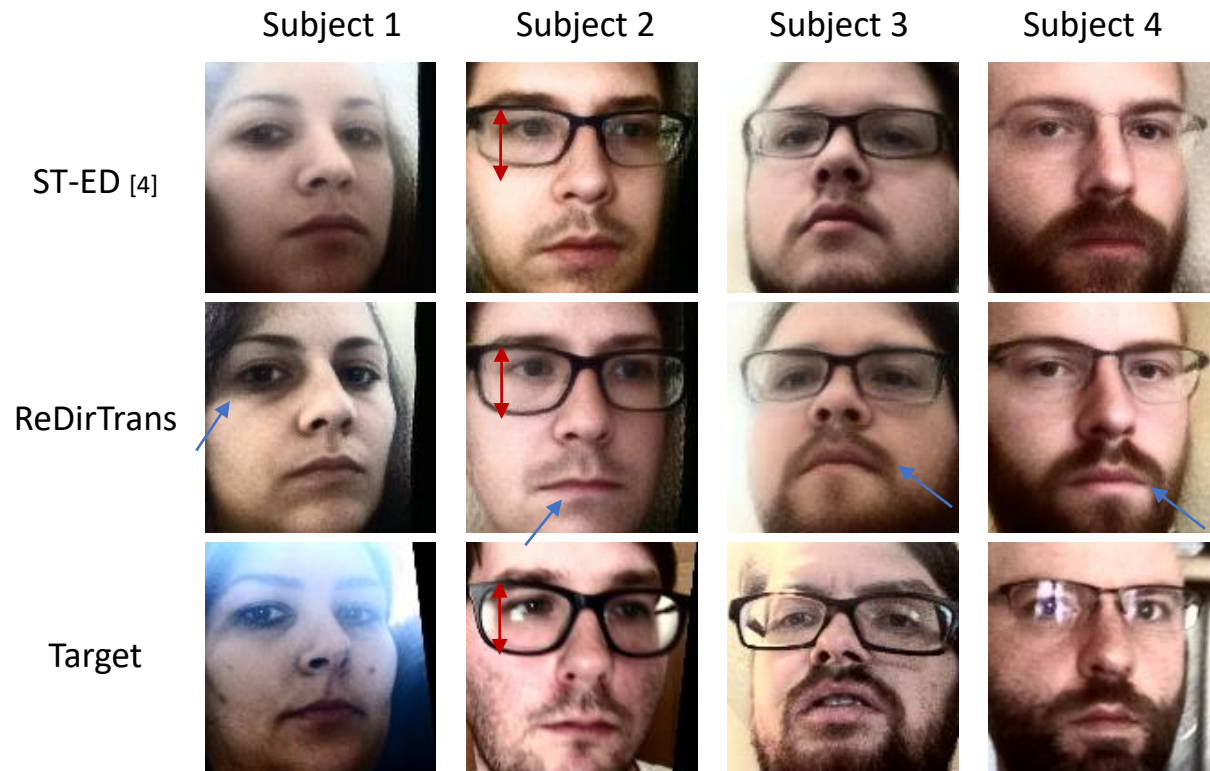
	Gaze Redir	Head Redir	Gaze Induce	Head Induce	LPIPS
StarGAN	4.488	3.031	0.786	2.783	0.260
He <i>et al.</i>	5.092	1.372	0.684	3.411	0.241
VecGAN	2.670	1.242	0.391	1.941	0.207
ST-ED	2.380	1.085	0.371	1.782	0.212
ReDirTrans	2.380	0.985	0.391	1.782	0.202

[8] K. Krafska, A. Khosla, P. Kellnhofer, H. Kannan, S. Bhandarkar, W. Matusik, and A. Torralba, "Eye tracking for everyone," in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2176–2184, 2016.

[9] X. Zhang, Y. Sugano, M. Fritz, and A. Bulling, "It's written all over your face: Full-face appearance-based gaze estimation," in Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp. 51–60, 2017.

Results

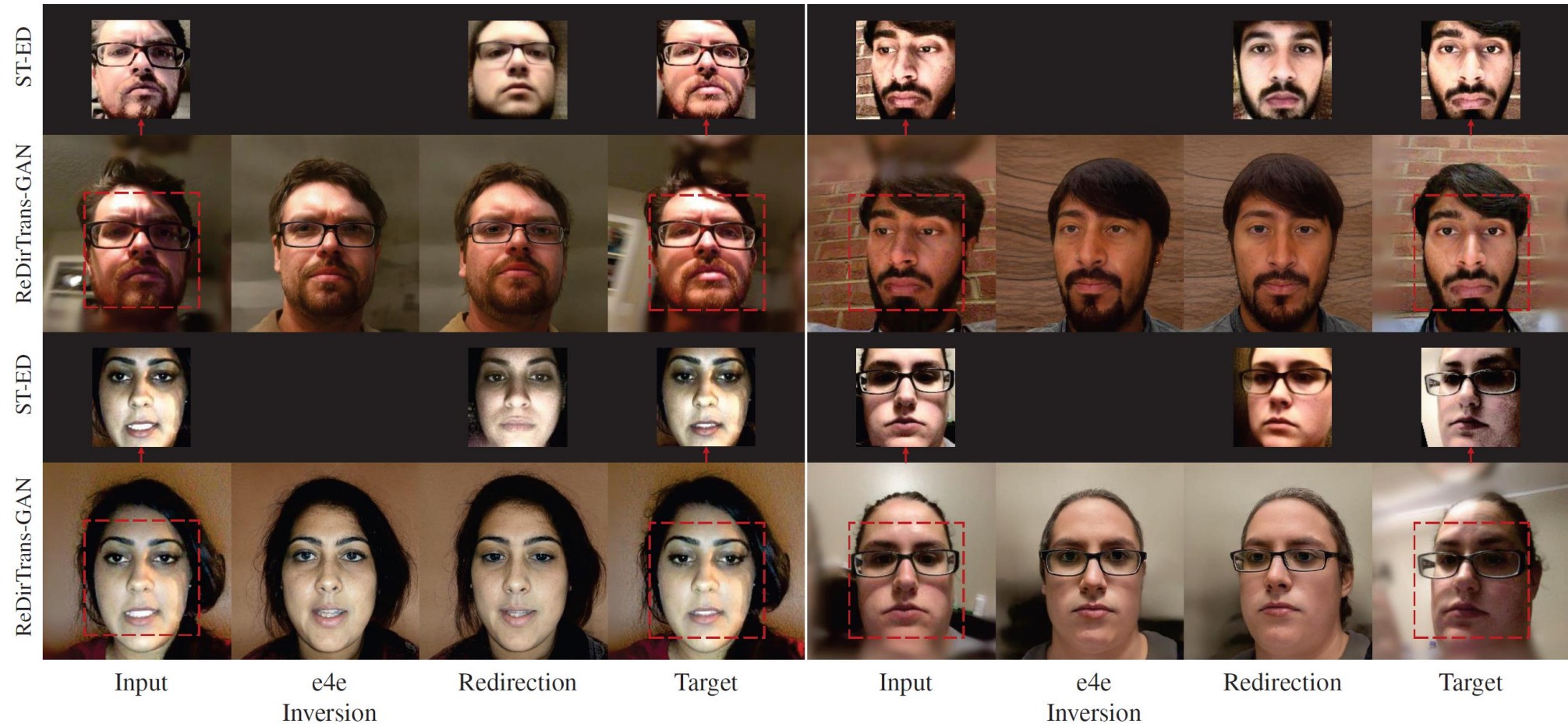
- Qualitative Comparison



[4] Y. Zheng, S. Park, X. Zhang, S. De Mello, and O. Hilliges, "Self-learning transformations for improving gaze and head redirection," Advances in Neural Information Processing Systems, vol. 33, pp. 13127–13138, 2020.

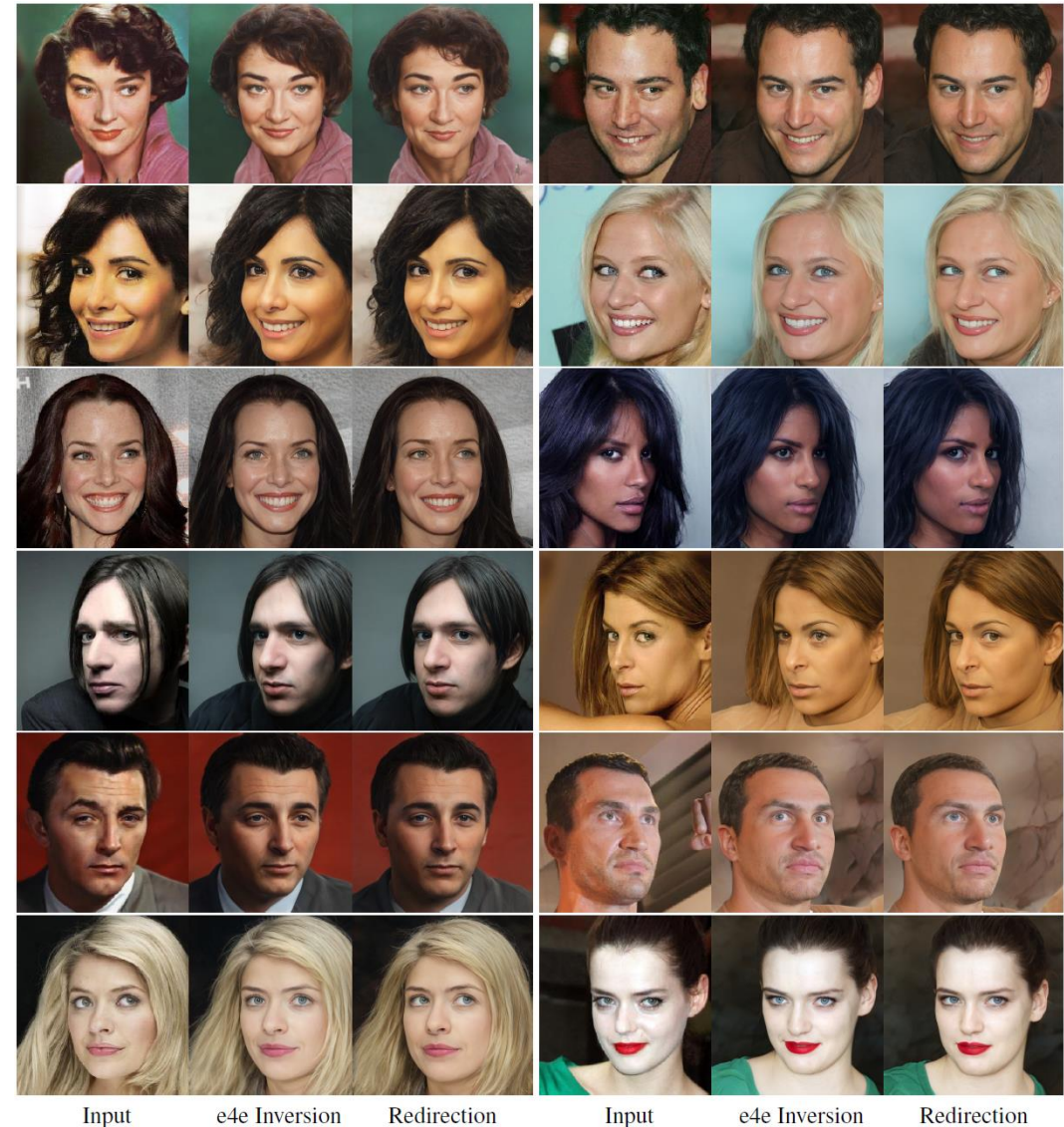
Results

- Qualitative Comparison



Gaze Correction

- Qualitative Comparison
 - CelebA-HQ [10]
- Pipeline
 - Using the same image as both input and target samples



[10] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of gans for improved quality, stability, and variation," arXiv preprint arXiv:1710.10196, 2017.

Gaze Correction

- Quantitative Performance

Within-dataset: GazeCapture Test Subset [8]

Method	Gaze Redir ↓	Head Redir ↓	LPIPS (I_t) ↓	ID (I_t) ↓	LPIPS (\hat{I}_t) ↓	ID (\hat{I}_t) ↓
GAN Inversion (e4e + StyleGAN2)	11.302	4.130	0.334	0.377	-	-
ReDirTrans-GAN	2.505	1.020	0.353	0.388	0.117	0.128

Cross-dataset: CelebA-HQ [10]

Method	Gaze Redir ↓	Head Redir ↓	LPIPS (I_t) ↓	ID (I_t) ↓	LPIPS (\hat{I}_t) ↓	ID (\hat{I}_t) ↓
GAN Inversion (e4e + StyleGAN2)	4.448	2.586	0.211	0.286	-	-
ReDirTrans-GAN	3.157	2.257	0.228	0.314	0.087	0.099

* I_t means the target image and \hat{I}_t means the inverted target image after GAN inversion.

[8] K. Krafcik, A. Khosla, P. Kellnhofer, H. Kannan, S. Bhandarkar, W. Matusik, and A. Torralba, "Eye tracking for everyone," in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2176–2184, 2016.

[10] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of gans for improved quality, stability, and variation," arXiv preprint arXiv:1710.10196, 2017.

Data Augmentation

- Downstream Gaze Estimation Task
 - 10,000 real images with annotations
 - Pick Q% real images (Subset A) for synthesizing the same number of images (Subset B)
 - Raw: Subset A, Aug: Subsets A & B

Q%	GazeCapture		MPIIFaceGaze	
	Raw ↓	Aug ↓	Raw ↓	Aug ↓
25	5.875	5.238	8.607	7.096
50	4.741	4.506	6.787	6.113
75	4.308	4.200	6.165	5.767

Thank You

Paper: <https://arxiv.org/pdf/2305.11452.pdf>