MyStyle++: A Controllable Personalized Generative Prior

Libing Zeng
libingzeng@tamu.edu
Texas A&M University
College Station, USA

Lele Chen
lele.chen@sony.com
Sony AI
New York, USA

Yi Xu
yi.xu@oppo.com
OPPO US Research Center
Palo Alto, USA

Nima Khademi Kalantari
nimak@tamu.edu
Texas A&M University
College Station, USA

ACM Reference Format:
Libing Zeng, Lele Chen, Yi Xu, and Nima Khademi Kalantari. 2023. MyStyle++: A Controllable Personalized Generative Prior. In *SIGGRAPH Asia 2023 Conference Papers (SA Conference Papers '23), December 12–15, 2023, Sydney, NSW, Australia.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3610548.3618171

In this supplementary materials, we offer additional visual and numerical comparisons against MyStyle_P and MyStyle_I for all individuals. Our method consistently outperforms MyStyle_P and MyStyle_I across all comparisons visually and numerically. Detailed visual results can be found in Sec. 1, while additional numerical results are provided in Sec. 2.

1 ADDITIONAL VISUAL RESULTS

In this section, we begin by presenting controlled image synthesis comparisons for all individuals. In the main paper, we have shown each attribute controlled image synthesis on different individuals. Here, we include image synthesis with all attributes controlled for each individual. as shown in Figs. 1 through 6 for Barack Obama, Michelle Obama, Oprah Winfrey, Taylor Swift, Emma Watson, and Leonardo DiCaprio, respectively. Our method consistently achieves attribute-controlled image synthesis, in contrast to MyStyle_I and MyStyle_P, which exhibit significant inaccuracies in attribute controllability.

Next, we provide more visual comparison results for sampled image editing in Figs. 7 through 10. For each individual, we present the results of editing all the attributes. Our method ensures the preservation of identity and consistency of non-edited attributes, while editing the expression, yaw, pitch, or age of the image.

Lastly, we showcase more visual comparisons for real image editing in Fig. 11. Our method demonstrates improved attribute disentanglement compared to MyStyle_P and MyStyle_I. It produces results that effectively preserve the unchanged attributes and identity.

SA Conference Papers '23, December 12–15, 2023, Sydney, NSW, Australia

ACM ISBN 979-8-4007-0315-7/23/12.
https://doi.org/10.1145/3610548.3618171

2 ADDITIONAL NUMERICAL RESULTS

We first further numerically evaluate attribute controllability by by fixing one attribute and randomly sampling the other ones to generate images, using the same strategy introduced in the main paper. Here, we shown more comparisons on individuals, such as Barack Obama, Joe Biden, Michelle Obama, Oprah Winfrey, Taylor Swift, and Emma Watson, as shown in Table 1 and Table 2. As seen, our approach consistently demonstrates the smallest standard deviation across all attributes for individuals.

Next, to ensure that the latent organization does not compromise the quality of the results, we evaluate our method against MyStyle using the ID metric [Nitzan et al. 2022] and diversity score [Ojha et al. 2021] on more individuals, such as Barack Obama, Emma Watson, Joe Bide, Michelle Obama, Oprah Winfrey, and Taylor Swift. The comparison, presented in Table 3, demonstrates that our method produces results that are comparable to those of MyStyle.

We further conduct a numerical comparison of our real image editing results with MyStyle_P and MyStyle_I on Michelle Obama, Leonardo DiCaprio, Barack Obama, Joe Biden, Oprah Winfrey, Scarlett Johansson, Taylor Swift, and Emma Watson, as presented in Table 4, Table 5, and Table 6. The results demonstrate the consistent superiority of our method over MyStyle_P and MyStyle_I across all metrics.

Figure 1: We present a comparison of our synthesis results with those of MyStyle_I and MyStyle_P on Barack Obama.

Figure 2: We present a comparison of our synthesis results with those of MyStyle_I and MyStyle_P on Michelle Obama.

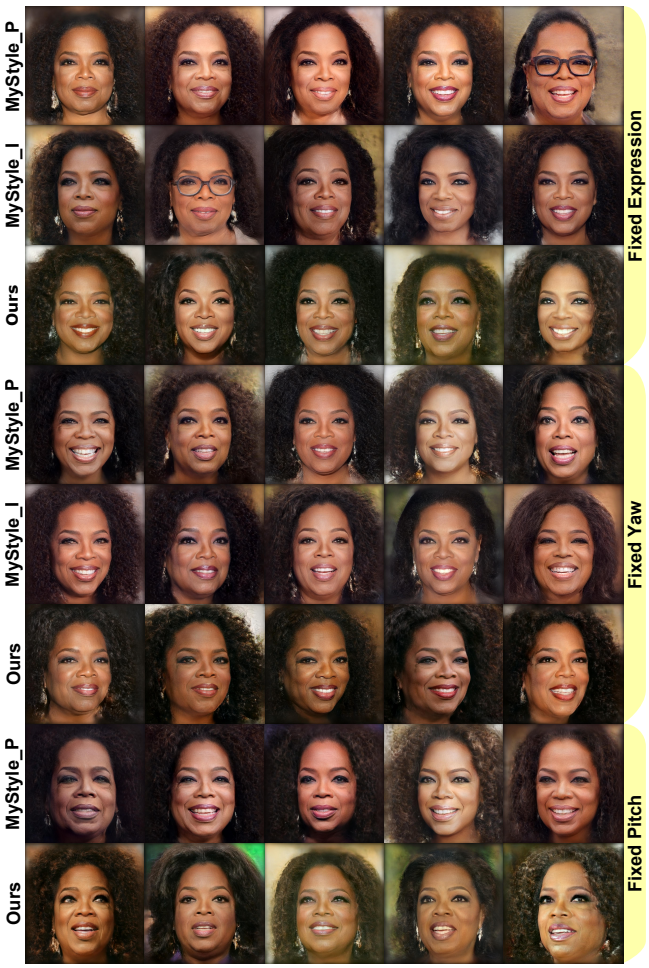

Figure 3: We present a comparison of our synthesis results with those of MyStyle_I and MyStyle_P on Oprah Winfrey.

Figure 4: We present a comparison of our synthesis results with those of MyStyle_I and MyStyle_P on Taylor Swift.

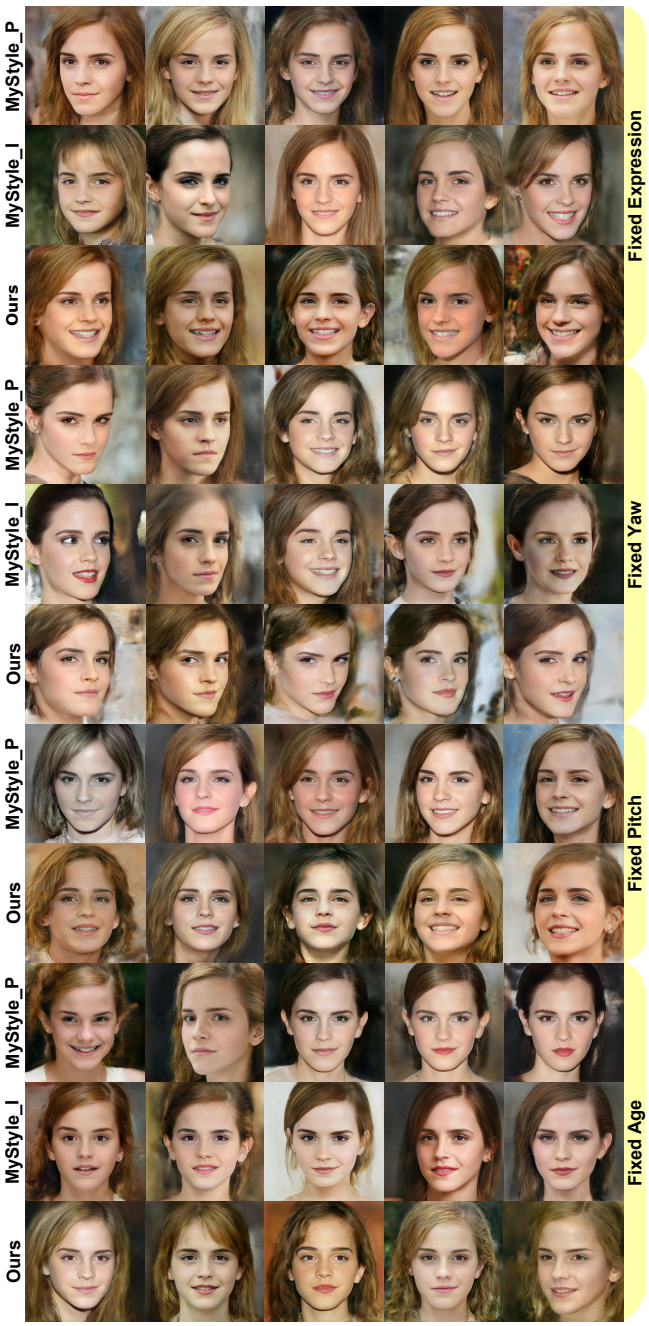

Figure 5: We present a comparison of our synthesis results with those of MyStyle_I and MyStyle_P on Emma Watson.

Figure 6: We present a comparison of our synthesis results with those of MyStyle_I and MyStyle_P on Leonardo Di-Caprio.

Figure 7: We show our results for editing expression, yaw, and pitch angles for sampled images against MyStyle_P and MyStyle_I on Barack Obama. Our method maintains the identity and consistency of other attributes while making modifications to the expression, yaw, or pitch of the image.

Figure 8: We show our results for editing expression, yaw, and pitch angles for sampled images against MyStyle_P and MyStyle_I on Joe Biden. Our method maintains the identity and consistency of other attributes while making modifications to the expression, yaw, or pitch of the image.

Figure 9: We show our results for editing expression, yaw, and pitch angles for sampled images against MyStyle_P and MyStyle_I on Scarlett Johansson. Our method maintains the identity and consistency of other attributes while making modifications to the expression, yaw, or pitch of the image.

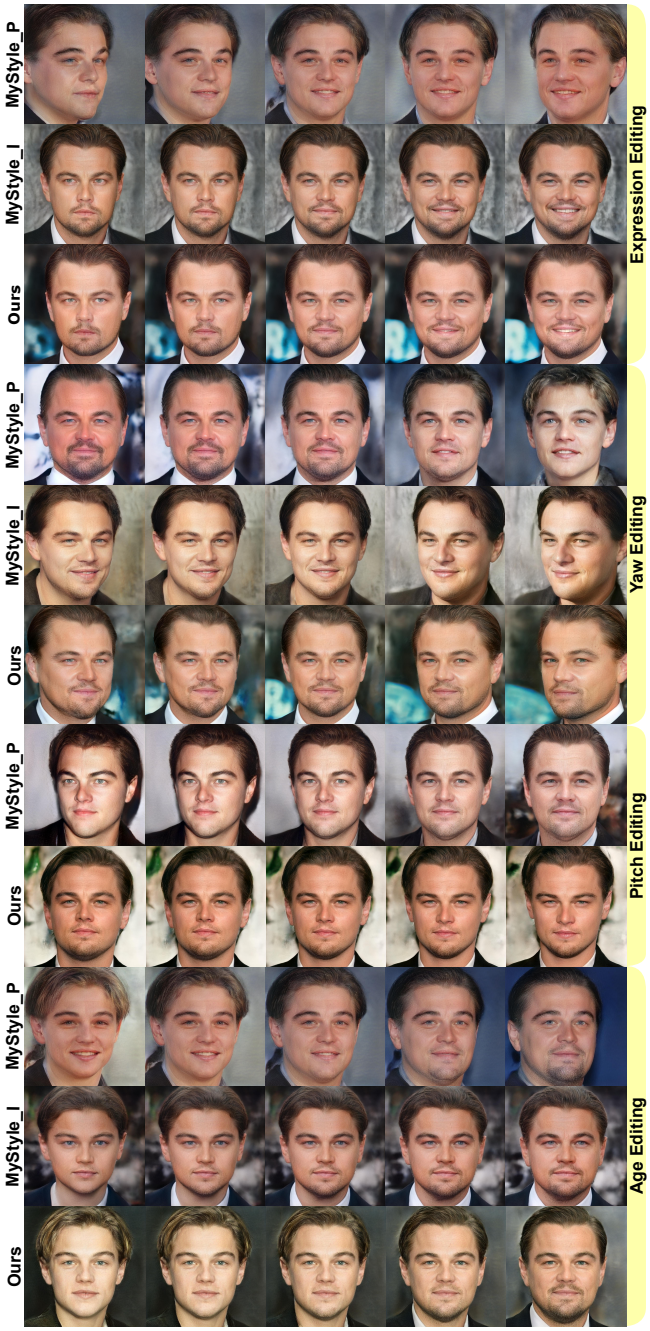

Figure 10: We show our results for editing expression, yaw, pitch, and age for sampled images against MyStyle_P and MyStyle_I on Leonardo DiCaprio. Our method maintains the identity and consistency of other attributes while making modifications to the expression, yaw, pitch, or age of the image.

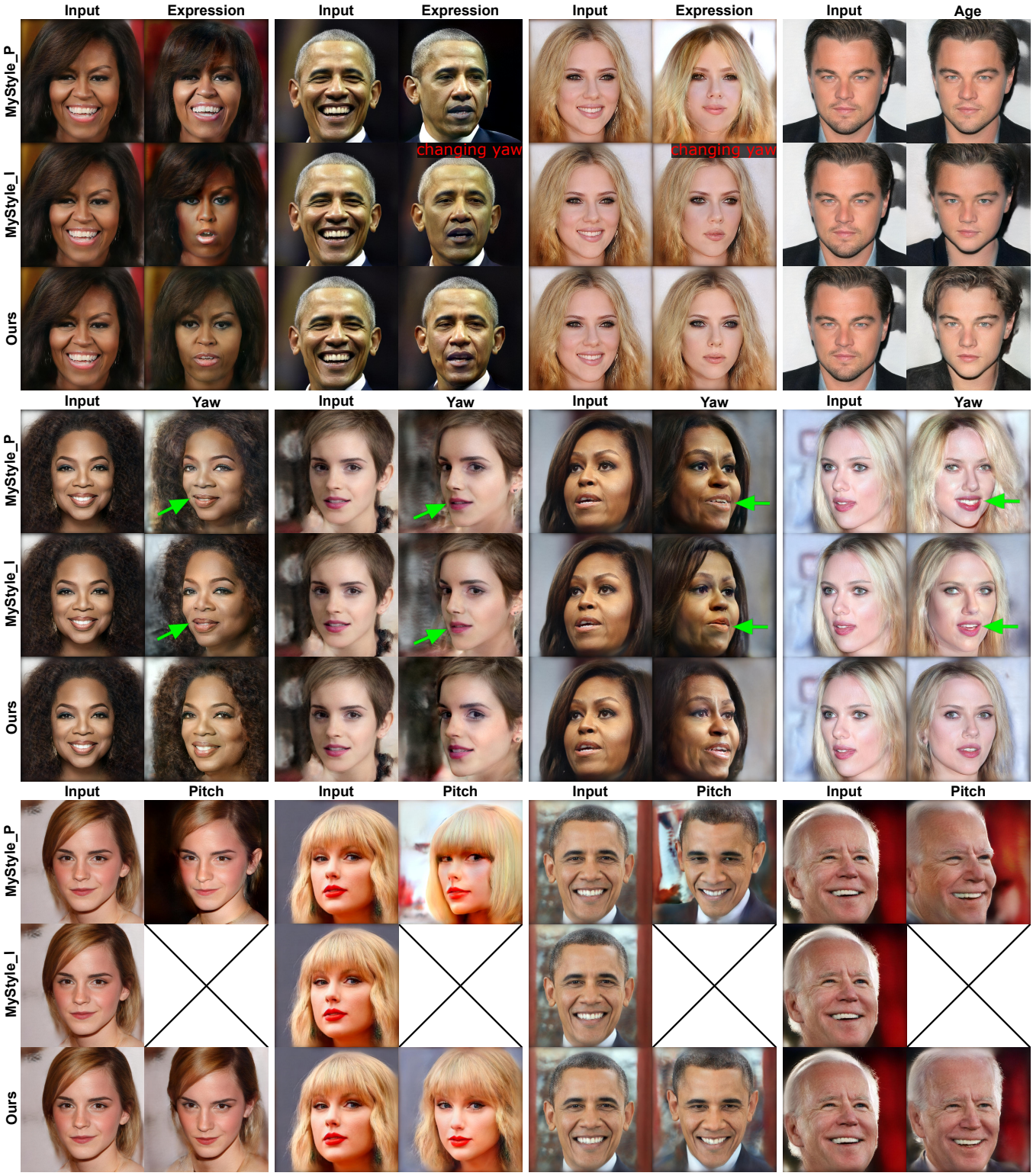

Figure 11: We show comparisons against both versions of MyStyle for semantic editing of real images. Input is the projected images into the latent space of MyStyle and our generators. Our method disentangles the attributes better than MyStyle_P and MyStyle_I, producing results that preserve the unchanged attributes and identity. InterFaceGAN [Shen et al. 2020] does not provide an edit direction for pitch, and thus we leave the corresponding area of MyStyle_I with a cross.

Table 1: We numerically compare our controlled synthesis results against MyStyle_P and MyStyle_I. We generate 100 images for each fixed attribute value and report the standard deviation of the estimated attribute of interest over the generated images. Note that the attribute values (e.g., 0.25) are in the normalized coordinate d_m. The best results are shown in bold.

		Barack Obama				
		0.0	0.25	0.5	0.75	1.0
Exp	MyStyle_P	0.664	0.822	0.846	0.915	0.899
	MyStyle_I	0.855	0.866	0.894	0.740	0.701
	Ours	**0.241**	**0.575**	**0.582**	**0.457**	**0.003**
Yaw	MyStyle_P	4.072	6.349	4.168	5.258	5.289
	MyStyle_I	5.858	5.530	5.249	3.364	5.360
	Ours	**1.951**	**1.952**	**1.846**	**2.158**	**1.900**
Pitch	MyStyle_P	5.748	4.061	7.127	6.295	5.611
	MyStyle_I	-	-	-	-	-
	Ours	**1.670**	**2.198**	**2.851**	**3.413**	**3.336**

		Joe Biden				
		0.0	0.25	0.5	0.75	1.0
Exp	MyStyle_P	0.330	0.408	0.415	0.375	0.370
	MyStyle_I	0.056	0.347	0.368	0.242	0.121
	Ours	**0.031**	**0.216**	**0.329**	**221**	**0.013**
Yaw	MyStyle_P	6.725	6.411	5.071	9.103	7.652
	MyStyle_I	5.578	3.777	3.474	4.832	6.687
	Ours	**2.331**	**2.585**	**1.571**	**3.503**	**5.947**
Pitch	MyStyle_P	1.612	3.901	4.158	7.854	7.681
	MyStyle_I	-	-	-	-	-
	Ours	**4.061**	**3.669**	**1.431**	**2.464**	**2.117**

		Michelle Obama				
		0.0	0.25	0.5	0.75	1.0
Exp	MyStyle_P	0.478	0.601	0.828	0.612	0.655
	MyStyle_I	0.872	0.798	0.521	0.361	0.001
	Ours	**0.169**	**0.378**	**0.217**	**0.038**	**0.000**
Yaw	MyStyle_P	3.239	4.464	2.757	5.195	4.835
	MyStyle_I	4.473	2.920	3.692	4.077	4.523
	Ours	**1.197**	**1.581**	**1.796**	**1.717**	**1.488**
Pitch	MyStyle_P	3.902	4.673	4.661	3.547	3.823
	MyStyle_I	-	-	-	-	-
	Ours	**2.426**	**2.264**	**1.699**	**1.823**	**1.687**

		Oprah Winfrey				
		0.0	0.25	0.5	0.75	1.0
Exp	MyStyle_P	0.425	0.509	0.039	0.0024	0.006
	MyStyle_I	0.067	0.371	0.003	0.009	0.002
	Ours	**0.048**	**0.293**	**0.002**	**0.003**	**0.001**
Yaw	MyStyle_P	5.905	4.397	7.275	3.844	6.130
	MyStyle_I	3.069	3.180	2.136	2.605	3.686
	Ours	**1.333**	**2.704**	**1.649**	**1.532**	**1.119**
Pitch	MyStyle_P	4.551	4.470	3.765	3.788	4.181
	MyStyle_I	-	-	-	-	-
	Ours	**1.509**	**2.520**	**1.622**	**1.729**	**1.797**

Table 2: We numerically compare our controlled synthesis results against MyStyle_P and MyStyle_I. We generate 100 images for each fixed attribute value and report the standard deviation of the estimated attribute of interest over the generated images. Note that the attribute values (e.g., 0.25) are in the normalized coordinate d_m. The best results are shown in bold.

		Taylor Swift				
		0.0	0.25	0.5	0.75	1.0
Exp	MyStyle_P	0.673	0.612	0.582	0.611	300
	MyStyle_I	0.361	0.602	0.355	0.690	0.677
	Ours	**0.296**	**0.415**	**0.284**	**0.405**	**0.414**
Yaw	MyStyle_P	5.287	4.799	5.436	6.135	6.131
	MyStyle_I	4.629	4.886	3.414	6.090	4.444
	Ours	**1.618**	**2.187**	**2.843**	**2.443**	**1.657**
Pitch	MyStyle_P	5.012	5.005	2.566	3.089	4.156
	MyStyle_I	-	-	-	-	-
	Ours	**3.513**	**2.189**	**2.304**	**2.395**	**2.261**

		Emma Watson				
		0.0	0.25	0.5	0.75	1.0
Exp	MyStyle_P	0.899	0.738	0.883	0.786	0.482
	MyStyle_I	0.273	0.0.583	0.777	0.313	0.002
	Ours	**0.005**	**0.004**	**0.531**	**0.213**	**0.001**
Yaw	MyStyle_P	4.353	5.834	5.156	4.268	3.948
	MyStyle_I	3.277	3.146	1.997	3.133	3.080
	Ours	**1.745**	**2.495**	**1.431**	**2.339**	**2.127**
Pitch	MyStyle_P	6.645	6.094	4.485	2.988	1.808
	MyStyle_I	-	-	-	-	-
	Ours	**4.887**	**2.714**	**1.758**	**2.374**	**1.426**
Age	MyStyle_P	2.760	2.016	0.899	1.041	1.062
	MyStyle_I	3.775	2.502	0.873	1.171	1.920
	Ours	**1.734**	**1.405**	**0.607**	**1.001**	**0.910**

Table 3: We compare our results against MyStyle in terms of the ID metric [Nitzan et al. 2022] and diversity score [Ojha et al. 2021]. Higher numbers are better. Our method produces similar results compared to MyStyles, which demonstrates that controllability does not hurt our system.

		ID ↑	Diversity ↑
Barack Obama	MyStyle	0.787±0.053	0.393±0.070
	Ours	0.796±0.050	0.411±0.026
Emma Watson	MyStyle	0.734±0.090	0.453±0.023
	Ours	0.742±0.061	0.449±0.018
Joe Biden	MyStyle	0.759±0.004	0.397±0.023
	Ours	0.761±0.001	0.402±0.044
Michelle Obama	MyStyle	0.752±0.003	0.391±0.039
	Ours	0.757±0.004	0.389±0.024
Oprah Winfrey	MyStyle	0.807±0.053	0.380±0.038
	Ours	0.821±0.058	0.380±0.057
Taylor Swift	MyStyle	0.775±0.065	0.393±0.049
	Ours	0.774±0.067	0.403±0.019

Table 4: We compare our editing results against MyStyle_I and MyStyle_P in terms of the mean standard deviation (STD) of the edited attribute to show editing consistency (marked with ∗), and of fixed attributes to demonstrate attribute disentanglement. We additionally report the ID metric to evaluate identity preservation ability. The best results are shown in bold.

		Michelle Obama			
		Exp*	Yaw	Pitch	ID ↑
Exp	MyStyle_P	0.547	0.716	4.436	0.786±0.057
	MyStyle_I	0.445	1.069	0.715	0.757±0.065
	Ours	**0.306**	**0.576**	**0.596**	**0.794±0.034**
		Yaw*	Exp	Pitch	ID ↑
Yaw	MyStyle_P	2.773	0.477	2.419	0.780±0.053
	MyStyle_I	4.424	0.510	2.181	0.715±0.110
	Ours	**0.876**	**0.298**	**1.137**	**0.792±0.047**
		Pitch*	Exp	Yaw	ID ↑
Pitch	MyStyle_P	6.306	0.463	2.242	0.727±0.047
	MyStyle_I	-	-	-	-
	Ours	**2.045**	**0.231**	**1.311**	**0.731±0.110**

		Leonardo DiCaprio				
		Exp*	Yaw	Pitch	Age	ID ↑
Exp	MyStyle_P	0.794	4.866	1.965	3.584	0.743±0.103
	MyStyle_I	0.527	1.538	1.227	3.179	0.731±0.108
	Ours	**0.268**	**1.204**	**0.999**	**2.201**	**0.752±0.107**
		Yaw*	Exp	Pitch	Age	ID ↑
Yaw	MyStyle_P	4.069	0.213	2.570	2.893	0.717±0.108
	MyStyle_I	3.925	0.111	2.426	2.700	0.716±0.117
	Ours	**2.097**	**0.075**	**2.108**	**2.212**	**0.728±0.115**
		Pitch*	Exp	Yaw	Age	ID ↑
Pitch	MyStyle_P	5.463	0.281	3.030	3.720	0.717±0.121
	MyStyle_I	-	-	-	-	-
	Ours	**3.591**	**0.071**	**1.786**	**3.023**	**0.726±0.114**
		Age*	Exp	Yaw	Pitch	ID ↑
Age	MyStyle_P	5.113	0.230	2.808	2.824	0.734±0.118
	MyStyle_I	7.152	0.134	1.294	2.095	0.723±0.120
	Ours	**3.473**	**0.087**	**0.467**	**1.217**	**0.739±0.113**

Table 5: We compare our editing results against MyStyle_I and MyStyle_P in terms of the mean standard deviation (STD) of edited attribute to show editing consistency (marked with ∗), and of fixed attributes to demonstrate attribute disentanglement. We additionally report the ID metric to evaluate identity preservation ability. The best results are shown in bold.

		Barack Obama			
		Exp*	Yaw	Pitch	ID ↑
Exp	MyStyle_P	0.156	5.502	2.265	0.700±0.087
	MyStyle_I	0.261	1.751	1.808	0.703±0.086
	Ours	**0.107**	**0.533**	**0.804**	**0.742±0.079**
		Yaw*	Exp	Pitch	ID ↑
Yaw	MyStyle_P	4.310	0.054	5.727	0.689±0.085
	MyStyle_I	3.117	0.043	3.770	0.687±0.087
	Ours	**1.524**	**0.028**	**2.349**	**0.709±0.083**
		Pitch*	Exp	Yaw	ID ↑
Pitch	MyStyle_P	5.131	0.031	6.285	0.692±0.087
	MyStyle_I	-	-	-	-
	Ours	**2.990**	**0.026**	**1.055**	**0.709±0.084**

		Joe Biden			
		Exp*	Yaw	Pitch	ID ↑
Exp	MyStyle_P	0.776	1.809	5.619	0.654±0.091
	MyStyle_I	0.446	1.707	1.054	0.693±0.072
	Ours	**0.332**	**1.095**	**1.254**	**0.703±0.076**
		Yaw*	Exp	Pitch	ID ↑
Yaw	MyStyle_P	10.42	0.739	10.50	0.666±0.084
	MyStyle_I	9.190	0.789	2.341	0.676±0.077
	Ours	**3.425**	**0.625**	**2.205**	**0.705±0.072**
		Pitch*	Exp	Yaw	ID ↑
Pitch	MyStyle_P	9.262	0.786	18.06	0.637±0.111
	MyStyle_I	-	-	-	-
	Ours	**3.958**	**0.603**	**2.656**	**0.693±0.080**

		Oprah Winfrey			
		Exp*	Yaw	Pitch	ID ↑
Exp	MyStyle_P	0.303	4.358	3.335	0.685±0.100
	MyStyle_I	0.140	1.738	1.523	0.690±0.092
	Ours	**0.116**	**0.531**	**0.401**	**0.712±0.088**
		Yaw*	Exp	Pitch	ID ↑
Yaw	MyStyle_P	6.338	0.052	5.284	0.695±0.075
	MyStyle_I	5.242	0.039	3.250	0.698±0.074
	Ours	**2.330**	**0.027**	**2.425**	**0.702±0.075**
		Pitch*	Exp	Yaw	ID ↑
Pitch	MyStyle_P	4.949	0.053	3.423	0.657±0.104
	MyStyle_I	-	-	-	-
	Ours	**2.001**	**0.031**	**1.445**	**0.698±0.070**

Table 6: We compare our editing results against MyStyle_I and MyStyle_P in terms of the mean standard deviation (STD) of edited attribute to show editing consistency (marked with $*$), and of fixed attributes to demonstrate attribute disentanglement. We additionally report the ID metric to evaluate identity preservation ability. The best results are shown in bold.

		Scarlett Johansson			
		Exp*	Yaw	Pitch	ID ↑
Exp	MyStyle_P	0.866	2.578	3.942	0.666±0.075
	MyStyle_I	0.328	1.131	0.577	0.660±0.090
	Ours	**0.270**	**0.607**	**0.510**	**0.675±0.074**
		Yaw*	Exp	Pitch	ID ↑
Yaw	MyStyle_P	9.216	0.869	1.671	0.686 ±0.078
	MyStyle_I	4.655	0.859	2.270	0.645±0.082
	Ours	**2.349**	**0.774**	**1.669**	**0.707±0.069**
		Pitch*	Exp	Yaw	ID ↑
Pitch	MyStyle_P	5.639	0.924	5.755	0.667±0.077
	MyStyle_I	-	-	-	-
	Ours	**4.130**	**0.793**	**1.316**	**0.684±0.076**

		Taylor Swift			
		Exp*	Yaw	Pitch	ID ↑
Exp	MyStyle_P	0.783	3.421	3.933	0.698±0.077
	MyStyle_I	0.314	2.444	1.237	0.683±0.079
	Ours	**0.230**	**1.532**	**0.783**	**0.713±0.077**
		Yaw*	Exp	Pitch	ID ↑
Yaw	MyStyle_P	7.343	0.783	3.210	0.688±0.087
	MyStyle_I	3.598	0.389	2.333	0.700±0.074
	Ours	**1.925**	**0.241**	**1.982**	**0.720±0.069**
		Pitch*	Exp	Yaw	ID ↑
Pitch	MyStyle_P	4.189	0.102	5.283	0.703±0.078
	MyStyle_I	-	-	-	-
	Ours	**2.374**	**0.041**	**2.338**	**0.721±0.079**

		Emma Watson				
		Exp*	Yaw	Pitch	Age	ID ↑
Exp	MyStyle_P	0.821	3.892	2.721	4.021	0.698±0.113
	MyStyle_I	0.482	1.732	2.034	3.248	0.710±0.096
	Ours	**0.240**	**1.330**	**1.218**	**2.017**	**0.748±0.094**
		Yaw*	Exp	Pitch	Age	ID ↑
Yaw	MyStyle_P	5.478	0.432	3.258	3.483	0.728±0.090
	MyStyle_I	4.130	0.384	2.832	3.003	0.713±0.094
	Ours	**2.783**	**0.104**	**2.203**	**2.438**	**0.733±0.092**
		Pitch*	Exp	Yaw	Age	ID ↑
Pitch	MyStyle_P	4.389	0.492	3.189	3.899	0.728±0.096
	MyStyle_I	-	-	-	-	-
	Ours	**3.294**	**0.085**	**1.839**	**2.893**	**0.746±0.094**
		Age*	Exp	Yaw	Pitch	ID ↑
Age	MyStyle_P	4.883	0.432	3.543	2.983	0.717±0.086
	MyStyle_I	4.238	0.233	2.899	2.339	0.725±0.114
	Ours	**2.938**	**0.183**	**1.032**	**1.384**	**0.742±0.082**

REFERENCES

Yotam Nitzan, Kfir Aberman, Qiurui He, Orly Liba, Michal Yarom, Yossi Gandelsman, Inbar Mosseri, Yael Pritch, and Daniel Cohen-Or. 2022. MyStyle: A Personalized Generative Prior. *arXiv preprint arXiv:2203.17272* (2022).

Utkarsh Ojha, Yijun Li, Jingwan Lu, Alexei A Efros, Yong Jae Lee, Eli Shechtman, and Richard Zhang. 2021. Few-shot image generation via cross-domain correspondence. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 10743–10752.

Yujun Shen, Ceyuan Yang, Xiaoou Tang, and Bolei Zhou. 2020. InterFaceGAN: Interpreting the Disentangled Face Representation Learned by GANs. *TPAMI* (2020).