Domain Adaptation of GANs

Aibek Alanov

Research scientist at Artificial Intelligence Research Institute (AIRI) and
Centre of Deep Learning and Bayesian Methods HSE University

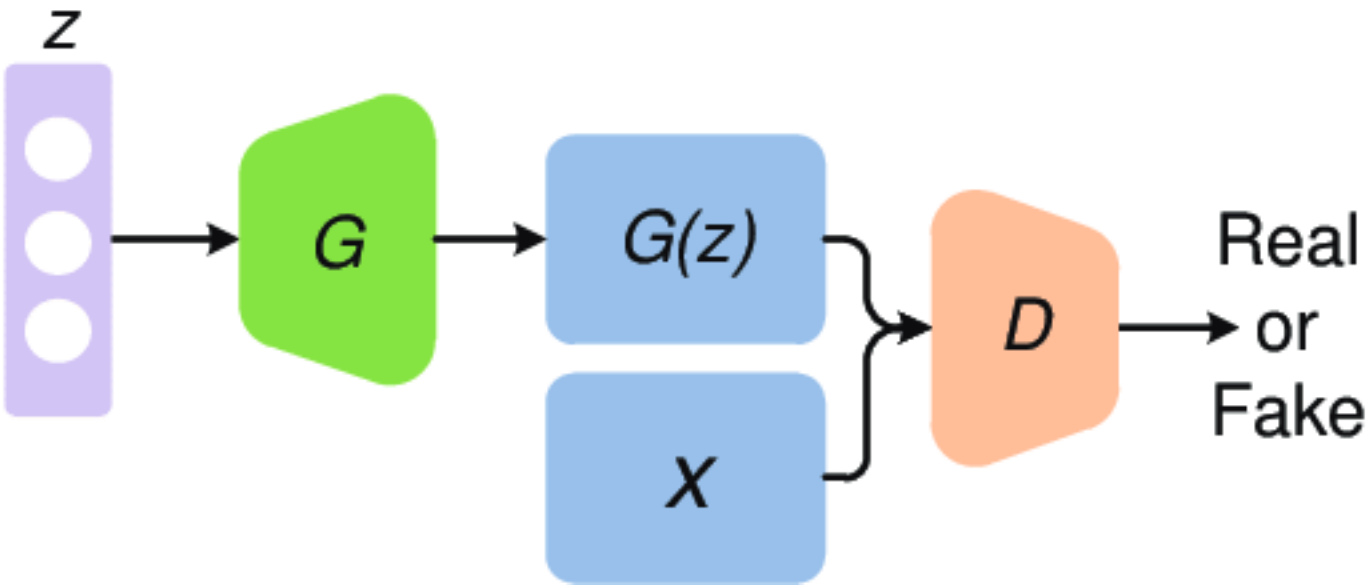
Our papers

“HyperDomainNet: Universal Domain Adaptation for Generative Adversarial Networks”, *Aibek Alanov*, Vadim Titov*, Dmitry Vetrov* (accepted to NeurIPS 2022)

“StyleDomain: Analysis of StyleSpace for Domain Adaptation of StyleGAN”,

Aibek Alanov, Vadim Titov*, Maksim Nakhodnov*, Dmitry Vetrov* (submitted to top-tier conference)

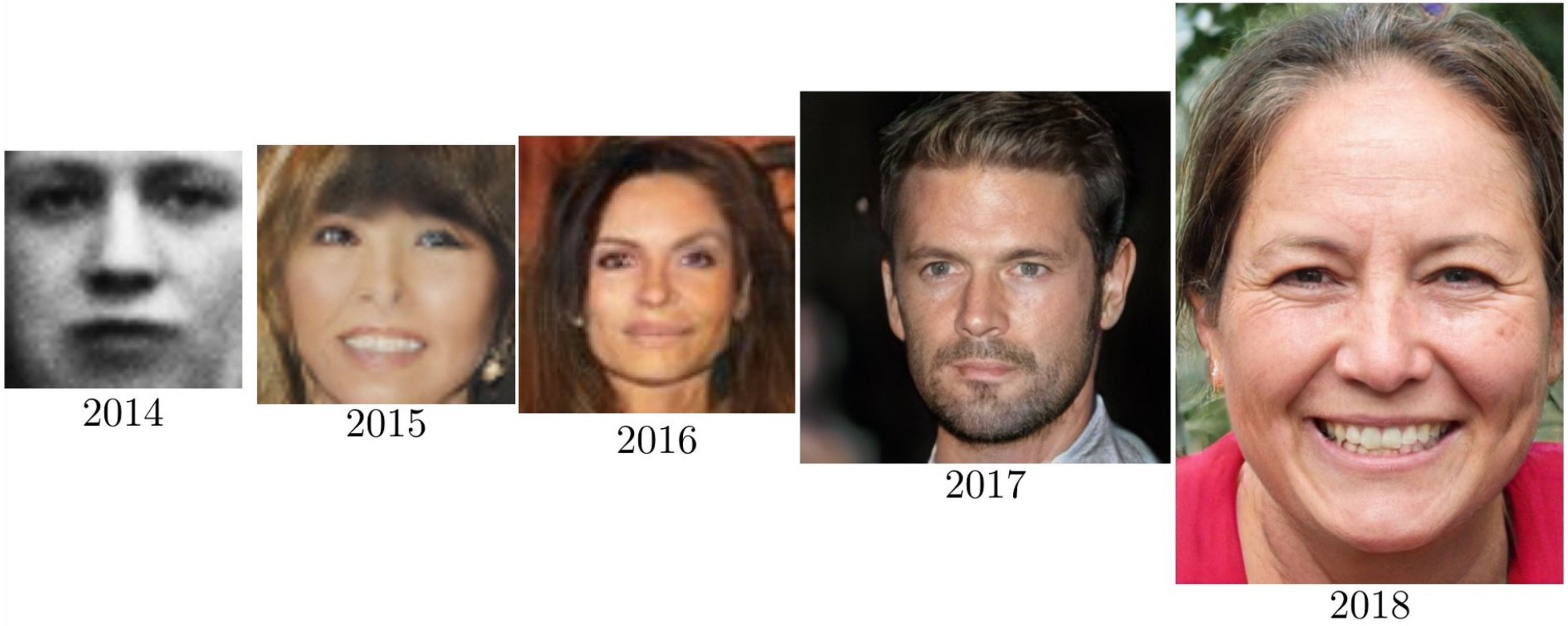
GAN framework



$$\min_G \max_D V(D, G) = \mathbb{E}_{\boldsymbol{x} \sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z} \sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))].$$

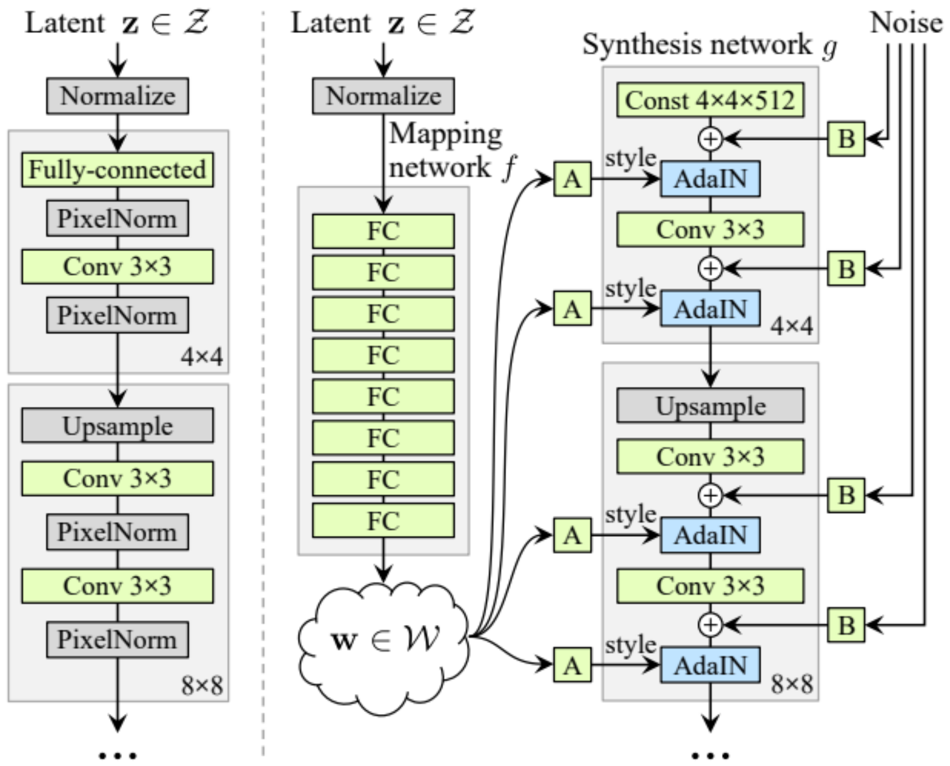
Pan et al. Recent Progress on Generative Adversarial Networks (GANs): A Survey. 2019
Goodfellow et al. Generative Adversarial Networks. NeurIPS 2014

Progress of GANs



Goodfellow et al. https://twitter.com/goodfellow_ian/status/1084973596236144640?s=20&t=xsPREVyAUQu6LSvY_Hzw7A. Twitter 2019

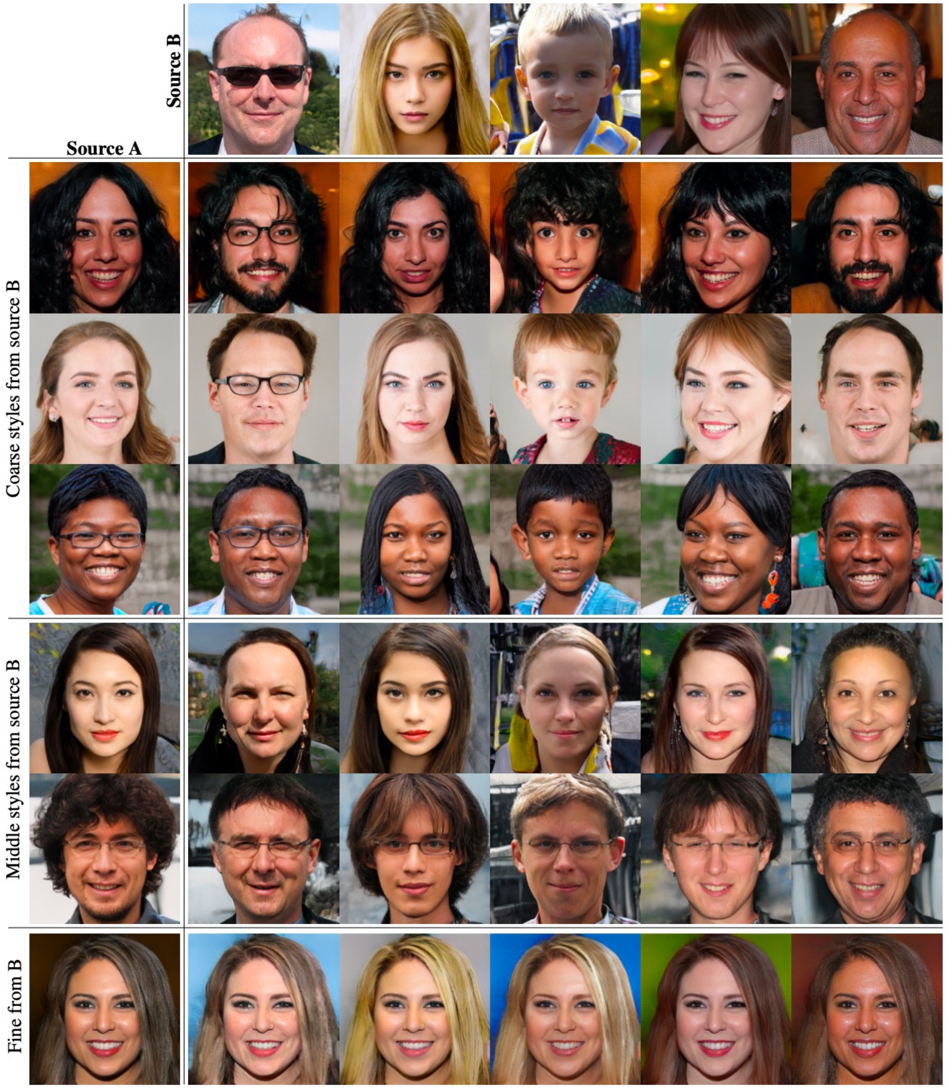
StyleGAN



(a) Traditional (b) Style-based generator

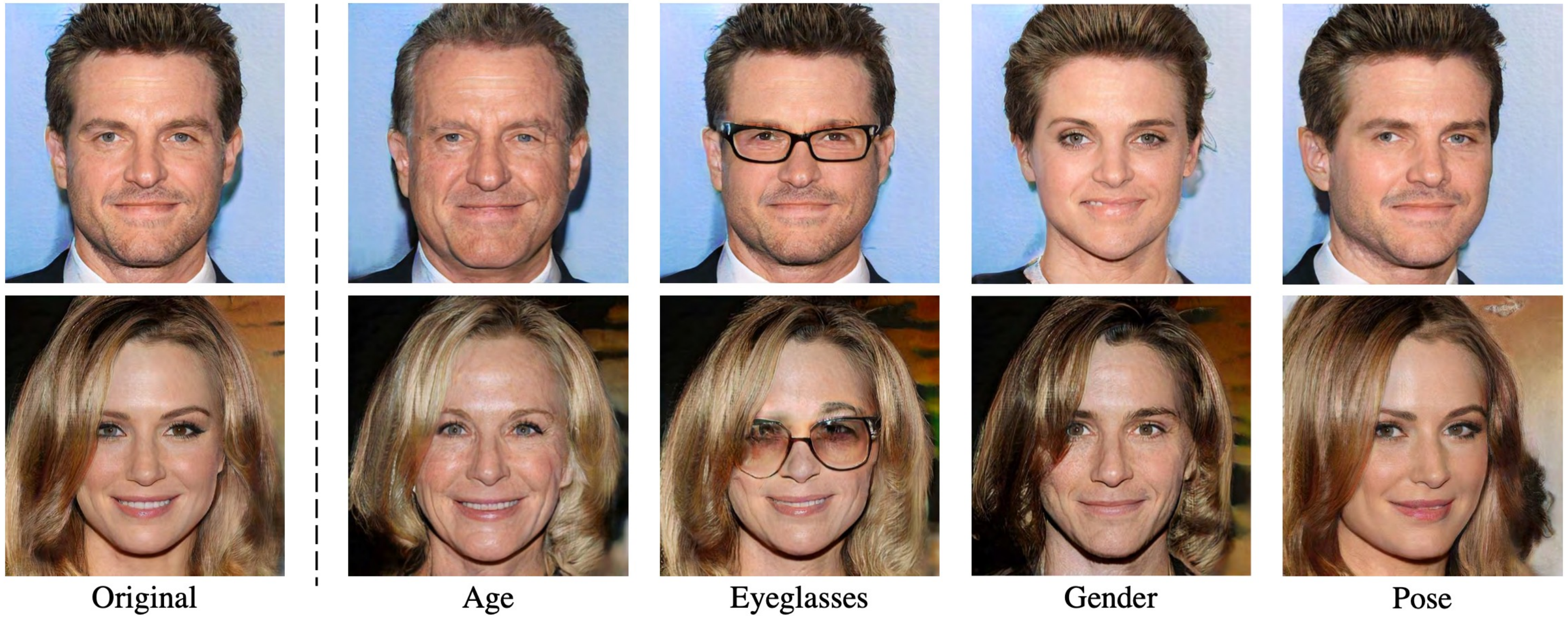
Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 4401–4410, 2019.

Properties of StyleGAN



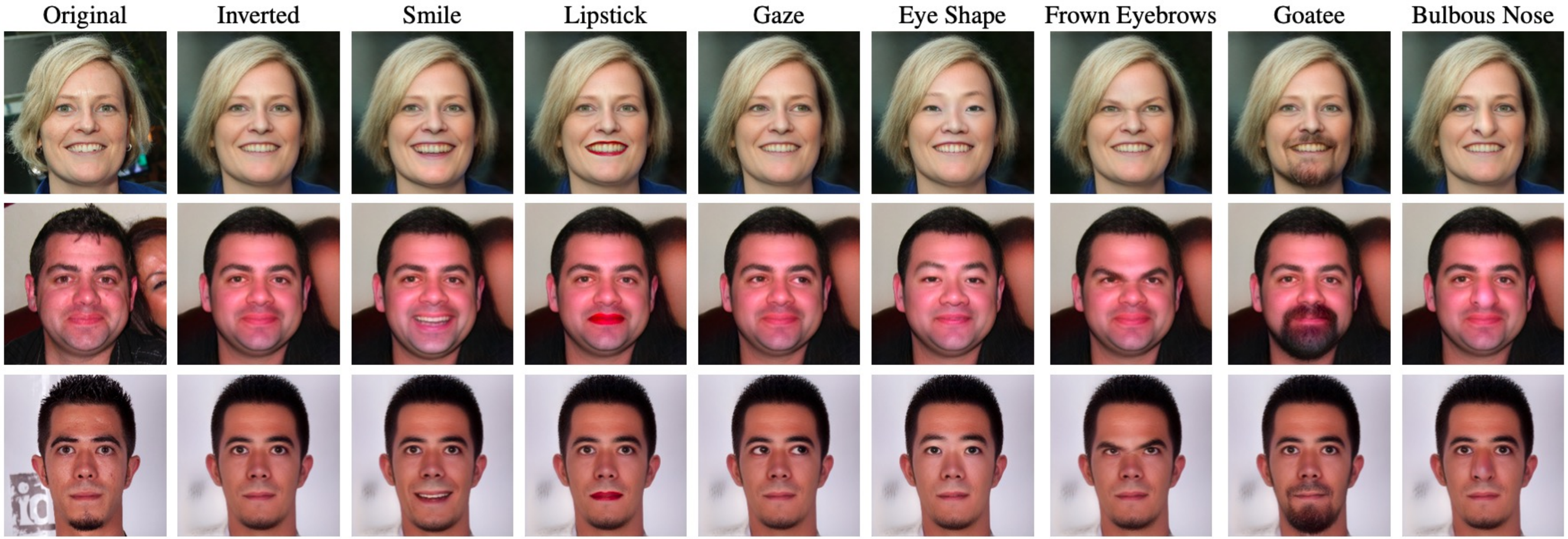
Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 4401–4410, 2019.

Semantic editing of generated images



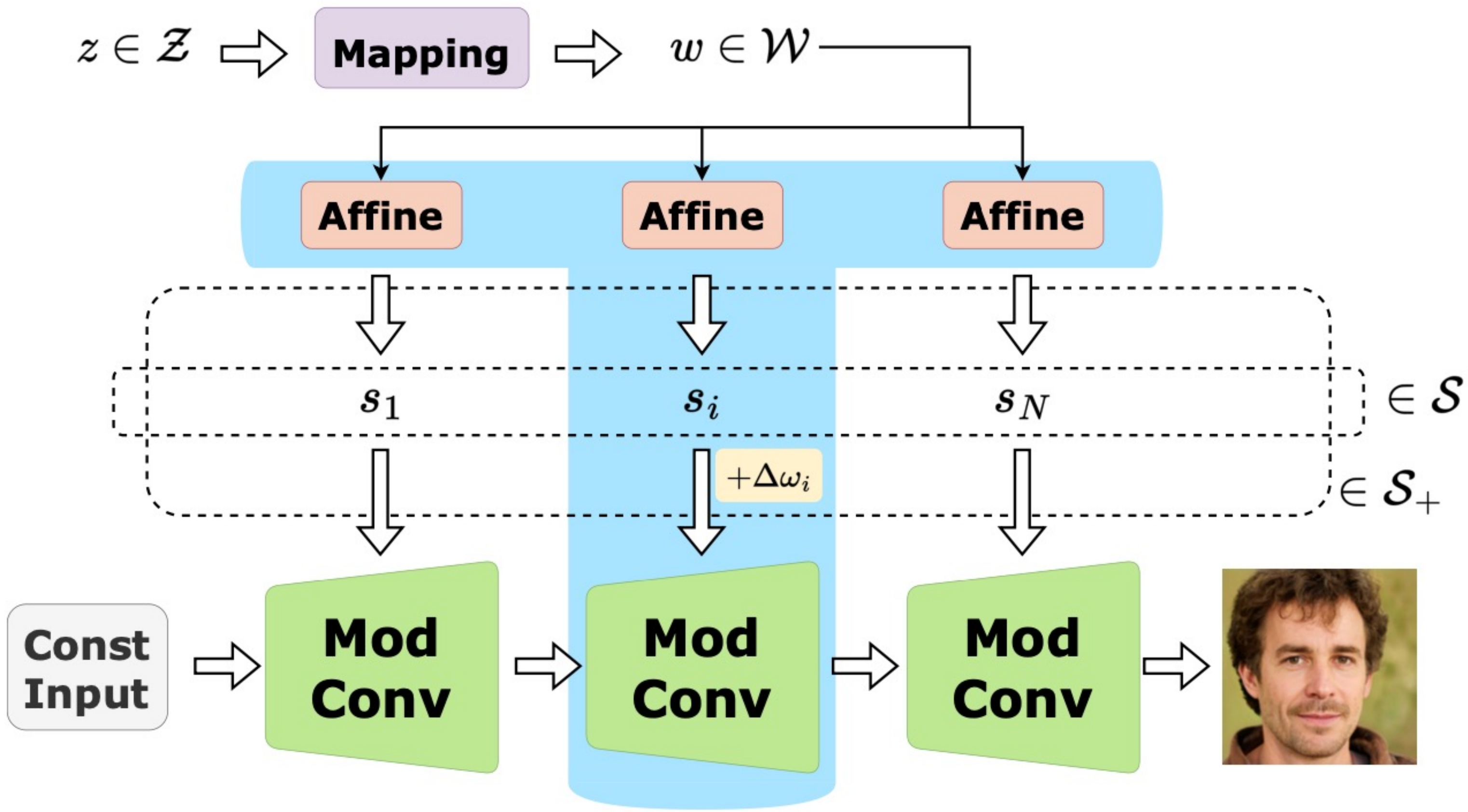
Yujun Shen, Jinjin Gu, Xiaoou Tang, and Bolei Zhou. Inter- preting the latent space of GANs for semantic face editing. In Proc. CVPR, pages 9243–9252, 2020.

Semantic editing of real images

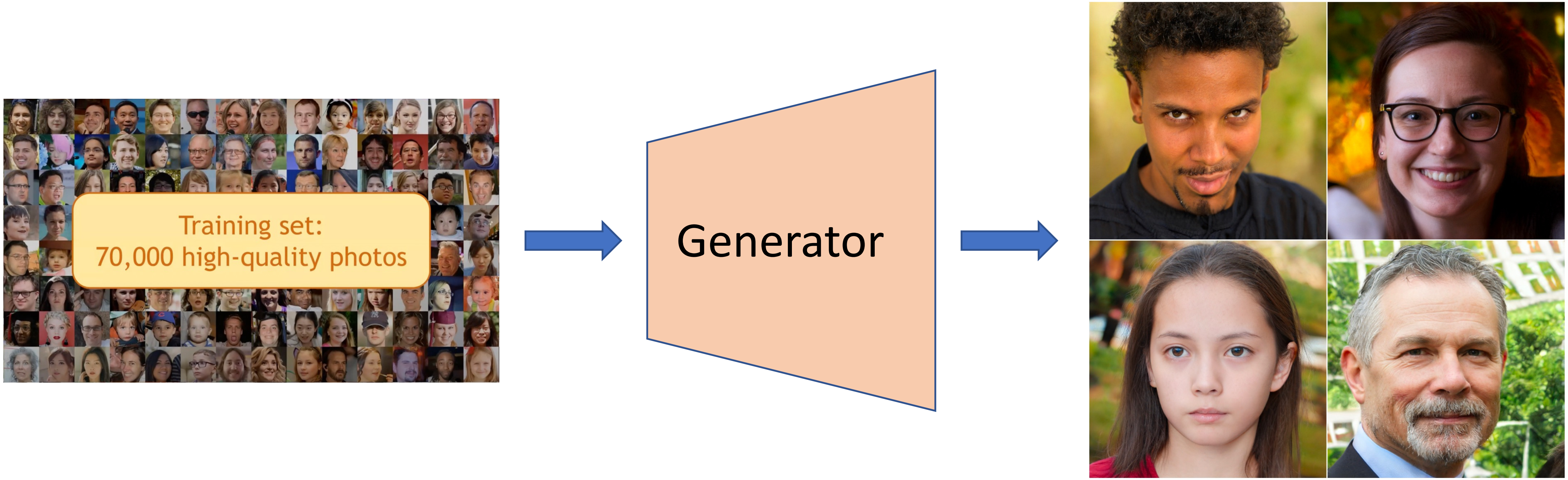


Zongze Wu, Dani Lischinski, and Eli Shechtman. Stylespace analysis: Disentangled controls for stylegan image generation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 12863–12872, 2021.

Latent spaces of StyleGAN



Training of StyleGAN requires lots of data



Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Analyzing and improving the image quality of stylegan. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 8110–8119, 2020.

Few-shot setting

MetFaces ~ 1k painting faces

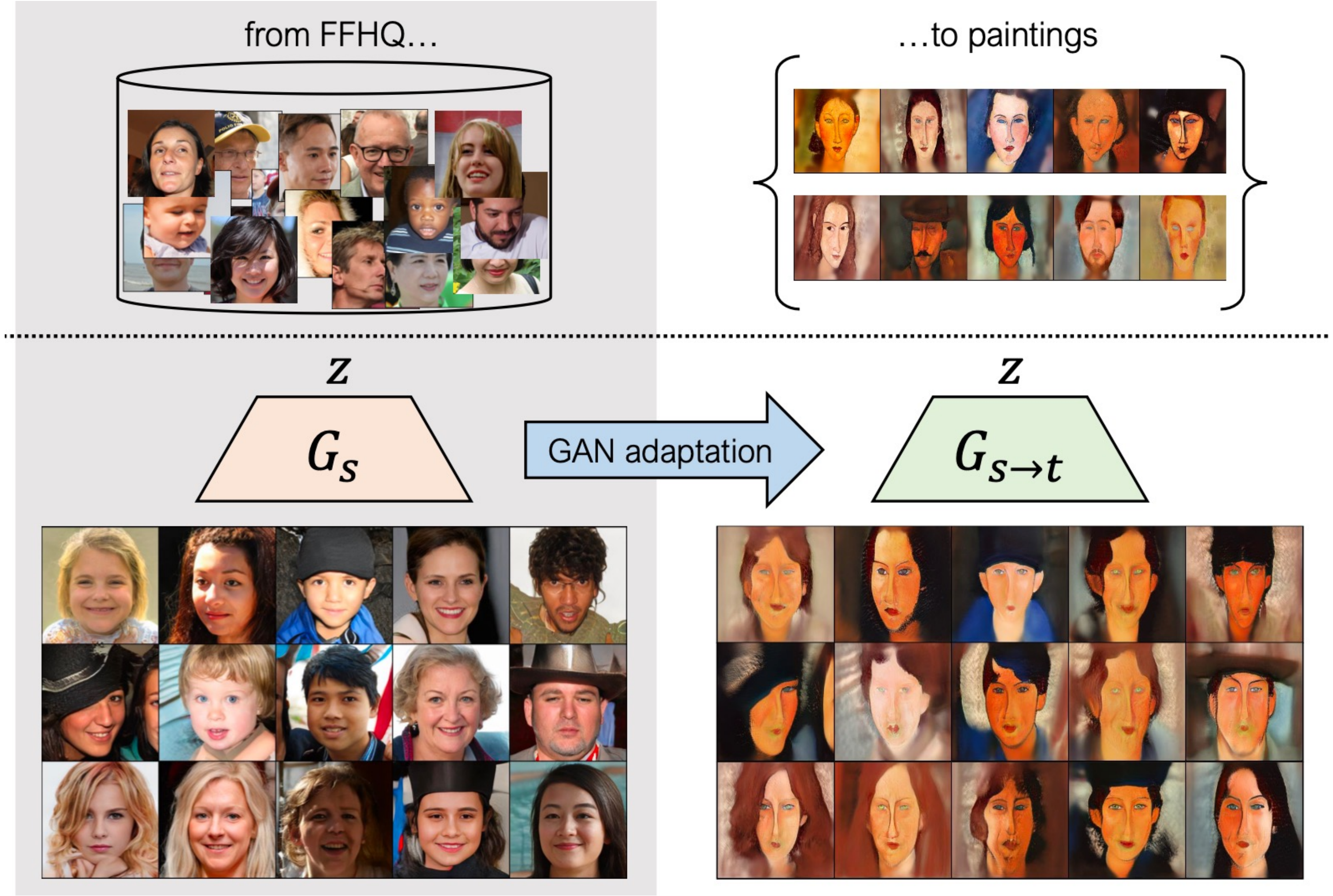
AFHQ ~ 5k cat faces

FaceSketches ~ 300 faces

Tero Karras, Miika Aittala, Janne Hellsten, Samuli Laine, Jaakko Lehtinen, and Timo Aila. Training generative adversarial networks with limited data. Advances in Neural Information Processing Systems, 33:12104–12114, 2020.

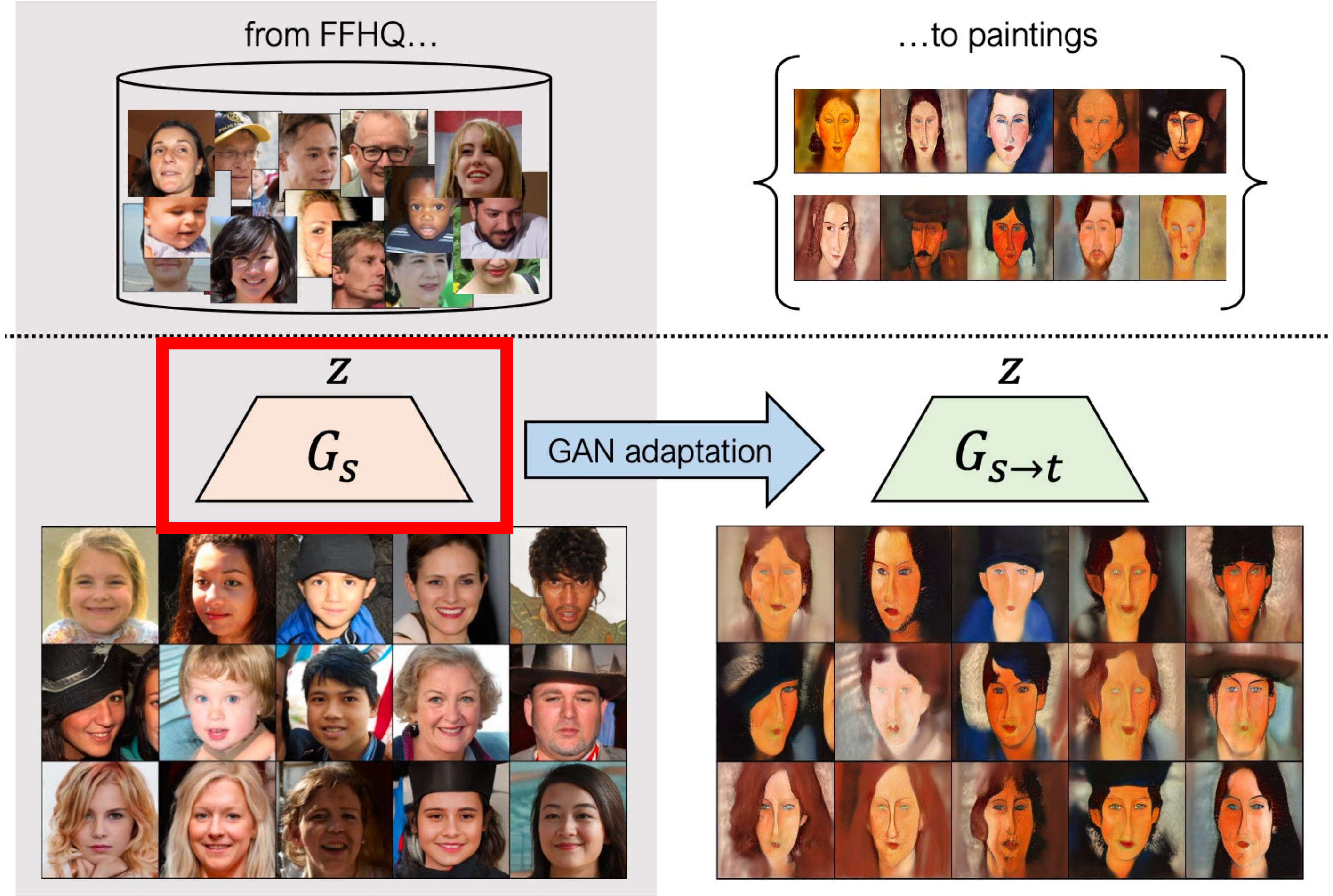
Utkarsh Ojha, Yijun Li, Jingwan Lu, Alexei A Efros, Yong Jae Lee, Eli Shechtman, and Richard Zhang. Few-shot image generation via cross-domain correspondence. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 10743–10752, 2021.

Domain adaptation of GANs



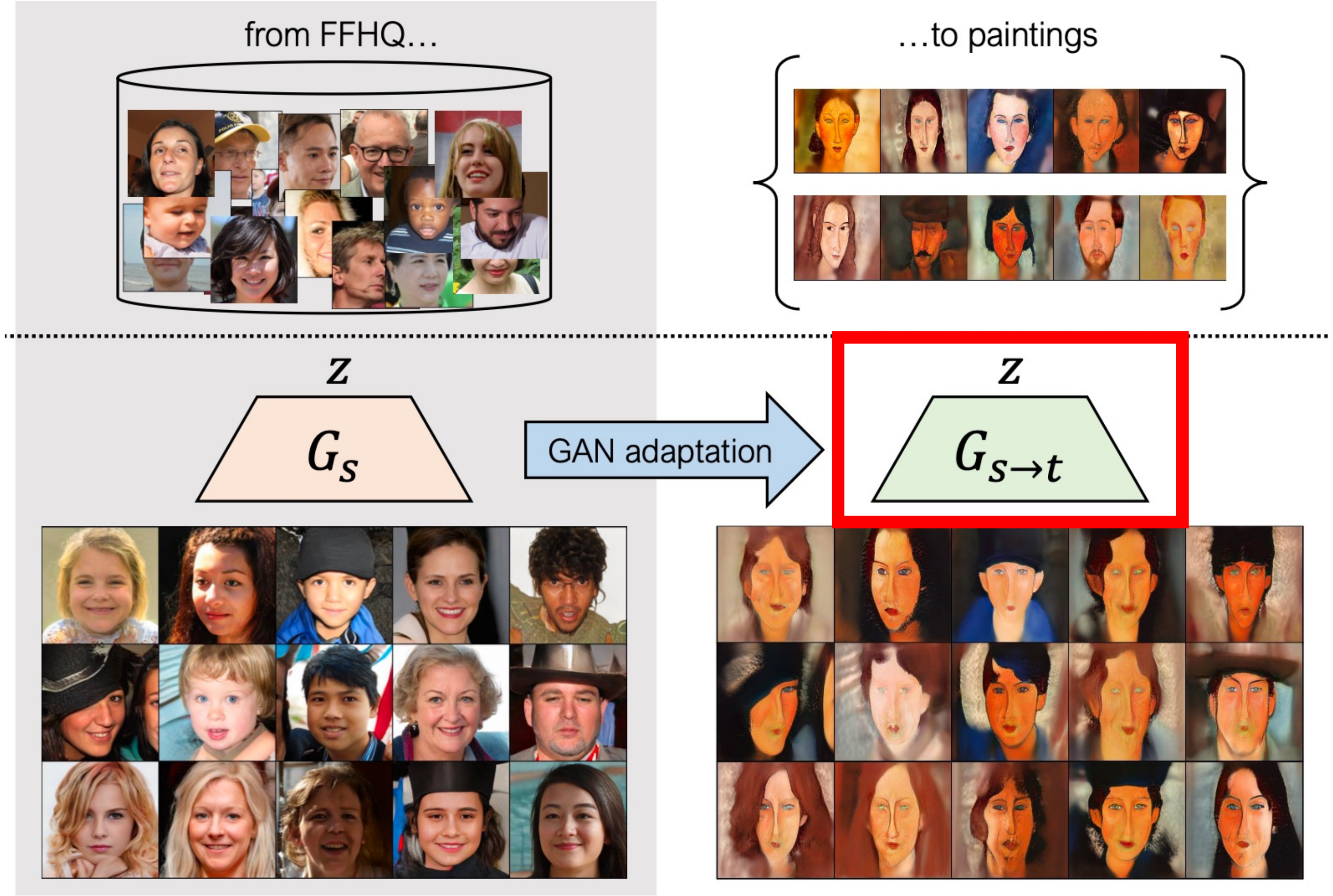
Utkarsh Ojha, Yijun Li, Jingwan Lu, Alexei A Efros, Yong Jae Lee, Eli Shechtman, and Richard Zhang. Few-shot image generation via cross-domain correspondence. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 10743–10752, 2021.

Domain adaptation of GANs



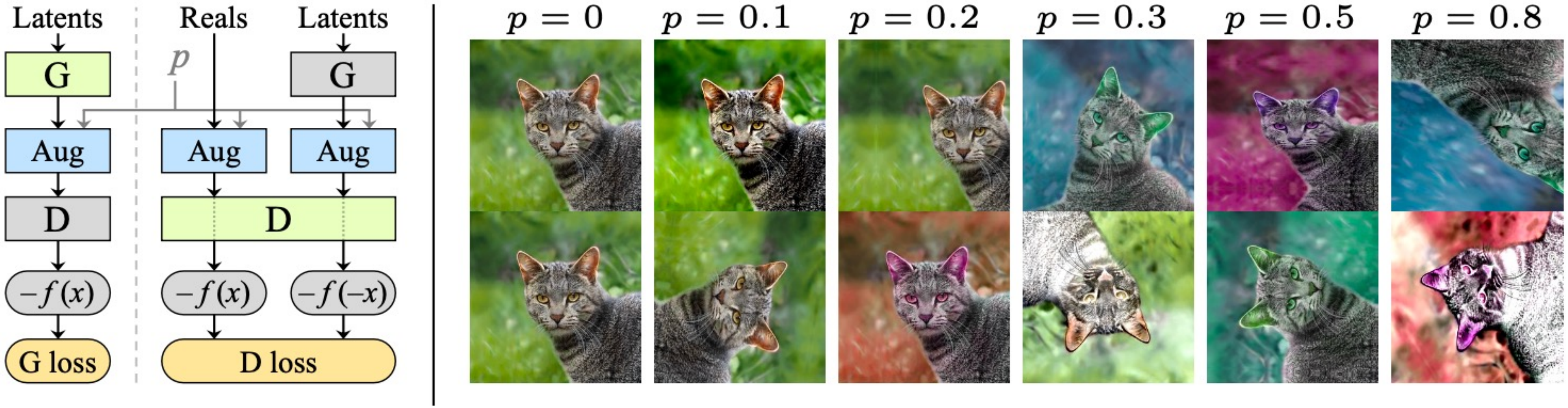
Utkarsh Ojha, Yijun Li, Jingwan Lu, Alexei A Efros, Yong Jae Lee, Eli Shechtman, and Richard Zhang. Few-shot image generation via cross-domain correspondence. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 10743–10752, 2021.

Domain adaptation of GANs



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

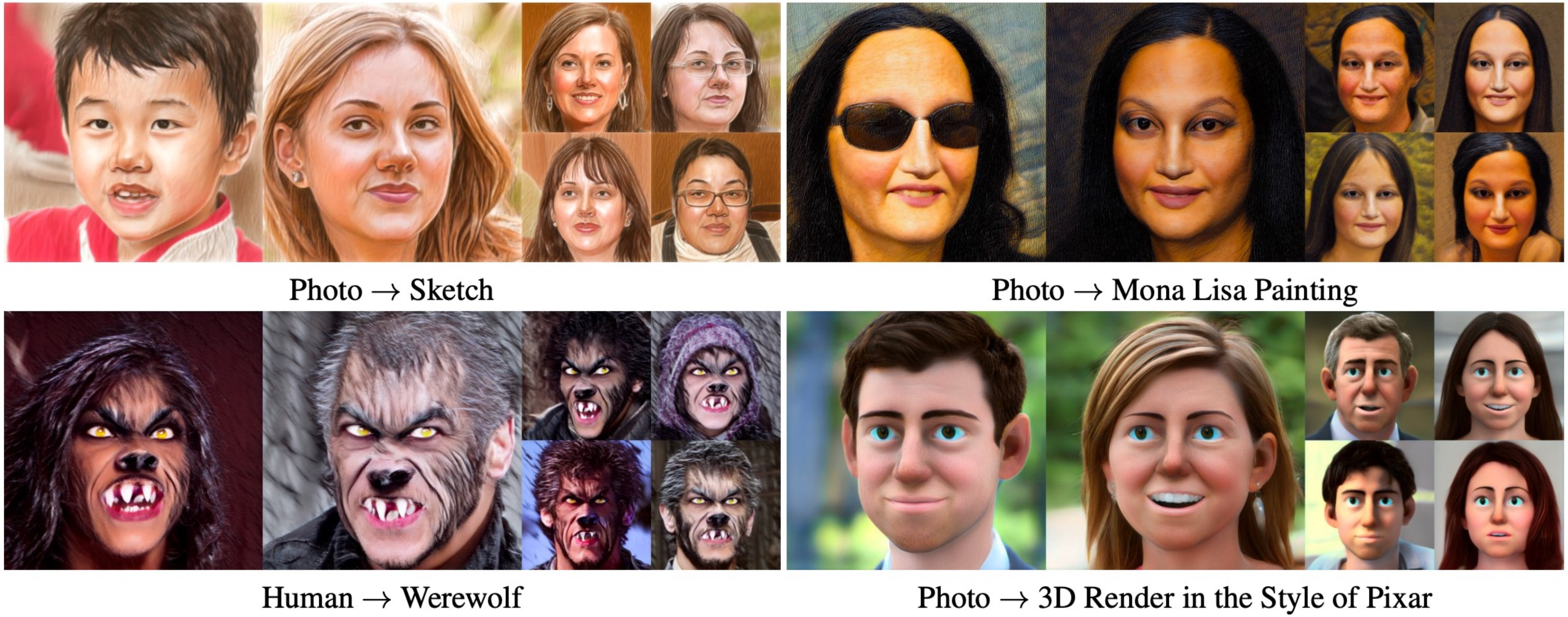
Fine-tuning StyleGAN with augmentations



$$L_D = \mathbb{E}_{\boldsymbol{x} \sim p_{\text{data}}}(\boldsymbol{x})[f_D(-D(T(\boldsymbol{x})))] + \mathbb{E}_{\boldsymbol{z} \sim p(\boldsymbol{z})}[f_D(D(T(G(\boldsymbol{z}))))],$$
$$L_G = \mathbb{E}_{\boldsymbol{z} \sim p(\boldsymbol{z})}[f_G(-D(T(G(\boldsymbol{z}))))].$$

Tero Karras, Miika Aittala, Janne Hellsten, Samuli Laine, Jaakko Lehtinen, and Timo Aila. Training generative adversarial networks with limited data. Advances in Neural Information Processing Systems, 33:12104–12114, 2020.
Shengyu Zhao, Zhijian Liu, Ji Lin, Jun-Yan Zhu, and Song Han. Differentiable augmentation for data- efficient gan training. Advances in Neural Information Processing Systems, 33:7559–7570, 2020.

Text-based domain adaptation methods



Rinon Gal et al. Stylegan-nada: Clip-guided domain adaptation of image generators. arXiv preprint arXiv:2108.00946, 2021.

Text-based domain adaptation methods

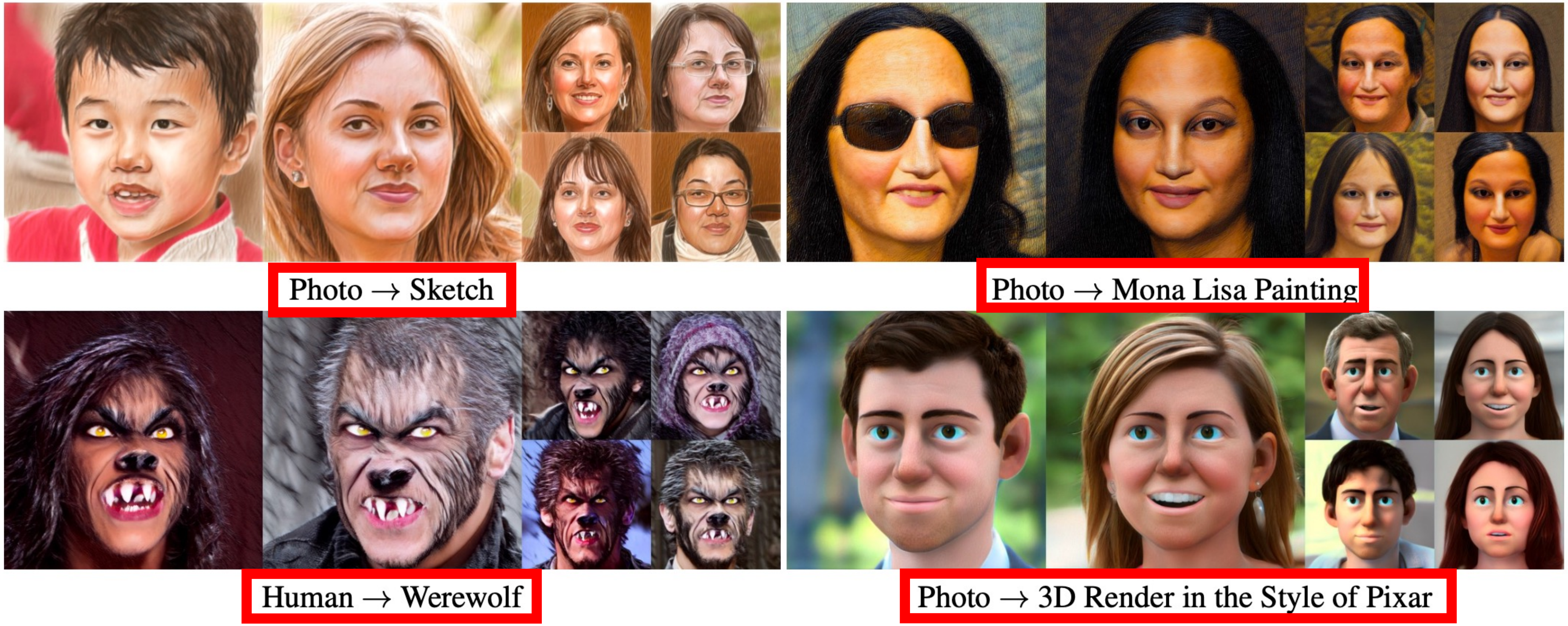
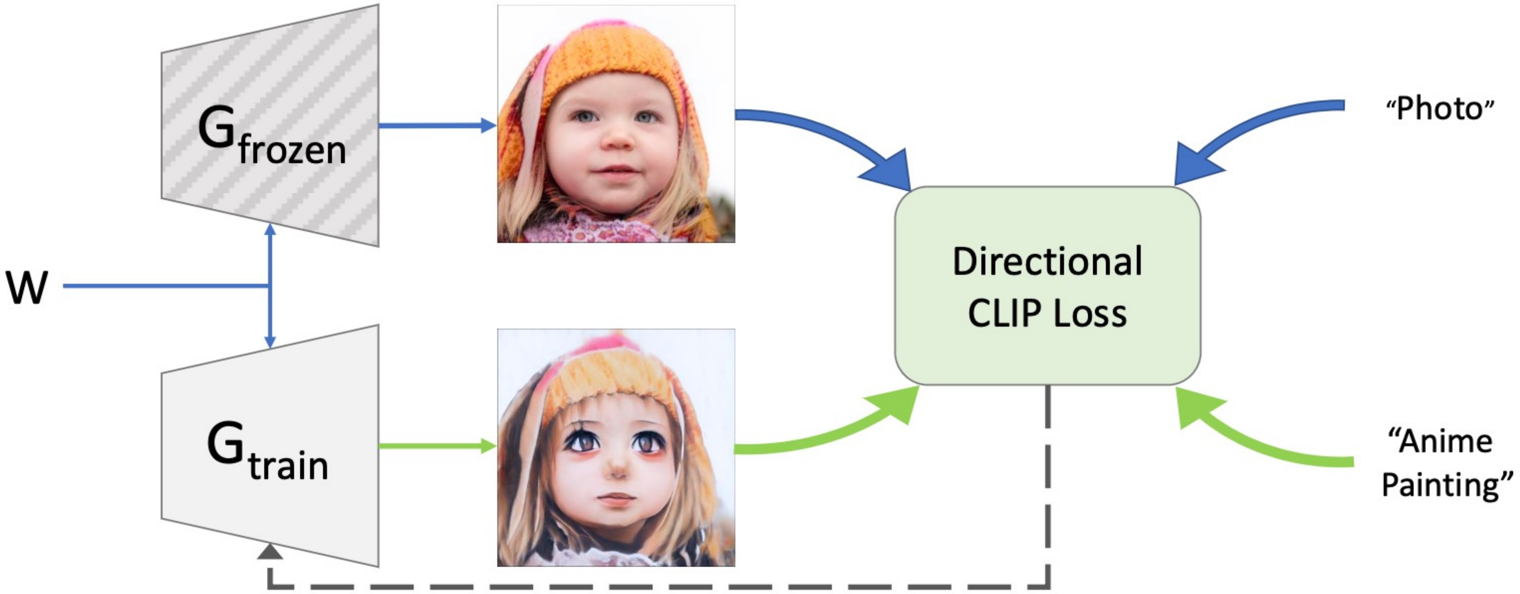


Image credit: StyleGAN-NADA paper [3]

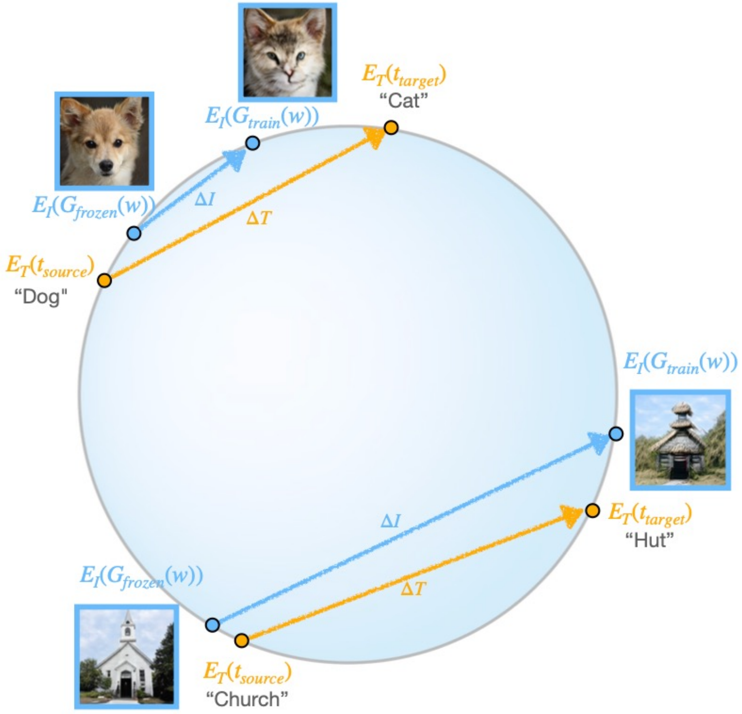
Rinon Gal et al. Stylegan-nada: Clip-guided domain adaptation of image generators. arXiv preprint arXiv:2108.00946, 2021.

CLIP-based domain adaptation loss



Rinon Gal et al. Stylegan-nada: Clip-guided domain adaptation of image generators. arXiv preprint arXiv:2108.00946, 2021.

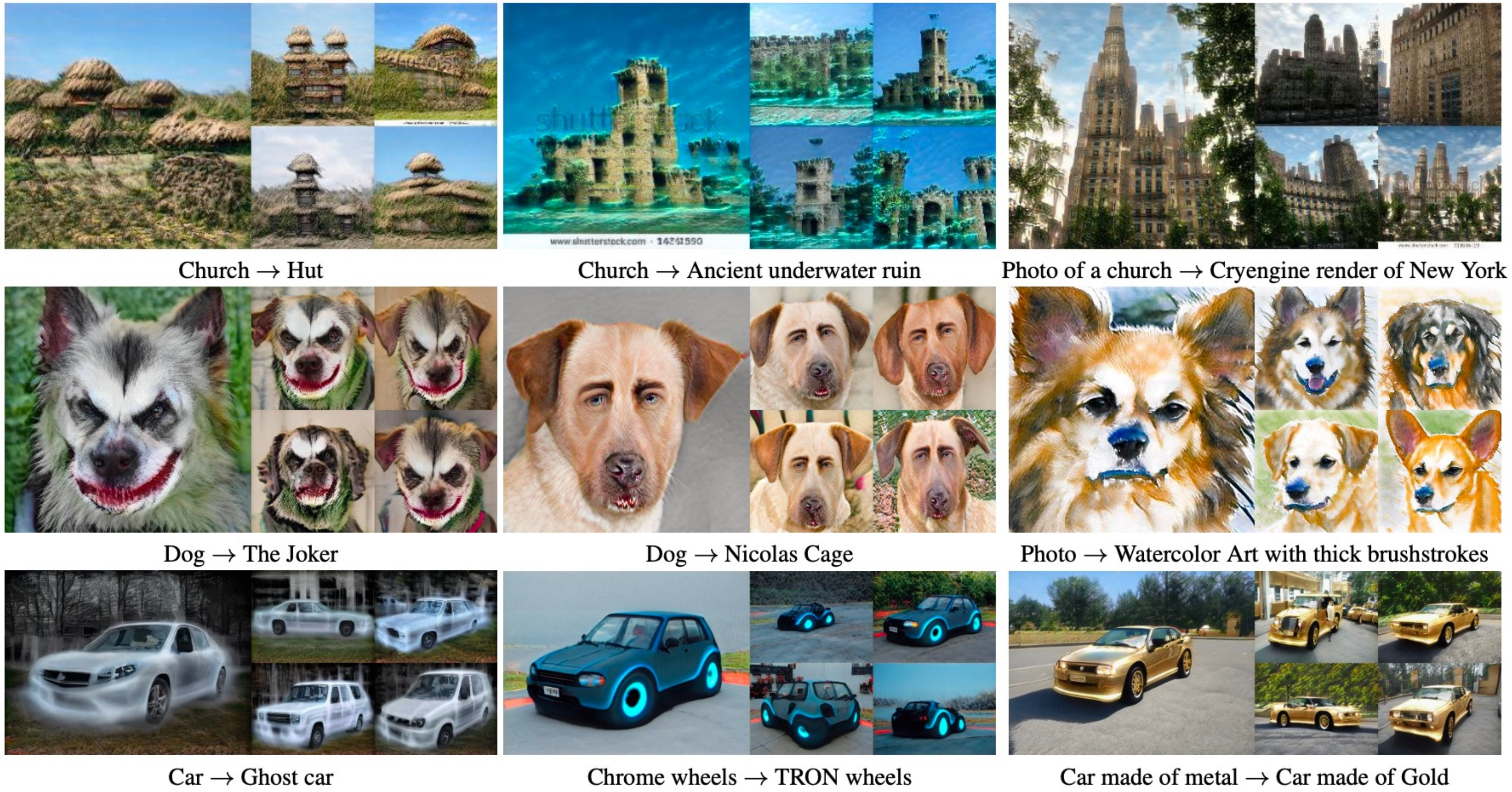
CLIP-based domain adaptation loss



$$\Delta T = E_T\left(t_{target}\right) - E_T\left(t_{source}\right) \ ,$$
$$\Delta I = E_I\left(G_{train}\left(w\right)\right) - E_I\left(G_{frozen}\left(w\right)\right) \ ,$$
$$\mathcal{L}_{direction} = 1 - \frac{\Delta I \cdot \Delta T}{|\Delta I|\,|\Delta T|} \ ,$$

Rinon Gal et al. Stylegan-nada: Clip-guided domain adaptation of image generators. arXiv preprint arXiv:2108.00946, 2021.

More examples

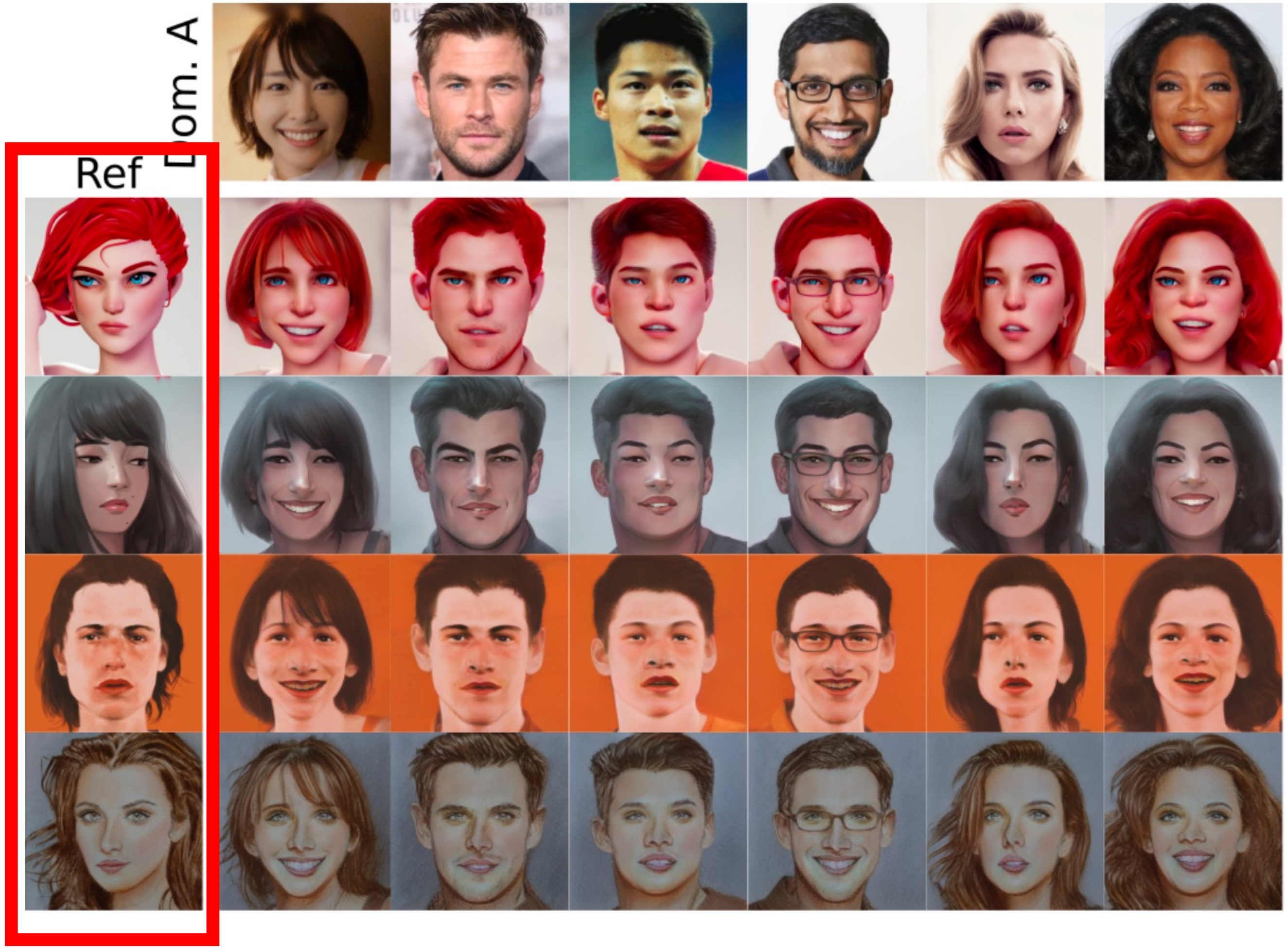


Rinon Gal et al. Stylegan-nada: Clip-guided domain adaptation of image generators. arXiv preprint arXiv:2108.00946, 2021.

One-shot image-based domain adaptation methods

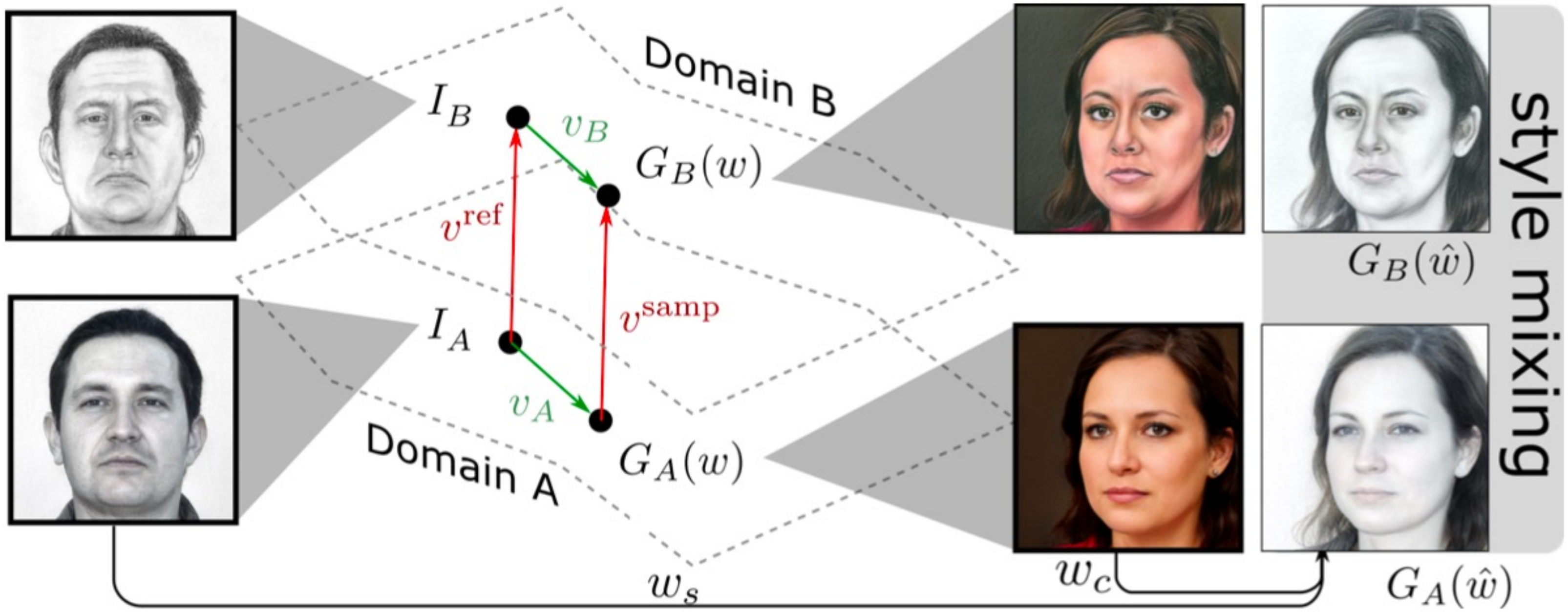
Peihao Zhu et al. Mind the gap: Domain gap control for single shot domain adaptation for generative adversarial networks. arXiv preprint arXiv:2110.08398, 2021.

One-shot image-based domain adaptation methods



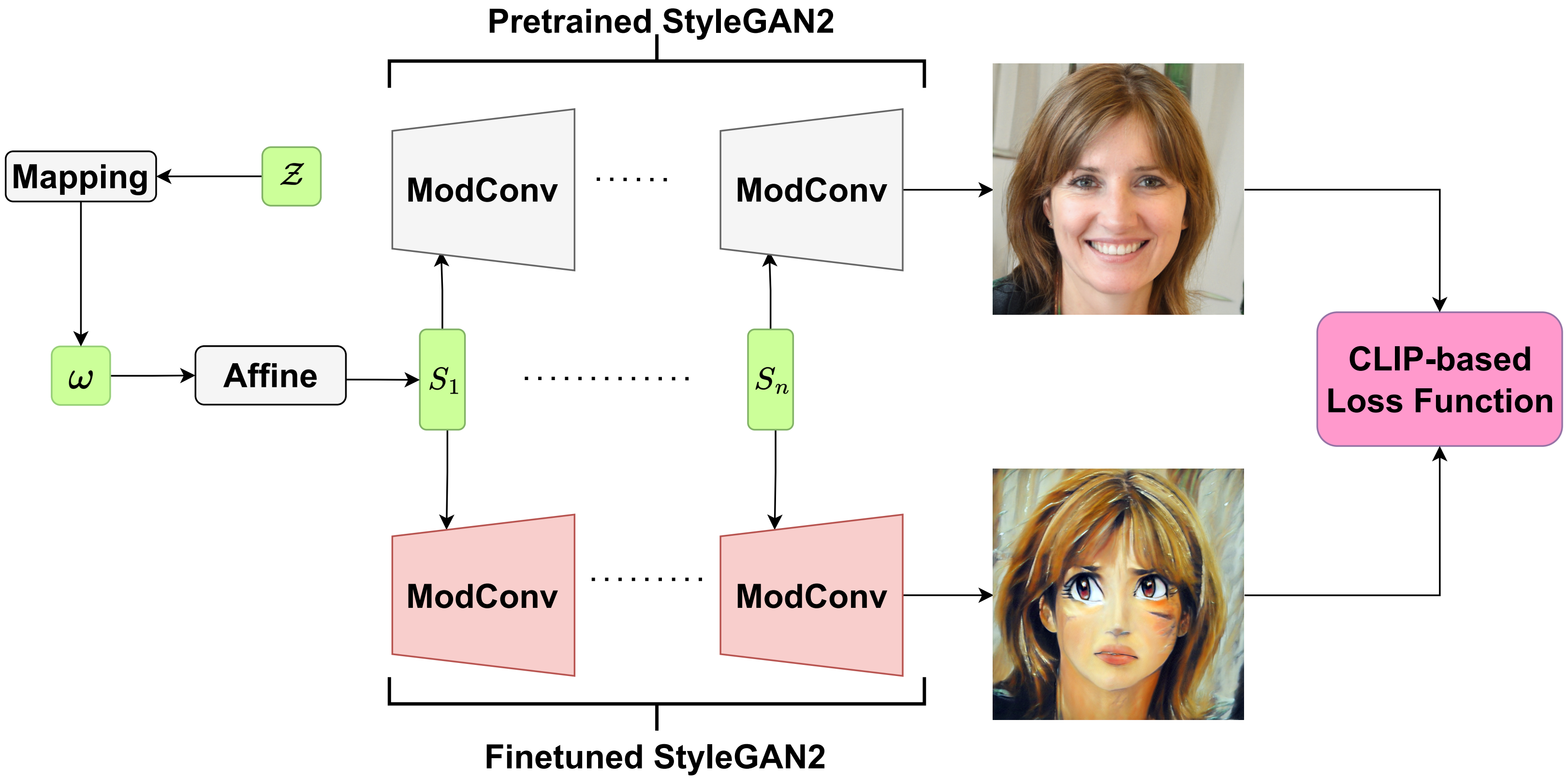
Peihao Zhu et al. Mind the gap: Domain gap control for single shot domain adaptation for generative adversarial networks. arXiv preprint arXiv:2110.08398, 2021.

Loss for one-shot image-based adaptation



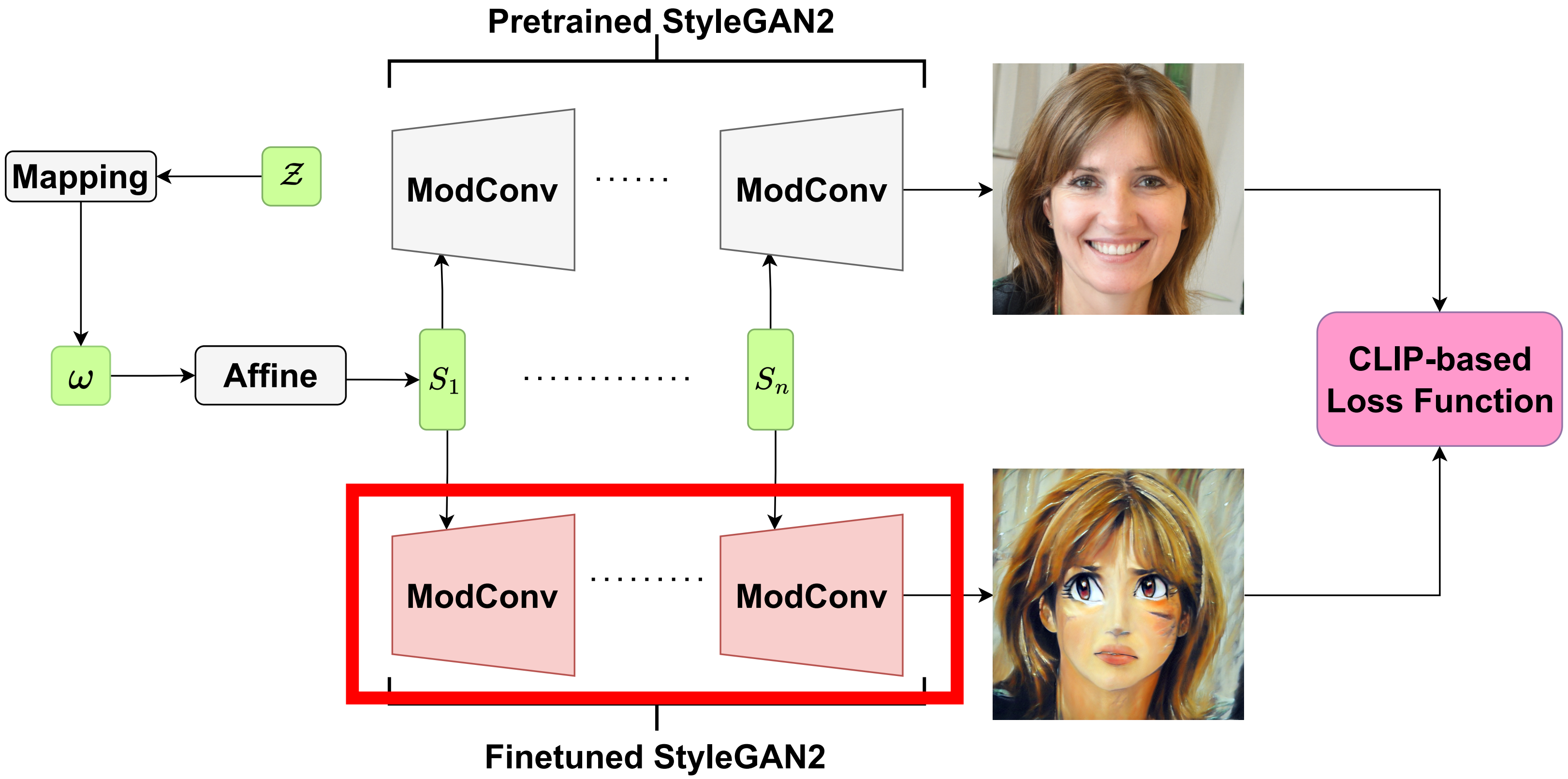
Peihao Zhu et al. Mind the gap: Domain gap control for single shot domain adaptation for generative adversarial networks. arXiv preprint arXiv:2110.08398, 2021.

Typical domain adaptation of StyleGAN2



Aibek Alanov, Vadim Titov, and Dmitry Vetrov. Hyperdomainnet: Universal domain adaptation for generative adversarial networks. arXiv preprint arXiv:2210.08884, 2022.

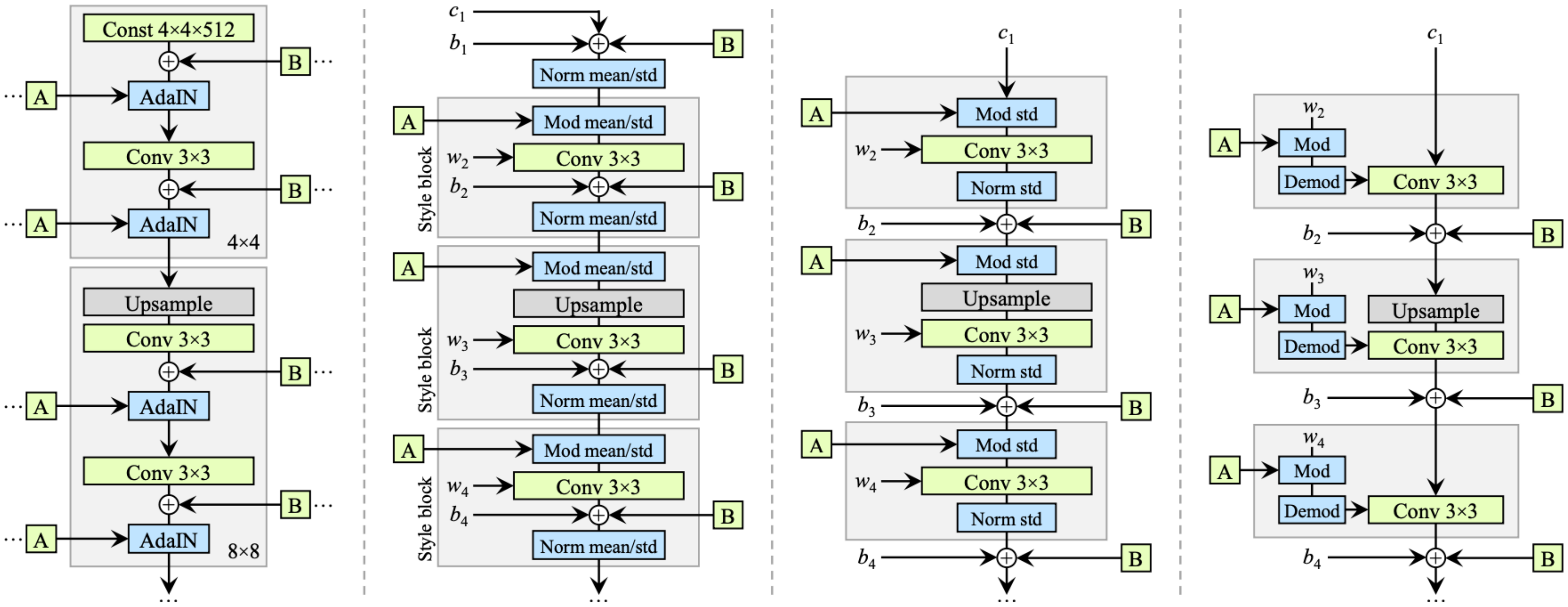
Typical domain adaptation of StyleGAN2



Almost all 30 million weights of StyleGAN2 are fune-tuned!

Aibek Alanov, Vadim Titov, and Dmitry Vetrov. Hyperdomainnet: Universal domain adaptation for generative adversarial networks. arXiv preprint arXiv:2210.08884, 2022.

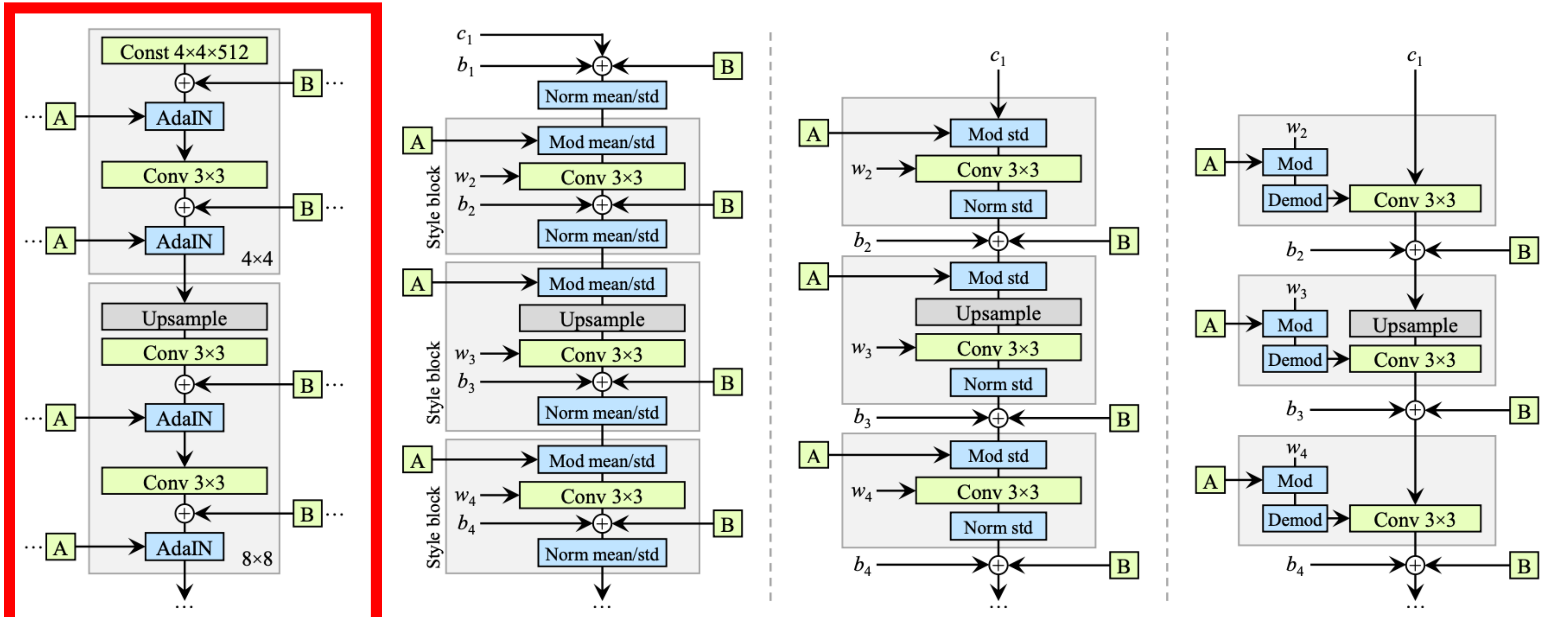
StyleGAN -> StyleGAN2



(a) StyleGAN (b) StyleGAN (detailed) (c) Revised architecture (d) Weight demodulation

Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Analyzing and improving the image quality of stylegan. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 8110–8119, 2020.

StyleGAN -> StyleGAN2

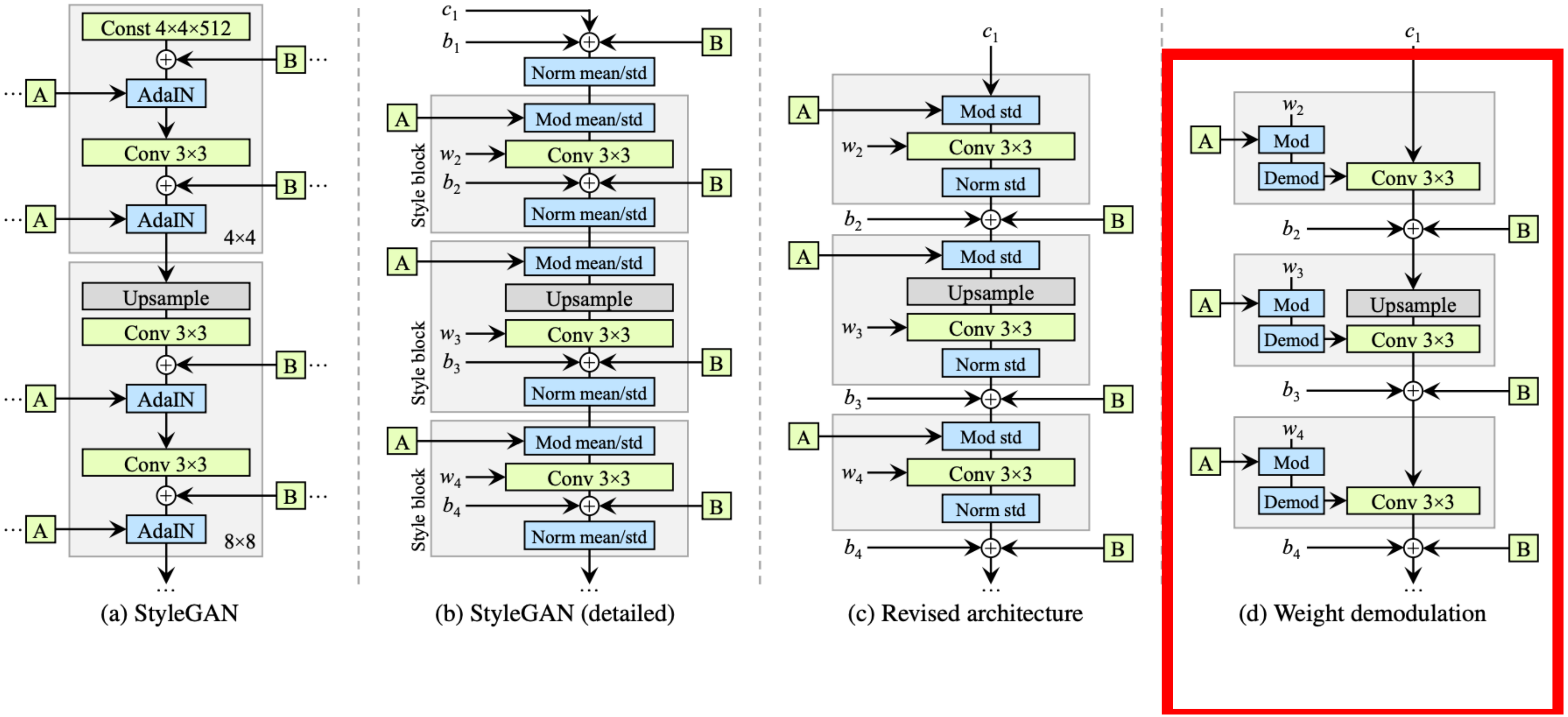


(a) StyleGAN (b) StyleGAN (detailed) (c) Revised architecture (d) Weight demodulation

$$\mathrm{AdaIN}(\mathbf{x}_i, \mathbf{y}) = \mathbf{y}_{s,i} \frac{\mathbf{x}_i - \mu(\mathbf{x}_i)}{\sigma(\mathbf{x}_i)} + \mathbf{y}_{b,i},$$

Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Analyzing and improving the image quality of stylegan. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 8110–8119, 2020.

StyleGAN -> StyleGAN2

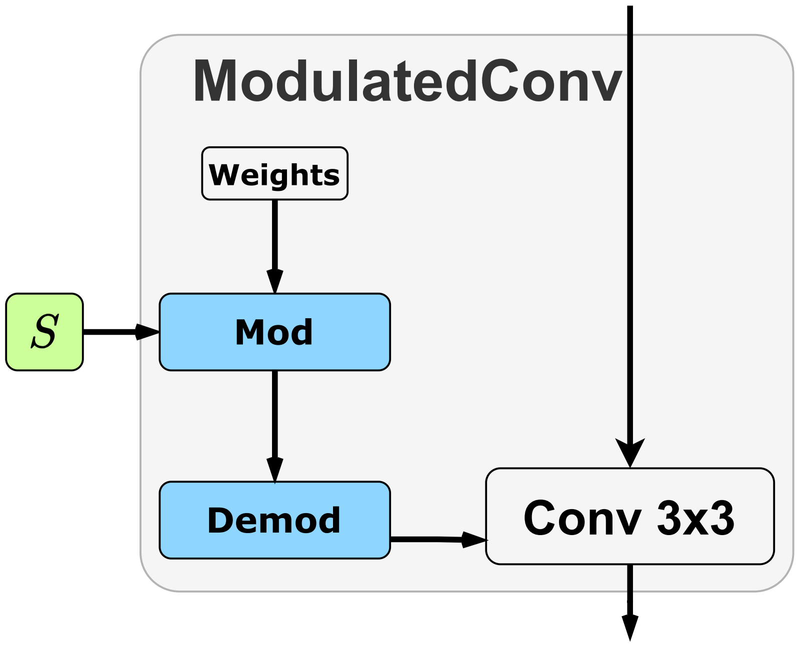


(a) StyleGAN (b) StyleGAN (detailed) (c) Revised architecture (d) Weight demodulation

$$\textbf{modulation: } w'_{ijk} = s_i \cdot w_{ijk}, \qquad \textbf{demodulation: } w''_{ijk} = \frac{w'_{ijk}}{\sqrt{\sum_{i,k} {w'_{ijk}}^2 + \varepsilon}}$$

Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Analyzing and improving the image quality of stylegan. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 8110–8119, 2020.

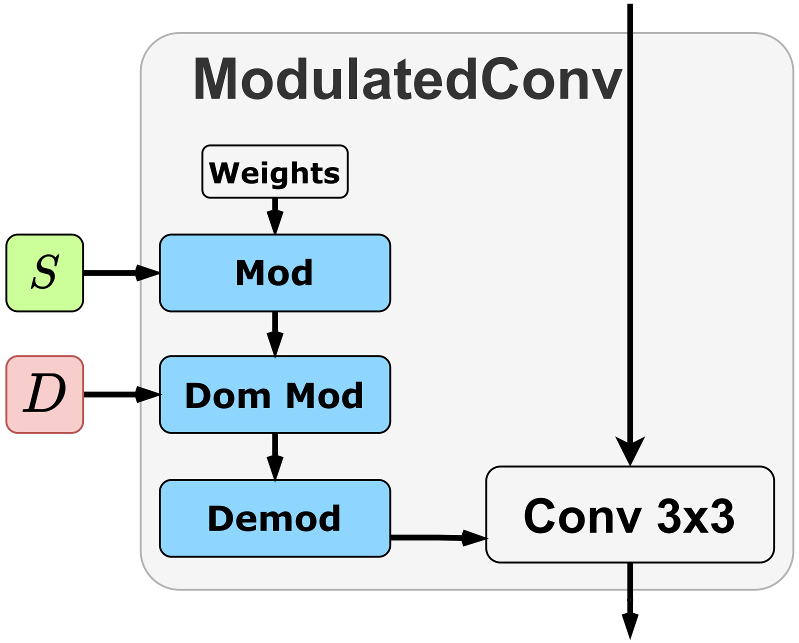
Modulated convolution of StyleGAN2



$$\text{modulation: } w'_{ijk} = s_i \cdot w_{ijk},$$

$$\text{demodulation: } w''_{ijk} = \frac{w'_{ijk}}{\sqrt{\sum_{i,k} {w'_{ijk}}^2 + \varepsilon}},$$

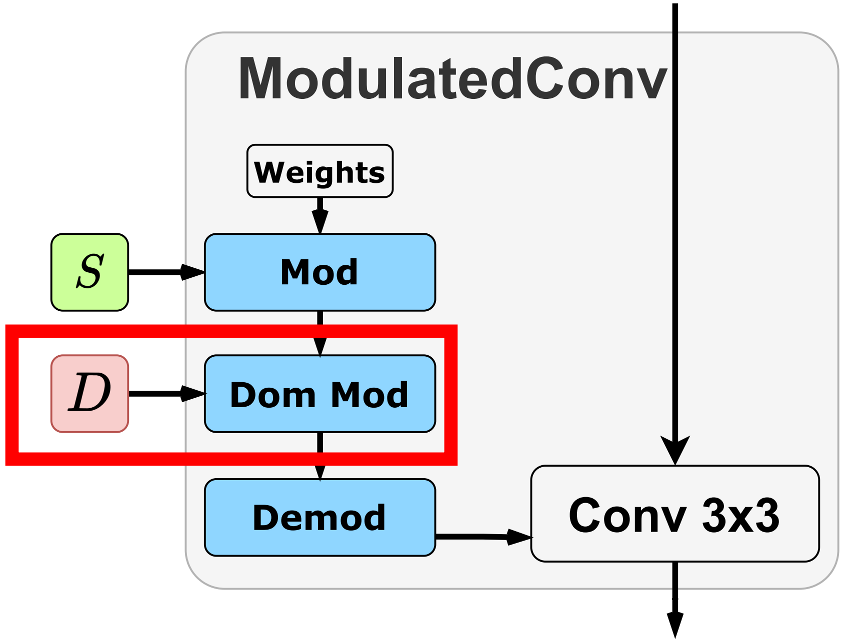
Proposed domain-modulation technique



domain-modulation: $w'_{ijk} = d_i \cdot w_{ijk},$

Aibek Alanov, Vadim Titov, and Dmitry Vetrov. Hyperdomainnet: Universal domain adaptation for generative adversarial networks. arXiv preprint arXiv:2210.08884, 2022.

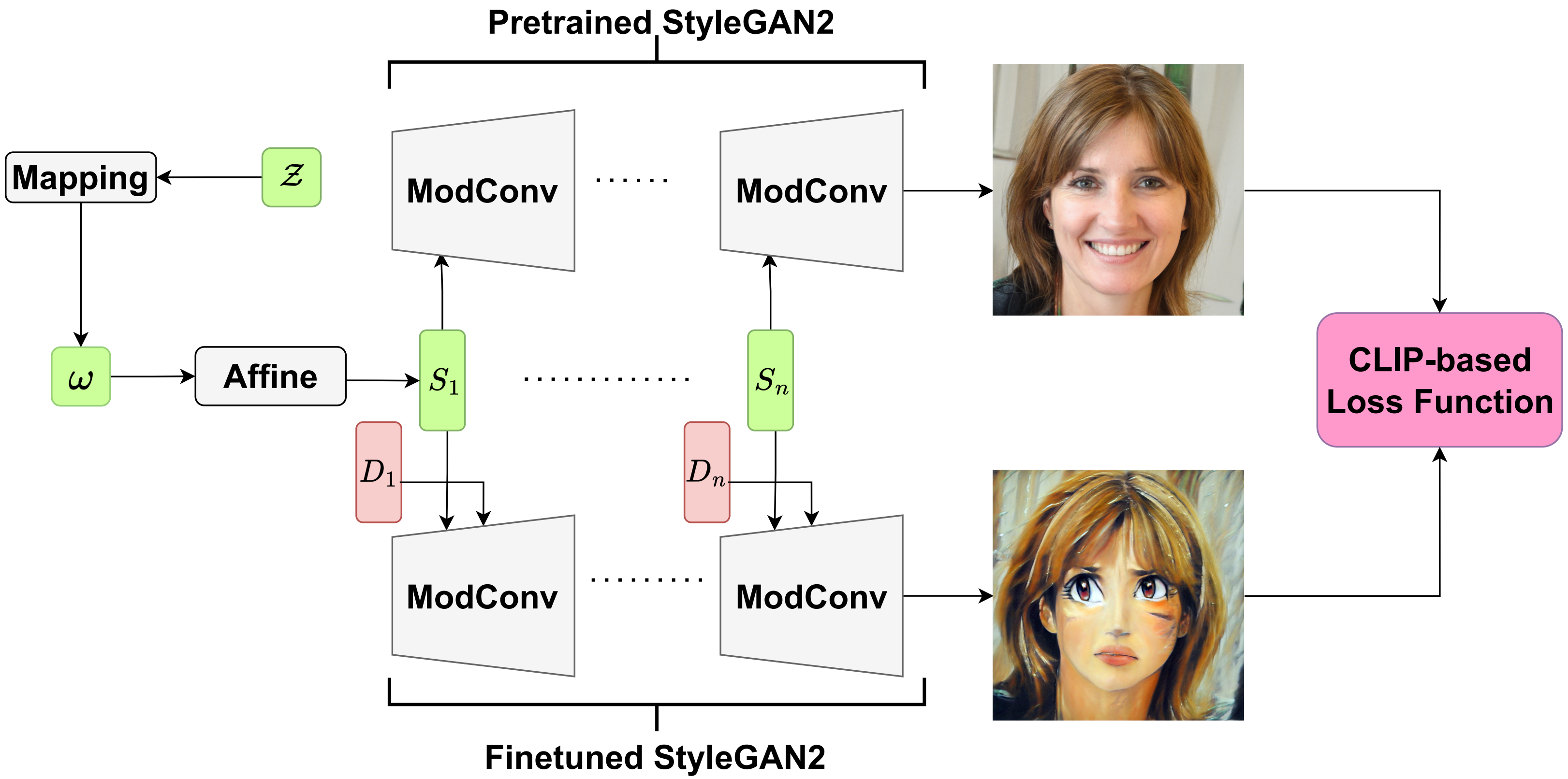
Proposed domain-modulation technique



$$\boxed{\text{domain-modulation: } w'_{ijk} = d_i \cdot w_{ijk},}$$

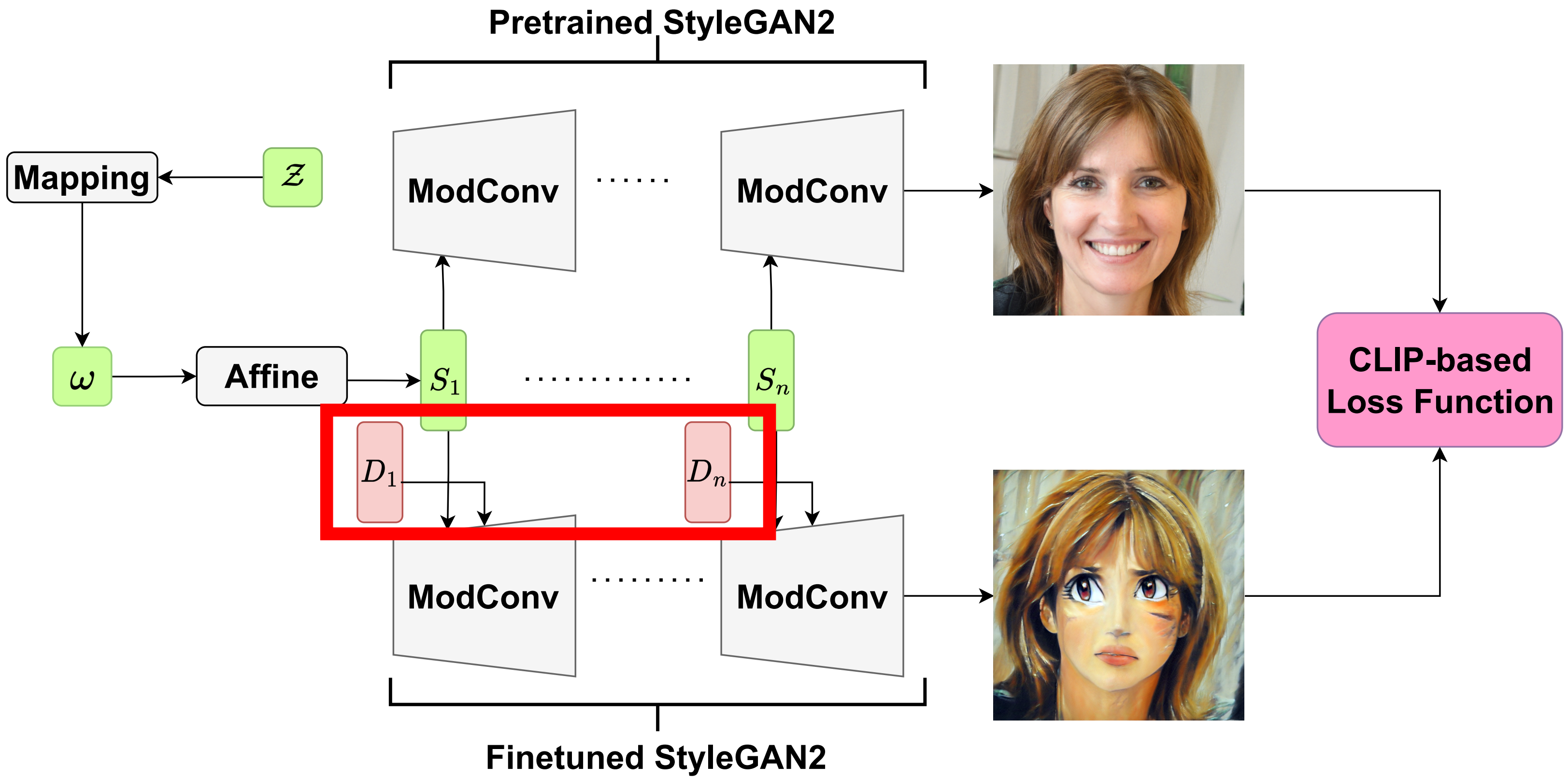
Aibek Alanov, Vadim Titov, and Dmitry Vetrov. Hyperdomainnet: Universal domain adaptation for generative adversarial networks. arXiv preprint arXiv:2210.08884, 2022.

Proposed method of domain adaptation



Aibek Alanov, Vadim Titov, and Dmitry Vetrov. Hyperdomainnet: Universal domain adaptation for generative adversarial networks. arXiv preprint arXiv:2210.08884, 2022.

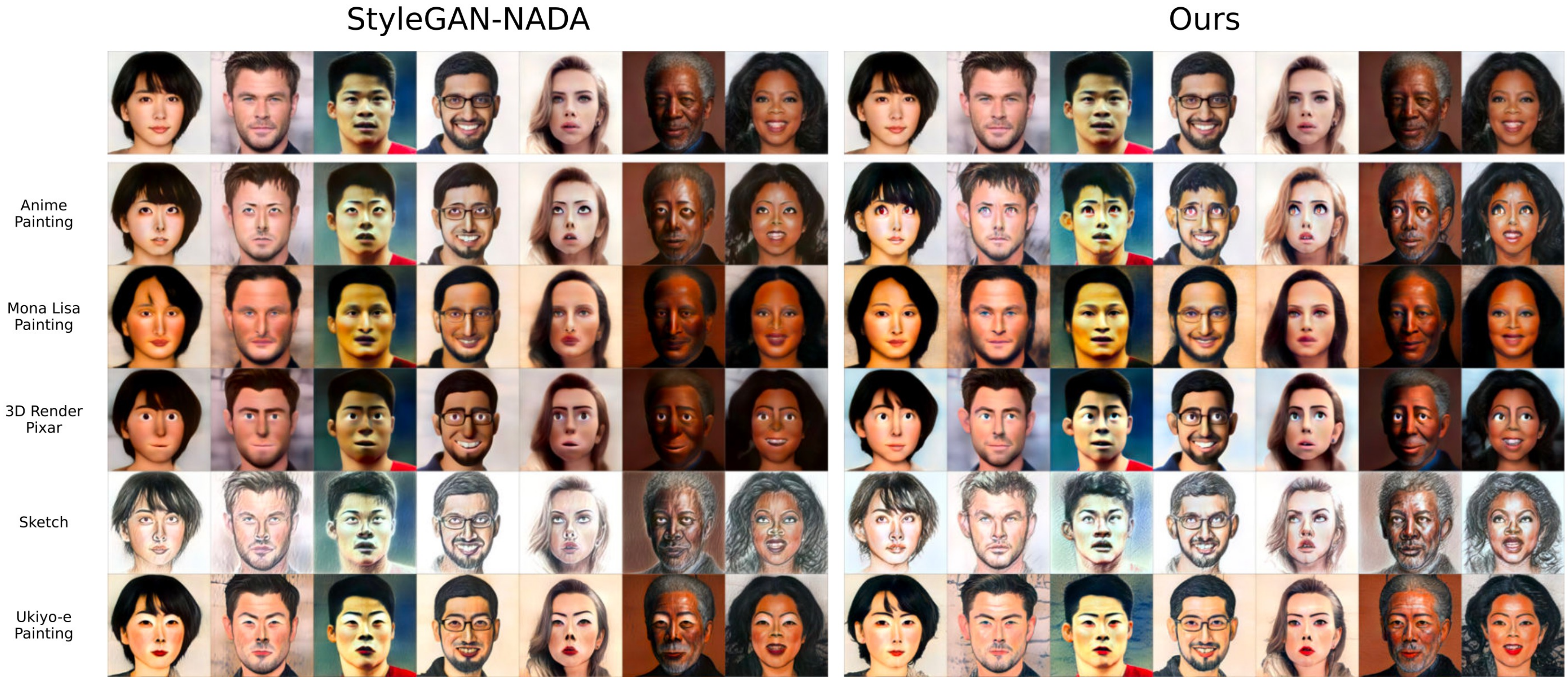
Proposed method of domain adaptation



Only 6 thousand-dimensional vector is fune-tuned!

Aibek Alanov, Vadim Titov, and Dmitry Vetrov. Hyperdomainnet: Universal domain adaptation for generative adversarial networks. arXiv preprint arXiv:2210.08884, 2022.

Qualitative comparison on text-based adaptation



Aibek Alanov, Vadim Titov, and Dmitry Vetrov. Hyperdomainnet: Universal domain adaptation for generative adversarial networks. arXiv preprint arXiv:2210.08884, 2022.

Metrics for text-based adaptation

$$\text{Quality} = \frac{1}{n}\sum_{i=1}^{n}\langle E_T(\text{target\_text}), E_I(I_i)\rangle, \text{ where}$$

n - number of the generated adapted images (we use 1000),
E_T - text CLIP encoder,
E_I - image CLIP encoder,
$I_1, \ldots, I_n$ - generated adapted images.

$$\text{Diversity} = \frac{2}{n(n-1)}\sum_{i<j}^{n}(1 - \langle E_I(I_i), E_I(I_j)\rangle), \text{ where}$$

n - number of the generated adapted images (we use 1000),
E_I - image CLIP encoder,
$I_1, \ldots, I_n$ - generated adapted images.

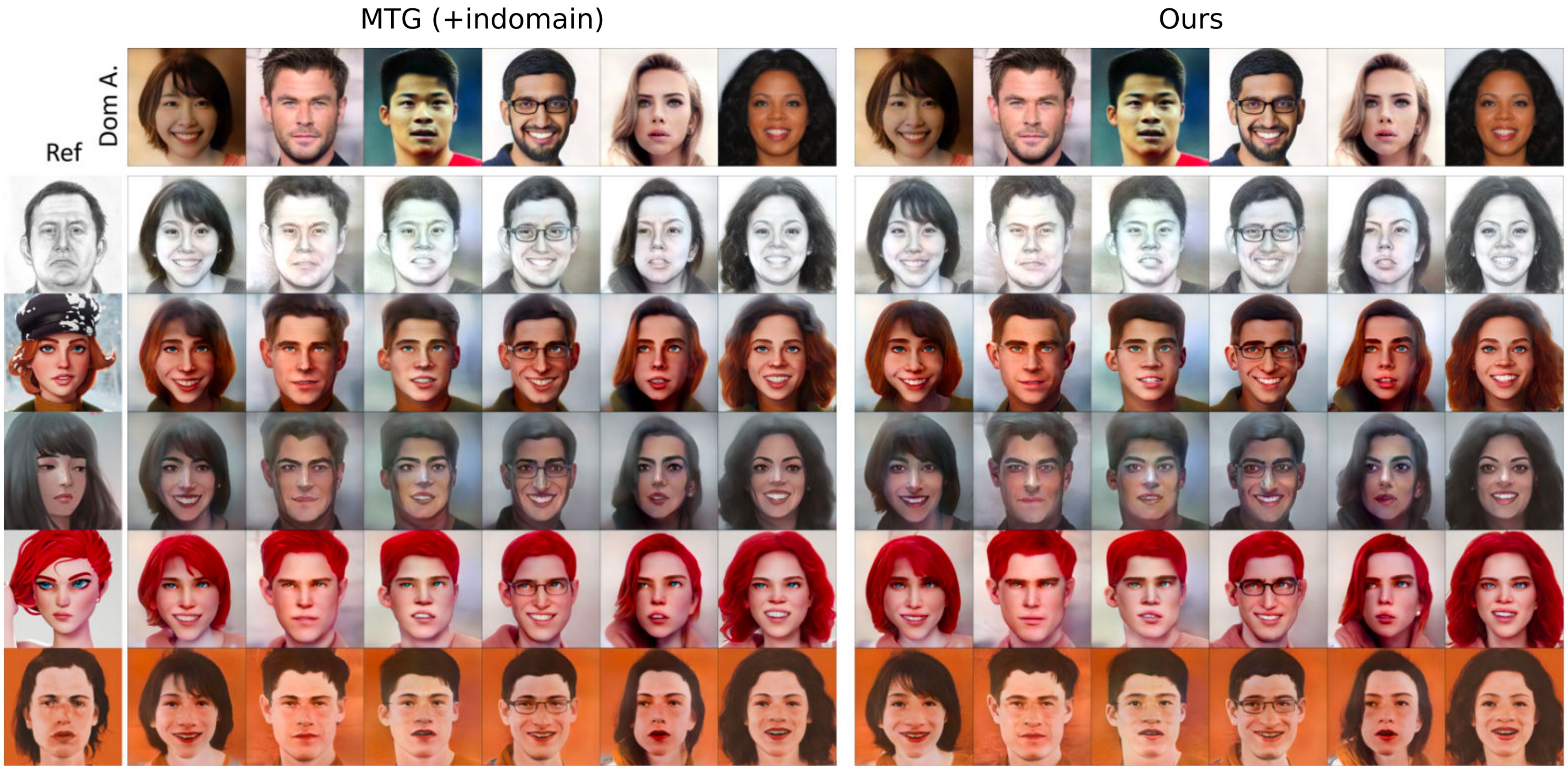
Aibek Alanov, Vadim Titov, and Dmitry Vetrov. Hyperdomainnet: Universal domain adaptation for generative adversarial networks. arXiv preprint arXiv:2210.08884, 2022.

Quantitative comparison on text-based adaptation

Model	Quality	Diversity
Anime Painting		
StyleGAN-NADA [6]	0.289	0.244
Ours	0.284	0.305
Across ten domains		
StyleGAN-NADA [6]	0.270 ± 0.032	0.196 ± 0.034
Ours	0.256 ± 0.019	0.306 ± 0.030

Aibek Alanov, Vadim Titov, and Dmitry Vetrov. Hyperdomainnet: Universal domain adaptation for generative adversarial networks. arXiv preprint arXiv:2210.08884, 2022.

Qualitative comparison on image-based adaptation



Aibek Alanov, Vadim Titov, and Dmitry Vetrov. Hyperdomainnet: Universal domain adaptation for generative adversarial networks. arXiv preprint arXiv:2210.08884, 2022.

Quantitative comparison on image-based adaptation

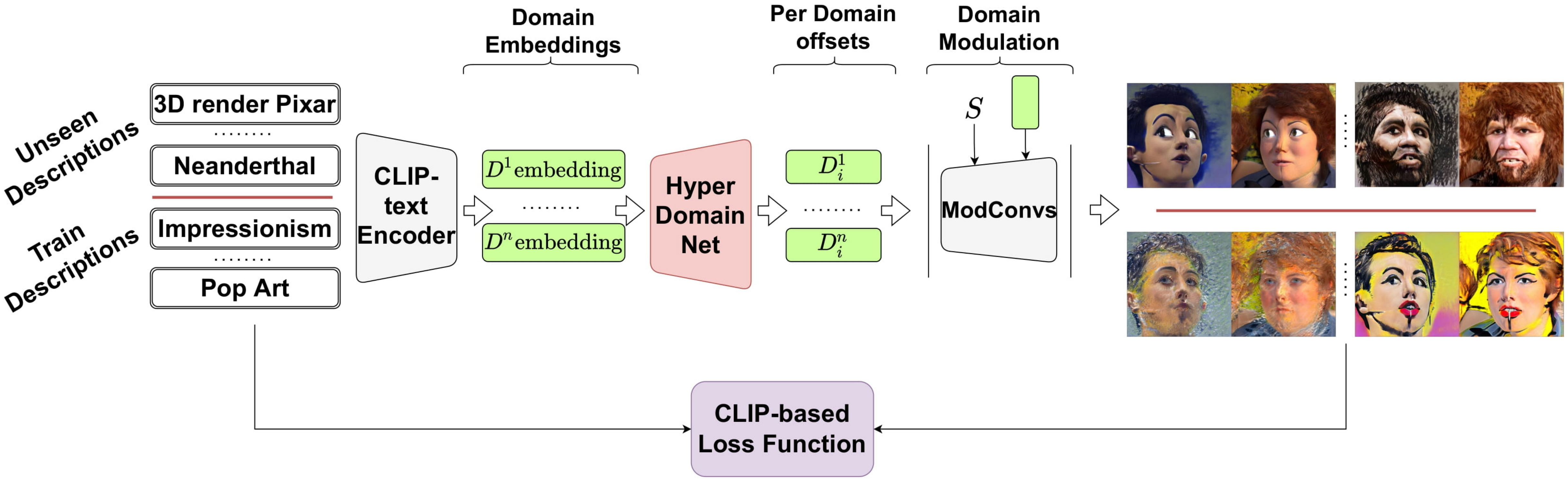
	Model quality			Model complexity
Model	FID	Precision	Recall	# trainable parameters
TargetCLIP [4]	199.33	0.000	0.293	30M
Cross-correspondence [24]	158.86	0.001	0	30M
StyleGAN-NADA [6]	124.55	0.118	0	24M
MindTheGap [48]	78.35	0.326	0.017	24M
MindTheGap (our param.)	79.83	0.452	0.017	6k
MindTheGap+indomain	71.46	0.503	0.014	24M
MindTheGap+indomain (our param.)	72.71	0.472	0.028	6k

Aibek Alanov, Vadim Titov, and Dmitry Vetrov. Hyperdomainnet: Universal domain adaptation for generative adversarial networks. arXiv preprint arXiv:2210.08884, 2022.

Quantitative comparison on image-based adaptation

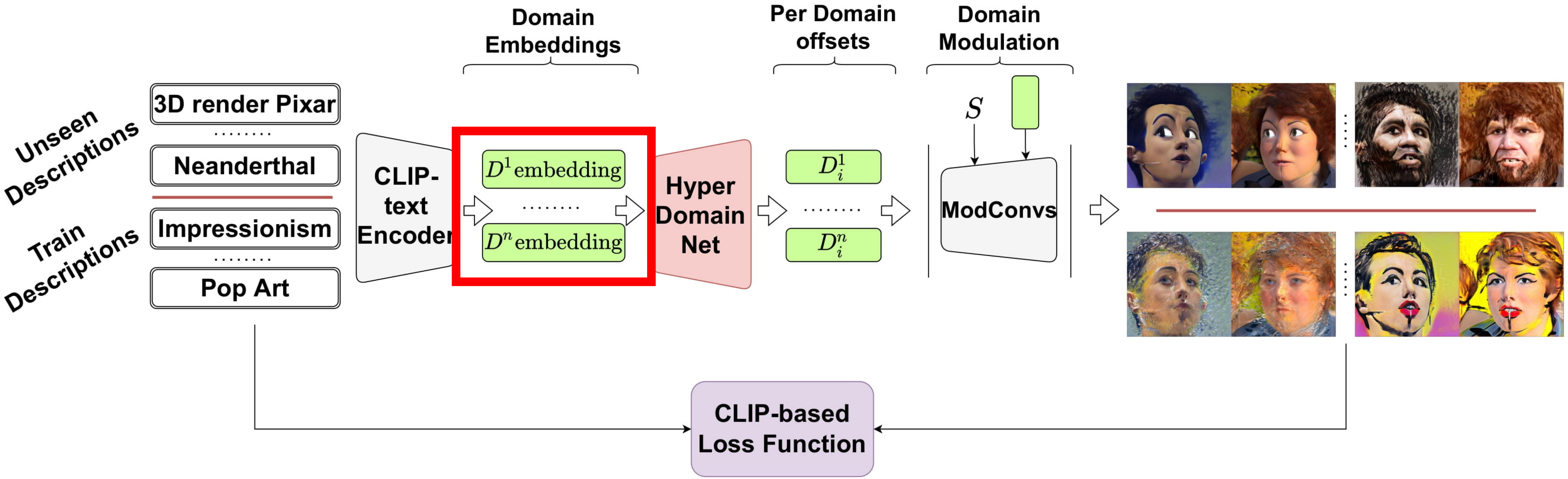
	Model quality			Model complexity
Model	FID	Precision	Recall	# trainable parameters
TargetCLIP [4]	199.33	0.000	0.293	30M
Cross-correspondence [24]	158.86	0.001	0	30M
StyleGAN-NADA [6]	124.55	0.118	0	24M
MindTheGap [48]	78.35	0.326	0.017	24M
MindTheGap (our param.)	79.83	0.452	0.017	6k
MindTheGap+indomain	71.46	0.503	0.014	24M
MindTheGap+indomain (our param.)	72.71	0.472	0.028	6k

HyperDomainNet



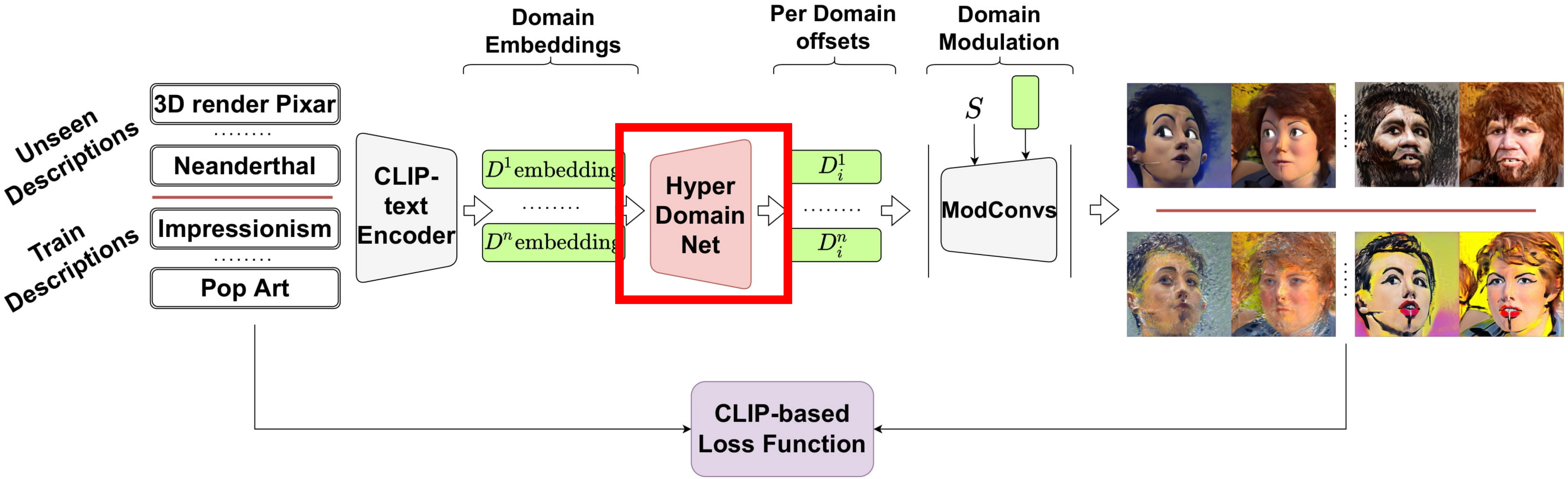
Aibek Alanov, Vadim Titov, and Dmitry Vetrov. Hyperdomainnet: Universal domain adaptation for generative adversarial networks. arXiv preprint arXiv:2210.08884, 2022.

HyperDomainNet for text-based adaptation



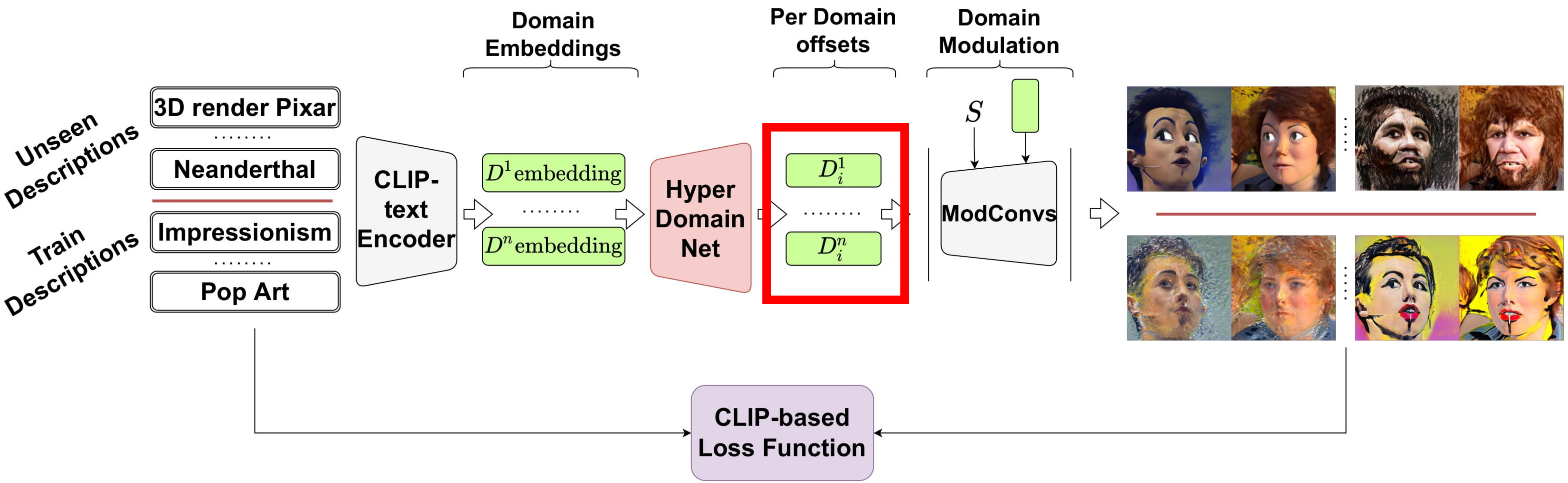
Aibek Alanov, Vadim Titov, and Dmitry Vetrov. Hyperdomainnet: Universal domain adaptation for generative adversarial networks. arXiv preprint arXiv:2210.08884, 2022.

HyperDomainNet for text-based adaptation



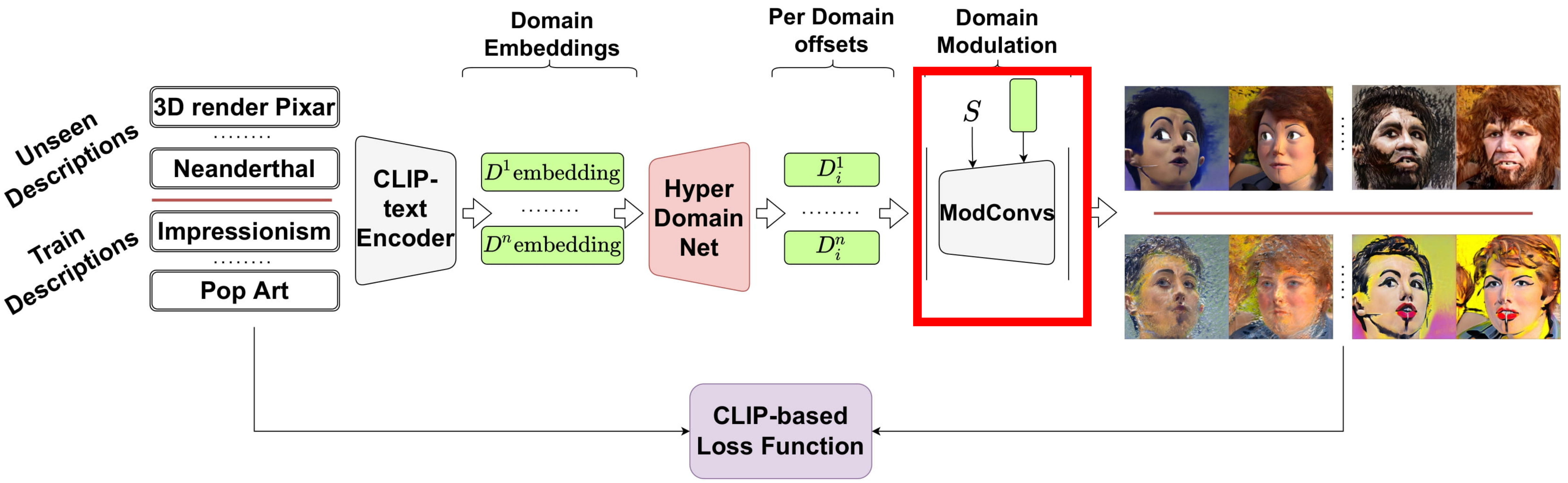
Aibek Alanov, Vadim Titov, and Dmitry Vetrov. Hyperdomainnet: Universal domain adaptation for generative adversarial networks. arXiv preprint arXiv:2210.08884, 2022.

HyperDomainNet for text-based adaptation



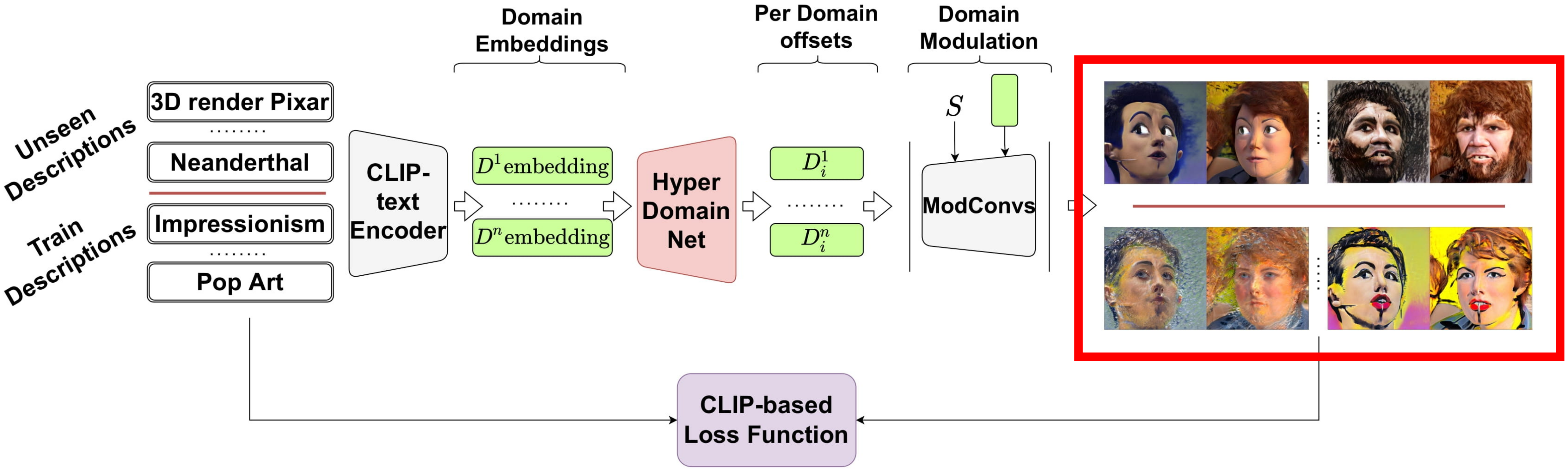
Aibek Alanov, Vadim Titov, and Dmitry Vetrov. Hyperdomainnet: Universal domain adaptation for generative adversarial networks. arXiv preprint arXiv:2210.08884, 2022.

HyperDomainNet for text-based adaptation



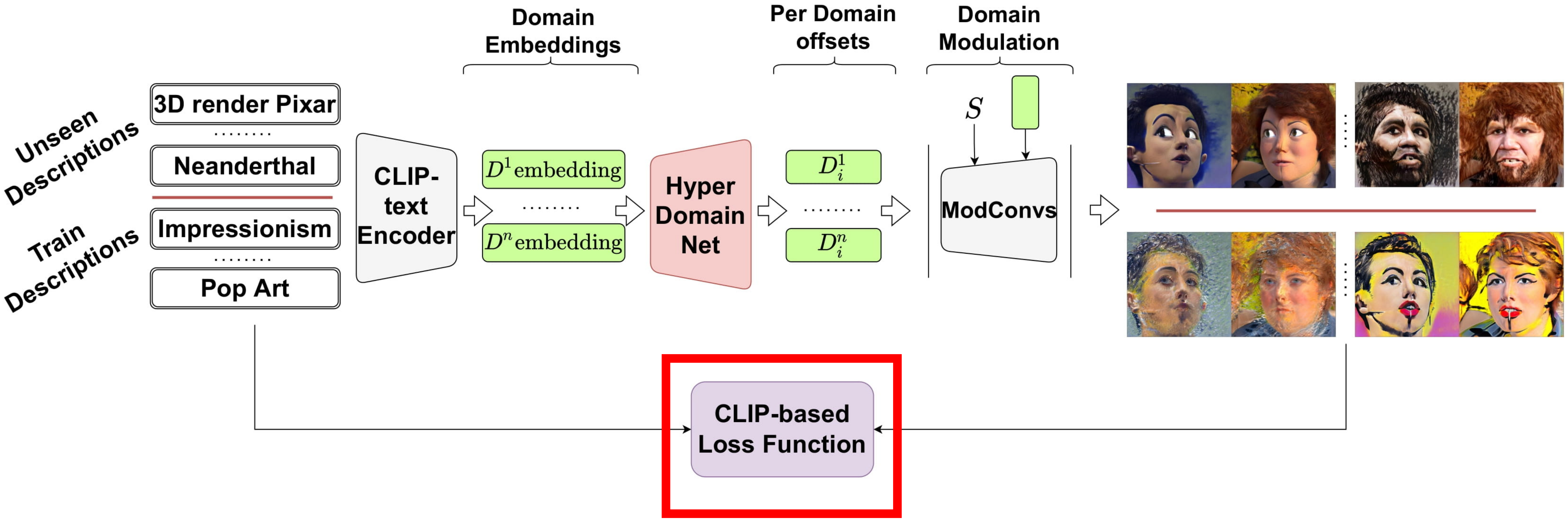
Aibek Alanov, Vadim Titov, and Dmitry Vetrov. Hyperdomainnet: Universal domain adaptation for generative adversarial networks. arXiv preprint arXiv:2210.08884, 2022.

HyperDomainNet for text-based adaptation



Aibek Alanov, Vadim Titov, and Dmitry Vetrov. Hyperdomainnet: Universal domain adaptation for generative adversarial networks. arXiv preprint arXiv:2210.08884, 2022.

HyperDomainNet for text-based adaptation



Aibek Alanov, Vadim Titov, and Dmitry Vetrov. Hyperdomainnet: Universal domain adaptation for generative adversarial networks. arXiv preprint arXiv:2210.08884, 2022.

Results on the train domains

Aibek Alanov, Vadim Titov, and Dmitry Vetrov. Hyperdomainnet: Universal domain adaptation for generative adversarial networks. arXiv preprint arXiv:2210.08884, 2022.

Quantitative results on the train domains

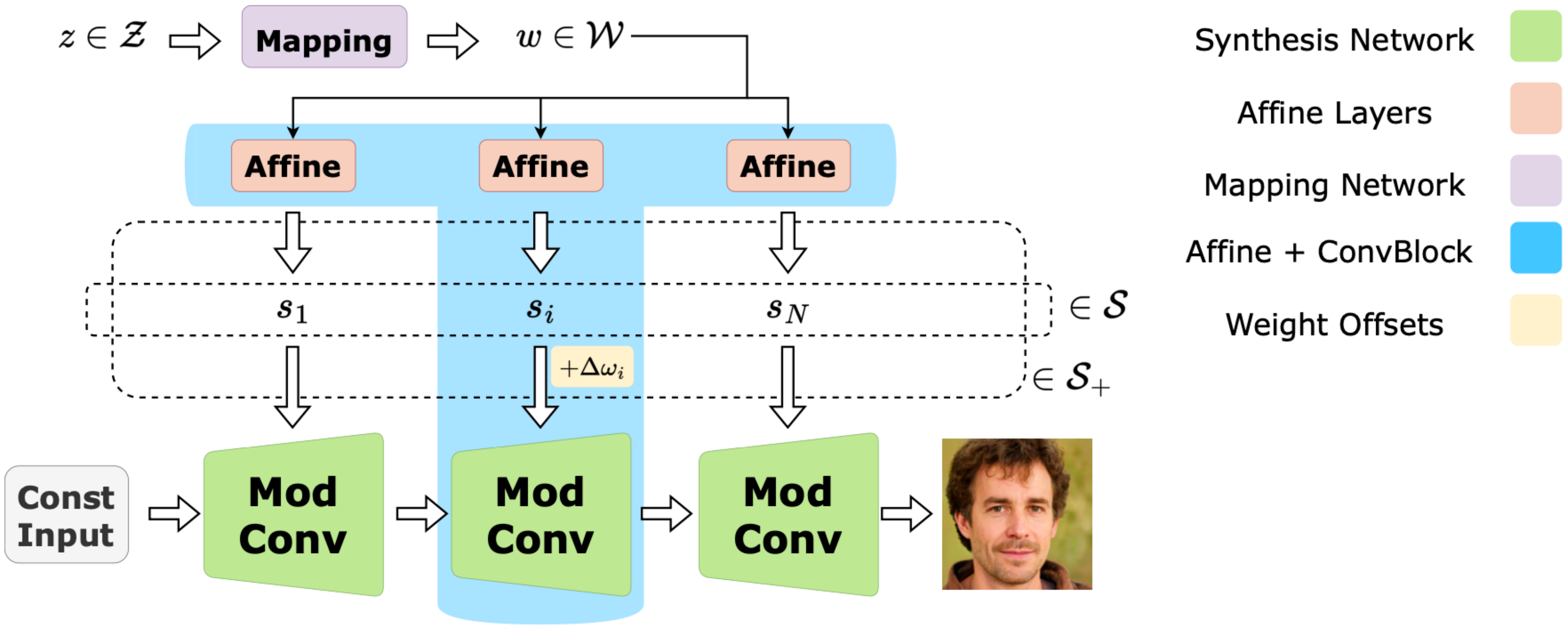
Model	Quality	Diversity
Across ten domains		
StyleGAN-NADA [6]	0.270 ± 0.032	0.196 ± 0.034
Ours	0.256 ± 0.019	0.306 ± 0.030
Multi-Domain+domain_norm+tt_direction	0.247 ± 0.026	0.250 ± 0.041

Aibek Alanov, Vadim Titov, and Dmitry Vetrov. Hyperdomainnet: Universal domain adaptation for generative adversarial networks. arXiv preprint arXiv:2210.08884, 2022.

Results on the unseen domains

Aibek Alanov, Vadim Titov, and Dmitry Vetrov. Hyperdomainnet: Universal domain adaptation for generative adversarial networks. arXiv preprint arXiv:2210.08884, 2022.

Analysis of StyleGAN components



Analysis of StyleGAN components

$$\mathcal{L}_D\left(\left\{G_\theta(s_1^i,\ldots,s_N^i)\right\}_{i=1}^K\right)\to\min_{\theta,f^A,f_M},$$

$\{\theta\}$ – *SyntConv*

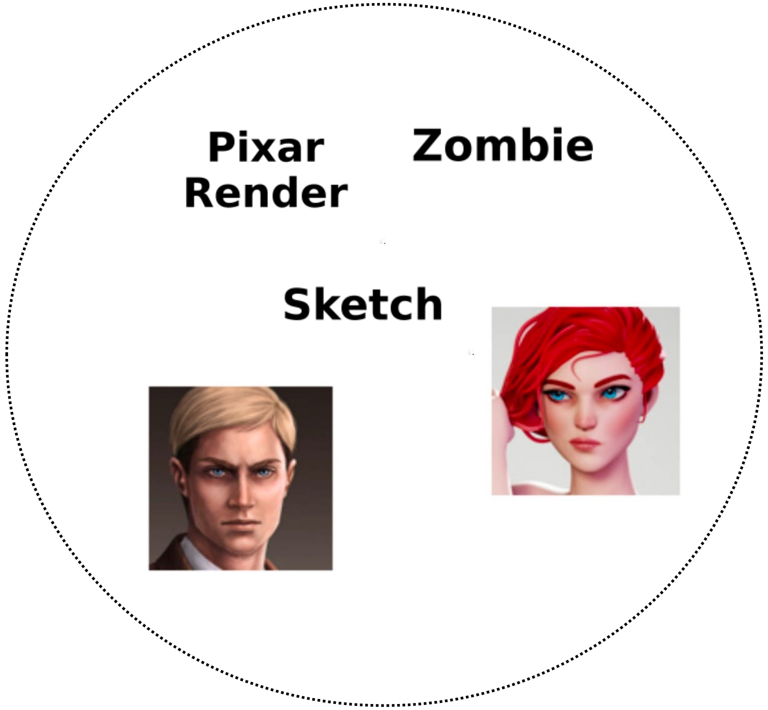
$\{f^A\}$ – *Affine*

$\{f_M\}$ – *Mapping*

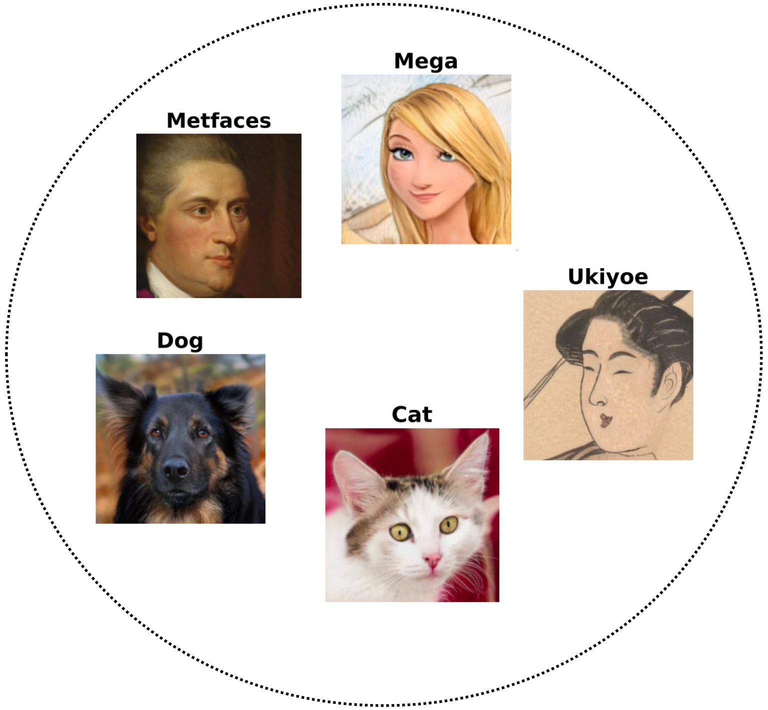
Affine+

Types of domains

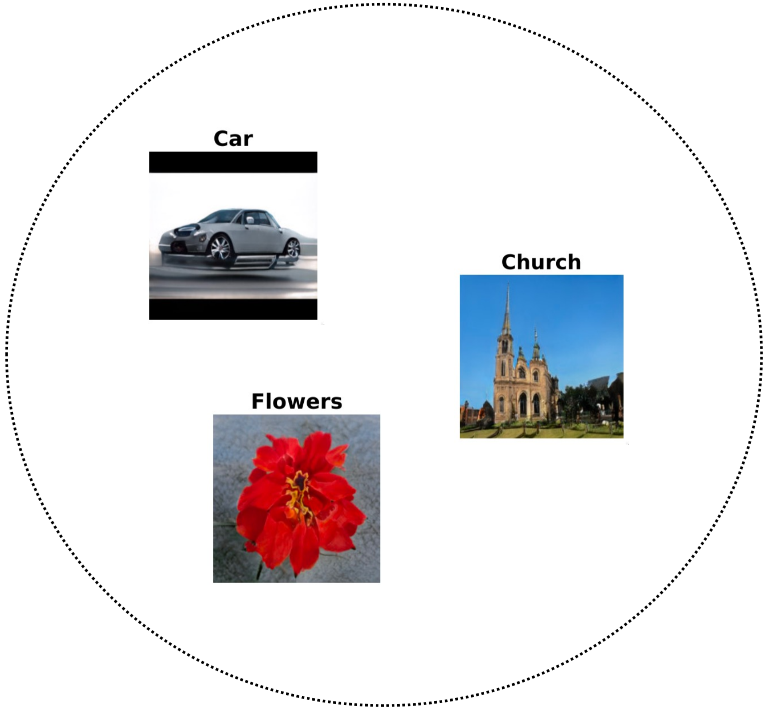
Similar domains



Moderately similar domains



Dissimilar domains

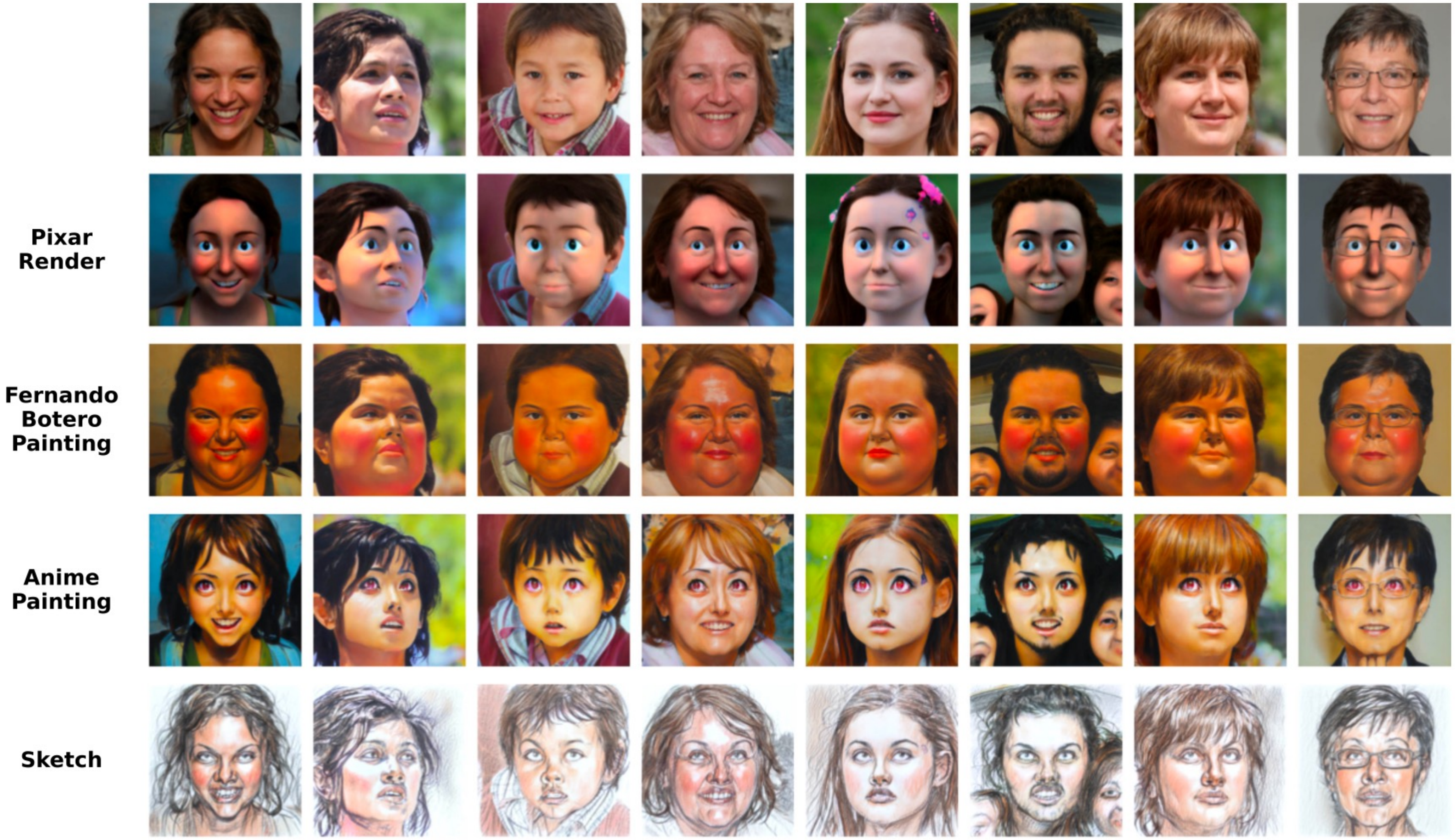


Similar domains

Similar domains

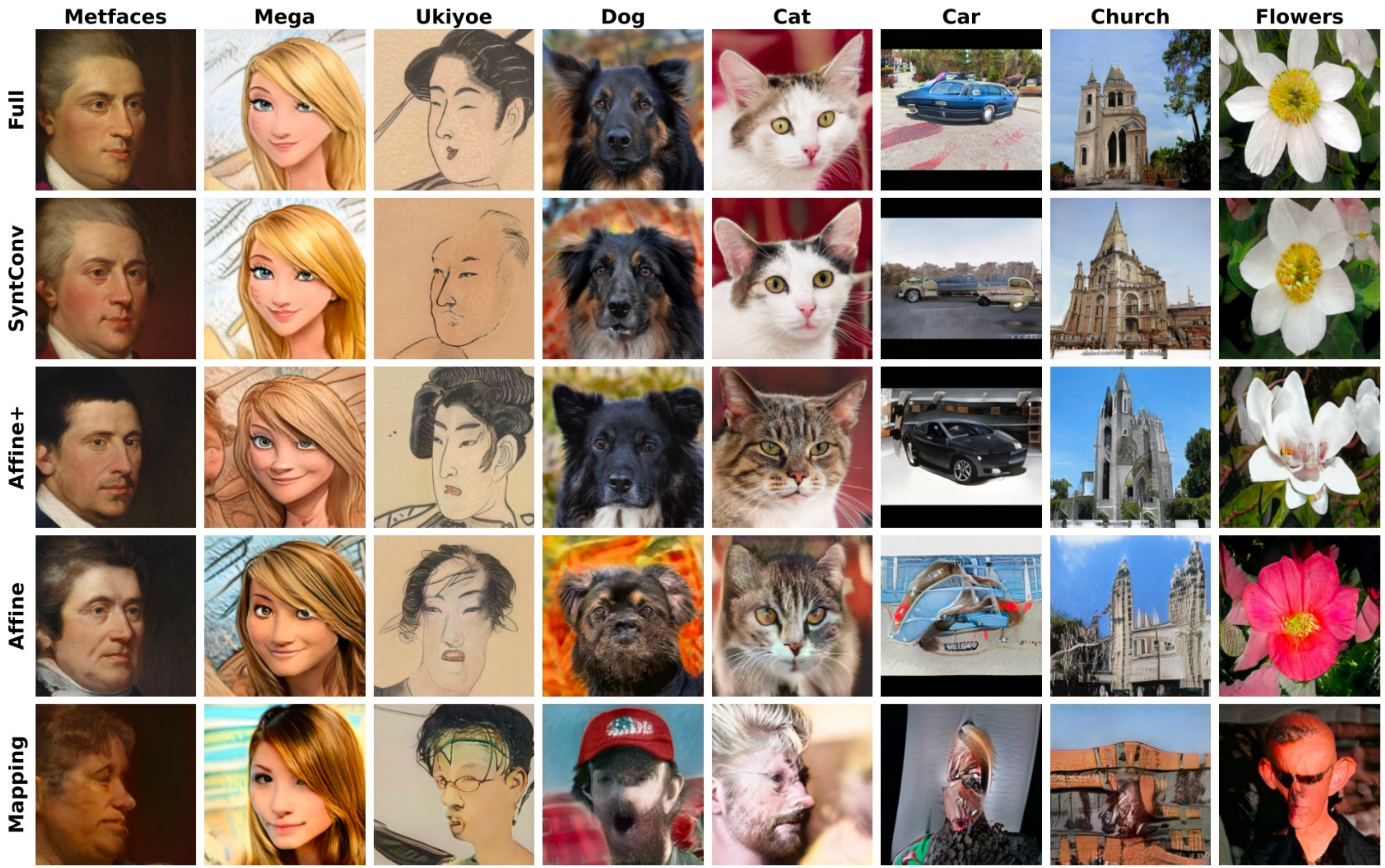
		Botero		Sketch		The Joker		Anastasia (image)	
Parameter Space	Size	Quality	Diversity	Quality	Diversity	Quality	Diversity	Quality	Diversity
Full	30.3M	0.312	0.228	0.208	0.296	0.246	0.167	0.640	0.276
SyntConv	23.6M	0.311	0.224	0.191	0.292	0.245	0.164	0.683	0.250
Affine	4.6M	0.298	0.221	0.194	0.296	0.244	0.168	0.642	0.269
Mapping	2.1M	0.226	0.115	0.182	0.143	0.246	0.136	0.523	0.202

Affine parameterization (more samples)



Mapping parameterization (more samples)

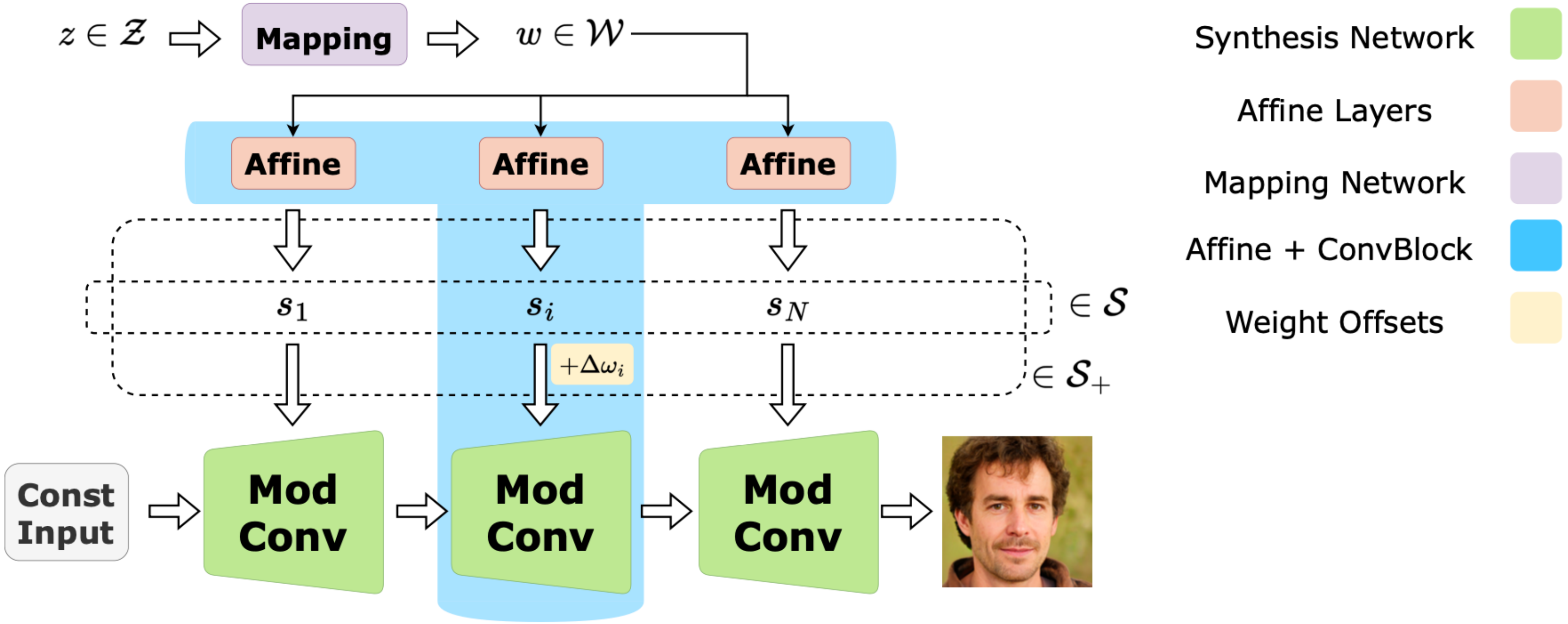
Moderately similar and dissimilar domains



Moderately similar and dissimilar domains

Parameter Space	Size	Metfaces	Mega	Ukiyoe	Dog	Cat	Car	Church	Flowers
Full	30.3M	20.0	80.4	19.6	18.9	5.9	20.9	15.0	14.8
SyntConv	23.6M	21.5	81.5	21.2	20.0	6.1	22.6	15.4	15.0
Affine+	5.1M	20.2	77.2	24.0	17.1	5.6	38.2	16.2	15.6
Affine	4.6M	24.3	107.9	60.3	68.0	27.0	110.4	85.9	40.4
Mapping	2.1M	53.4	136.2	155.6	206.2	225.6	237.1	263.5	234.3

StyleSpace for domain adaptation



$$\mathcal{L}_D\left(\{G_\theta(s_1^i + \Delta s_1^D, \ldots, s_N^i + \Delta s_N^D)\}_{i=1}^K\right) \to \min_{\Delta s^D}$$

Δs^D – *StyleDomain* direction

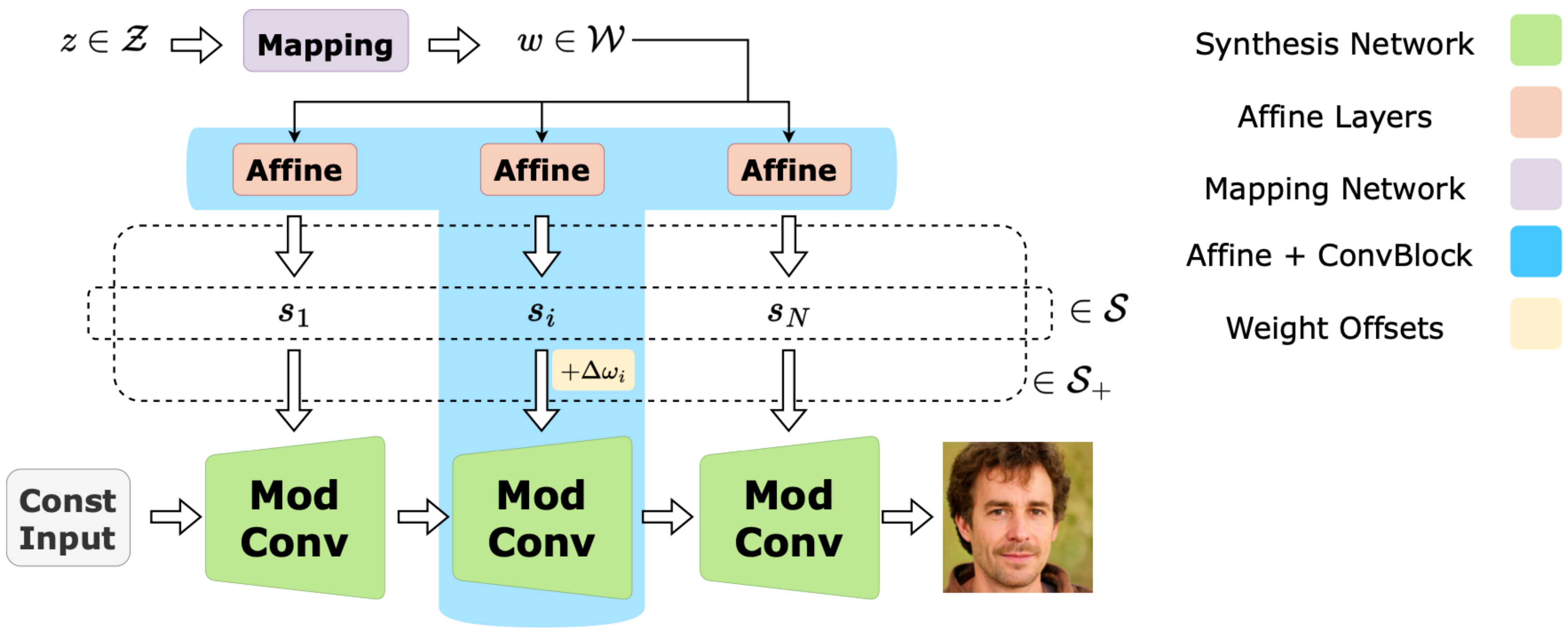
StyleSpace for similar domains

StyleSpace for similar domains

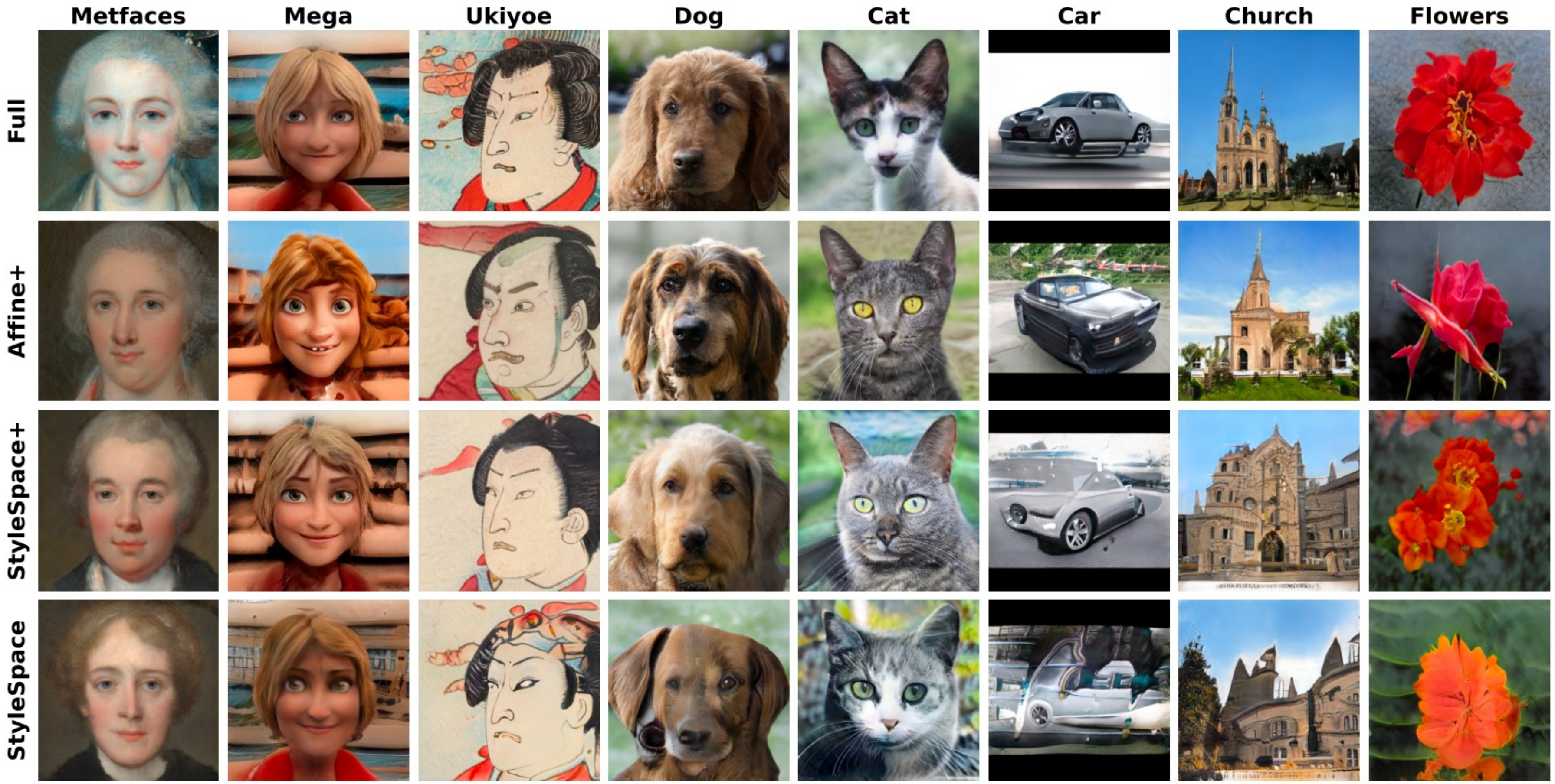
		Botero		Sketch		The Joker		Anastasia (image)	
Parameter Space	Size	Quality	Diversity	Quality	Diversity	Quality	Diversity	Quality	Diversity
Full	30.3M	0.312	0.228	0.208	0.296	0.246	0.167	0.640	0.276
SyntConv	23.6M	0.311	0.224	0.191	0.292	0.245	0.164	0.683	0.250
Affine	4.6M	0.298	0.221	0.194	0.296	0.244	0.168	0.642	0.269
Mapping	2.1M	0.226	0.115	0.182	0.143	0.246	0.136	0.523	0.202
StyleSpace	9K	0.309	0.23	0.193	0.306	0.242	0.187	0.616	0.296

StyleSpace for similar domains (more samples)

StyleSpace+ for domain adaptation



Moderately similar and dissimilar domains



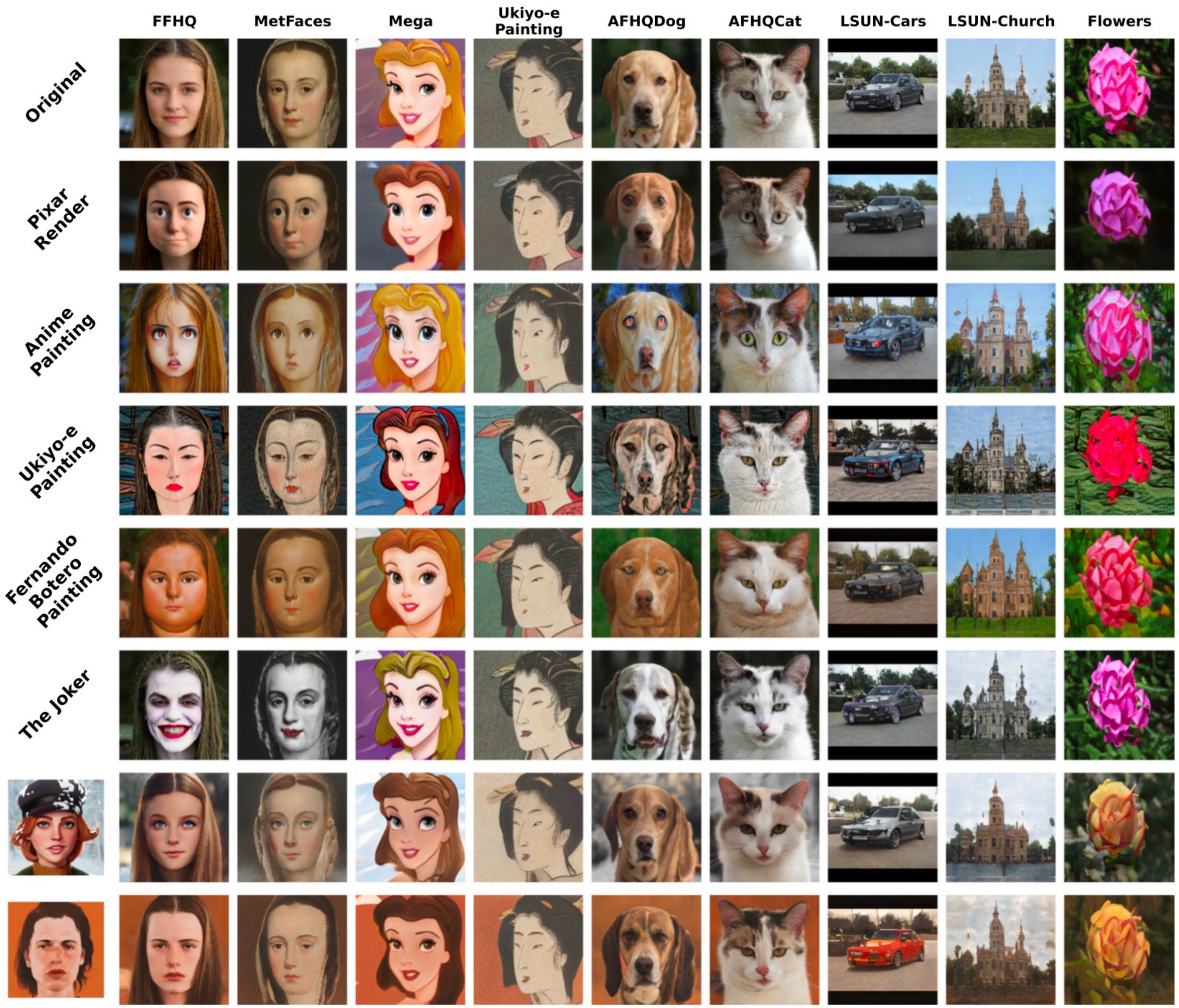
Moderately similar and dissimilar domains

Parameter Space	Size	Metfaces	Mega	Ukiyoe	Dog	Cat	Car	Church	Flowers
Full	30.3M	20.0	80.4	19.6	18.9	5.9	20.9	15.0	14.8
Affine+	5.1M	20.2	77.2	24.0	17.1	5.6	38.2	16.2	15.6
StyleSpace+	0.5M	23.5	82.5	29.1	28.0	7.8	41.7	25.7	19.9
StyleSpace	9.0K	28.6	102.0	48.8	74.2	20.6	132.3	63.9	36.8

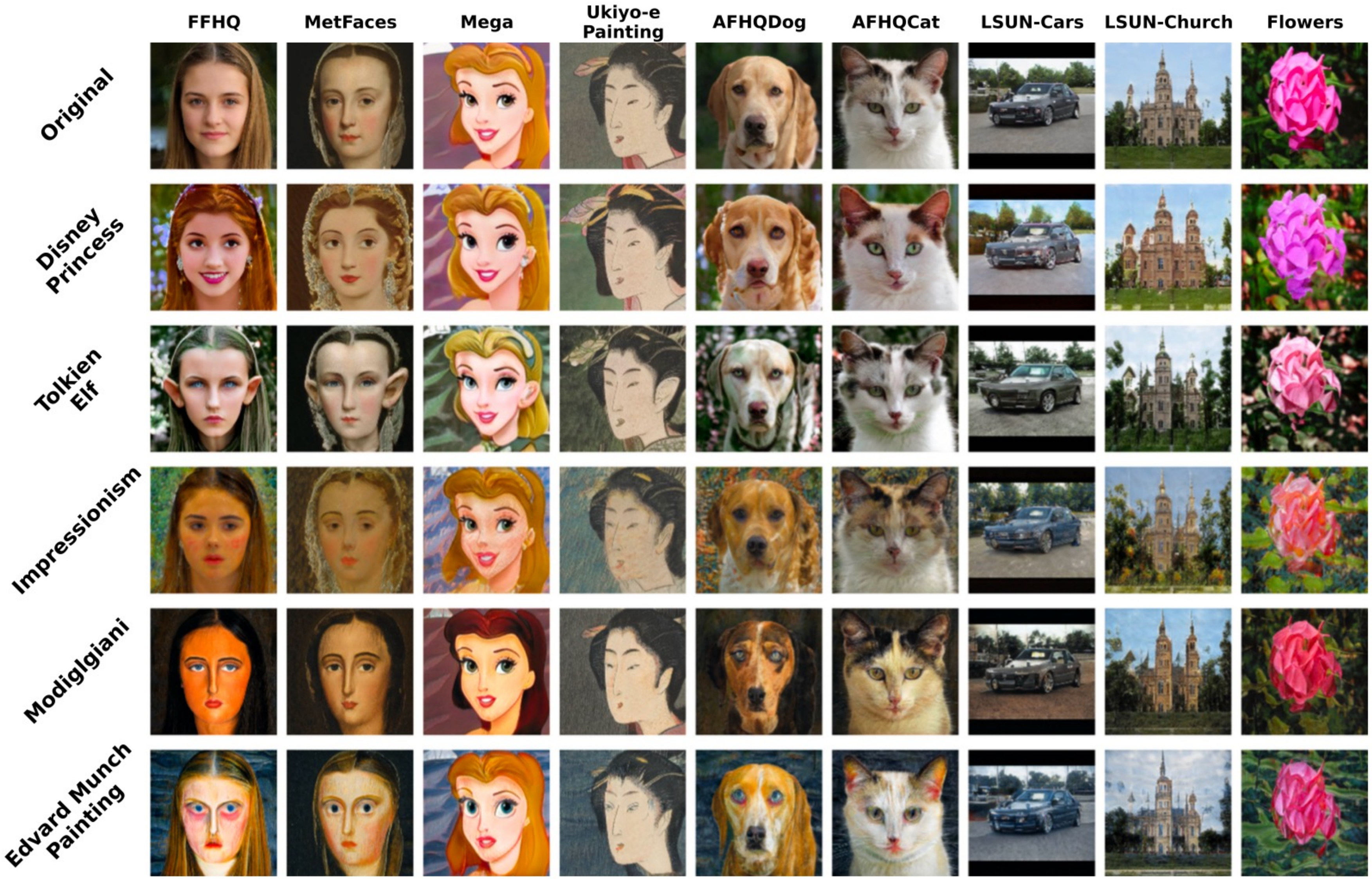
Properties of StyleDomain directions

- **Transferability of StyleDomain directions**
- **Different StyleDomain directions can be mixed**
- **StyleDomain directions can be combined with semaintic editing**

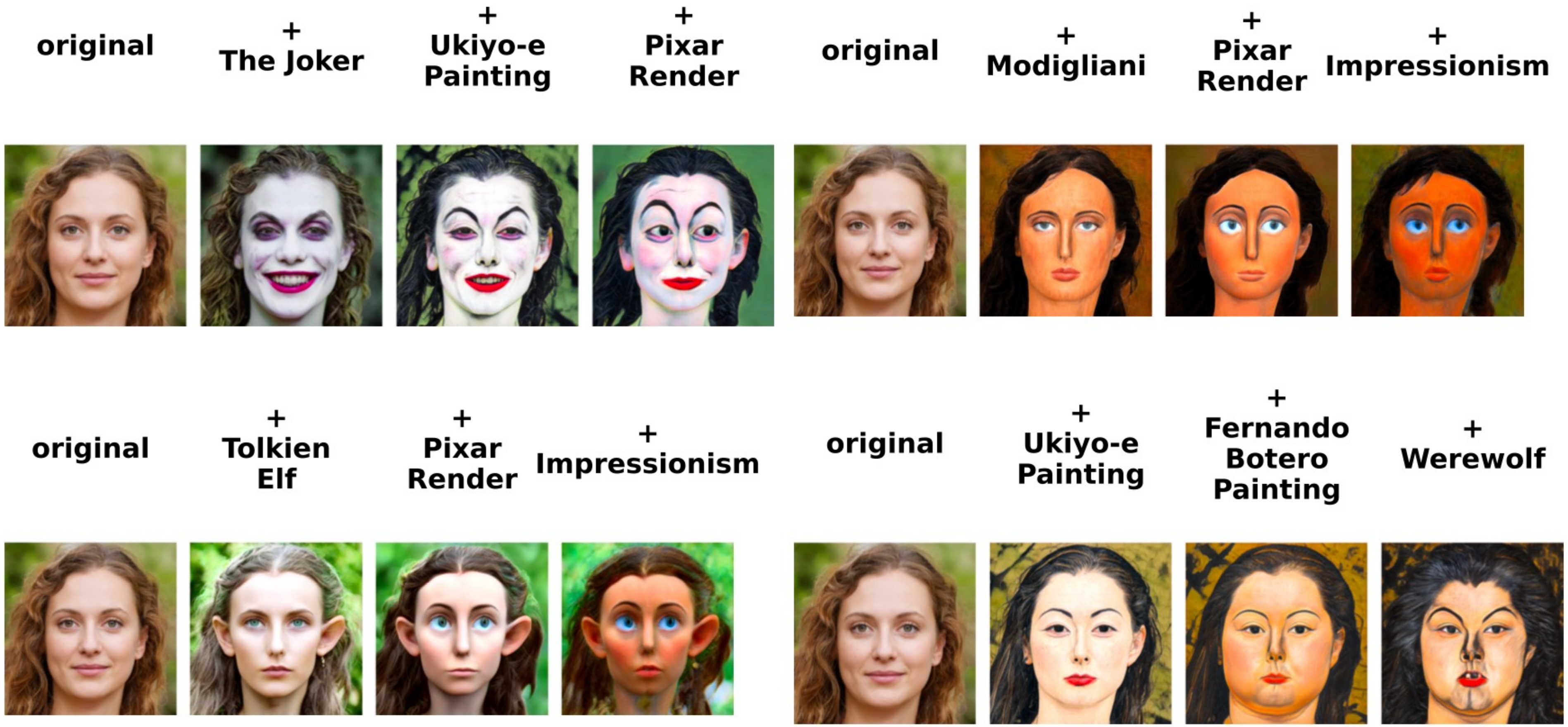
Transferability of StyleDomain directions



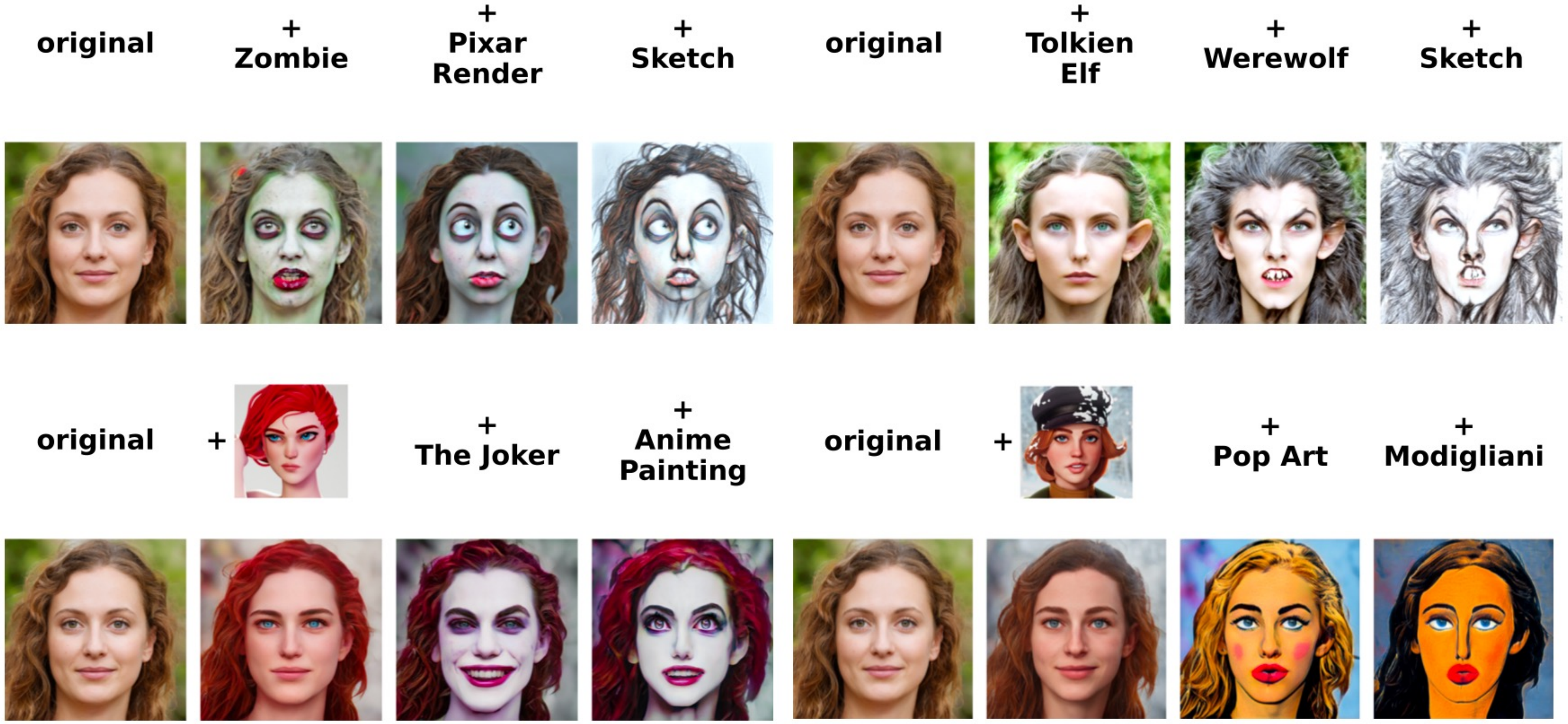
Transferability of StyleDomain directions



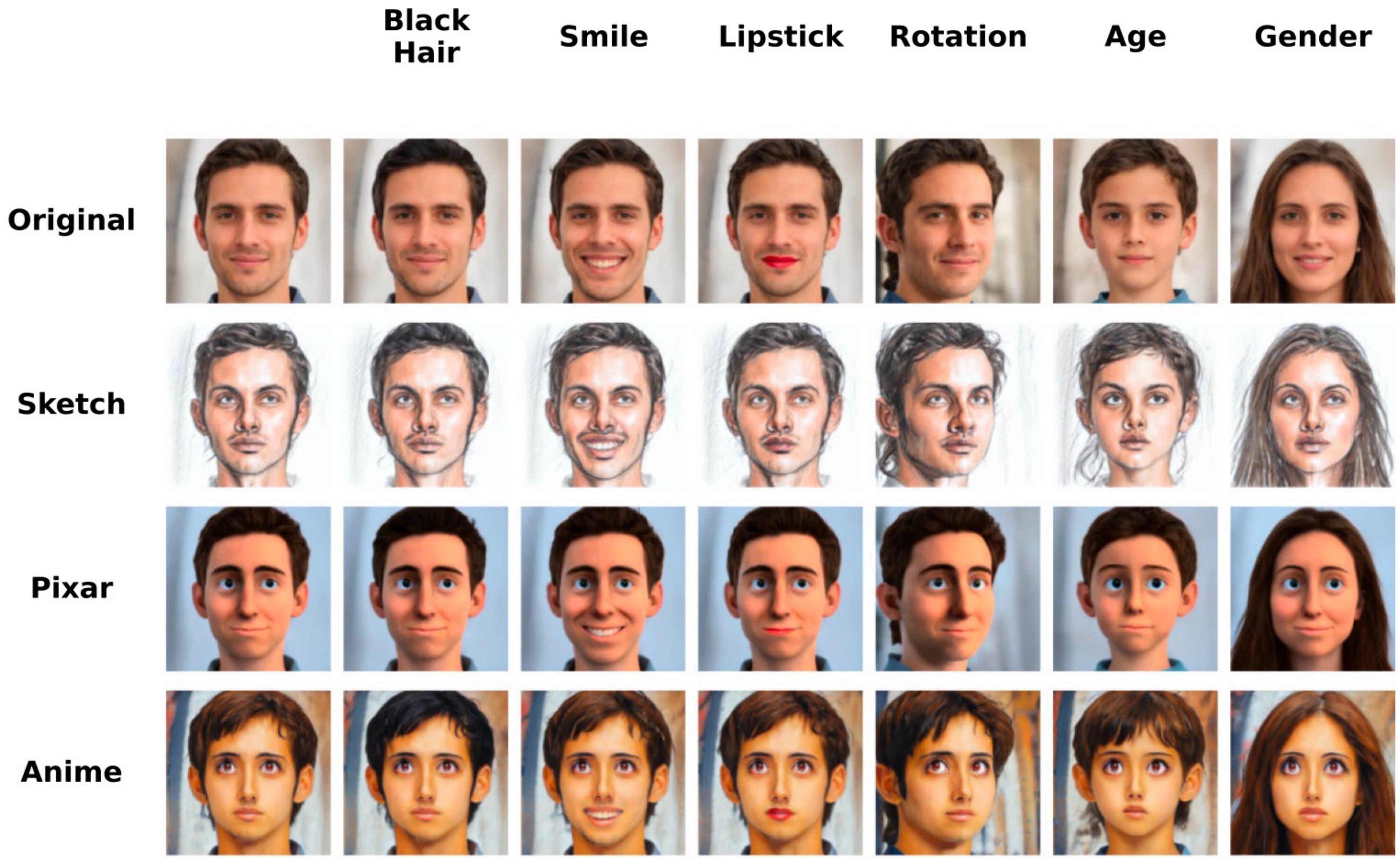
Mixing of StyleDomain directions



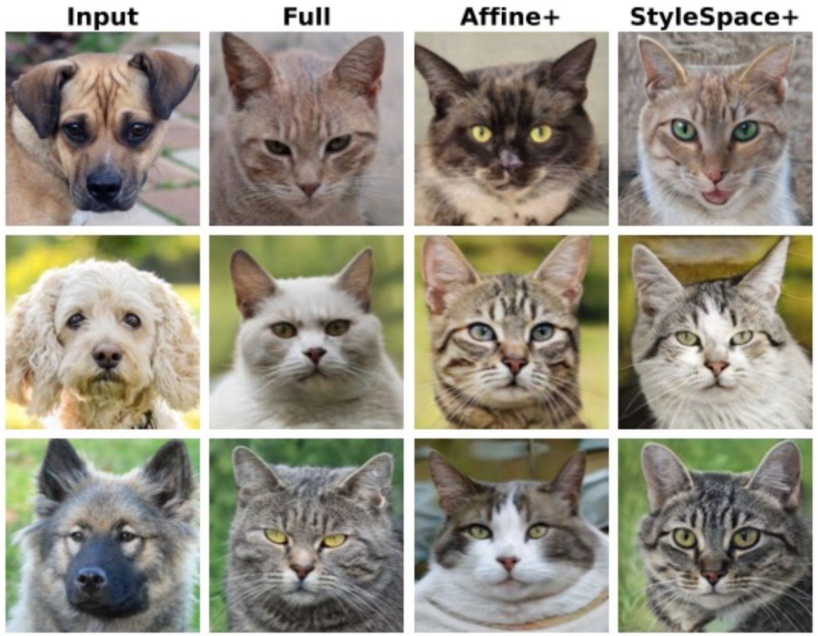
Mixing of StyleDomain directions

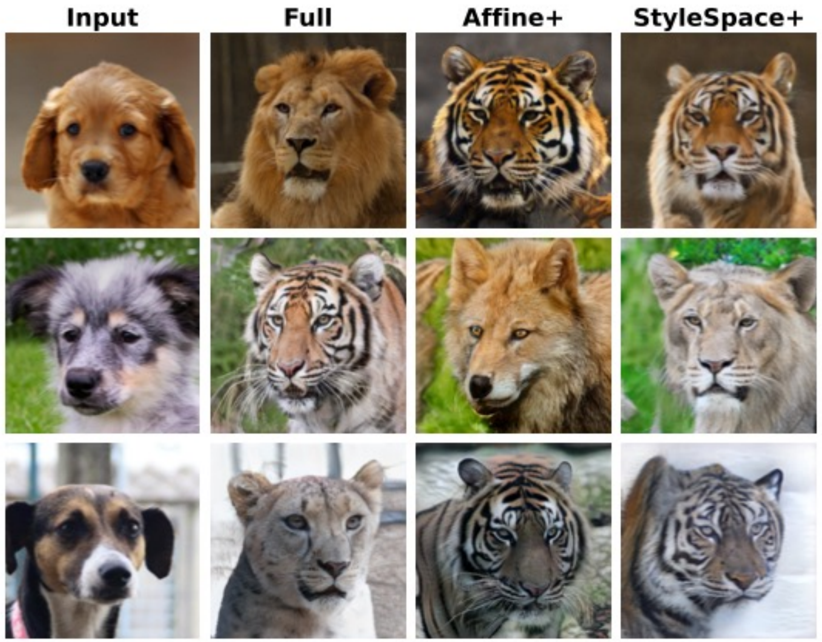


Combination with semantic editing



Cross-domain image translation

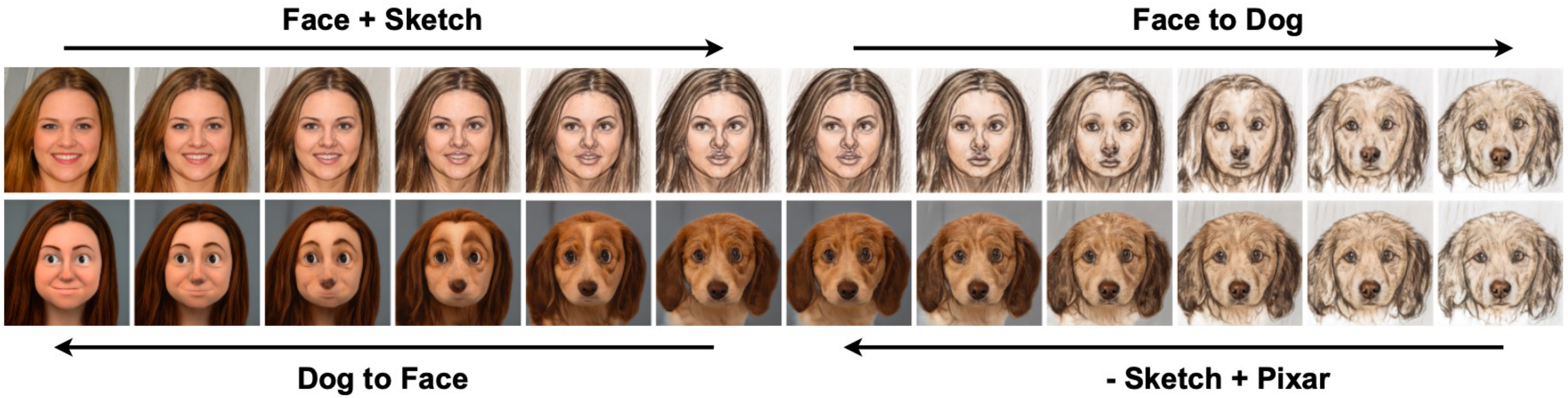




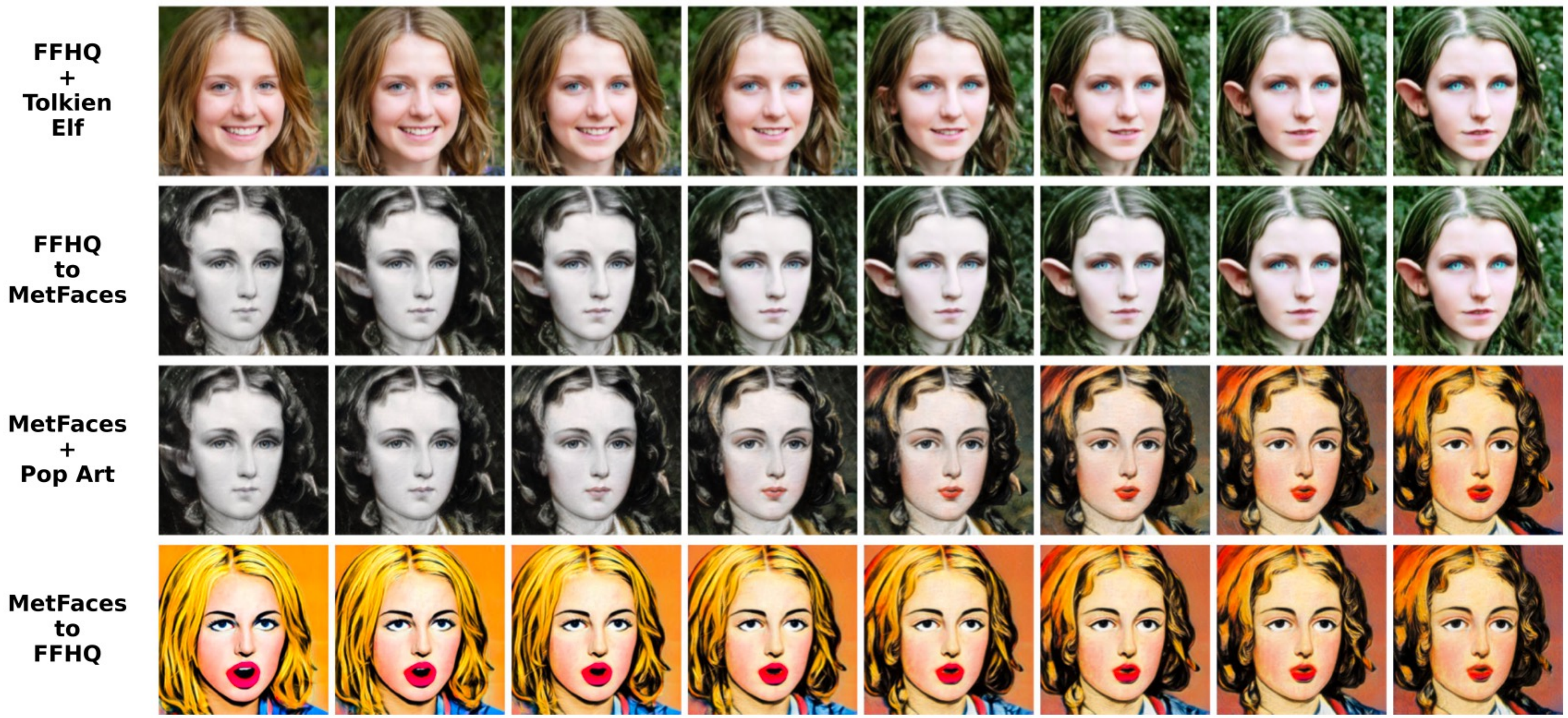
Cross-domain image translation

		cat2dog		cat2wild		dog2cat		dog2wild		wild2cat		wild2dog	
	Size	FID	KID$\times 10^3$	FID	KID$\times 10^3$	FID	KID$\times 10^3$	FID	KID$\times 10^3$	FID	KID$\times 10^3$	FID	KID$\times 10^3$
Full	60.6M	40.5	11.6	12.8	2.44	18.6	2.78	17.5	4.64	18.6	3.28	42.6	12.7
Affine+	10.2M	42.3	12.9	16.3	4.92	18.4	2.37	13.2	3.14	21.5	3.18	41.4	10.6
StyleSpace+	1.1M	54.8	24.9	25.4	11.2	21.5	5.00	19.9	6.09	21.5	3.49	53.8	22.0

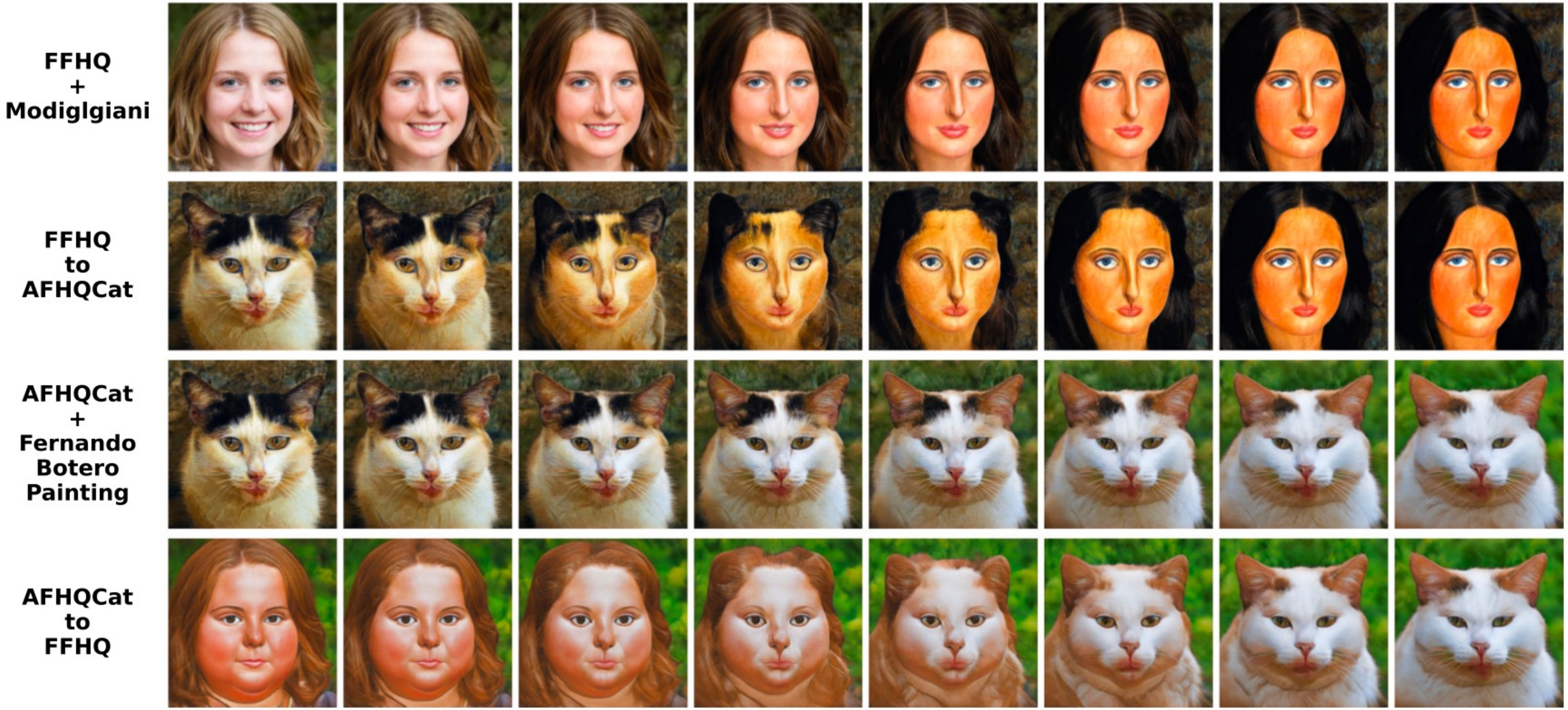
Cross-domain image morphing



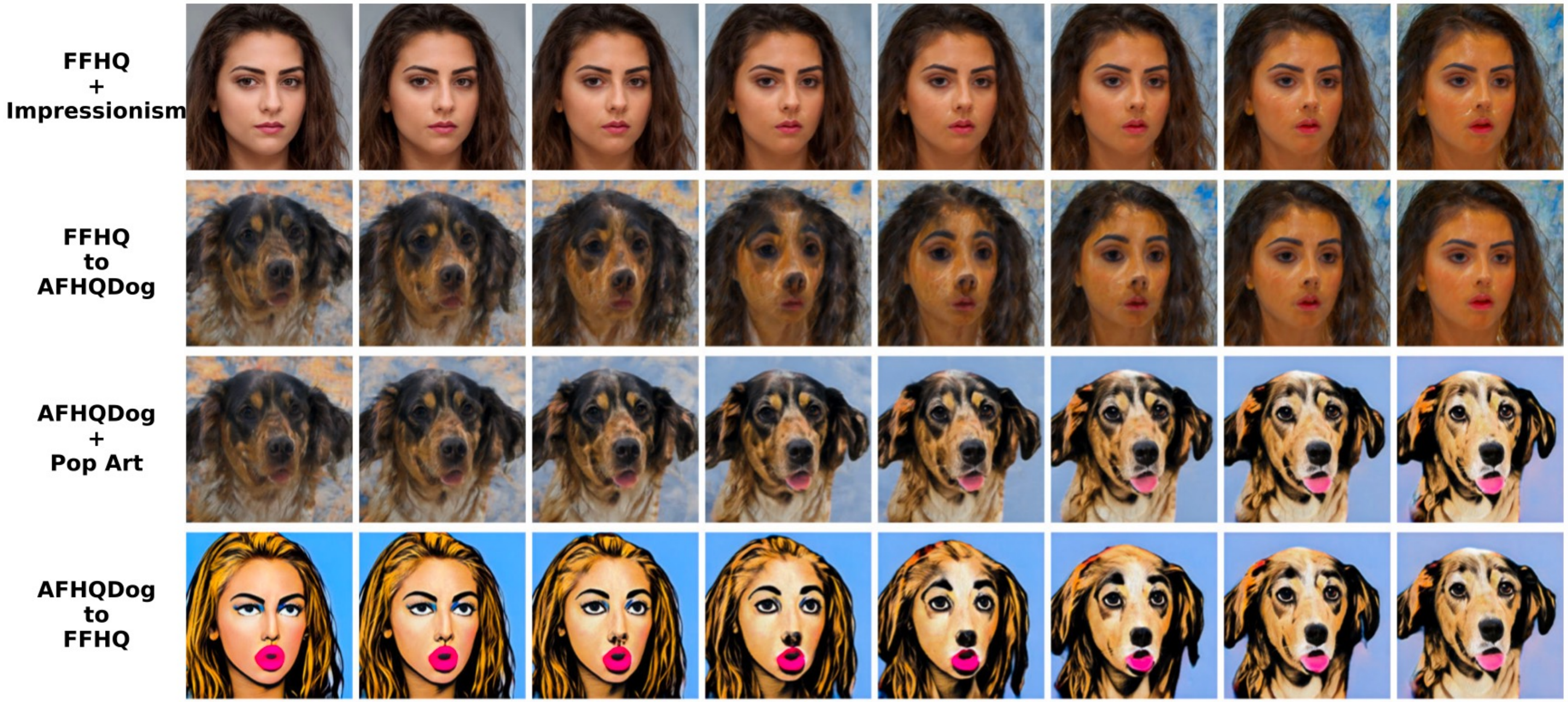
Cross-domain image morphing



Cross-domain image morphing



Cross-domain image morphing



Conclusion

- We propose the **domain-modulation technique**: allows to fune-tune 6 thousand-dimensional vector instead of 30 million weights!
- We propose the **HyperDomainNet:** predicts the domain vector given the input. It shows inspiring generalization results on unseen domains.
- We provide the **comprehensive analysis** of domain adaptation problem for StyleGAN and explore new insights.
- We propose **StyleDomain directions** and discover its surprising properties.

Thanks!

https://github.com/MACderRu/HyperDomainNet