Dual-Domain Image Synthesis Using Segmentation-Guided GAN

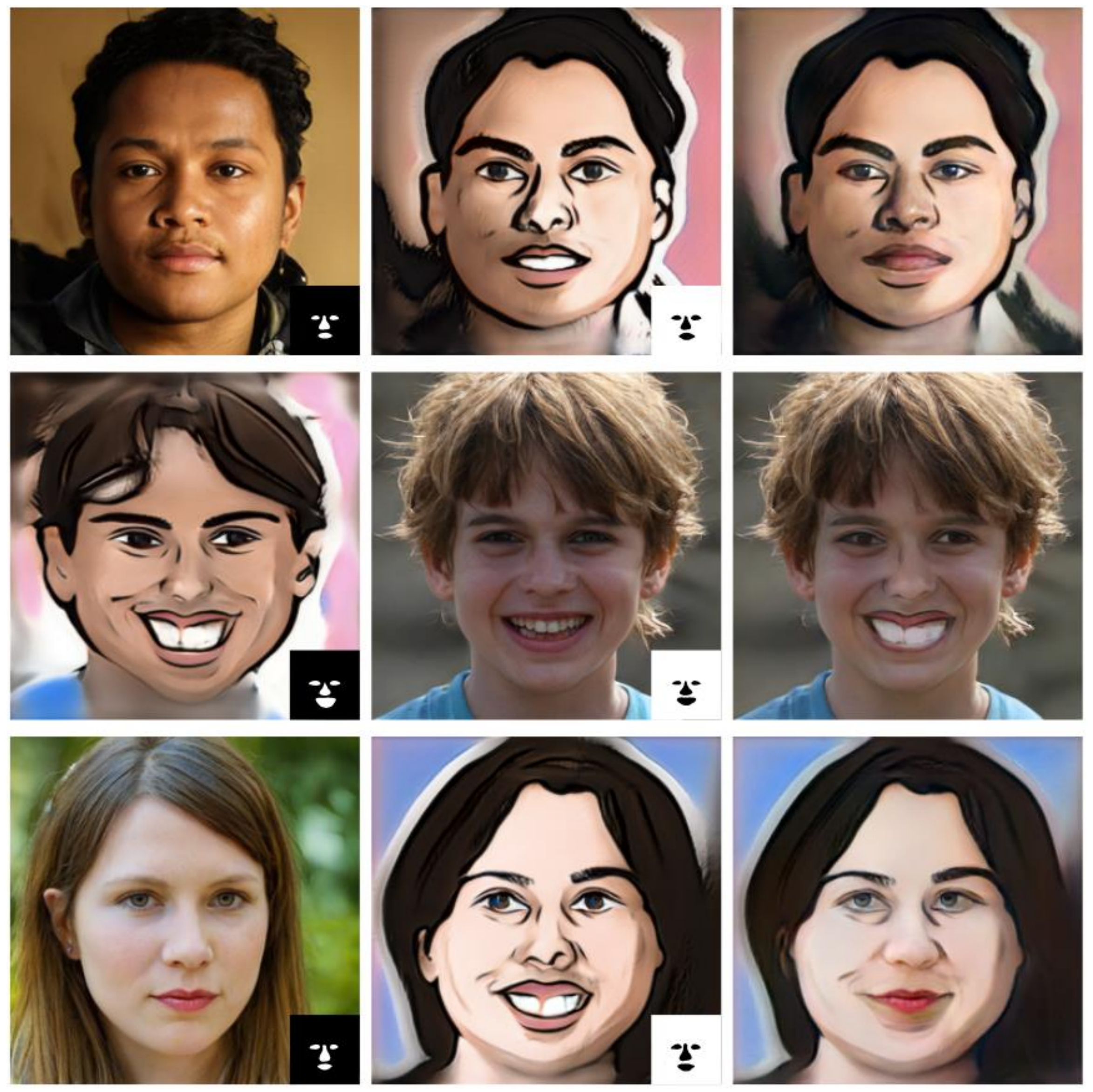
Dena Bazazian Andrew Calway Dima Damen
Department of Computer Science, University of Bristol
{dena.bazazian, andrew.calway, dima.damen}@bristol.ac.uk

CVPR JUNE 19-24 2022 NEW ORLEANS • LOUISIANA

Overview

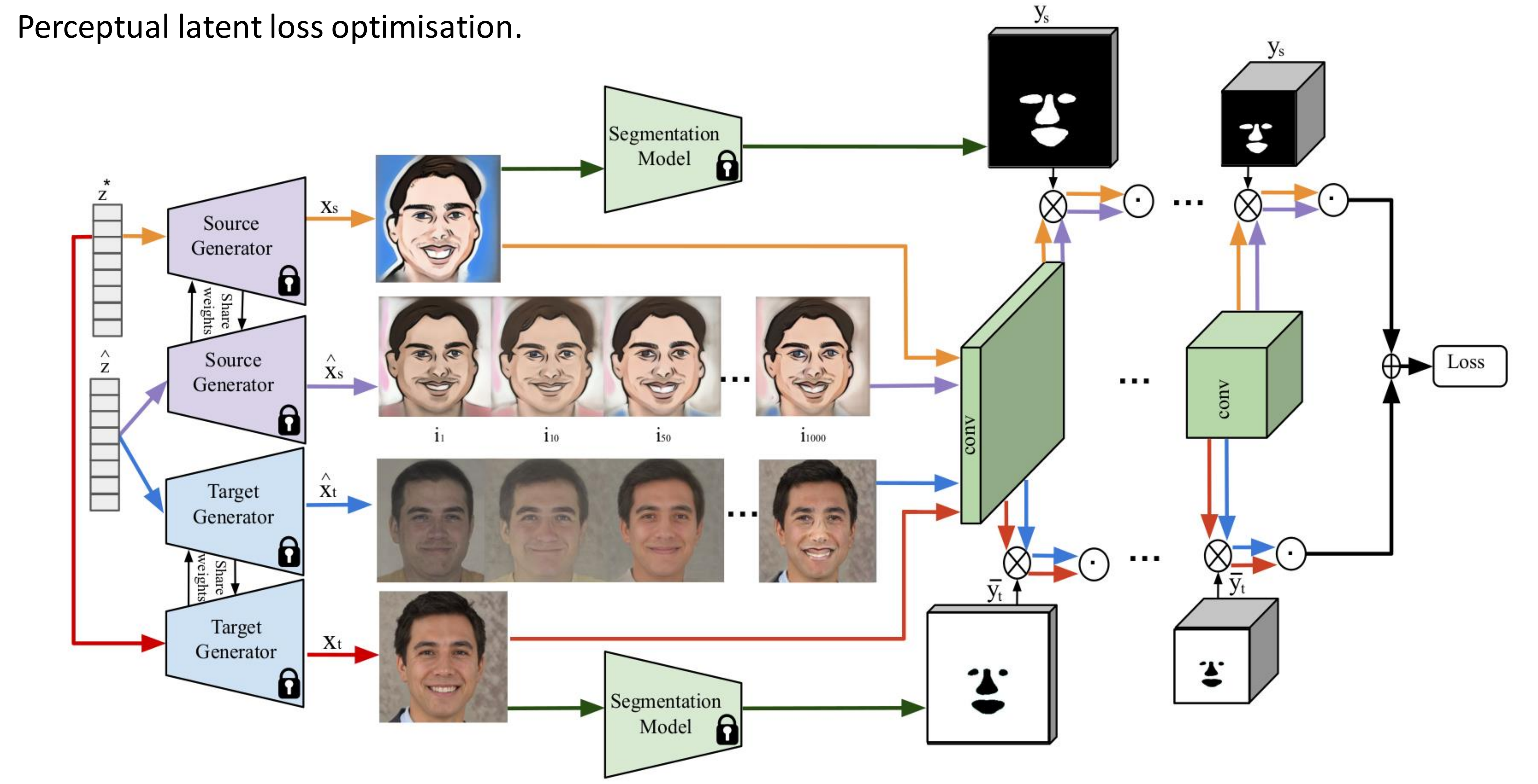
- A segmentation-guided approach to synthesise images that integrate features from two distinct domains.

- Images synthesised by our dual-domain model belong to one domain within the semantic-mask, and to another in the rest of the image - smoothly integrated.

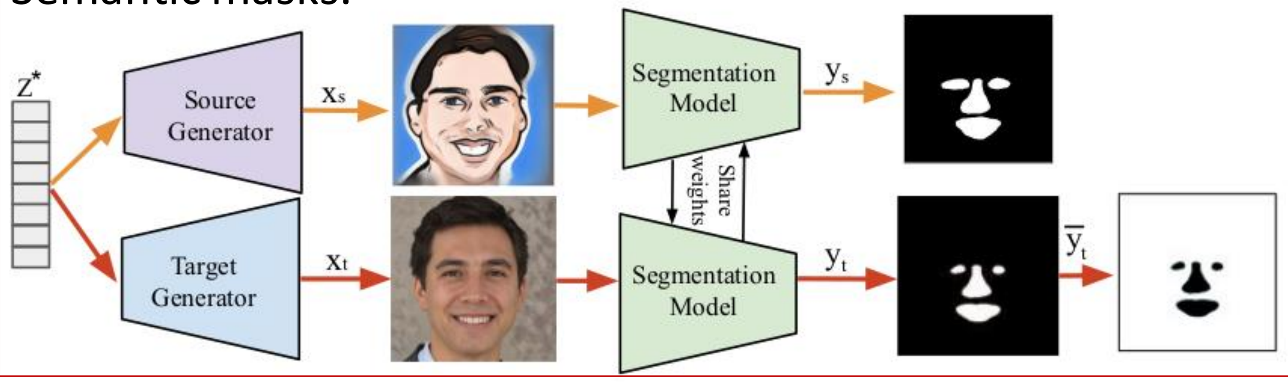


Method

Perceptual latent loss optimisation.



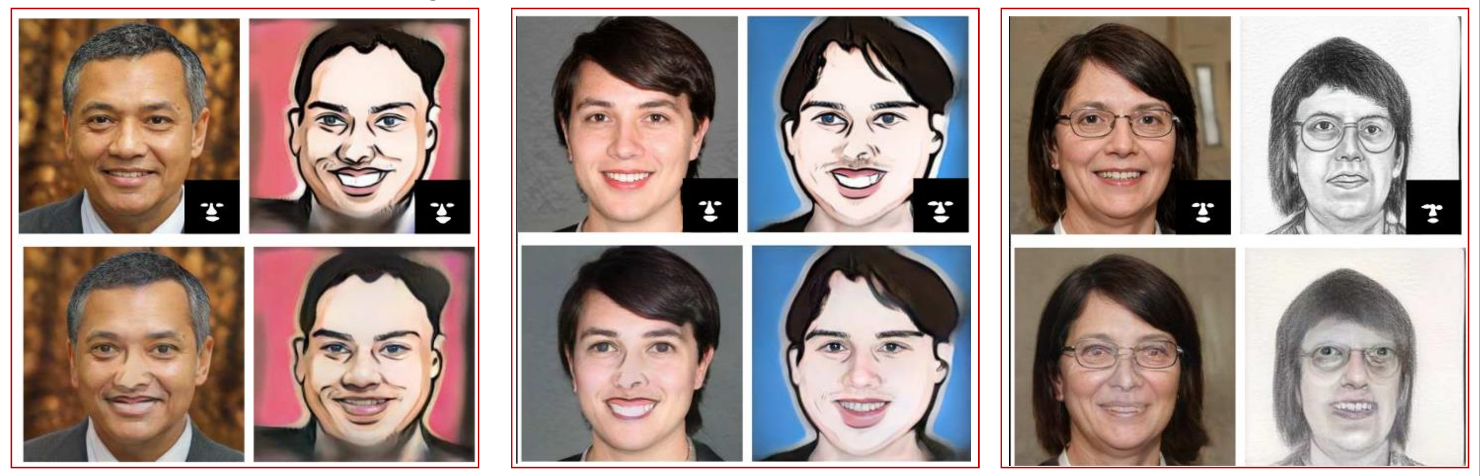
Semantic masks.



Naive crossover image.

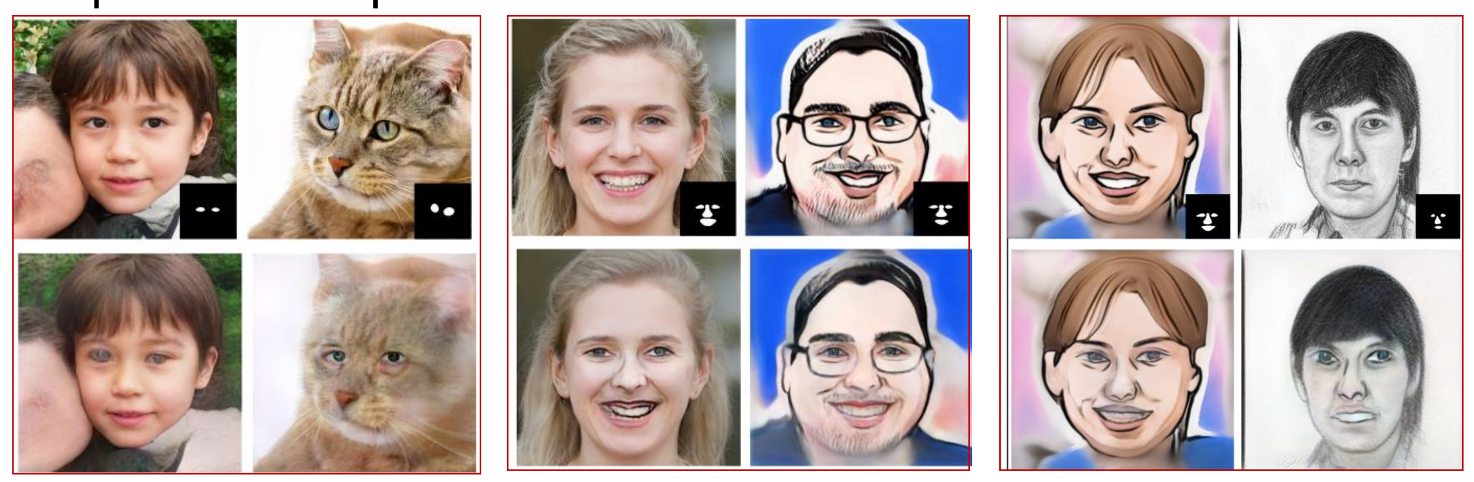
Qualitative Results

Semantic mask: eye/nose/mouth.



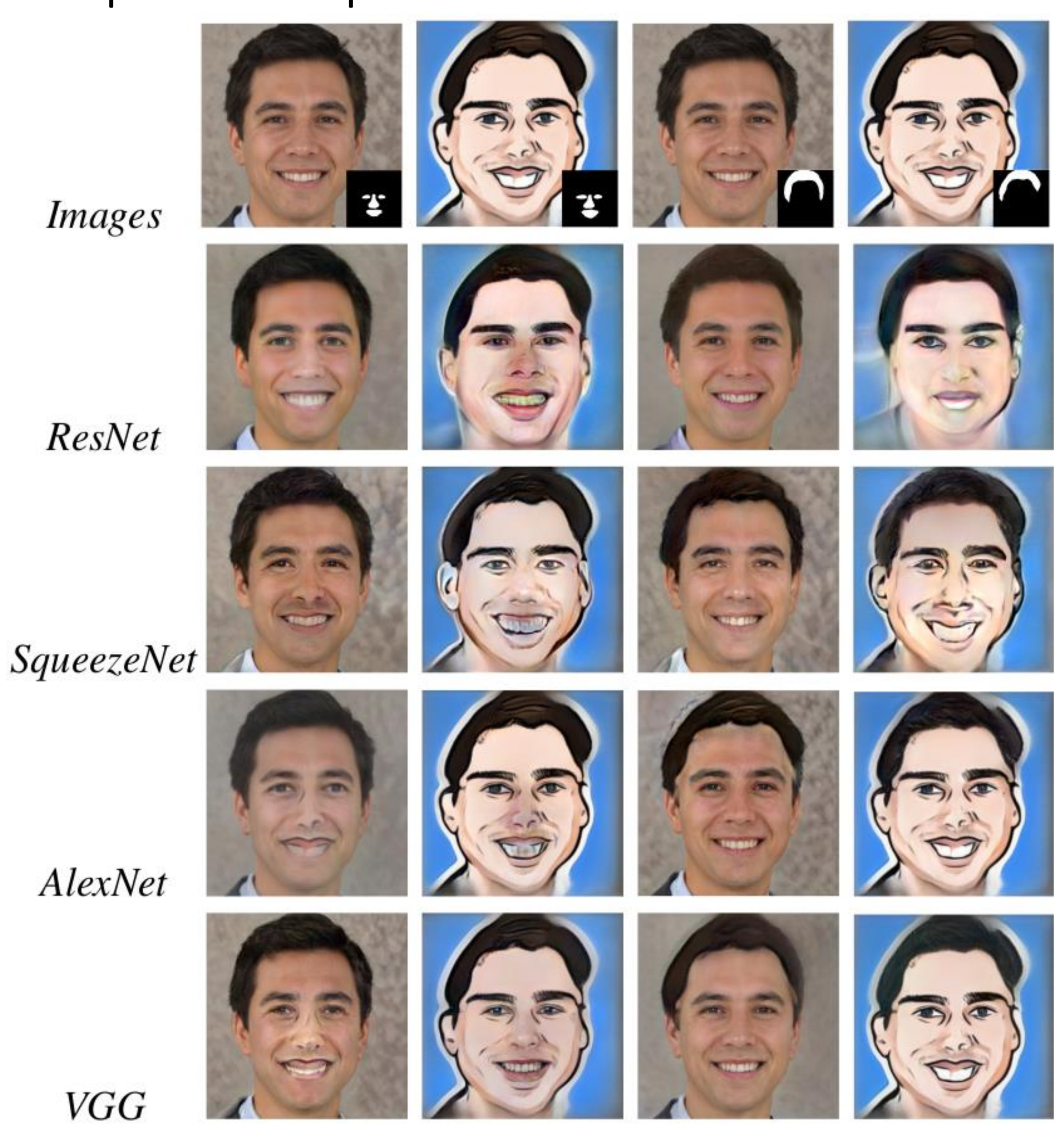
Semantic mask: hair.

Unpaired examples.



Ablation Study

The impact of four pre-trained backbones.



Quantitative Results

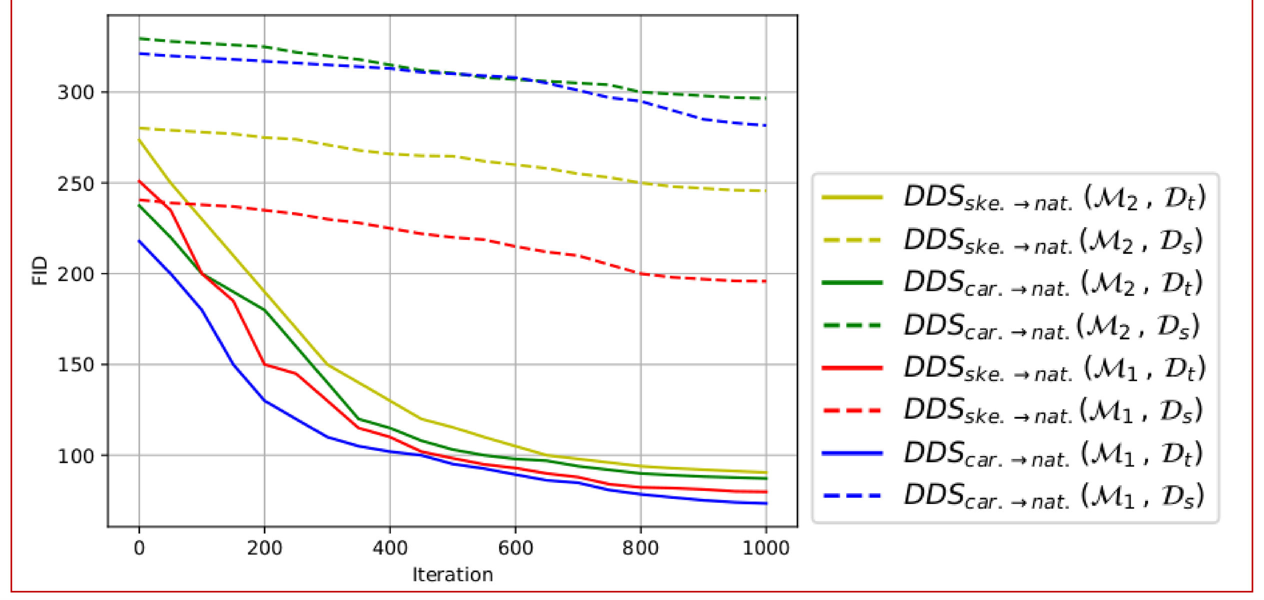
Metric comparisons on the face domains.

		$DDS_{caricature \to natural}$			$DDS_{natural \to caricature}$			$DDS_{sketch \to natural}$			$DDS_{natural \to sketch}$		
		FID↓	SSIM↑	PSNR↑	FID↓	SSIM↑	PSNR↑	FID↓	SSIM↑	PSNR↑	FID↓	SSIM↑	PSNR↑
$\mathcal{M}_1$	$\mathcal{D}_s$	281.61	0.39	27.98	259.12	0.38	27.95	195.80	0.31	27.87	225.31	0.33	27.81
	$\{x_c\}$	114.03	0.69	29.45	121.21	0.75	28.75	200.35	0.71	29.81	207.41	0.65	29.28
	$\mathcal{D}_t$	**73.51**	**0.70**	**29.51**	**74.51**	**0.76**	**28.76**	**79.87**	**0.73**	**29.84**	**97.65**	**0.66**	**29.29**
$\mathcal{M}_2$	$\mathcal{D}_s$	296.61	0.47	27.99	290.63	0.42	27.92	245.61	0.38	27.87	262.90	0.36	27.83
	$\{x_c\}$	155.91	0.73	29.63	95.30	0.71	28.43	177.03	**0.66**	29.25	152.41	0.61	**28.54**
	$\mathcal{D}_t$	**87.28**	**0.74**	**29.65**	**76.96**	**0.72**	**28.45**	**90.58**	0.64	**29.31**	**94.83**	**0.62**	28.49

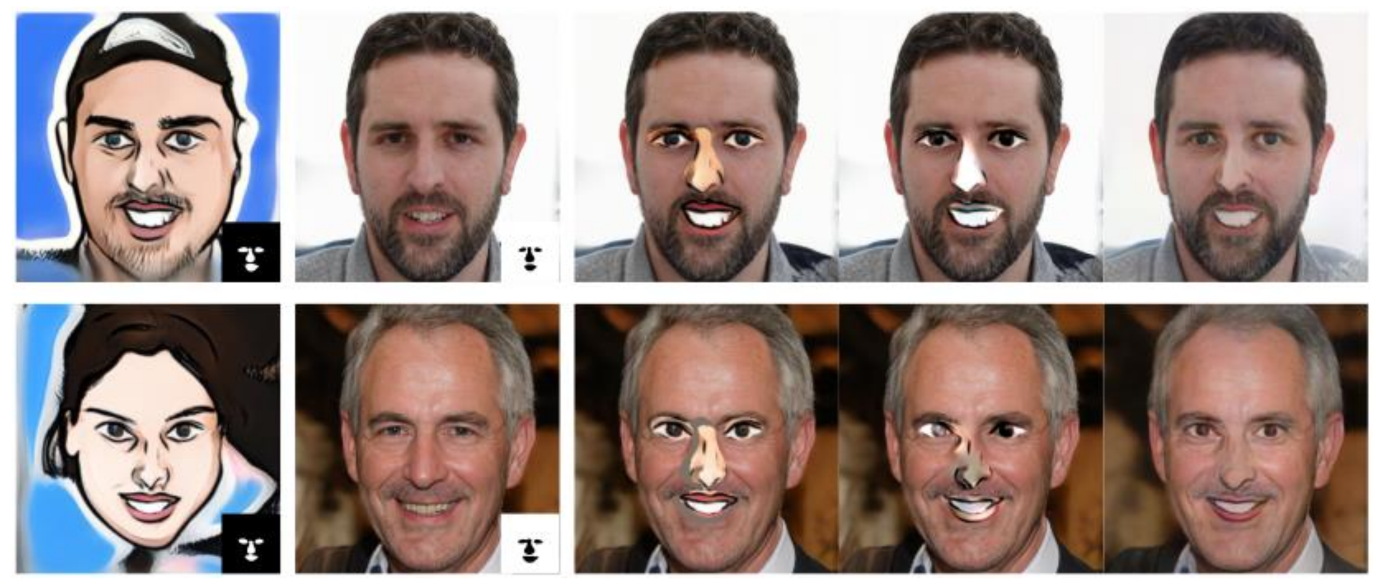
The impact of backbones.

		$DDS_{car. \to nat.}$			$DDS_{nat. \to car.}$		
		FID↓	SSIM↑	PSNR↑	FID↓	SSIM↑	PSNR↑
ResNet	$\mathcal{D}_t$	95.66	0.55	28.89	98.71	0.47	27.93
SqueezNet	$\mathcal{D}_t$	89.48	0.62	29.15	94.29	0.52	28.05
AlexNet	$\mathcal{D}_t$	88.63	0.66	29.36	93.27	0.71	28.11
VGG	$\mathcal{D}_t$	**87.28**	**0.71**	**29.58**	**76.96**	**0.72**	**28.39**

FID during latent code optimisation iterations.



Dual-Domain Synthesis vs. Image Blending



Source Target [4] [5] ours

References

[1] U. Ojha, et al., Few-shot Image Generation via Cross-domain Correspondence. CVPR, 2021.
[2] T. Karras, et al., Analyzing and Improving the Image Quality of StyleGAN. CVPR, 2020.
[3] N. Tritrong, et al., Repurposing GANs for One-shot Semantic Part Segmentation. CVPR, 2021.
[4] K. Sofiiuk et al., Foreground-aware semantic representations for image harmonization. WACV, 2021.
[5] L. Zhang et al., Deep image blending. WACV, 2020.

Code