

Jurassic World Remake

Bringing Ancient Fossils Back to Life Via Zero-Shot Image-to-Image
Translation

Project page: <https://tinyurl.com/skull2animal>

Alexander Martin, Haitian Zheng, Jie An, Jiebo Luo

Department of Computer Science, University of Rochester

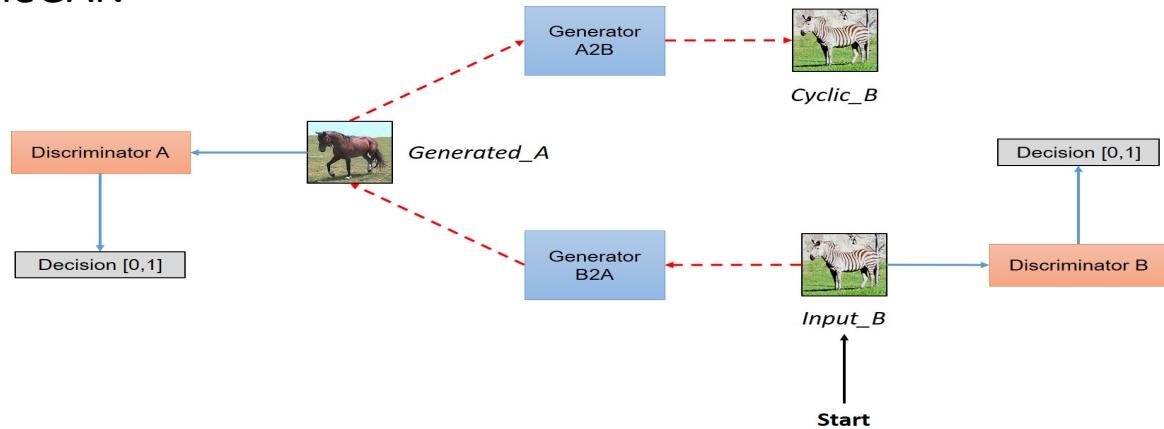
Image-to-Image Translation

- Learn a mapping Domain A → Domain B

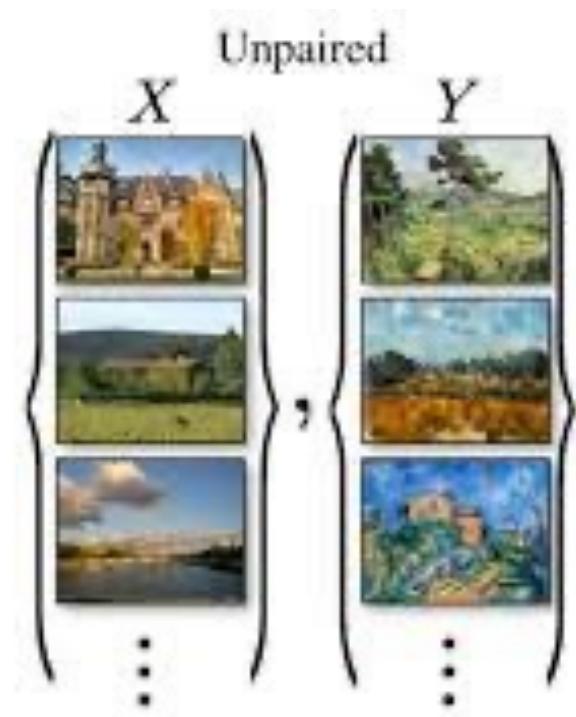
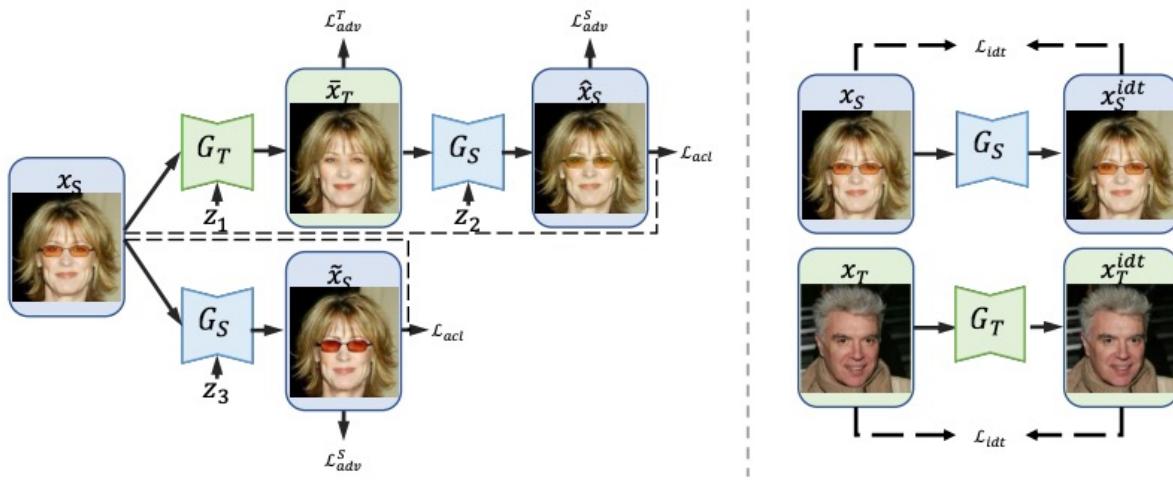


Unguided Generative Adversarial Networks

CycleGAN



ACL-GAN



Guided Diffusion Models

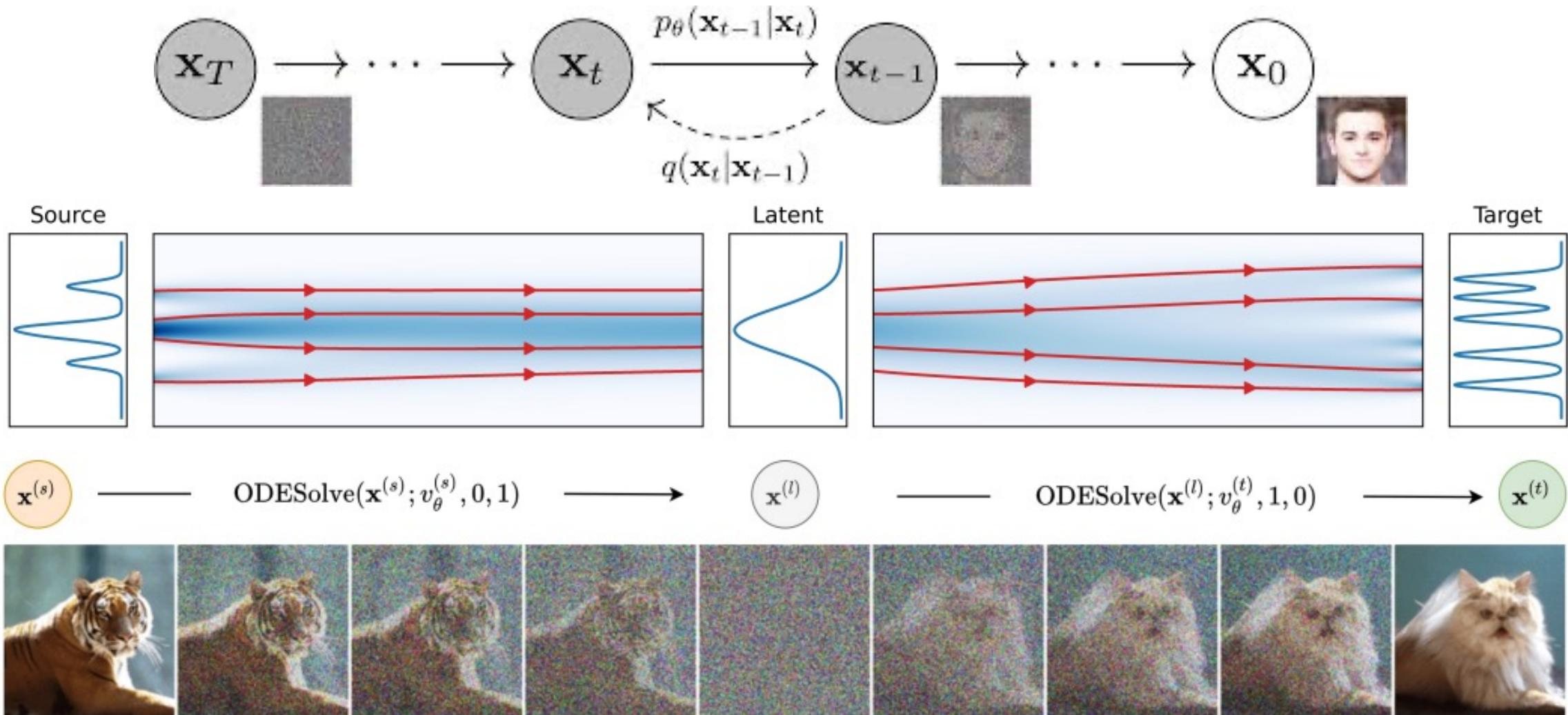


Figure From: Su et al. Dual Diffusion Implicit Bridges for Image-to-Image Translation (<https://arxiv.org/abs/2203.08382>); Song et al, Denoising Diffusion Implicit Models (<https://arxiv.org/pdf/2010.02502.pdf>)

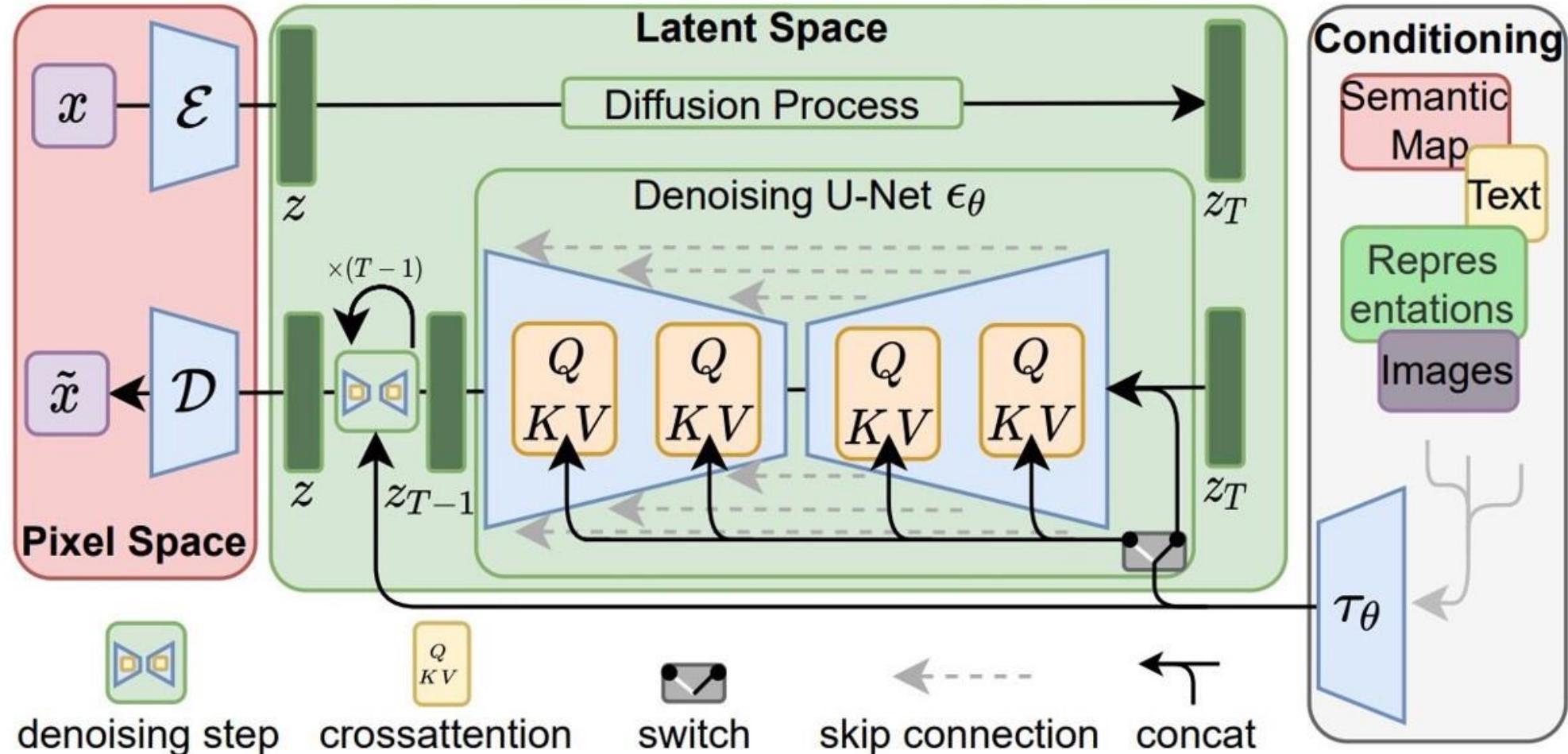
Image Editing



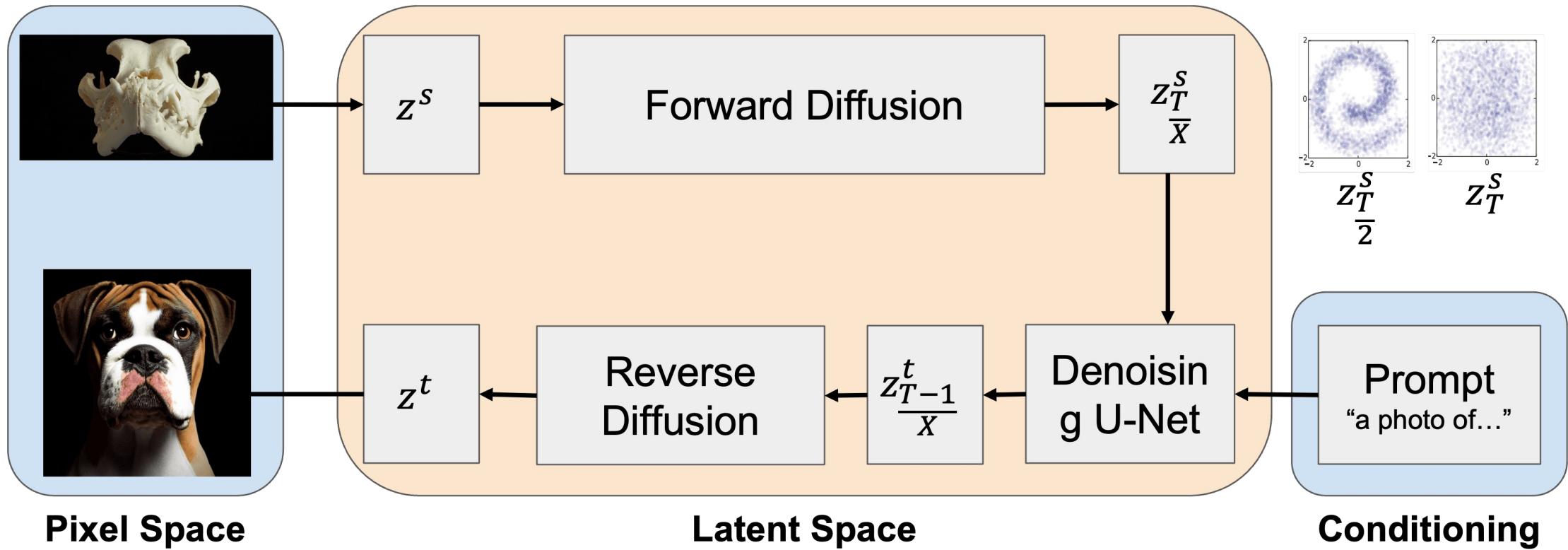
Figure From: Mokady et al, Null-text Inversion for Editing Real Images using Guided Diffusion Models, (<https://arxiv.org/pdf/2211.09794.pdf>)



Latent Diffusion Models

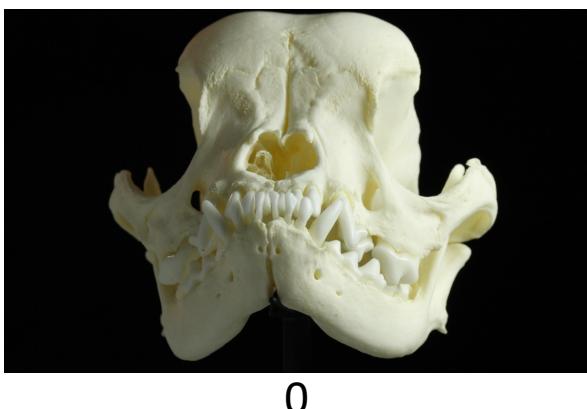
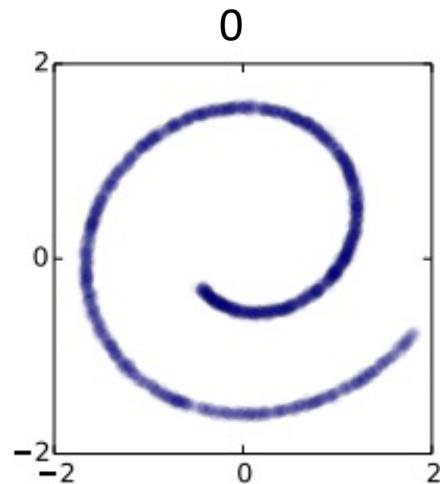


Full Process

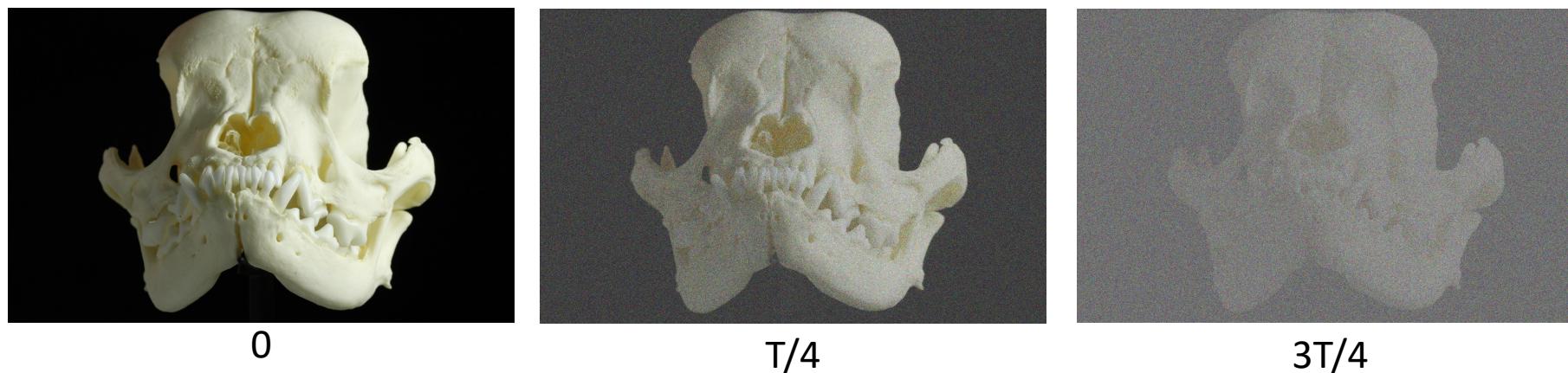
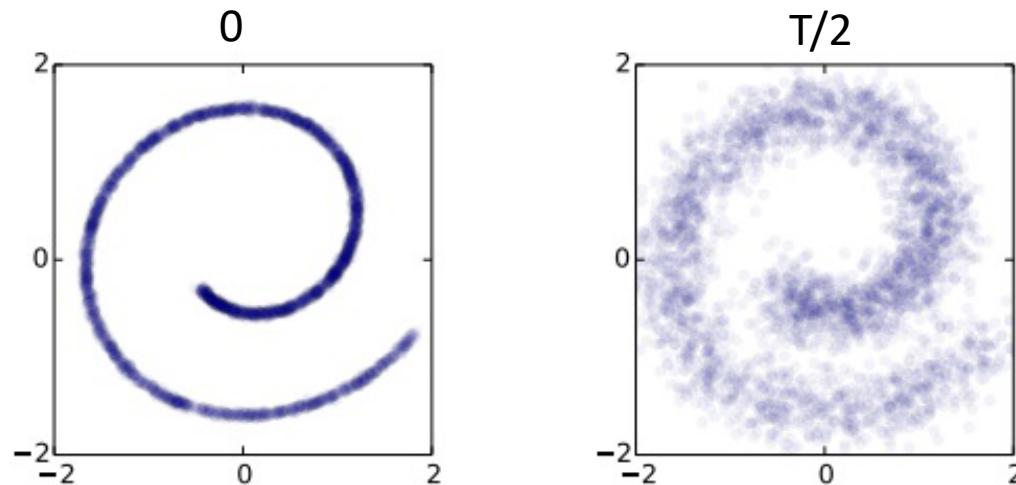


Prompt: "A photo of the head of a <Class Label>"

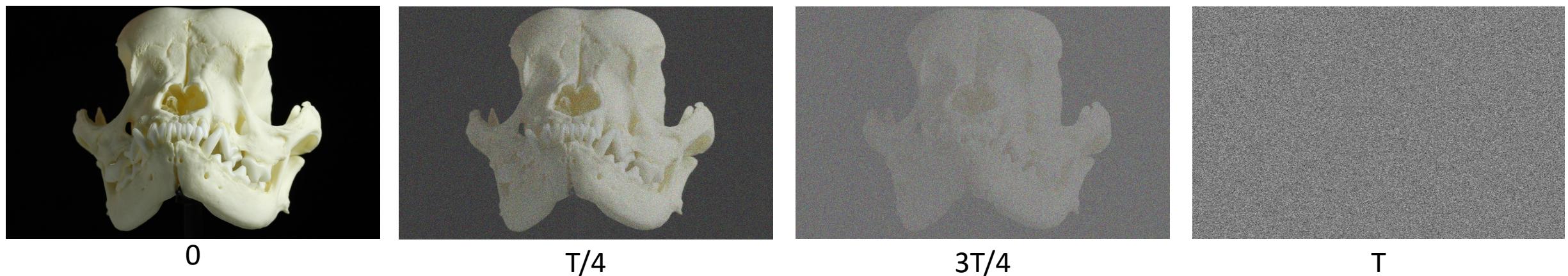
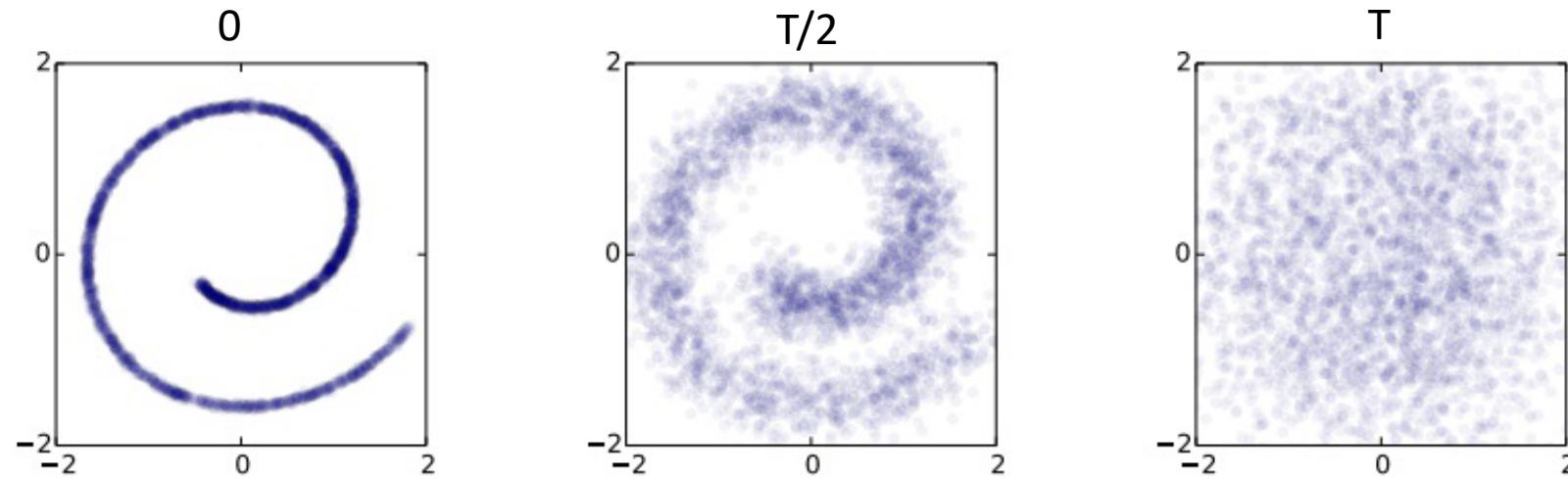
Partial Diffusion



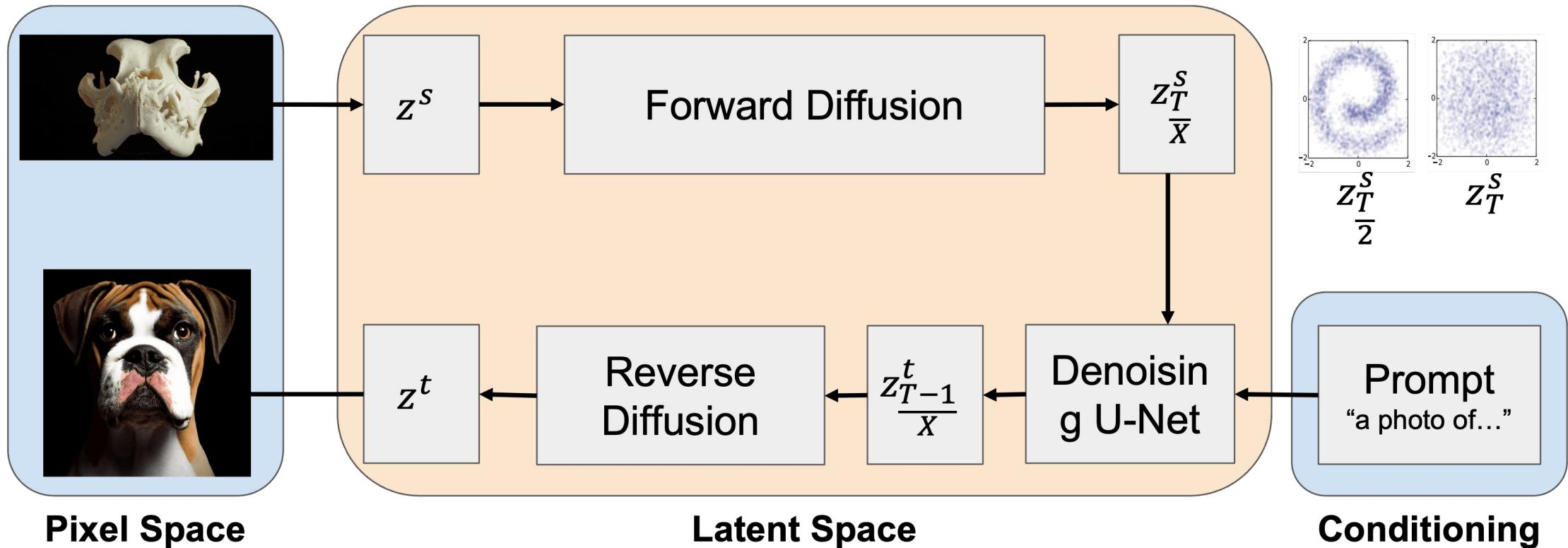
Partial Diffusion



Partial Diffusion

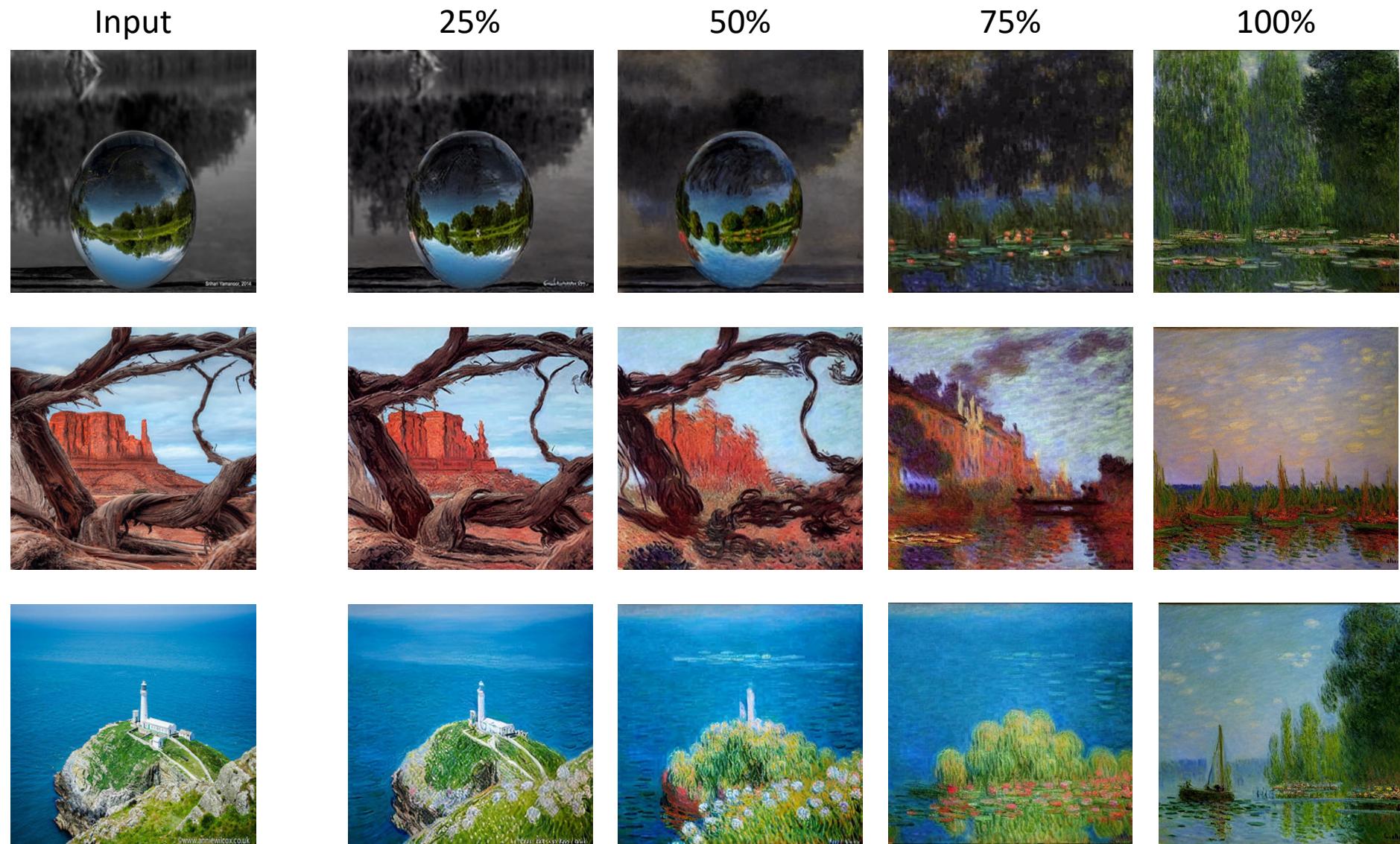


Full Process

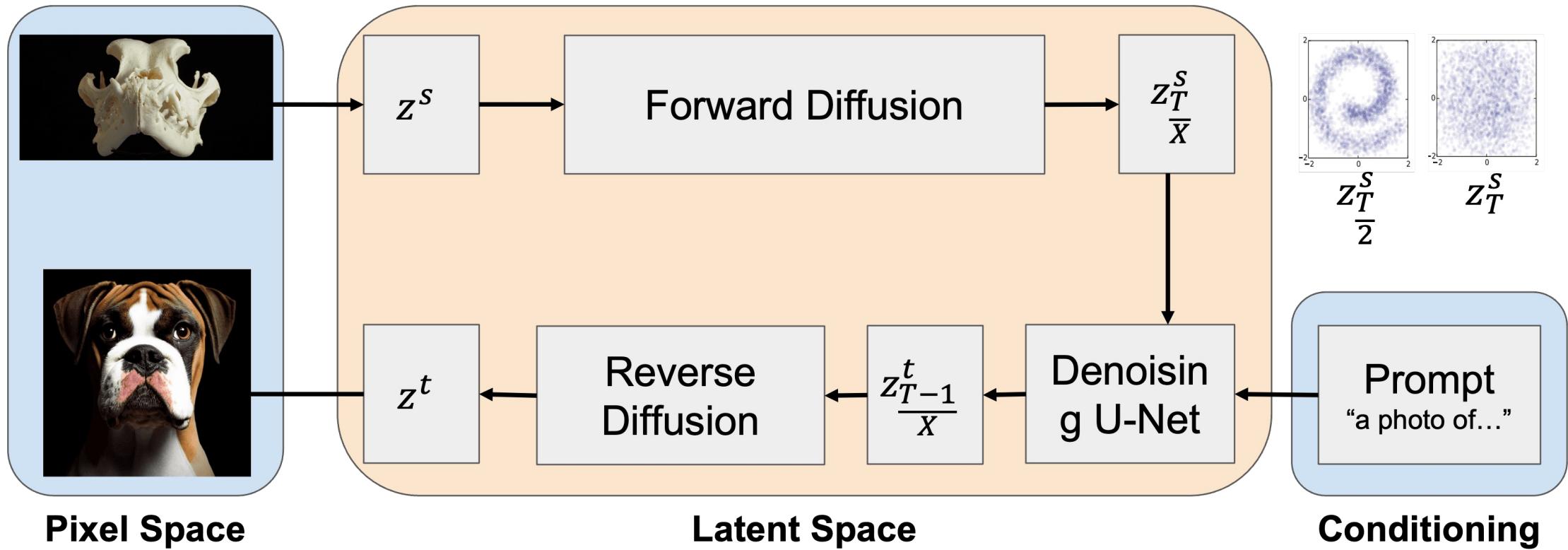


Prompt: "A photo of the head of a <Class Label>"

Results of Partial Diffusions



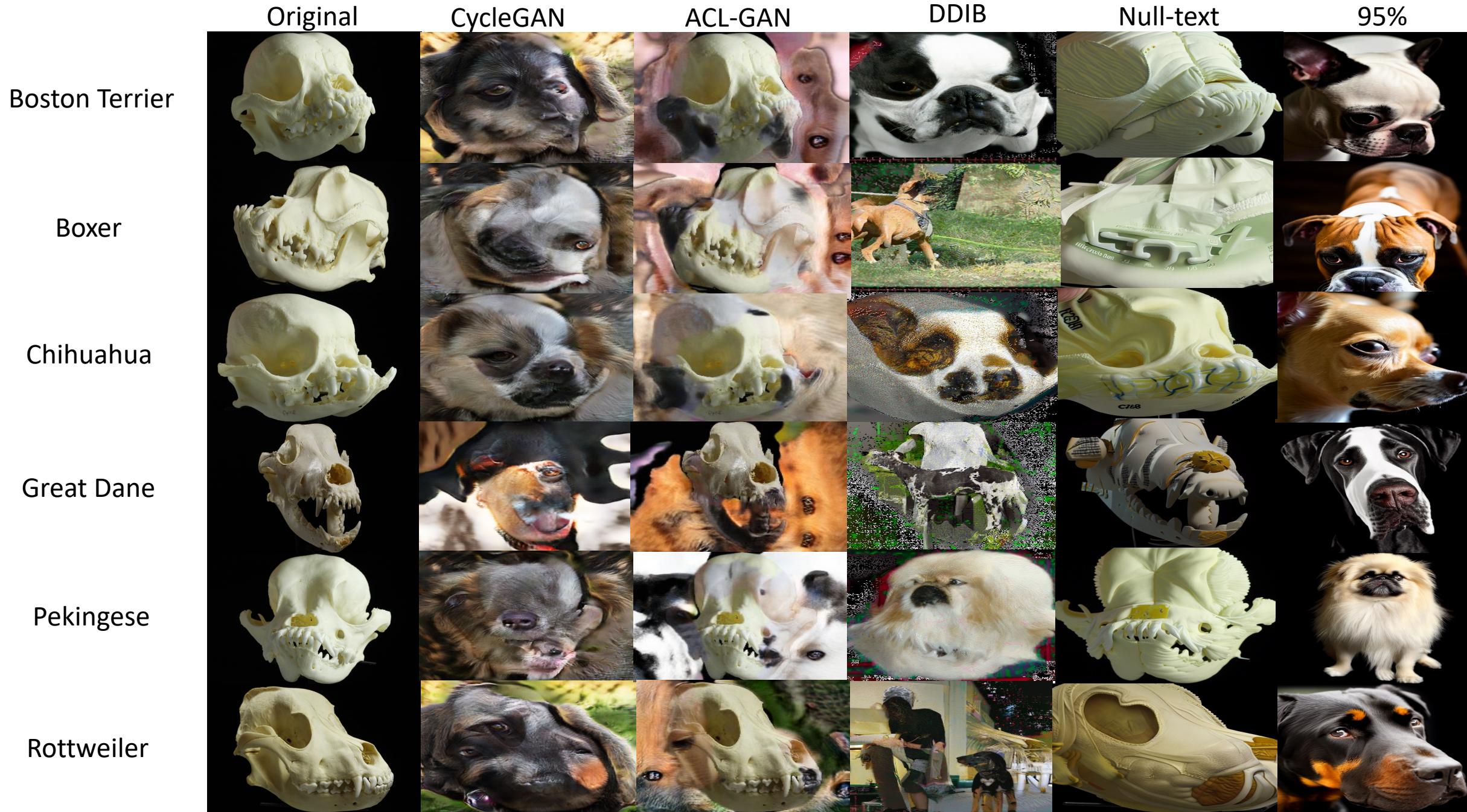
Full Process



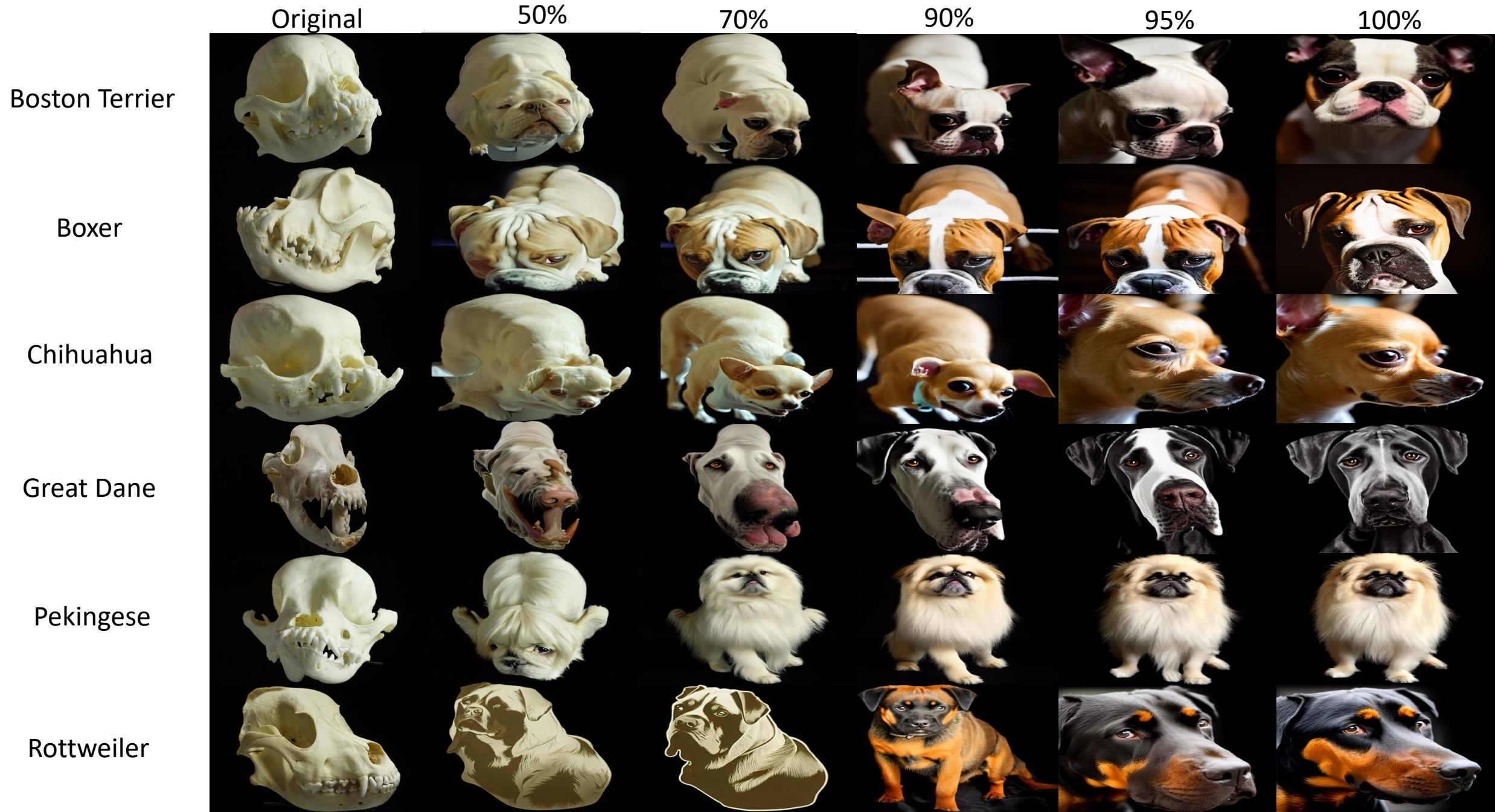
Prompt: “A photo of the head of a <Class Label>”

Table 1: Quantitative evaluation on Skull2Dog. Lower scores are better for metrics with down arrows (\downarrow), and vice versa.

	FID \downarrow	KID \downarrow	All @1 \uparrow	Class @1 \uparrow
CycleGAN [48]	212.66	0.1325	77.00	5.70
ACL-GAN [45]	342.16	0.3134	38.84	1.65
DDIB [38]	204.42	0.0698	73.68	56.14
Null-text [26]	308.19	0.1559	1.79	1.79
Revive-2I_{95%}	143.29	0.0345	100	92.56



	FID ↓	KID↓	All @1↑	Class @1↑
Revive-2I _{50%}	236.59	0.118	43.80	20.66
Revive-2I _{60%}	177.30	0.0584	70.25	34.70
Revive-2I _{70%}	155.066	0.0437	93.89	59.50
Revive-2I _{80%}	147.85	0.0368	100	67.77
Revive-2I _{90%}	149.42	0.0365	100	89.26
Revive-2I_{95%}	143.29	0.0345	100	92.56
Revive-2I _{100%}	155.9	0.0391	100	92.56





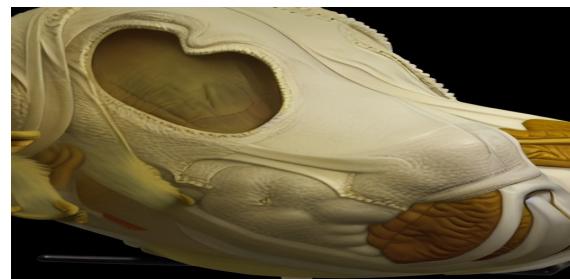
Constraining the Output



Constraining the Output



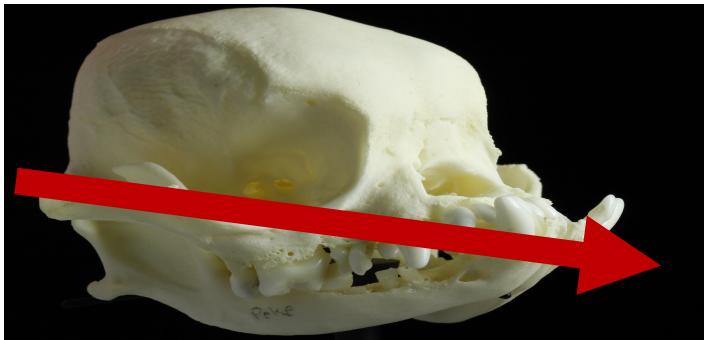
Constraining the Output



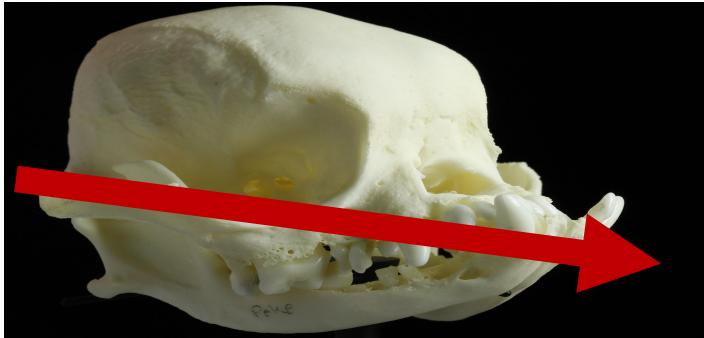
Faithfulness to the Source



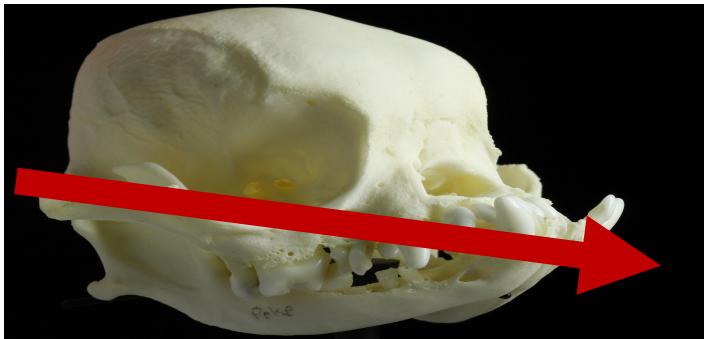
Faithfulness to the Source



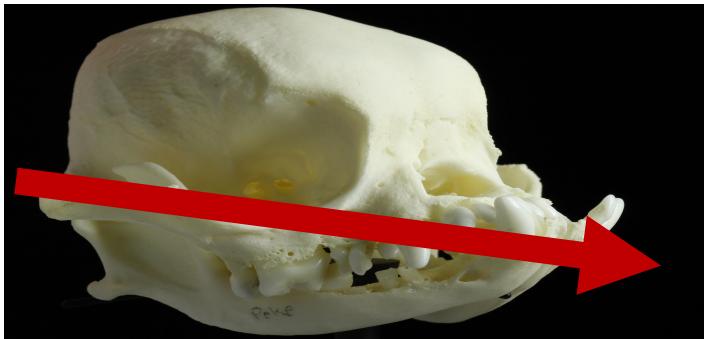
Faithfulness to the Source



Faithfulness to the Source



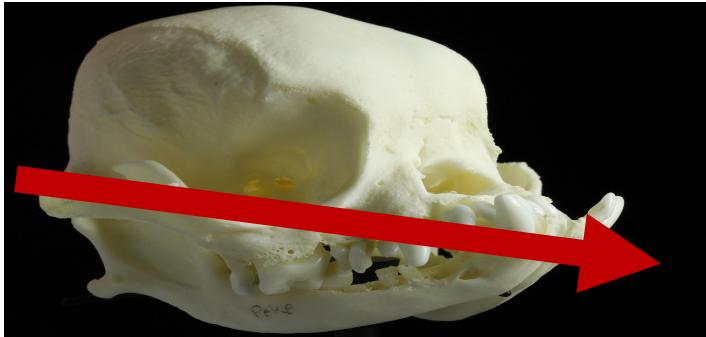
Faithfulness to the Source



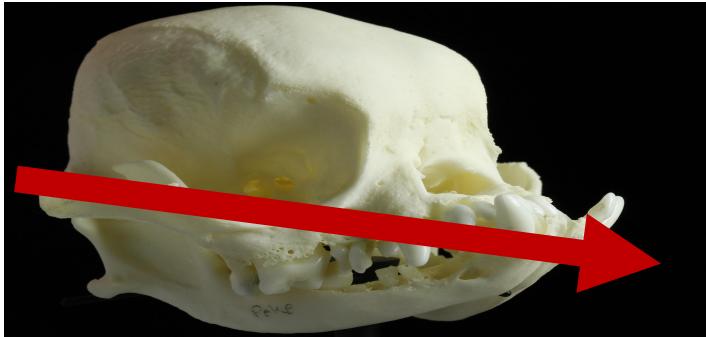
Faithfulness to the Source



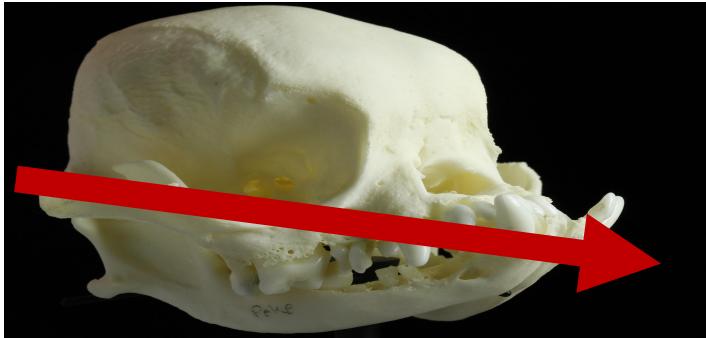
Faithfulness to the Source



Faithfulness to the Source



Faithfulness to the Source







Alexander Martin
amart50@u.rochester.edu
<https://tinyurl.com/skull2animal>

Appendix

What's the difference between 95% and 100% model?

- 95% is more faithful to the source content than the 100%
- Faithfulness is determined by looking the same direction as the original and not hallucinating new features. (Hallucinating new features isn't a big difference because both models hallucinate whole body dogs at different points)

Original

100%

95%

Boston Terrier



Boxer



Great Dane



Rottweiler



Table 1: Quantitative evaluation on Skull2Dog. Lower scores are better for metrics with down arrows (\downarrow), and vice versa.

	FID \downarrow	KID \downarrow	All @1 \uparrow	Class @1 \uparrow
CycleGAN [48]	212.66	0.1325	77.00	5.70
ACL-GAN [45]	342.16	0.3134	38.84	1.65
DDIB [38]	204.42	0.0698	73.68	56.14
Null-text [26]	308.19	0.1559	1.79	1.79
Revive-2I _{50%}	236.59	0.118	43.80	20.66
Revive-2I _{60%}	177.30	0.0584	70.25	34.70
Revive-2I _{70%}	155.066	0.0437	93.89	59.50
Revive-2I _{80%}	147.85	0.0368	100	67.77
Revive-2I _{90%}	149.42	0.0365	100	89.26
Revive-2I_{95%}	143.29	0.0345	100	92.56
Revive-2I _{100%}	155.9	0.0391	100	92.56