

# Distilling Diffusion Models into Conditional GANs

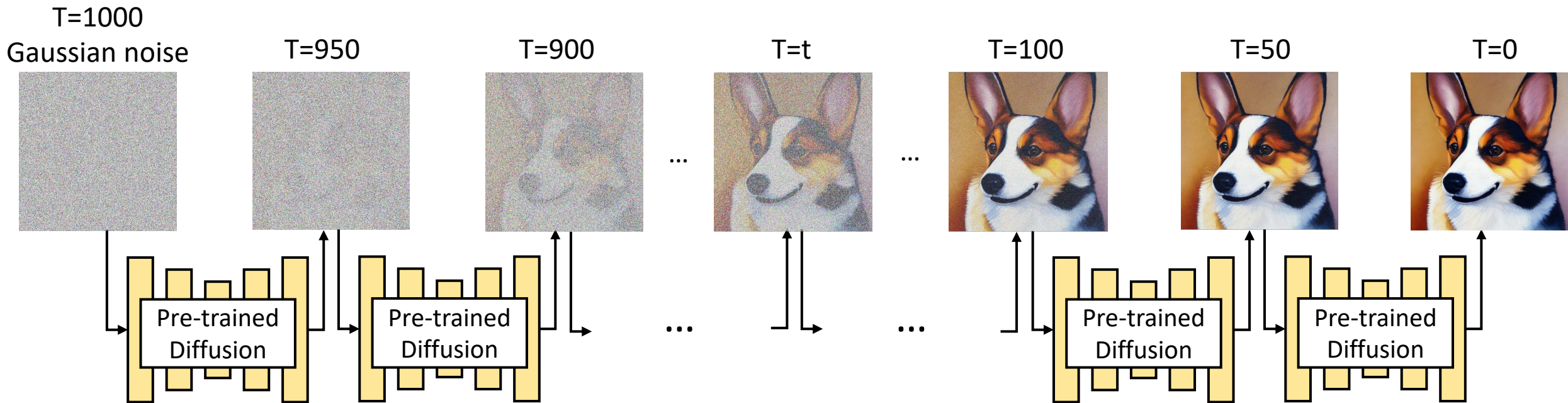
Minguk Kang, Richard Zhang, Connelly Barnes, Sylvain Paris, Suha Kwak,  
Jaesik Park, Eli Shechtman, Jun-Yan Zhu, and Taesung Park



**Carnegie  
Mellon  
University**



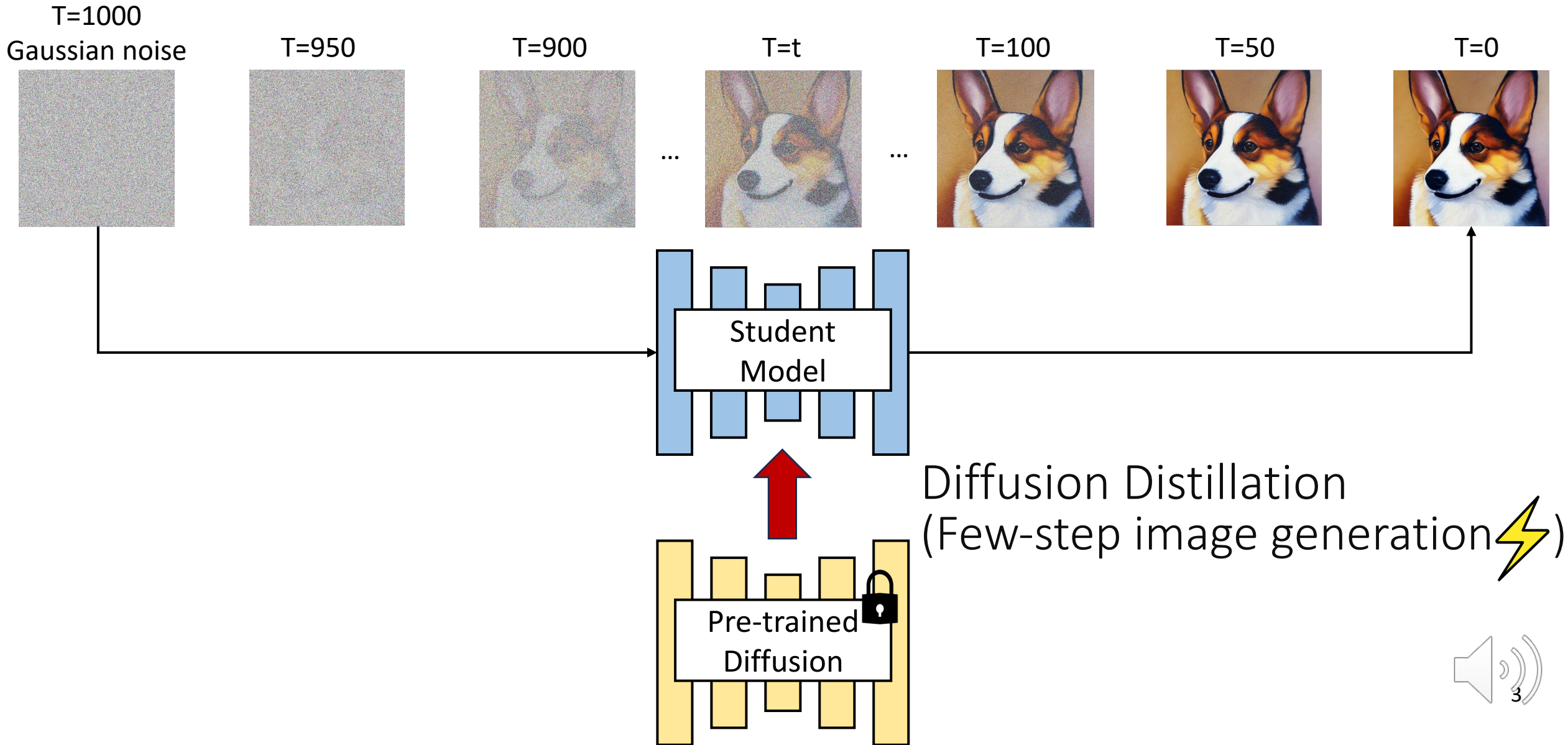
# What is diffusion model distillation



Requires 20~50 steps of denoising to generate a single image.

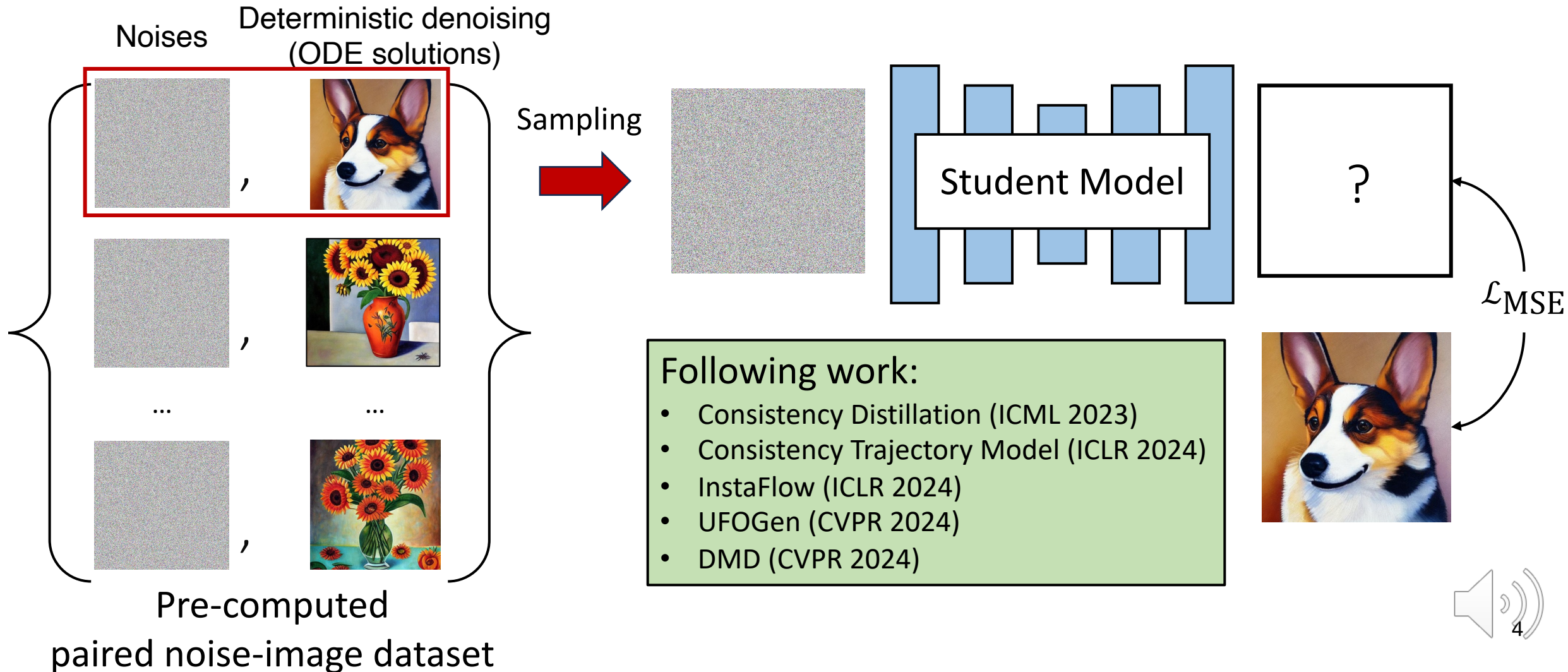


# What is diffusion model distillation



# How can we distill a diffusion model?

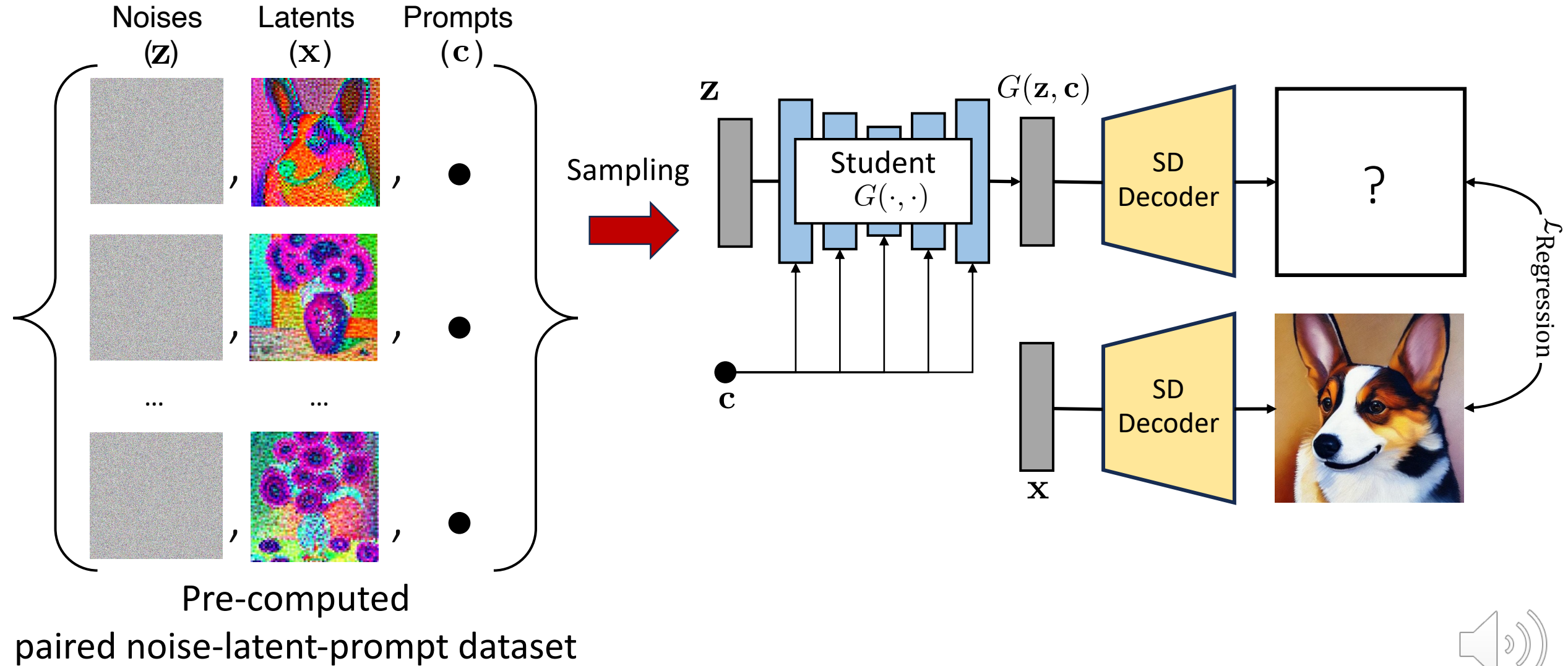
Direct regression approach: Luhman et al., 2021.



Is direct regression no longer not  
necessary?

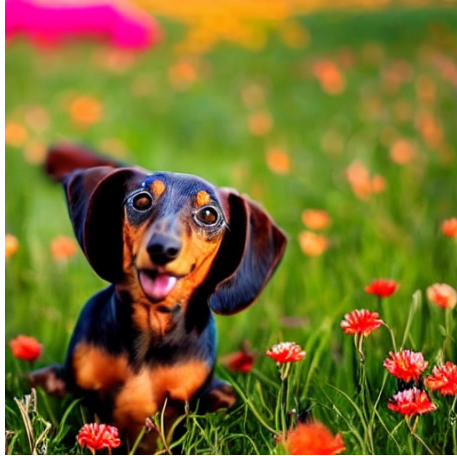


# Re-discovery of direct regression training

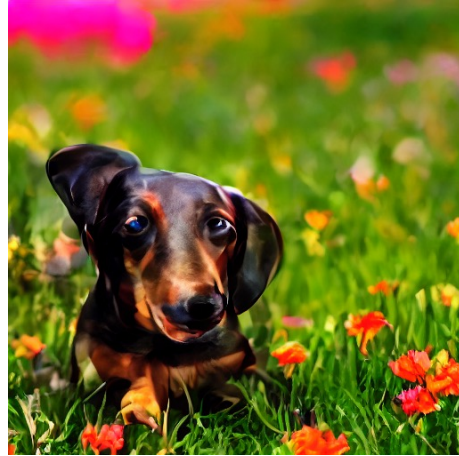


# Perceptual loss is what we need

Target



LPIPS Outputs

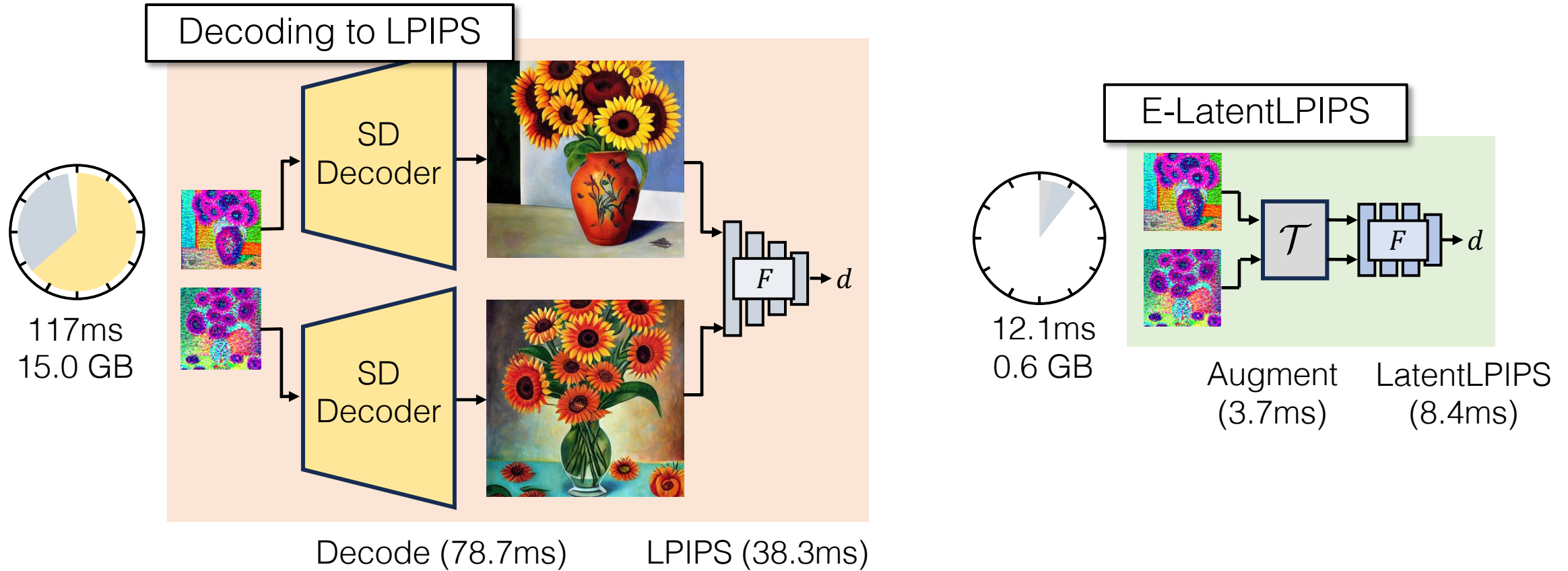


Loss function	img/sec	FID	CLIP-score
MSE	138.4	110.55	0.222
<b>LPIPS</b>	<b>40.0</b>	<b>25.94</b>	<b>0.288</b>

- 😊 Direct regression using LPIPS = Good
- 😭 Pixel-space LPIPS is expensive.
- 😭 Decoding can introduce errors in LPIPS.



# E-LatentLPIPS for efficient LPIPS

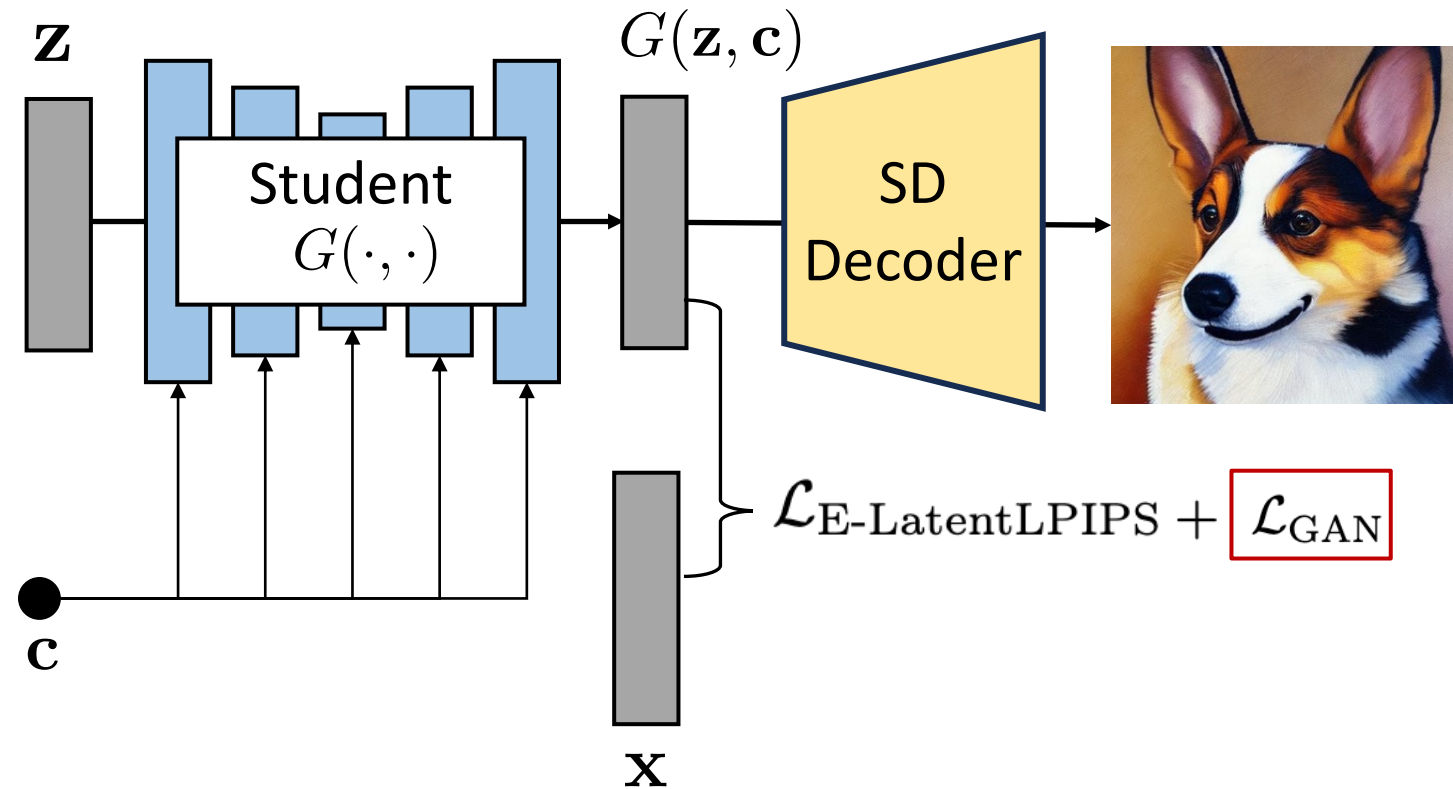


9.7× speedup achieved with E-LatentLPIPS! ⚡





# View the distillation as image-to-image translation

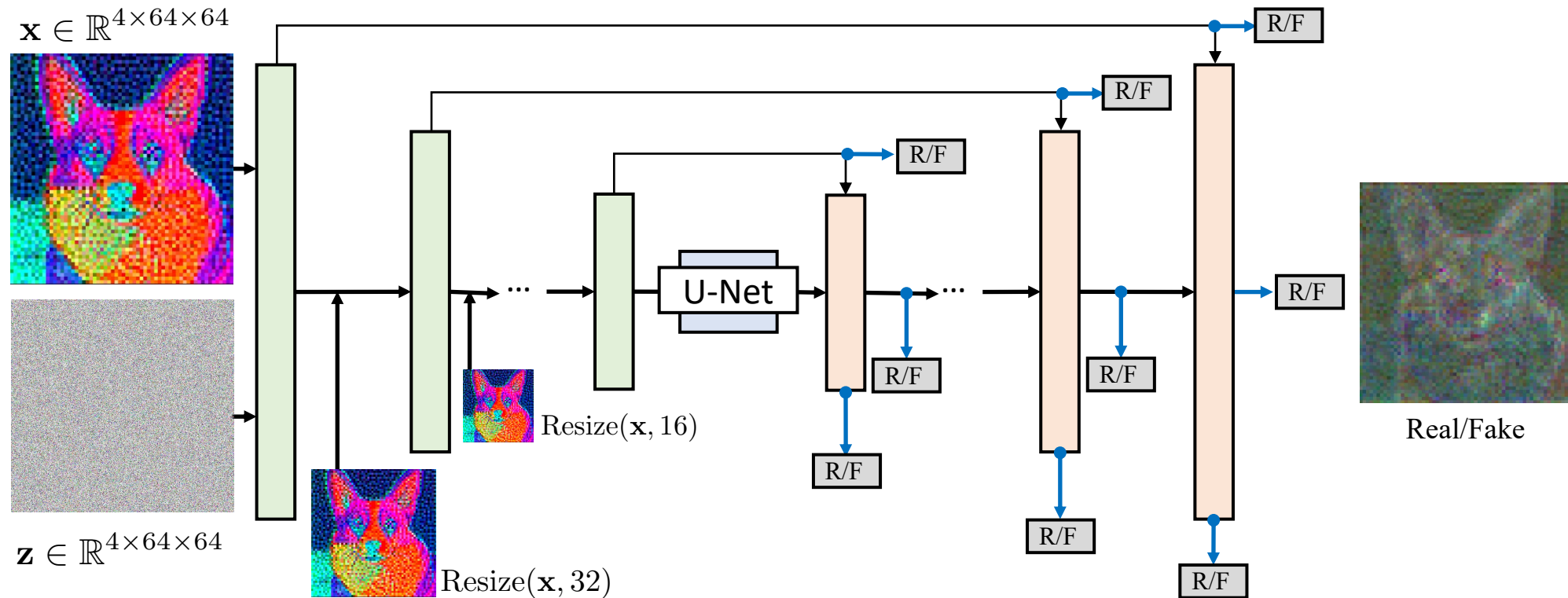


In practice, paired image translation yields better results when combined with conditional GAN losses.

(sources = **noise & prompt**, target = **ODE solution**)



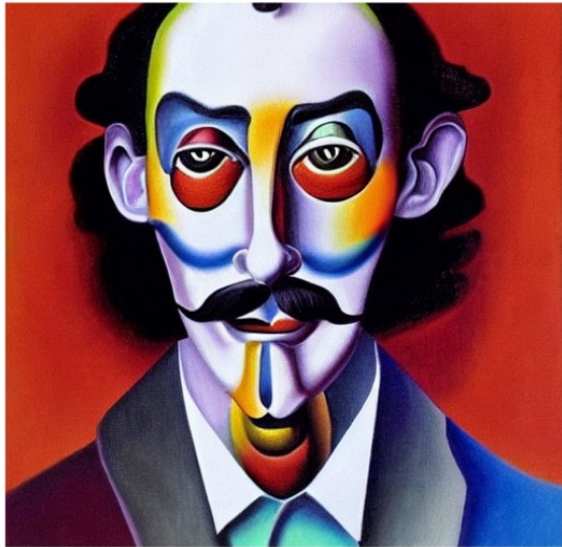
# Pre-trained diffusion as a strong discriminator



Multi-scale Input/output discriminator following GigaGAN.

# Visual comparison with previous models

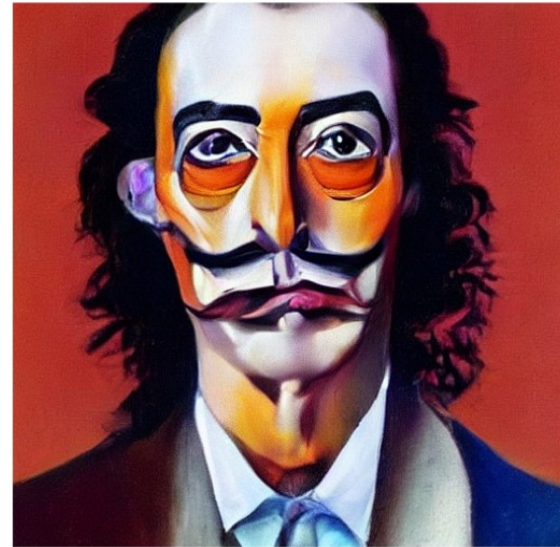
Stable Diffusion 1.5



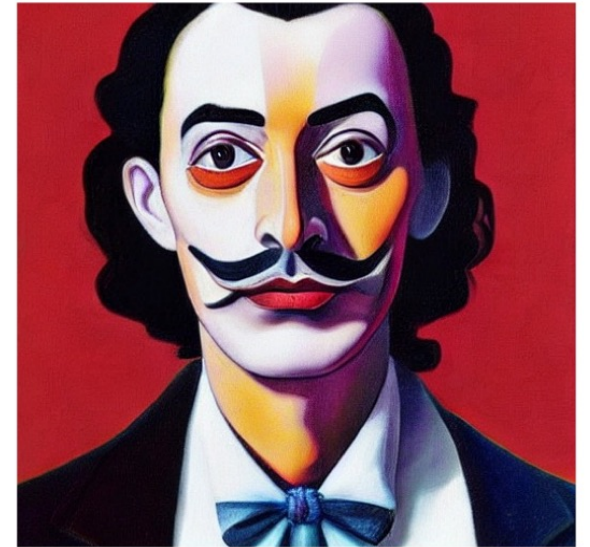
GigaGAN



InstaFlow-0.9B



Diffusion2GAN



# Visual comparison with latest work



SDXL-Turbo (arXiv Nov., 2023)



Diffusion2GAN

“A woman with bright blue eyes and curly blonde hair smiles warmly in a softly lit café.”



# Visual comparison with latest work



SDXL-Lightening (arXiv Feb., 2024)



Diffusion2GAN

"A woman with bright blue eyes and curly blonde hair smiles warmly in a softly lit café."



# Visual comparison with latest work



SDXL-Turbo (arXiv Nov., 2023)



Diffusion2GAN

“A bald eagle made of chocolate powder, mango, and whipped cream”



# Visual comparison with latest work



SDXL-Lightening (arXiv Feb., 2024)



Diffusion2GAN

“A bald eagle made of chocolate powder, mango, and whipped cream”



Thank you!

Please visit our webpage for more details.



SCAN ME

