SC395:

Image Generative Models in Computer Vision

Viraj Shah

Lecture S1

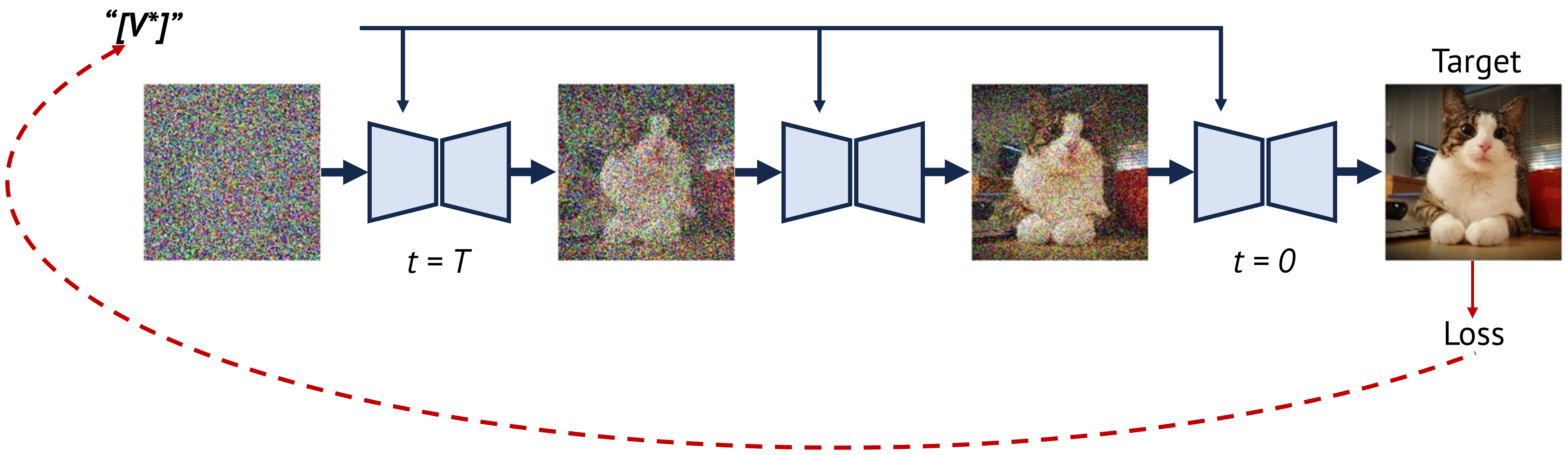
Jan 19th, 2026

sc395.virajshah.com

Diffusion Models: Applications

Controllable Image Synthesis

Textual Inversion



- **Find the text embedding that generates the given image**
- **Re-use the text embedding with new text phrases to obtain editing**

Ref: Textual Inversion, Gal et al. ICLR'23

Customizing DMs: Textual inversion

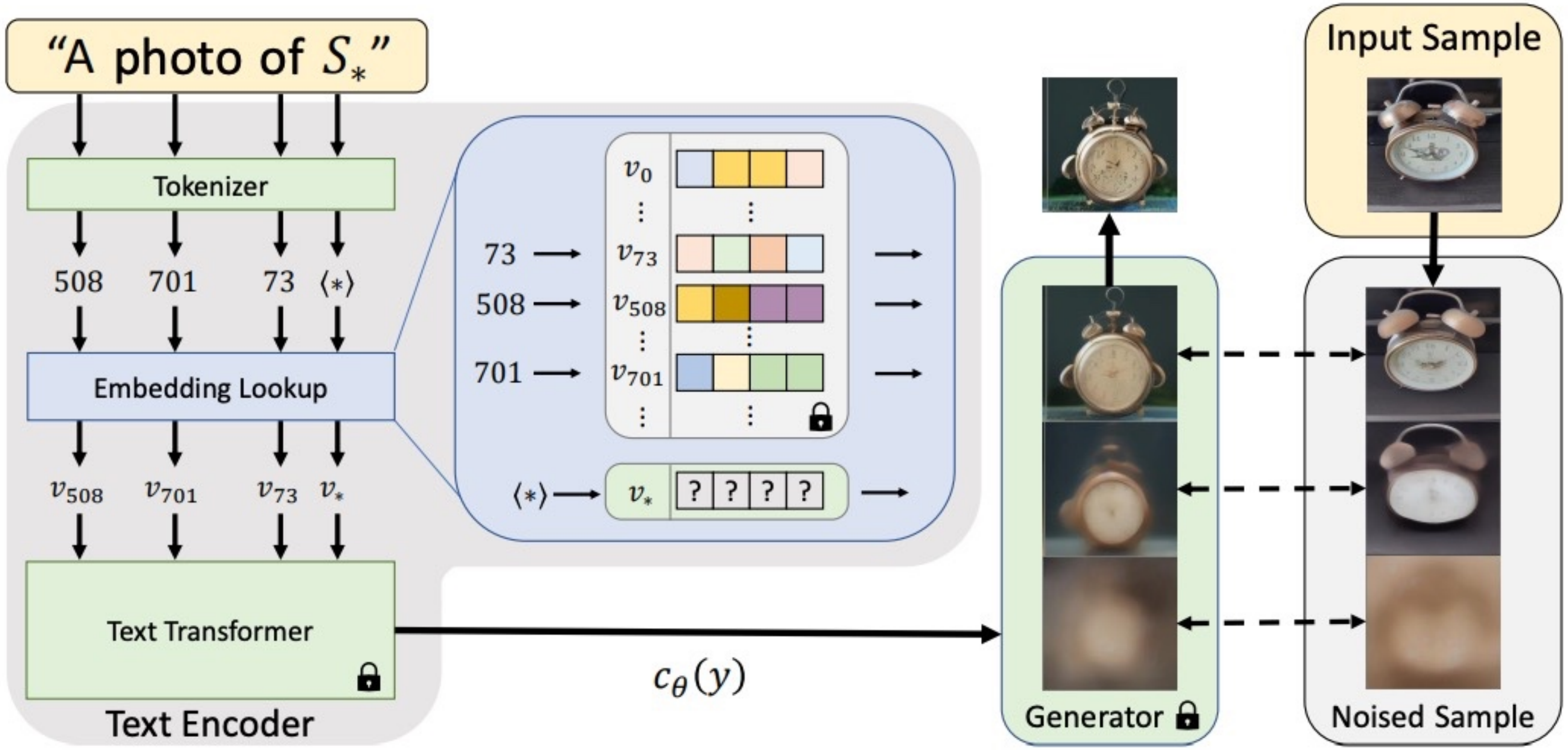


Figure 2: Outline of the text-embedding and inversion process. A string containing our placeholder word is first converted into tokens (*i.e.* word or sub-word indices in a dictionary). These tokens are converted to continuous vector representations (the "embeddings", v). Finally, the embedding vectors are transformed into a conditioning code $c_\theta(y)$ that guides the generation. We optimize the embedding vector v_* associated with our pseudo-word S_*, using a reconstruction objective.

R. Gal et al. [An Image is Worth One Word: Personalizing Text-to-Image Generation using Textual Inversion](). ICLR 2023

Customizing DMs: Textual inversion

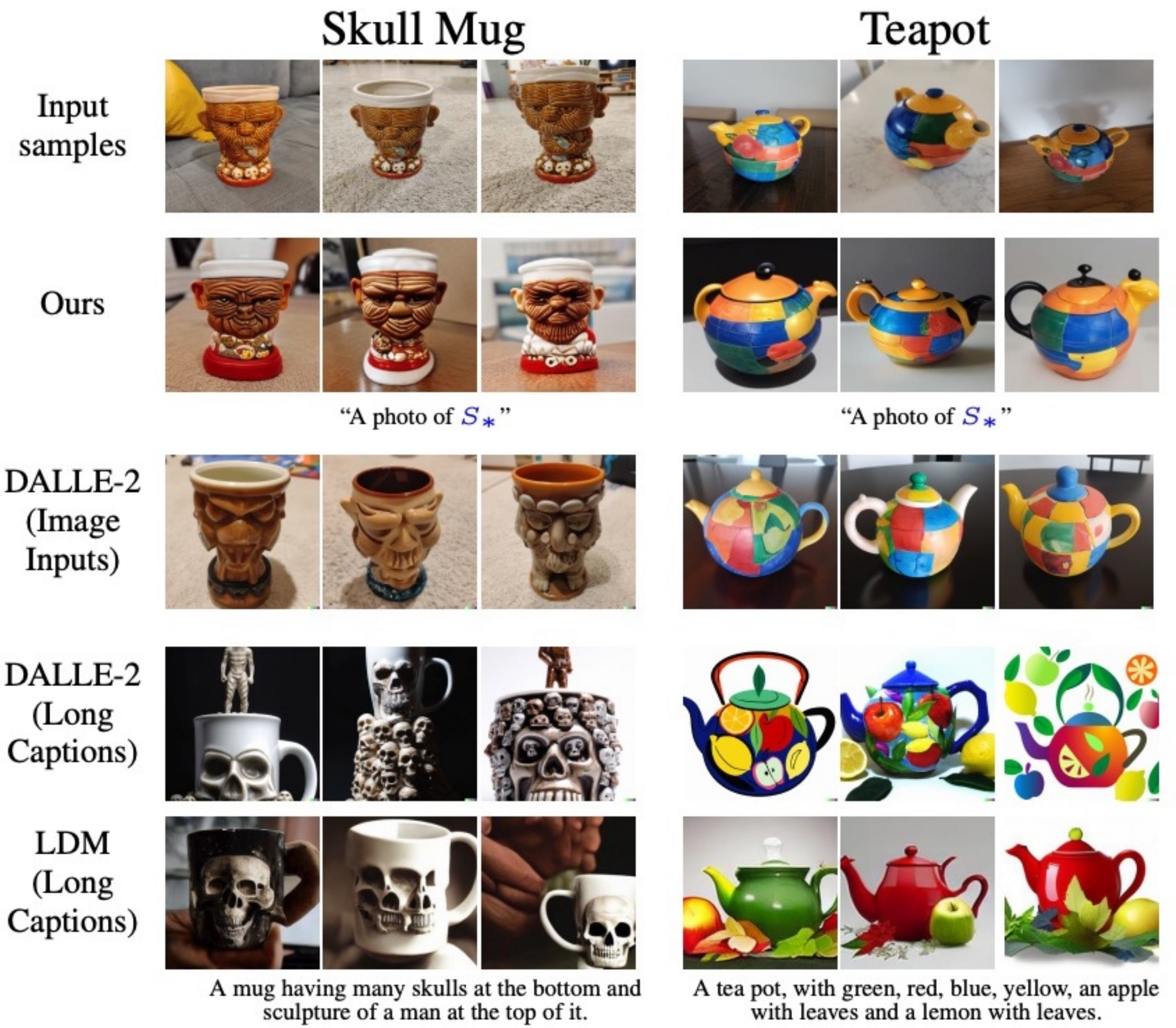
Figure 1: (left) We find new pseudo-words in the embedding space of pre-trained text-to-image models which describe specific concepts. (right) These pseudo-words are composed into new sentences, placing our targets in new scenes, changing their style or ingraining them into new products.

R. Gal et al. [An Image is Worth One Word: Personalizing Text-to-Image Generation using Textual Inversion](). ICLR 2023

Textual inversion: Results

Textual inversion: Results

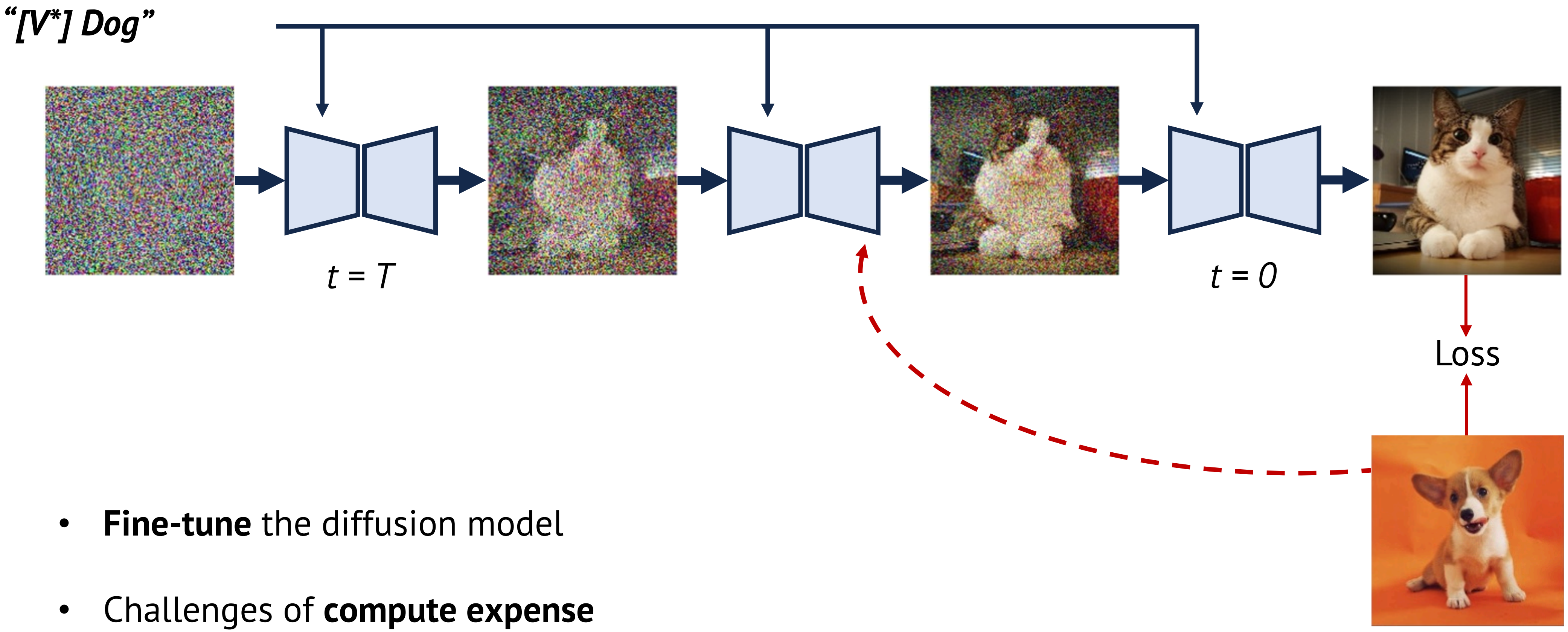
Textual inversion: Comparisons



Customizing DMs: DreamBooth

N. Ruiz et al. [DreamBooth: Fine Tuning Text-to-Image Diffusion Models for Subject-Driven Generation](). arXiv 2022

DreamBooth: Local Customization of Diffusion Model



- **Fine-tune** the diffusion model
- Challenges of **compute expense**

Ref: DreamBooth, Ruiz et al., CVPR '23

Customizing DMs: DreamBooth

Prompt: "retro style yellow alarm clock with a white clock face and a yellow number three on the right part of the clock face in the jungle"

N. Ruiz et al. [DreamBooth: Fine Tuning Text-to-Image Diffusion Models for Subject-Driven Generation](). arXiv 2022

Customizing DMs: DreamBooth

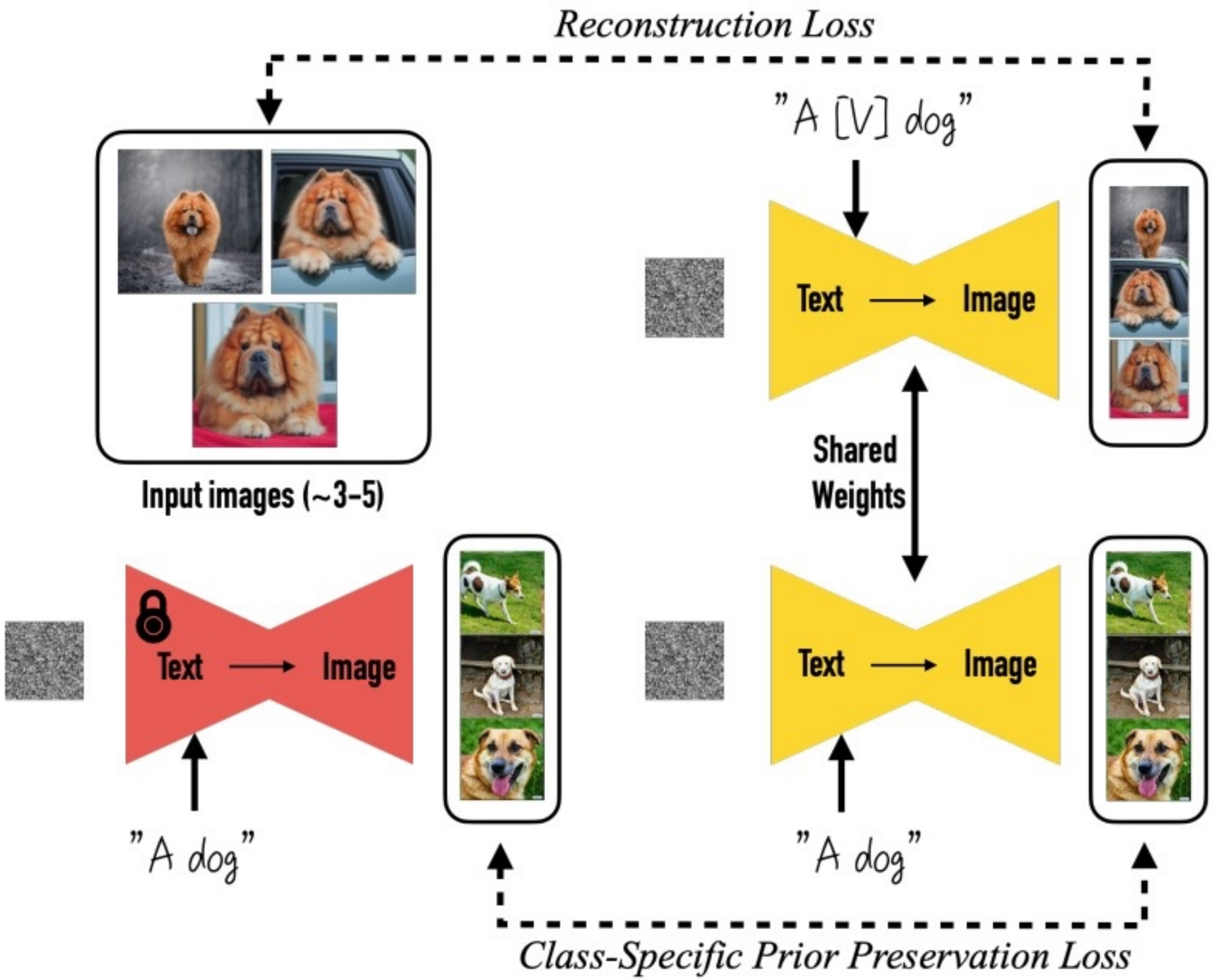
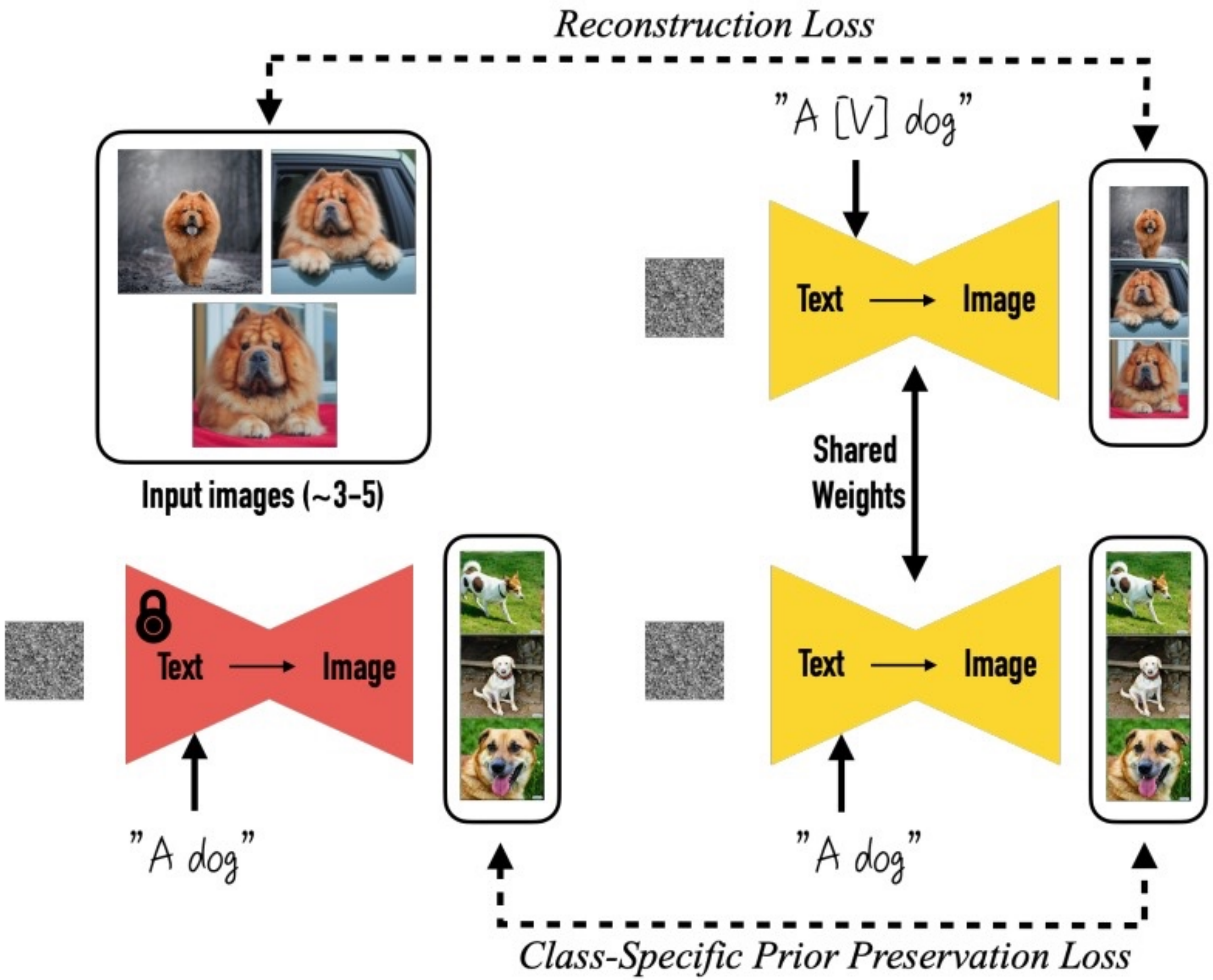


Figure 3. **Fine-tuning.** Given $\sim 3-5$ images of a subject we fine-tune a text-to-image diffusion model with the input images paired with a text prompt containing a unique identifier and the name of the class the subject belongs to (e.g., "A [V] dog"), in parallel, we apply a class-specific prior preservation loss, which leverages the semantic prior that the model has on the class and encourages it to generate diverse instances belong to the subject's class using the class name in a text prompt (e.g., "A dog").

N. Ruiz et al. [DreamBooth: Fine Tuning Text-to-Image Diffusion Models for Subject-Driven Generation](). arXiv 2022

Customizing DMs: DreamBooth



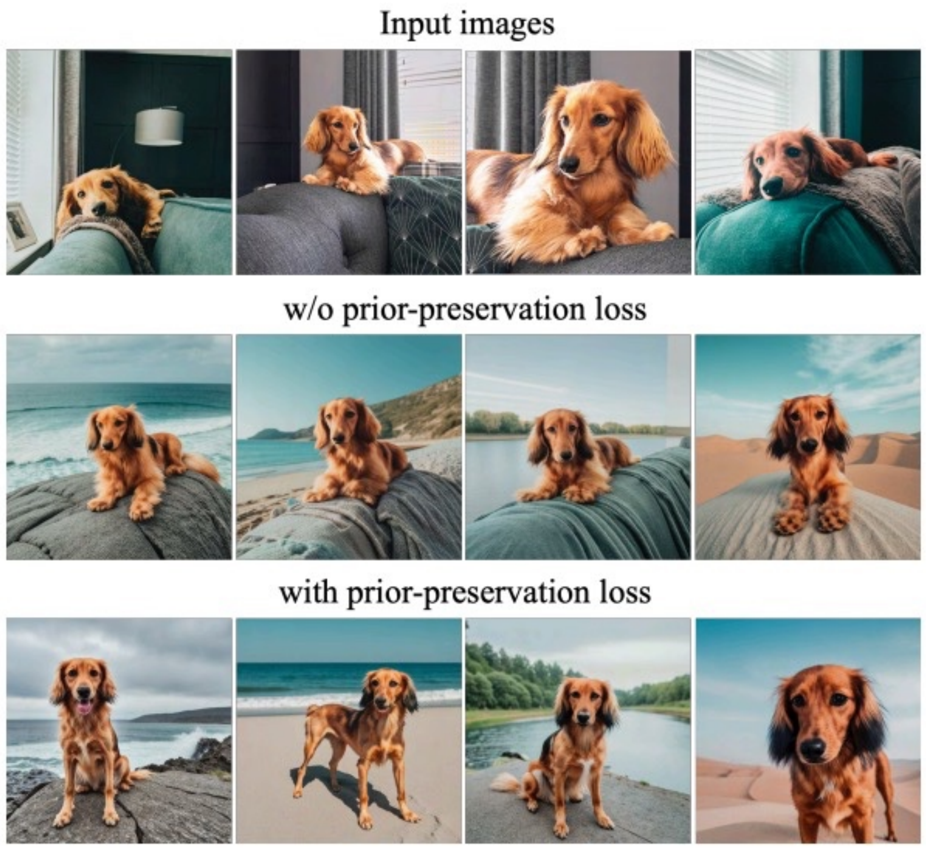
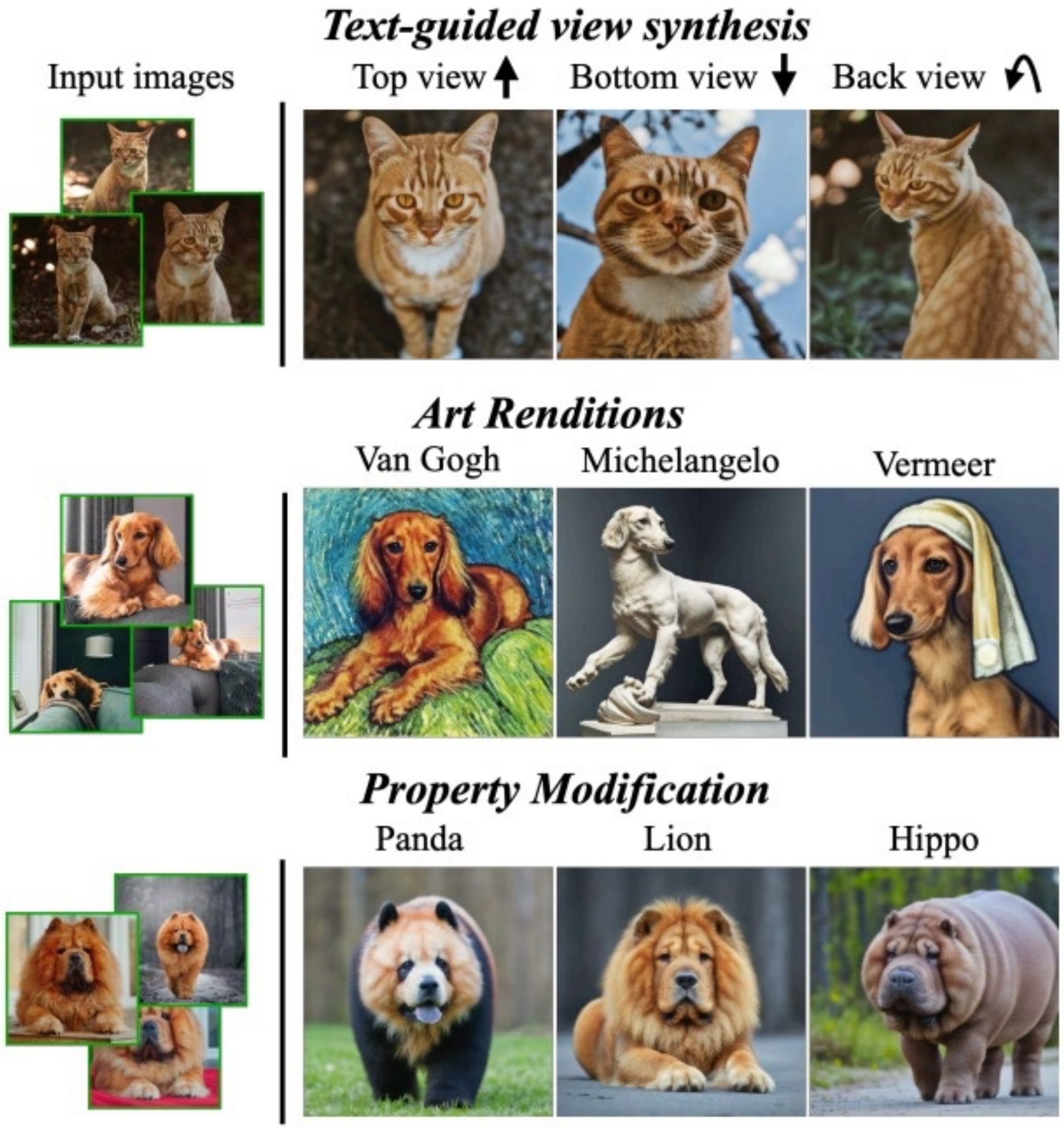


Figure 6. **Encouraging diversity with prior-preservation loss.** Naive fine-tuning can result in overfitting to input image context and subject appearance (e.g. pose). PPL acts as a regularizer that alleviates overfitting and encourages diversity, allowing for more pose variability and appearance diversity.

N. Ruiz et al. [DreamBooth: Fine Tuning Text-to-Image Diffusion Models for Subject-Driven Generation](). arXiv 2022

DreamBooth: Results

DreamBooth: Results



"a [painting/sculpture] of a [V] [class noun] in the style of [famous artist]"

"a cross of a [V] dog and a [target species]"

DreamBooth: Comparison with textual inversion

“a [V] vase in the snow” “a [V] vase on the beach” “a [V] vase in the jungle” “a [V] vase with Eiffel Tower in the background

Method	Subject Fidelity ↑	Prompt Fidelity ↑
DreamBooth (Stable Diffusion)	**68%**	**81%**
Textual Inversion (Stable Diffusion)	22%	12%
Undecided	10%	7%

DreamBooth: Limitations

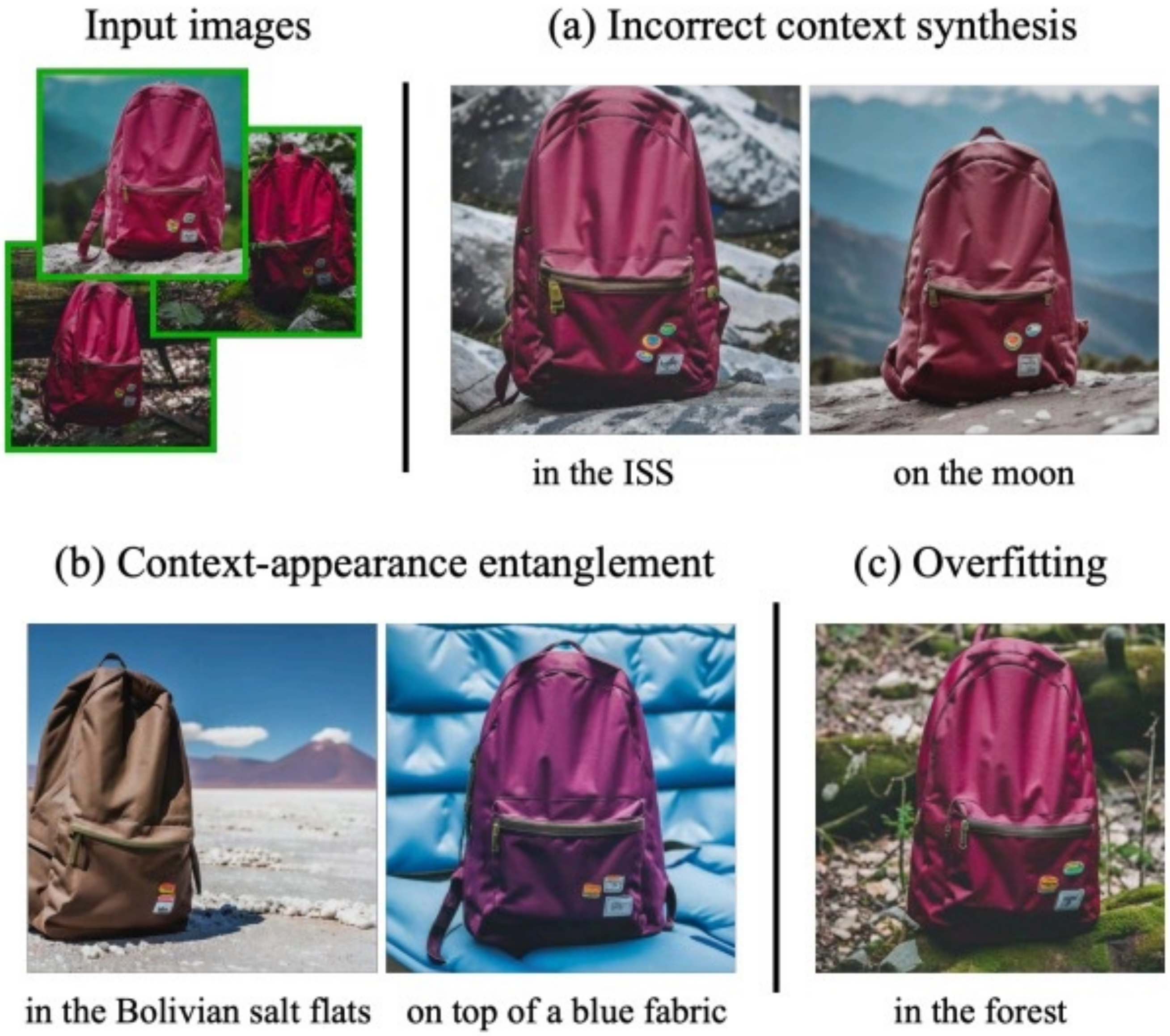
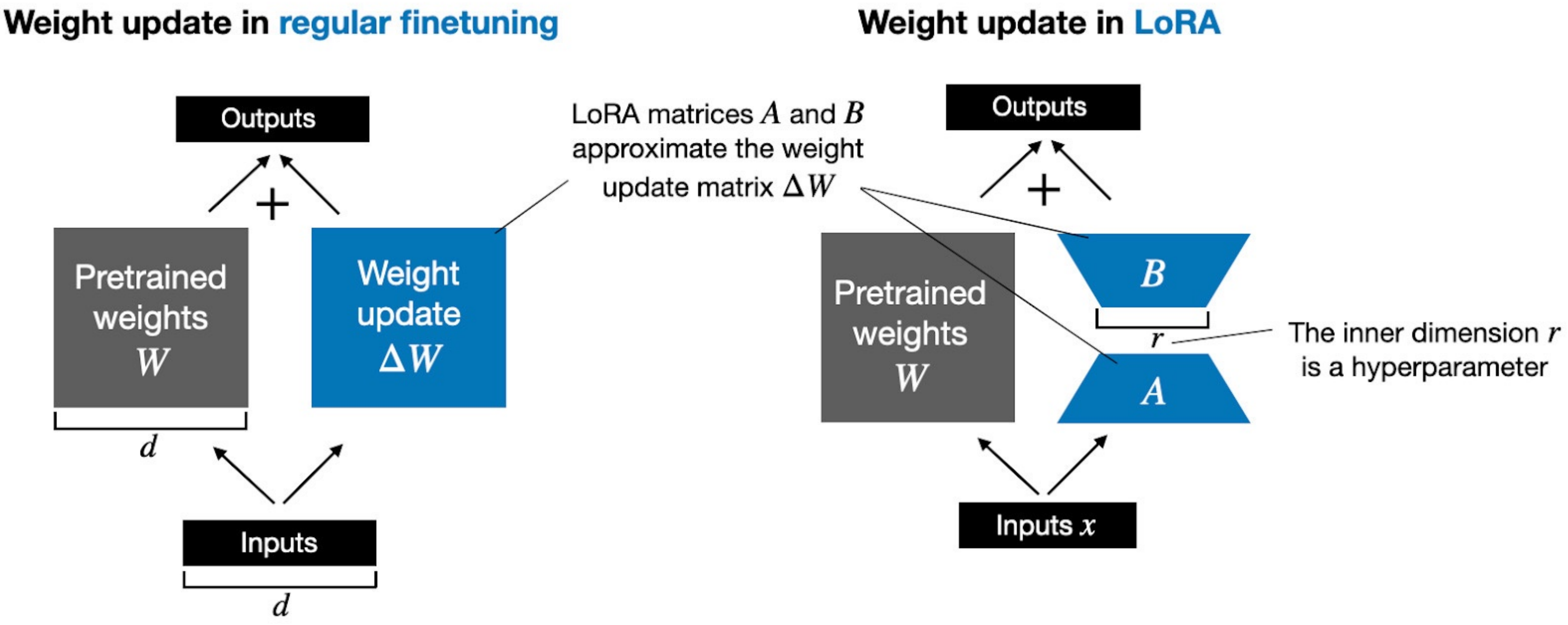


Figure 9. **Failure modes.** Given a rare prompted context the model might fail at generating the correct environment (a). It is possible for context and subject appearance to become entangled (b). Finally, it is possible for the model to overfit and generate images similar to the training set, especially if prompts reflect the original environment of the training set (c).

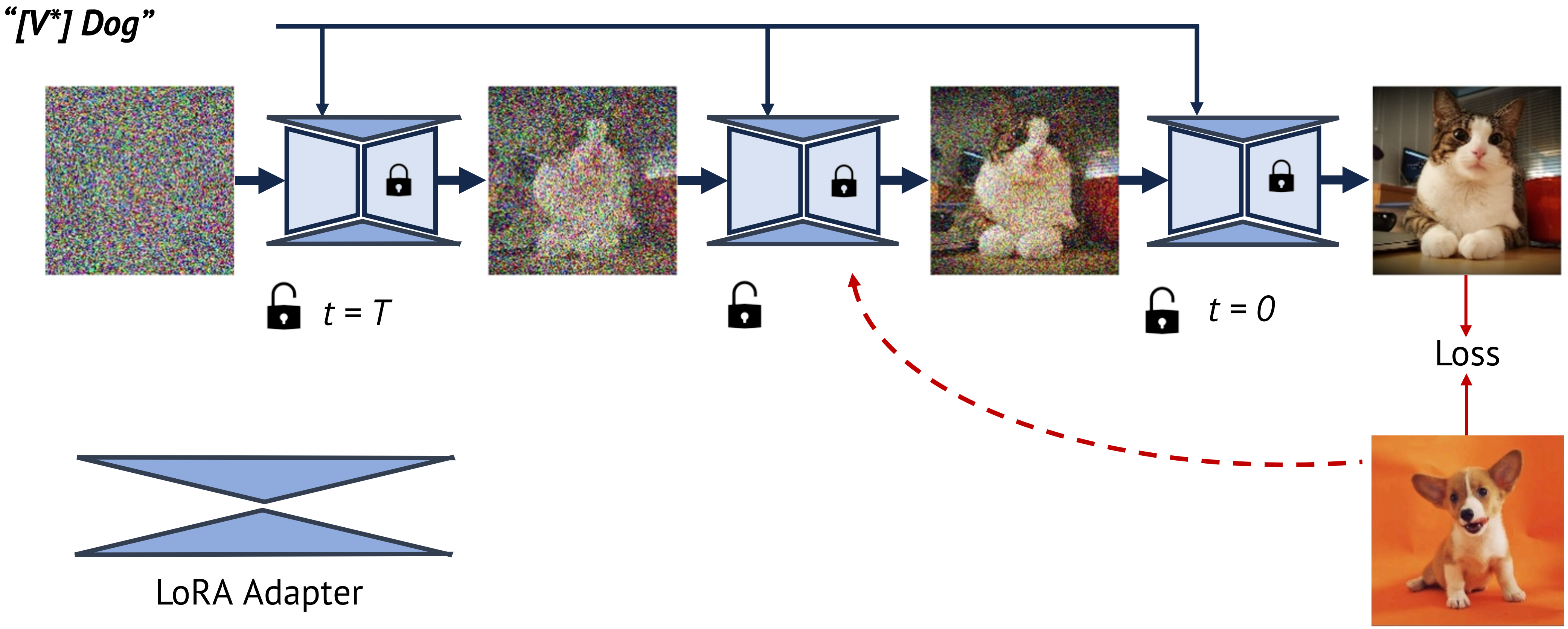
Efficient Customization: DreamBooth with LoRA



$\boldsymbol{A}$ is initialized with standard normal; $\boldsymbol{B}$ is initialized with zeros

Hu et al., LoRA: Low-Rank Adaptation of Large Language Models, ICLR'22; https://magazine.sebastianraschka.com/p/practical-tips-for-finetuning-llms

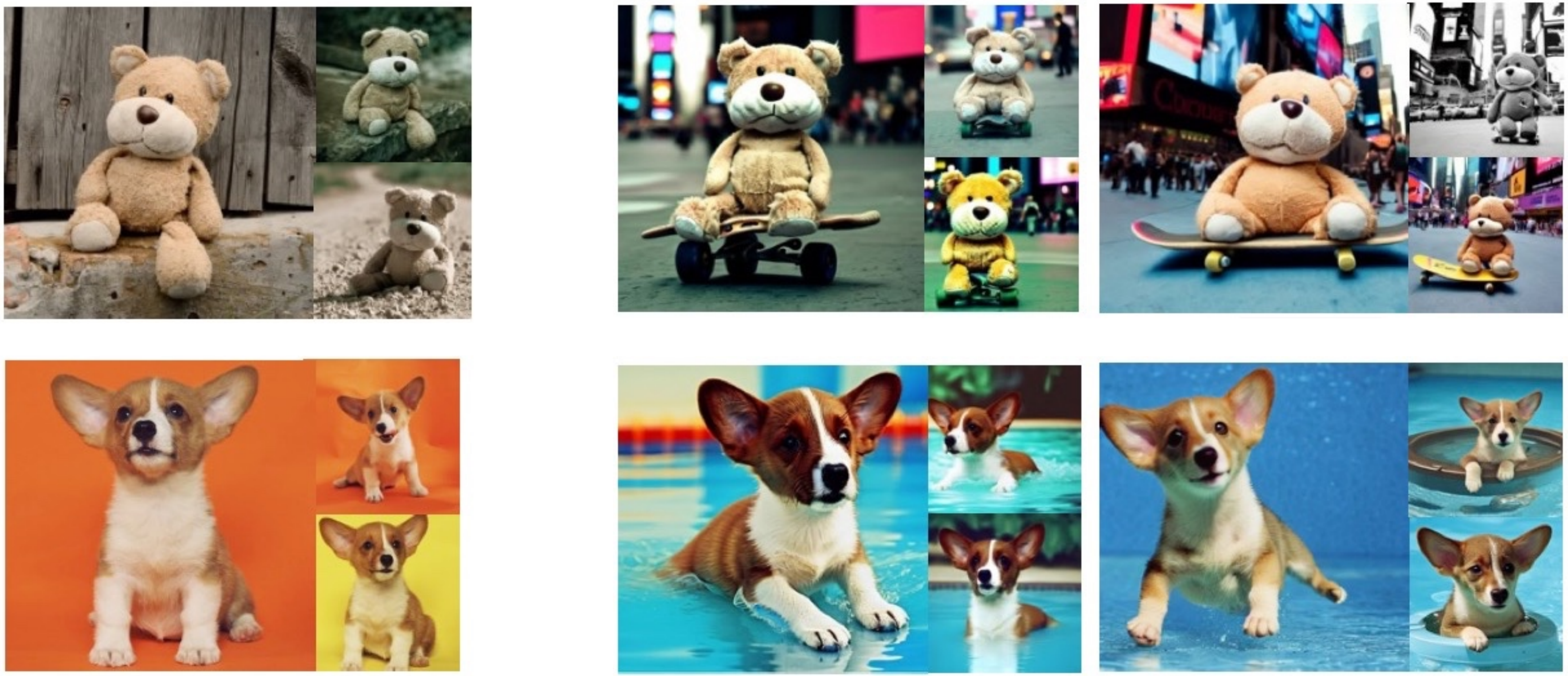
DreamBooth with Low Rank Approximation (LoRA)



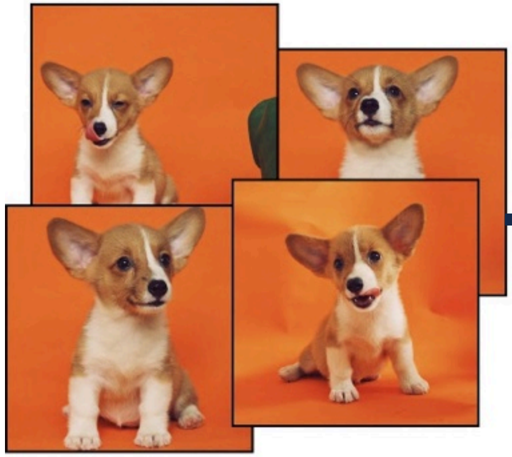
Ref: DreamBooth, Ruiz et al., CVPR ’23
LoRA, Hu et al., ‘21

LoRA DreamBooth: Results

Input Images | LoRA DreamBooth (r=4) | DreamBooth



DreamBooth with Low Rank Approximation (LoRA)



A [V] dog

DreamBooth using LoRA on SD v1.5

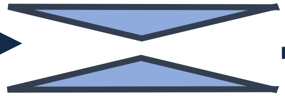
Fine-tuning on Subject

LoRA Adapter

Style Learning using LoRA on SDXL

A statue in [S] style
[S]: "matte black sculpture"

DreamBooth using LoRA on **SDXL**



LoRA Adapter

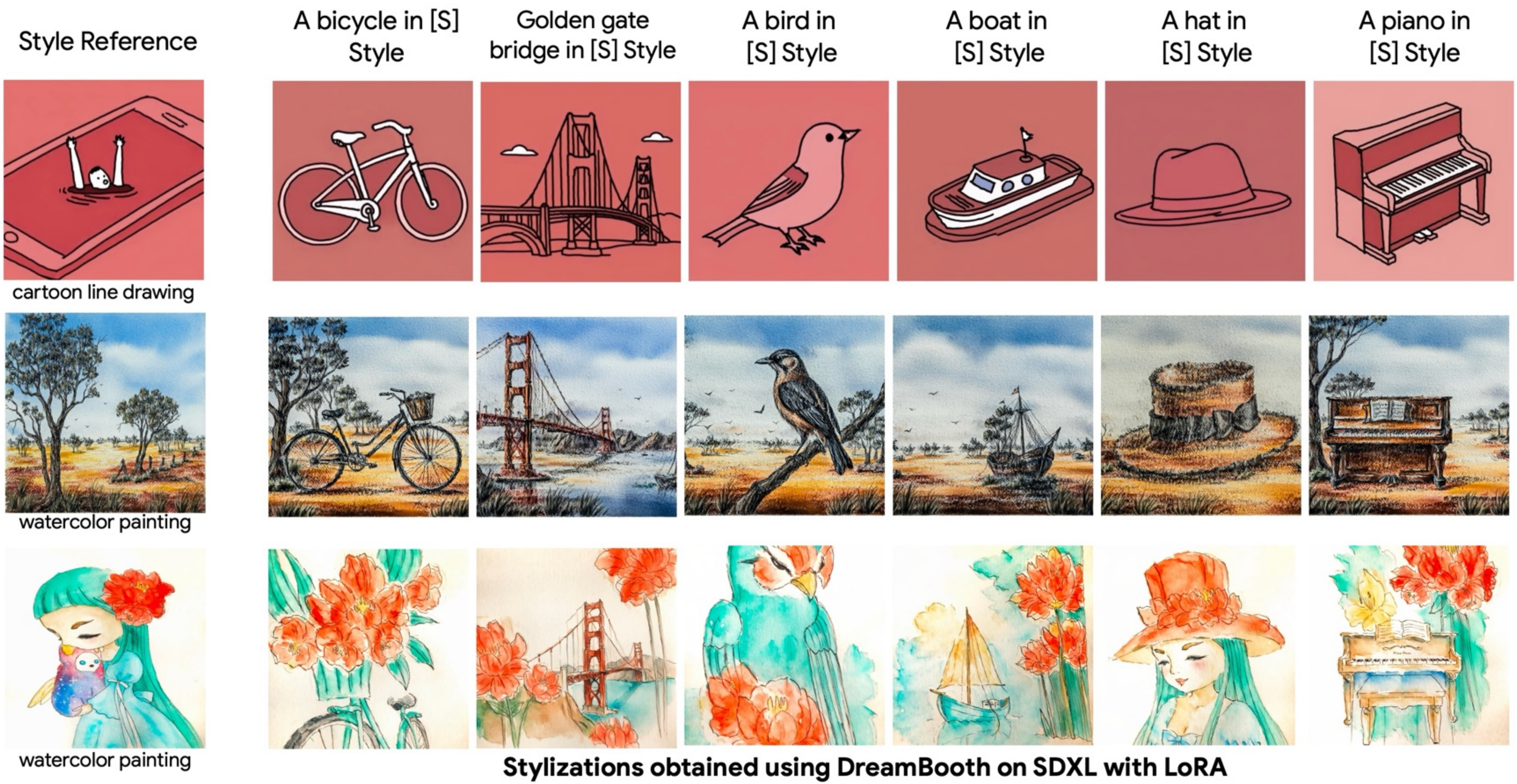
A cat in [S] style; A men in [S] style; Old lady in [S] style

Works with SDXL!

Ref: SDXL, Podell et al.'23

Slide from Jia-Bin Huang

LoRA DreamBooth for Stylizations (on SDXL)

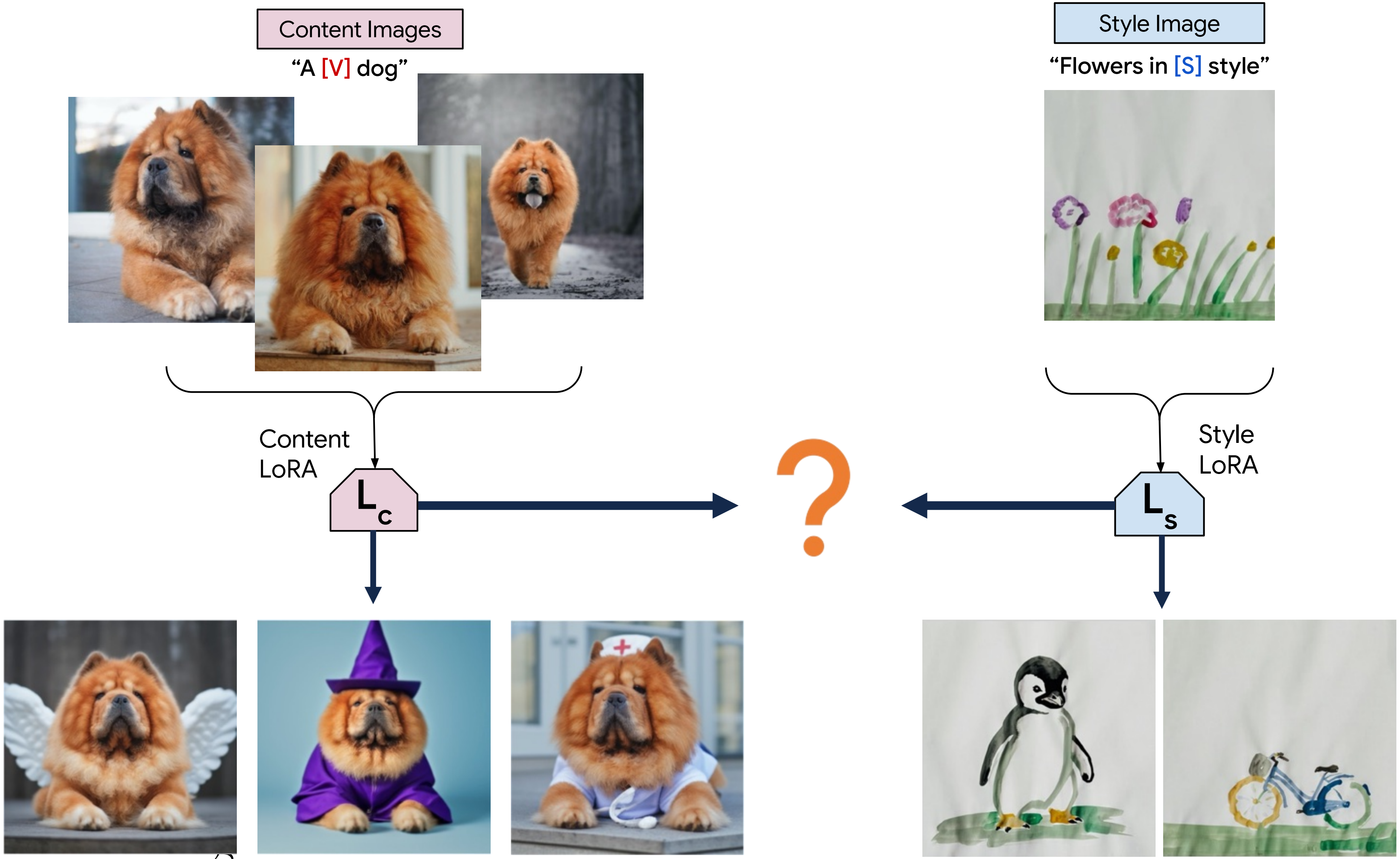


Stylizations obtained using DreamBooth on SDXL with LoRA

Shah et al., [ZipLoRA: Any Subject in Any Style by Effectively Merging LoRAs](), arxiv'23

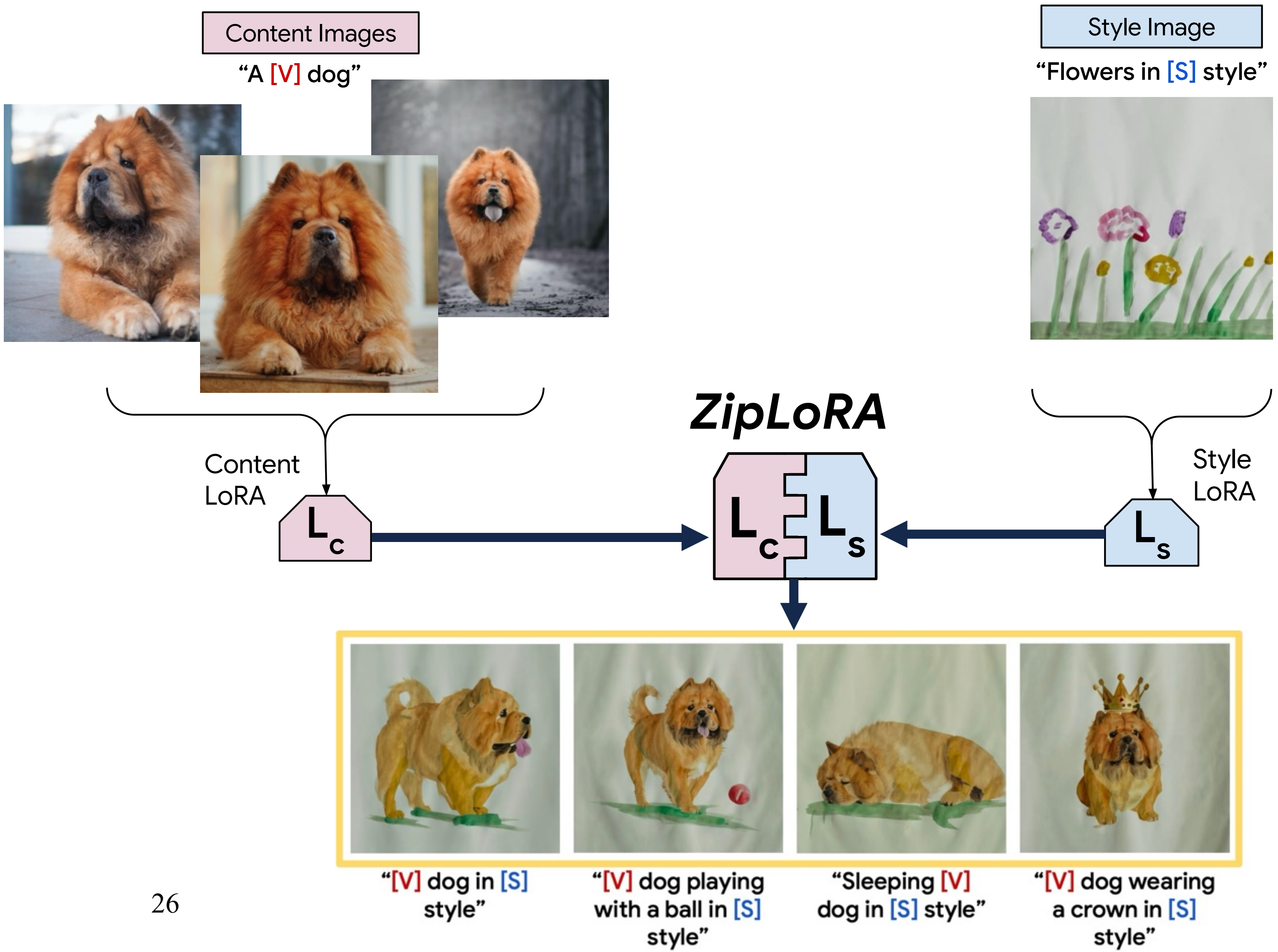
Slide from Jia-Bin Huang

Can we Merge Content and Style LoRAs?



[V] with angel wings **[V] in wizard outfit** **[V] as nurse** **penguin in [S] style** **bike in [S] style**

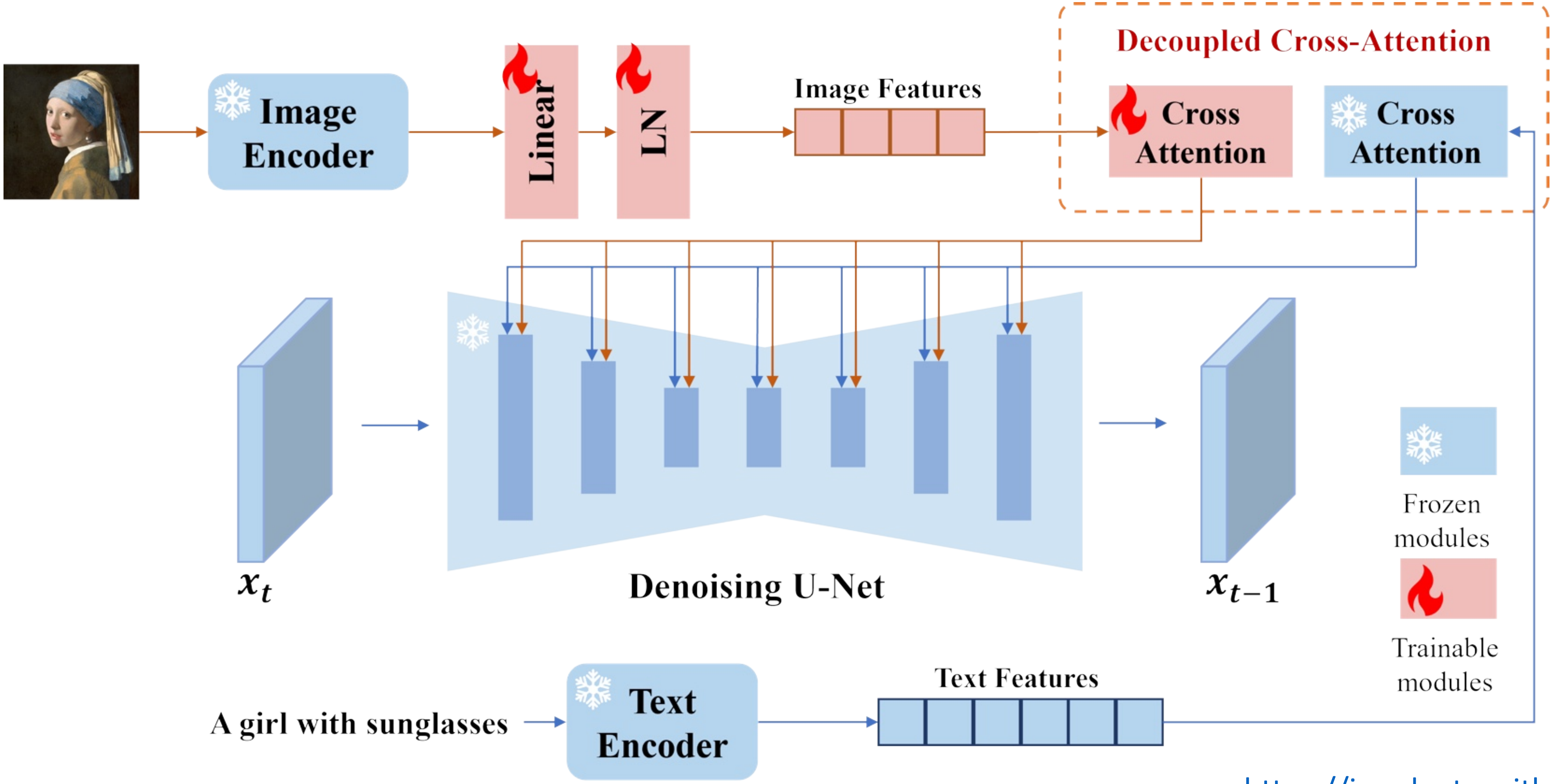
Can we Merge Content and Style LoRAs?




A [V] toy in
watercolor painting style
kid line drawing style
flat cartoon illustration style
Direct arithmetic merge
Joint Training
StyleDrop
Ours
A [V] stuffed animal in
watercolor painting style
flat cartoon illustration style
watercolor painting style
Direct arithmetic merge
Joint Training
StyleDrop
Ours

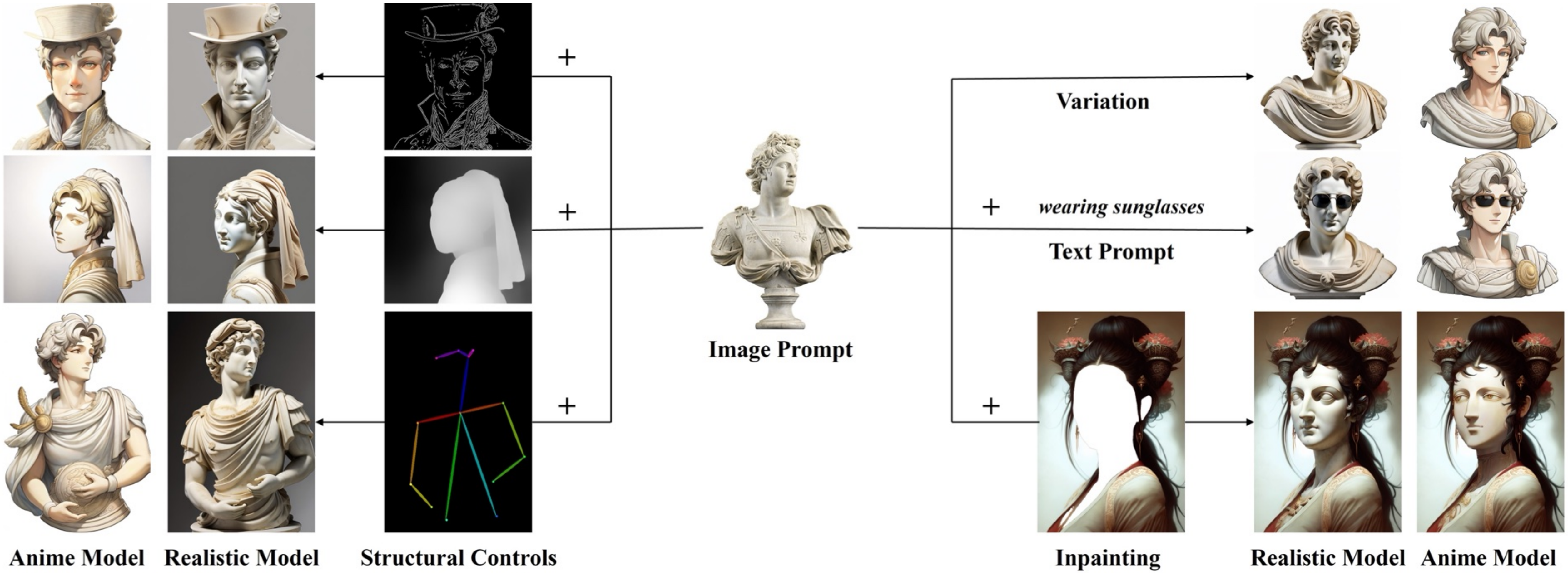
ZipLoRA produces successful recontextualizations

IP-Adapter



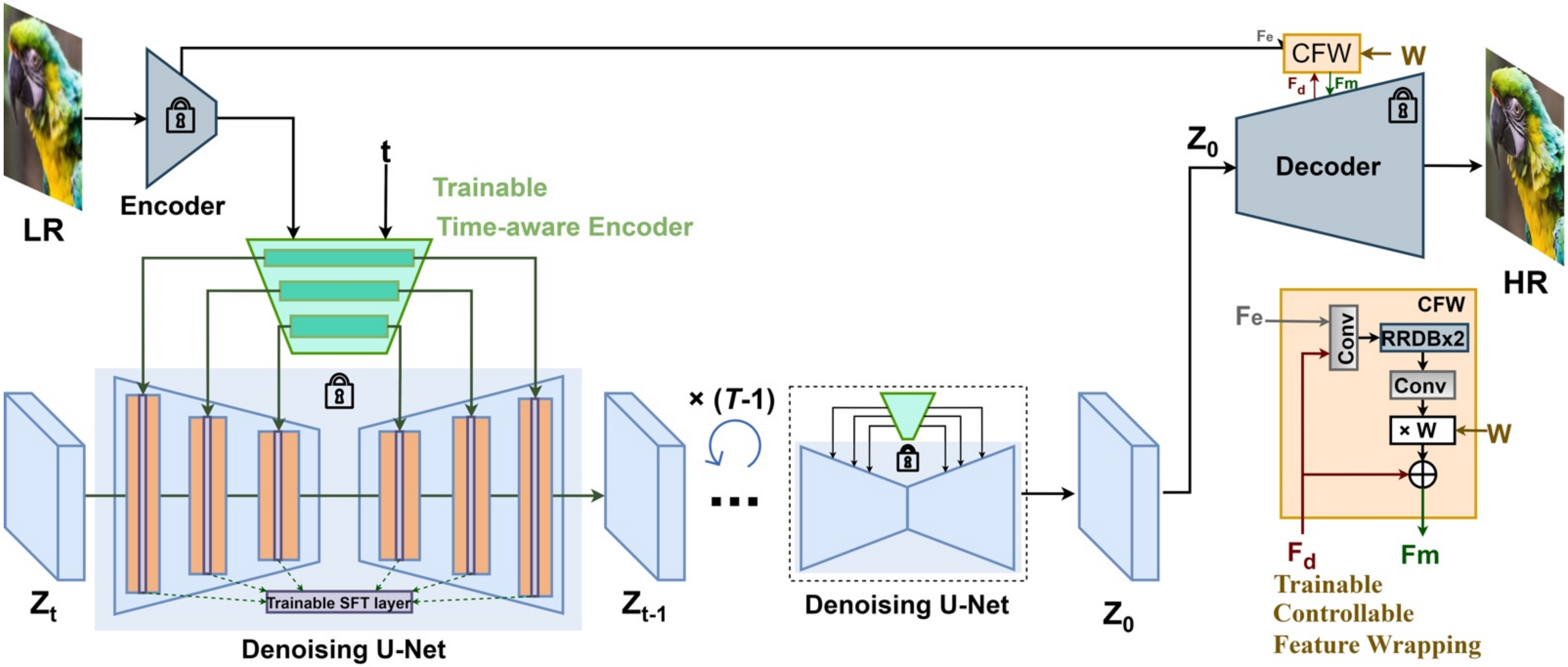
https://ip-adapter.github.io

IP-Adapter



https://ip-adapter.github.io

Stable-SR: Super-resolution

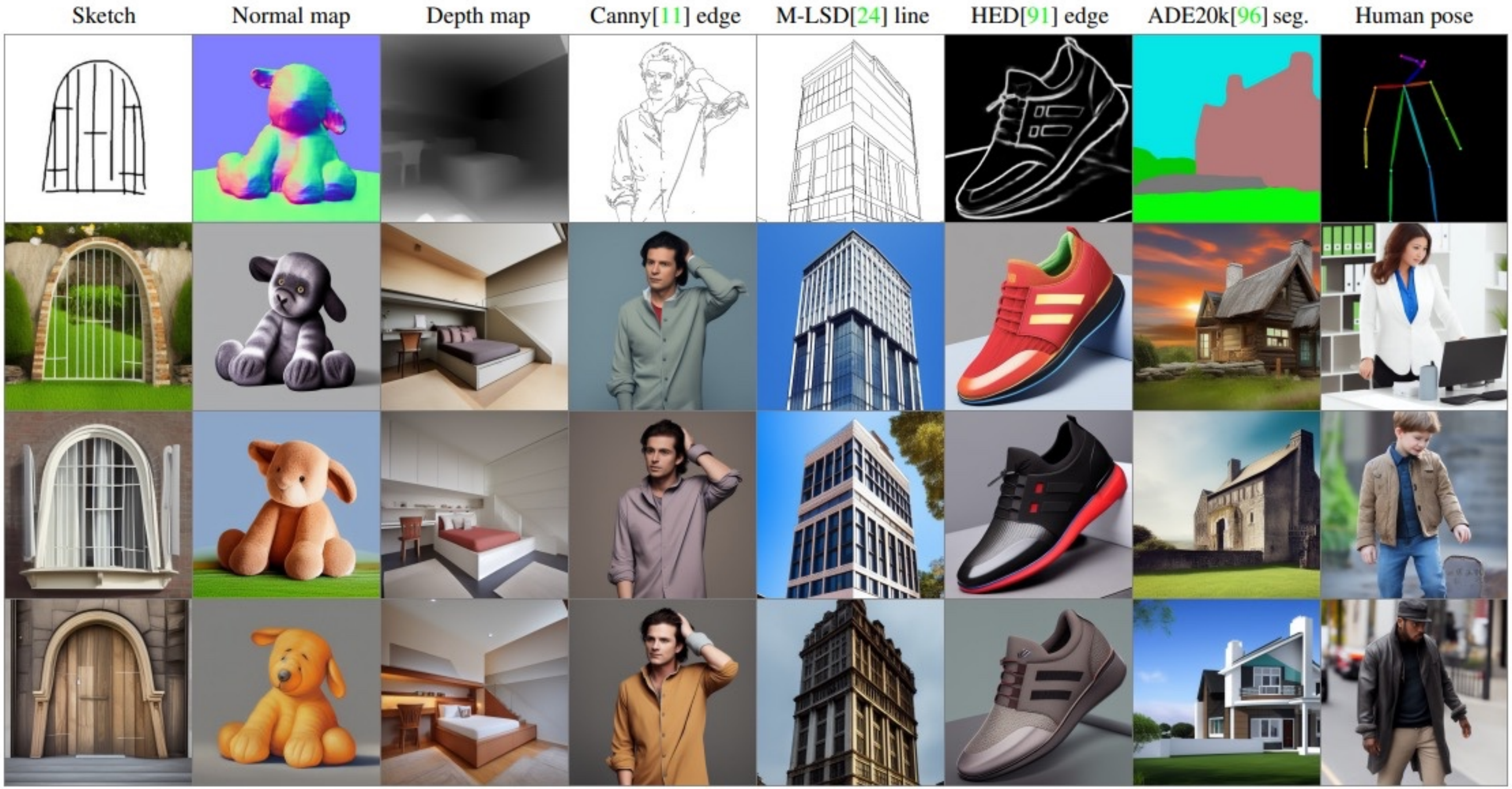


Results: https://iceclear.github.io/projects/stablesr/

Outline

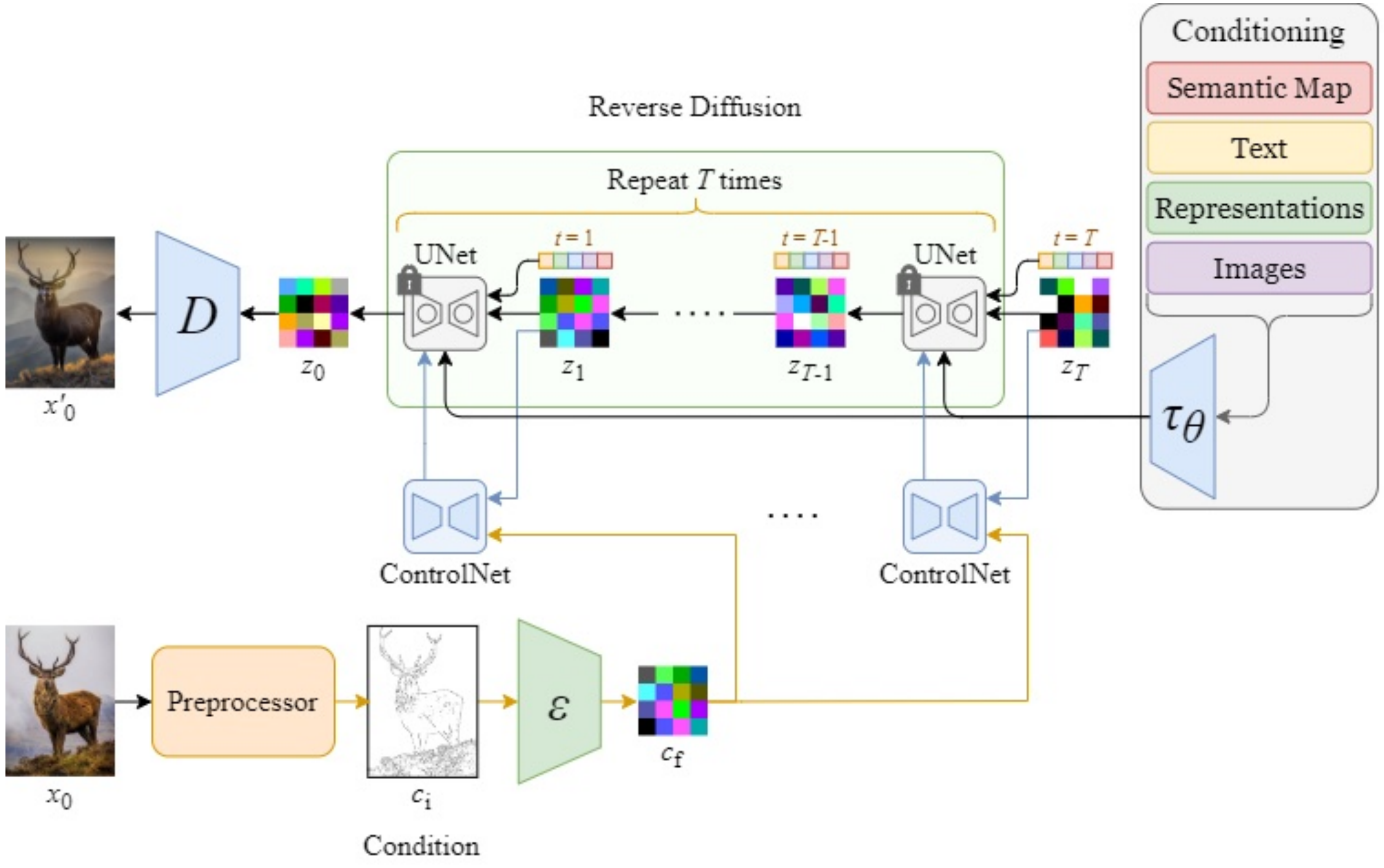
- **Part 3: Applications and Implementation; Ethical Issues**
- Customizing Diffusion Models
  - Textual Inversion
  - DreamBooth
  - Low Rank Approximation (LoRA)
  - ZipLoRA
- ControlNet

ControlNet


Sketch
Normal map
Depth map
Canny[11] edge
M-LSD[24] line
HED[91] edge
ADE20k[96] seg.
Human pose

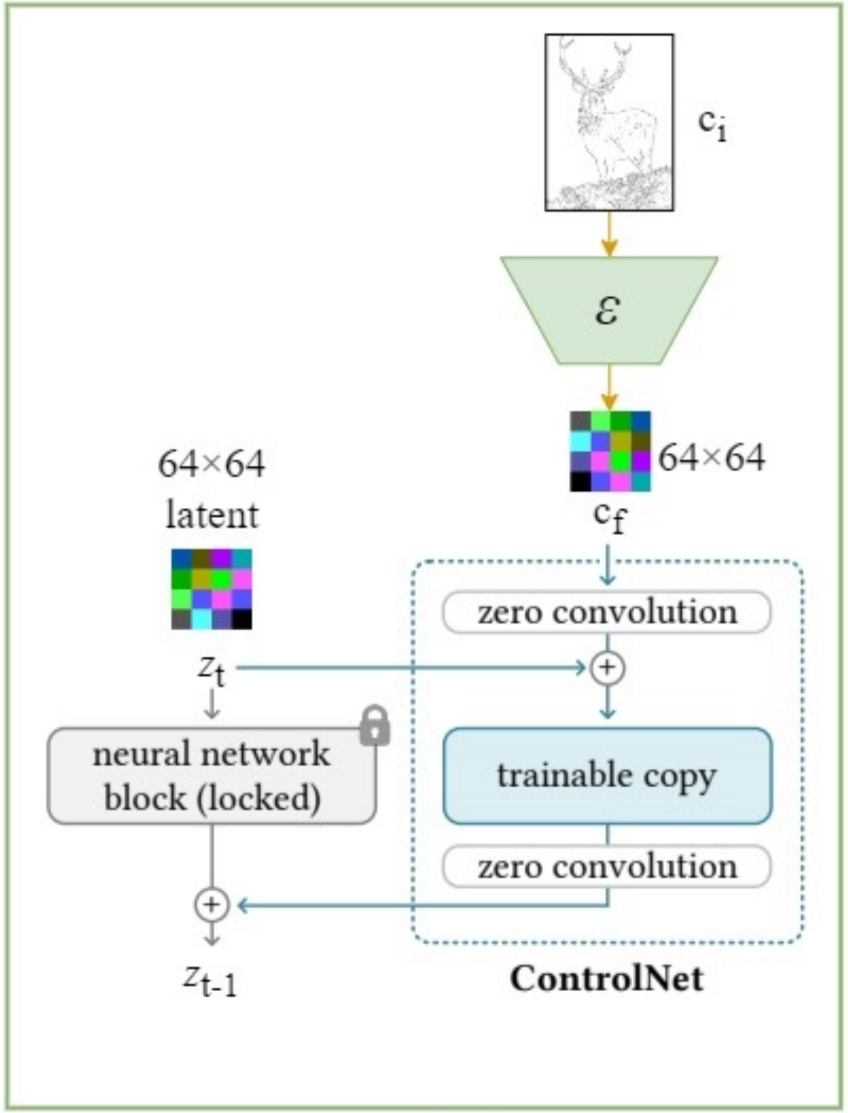
ControlNet

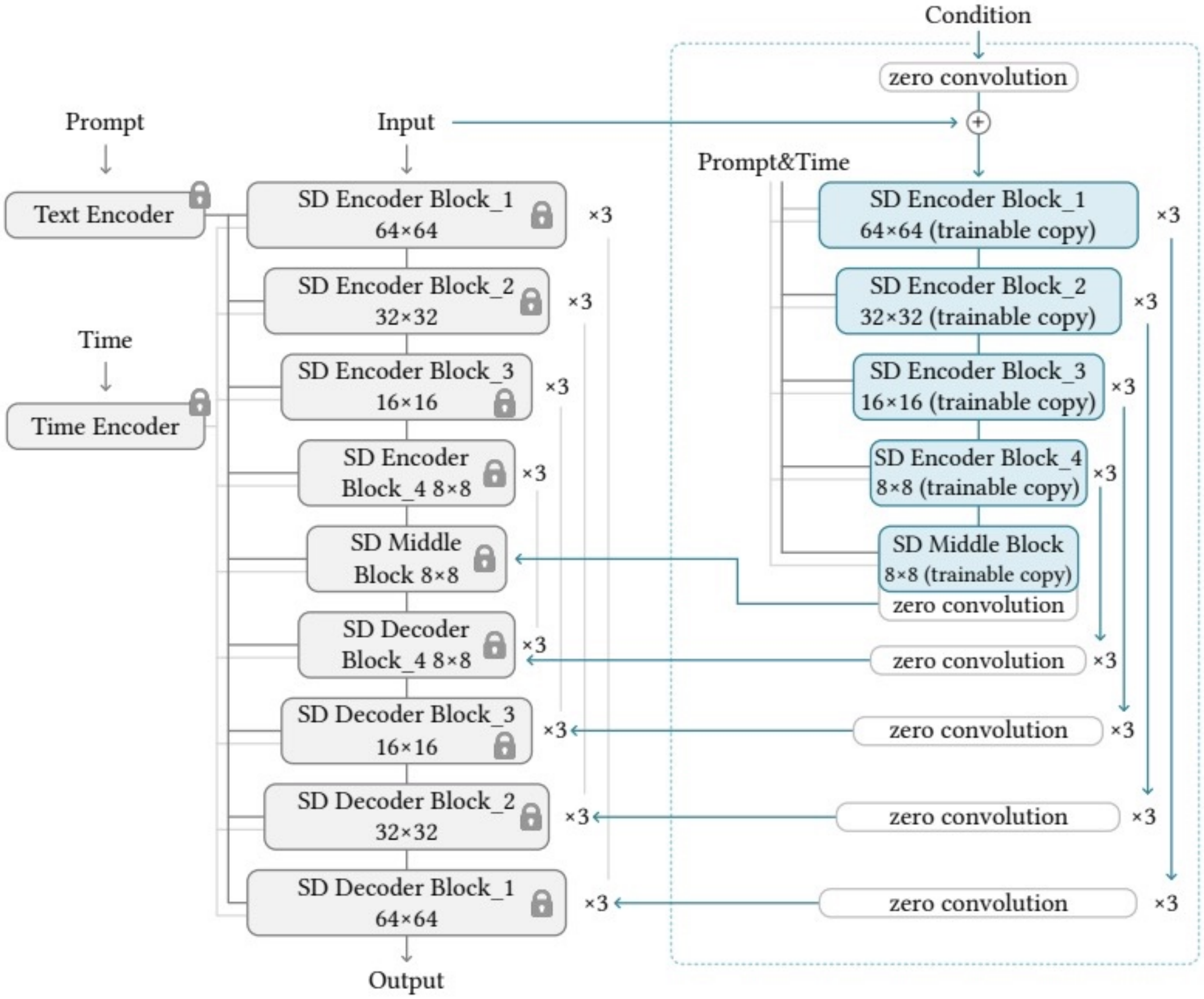
- Add a trainable “wrapper” around a pre-trained DM to fine-tune it for pix2pix tasks



L. Zhang and M. Agrawala. [Adding Conditional Control to Text-to-Image Diffusion Models](). arXiv 2023

ControlNet



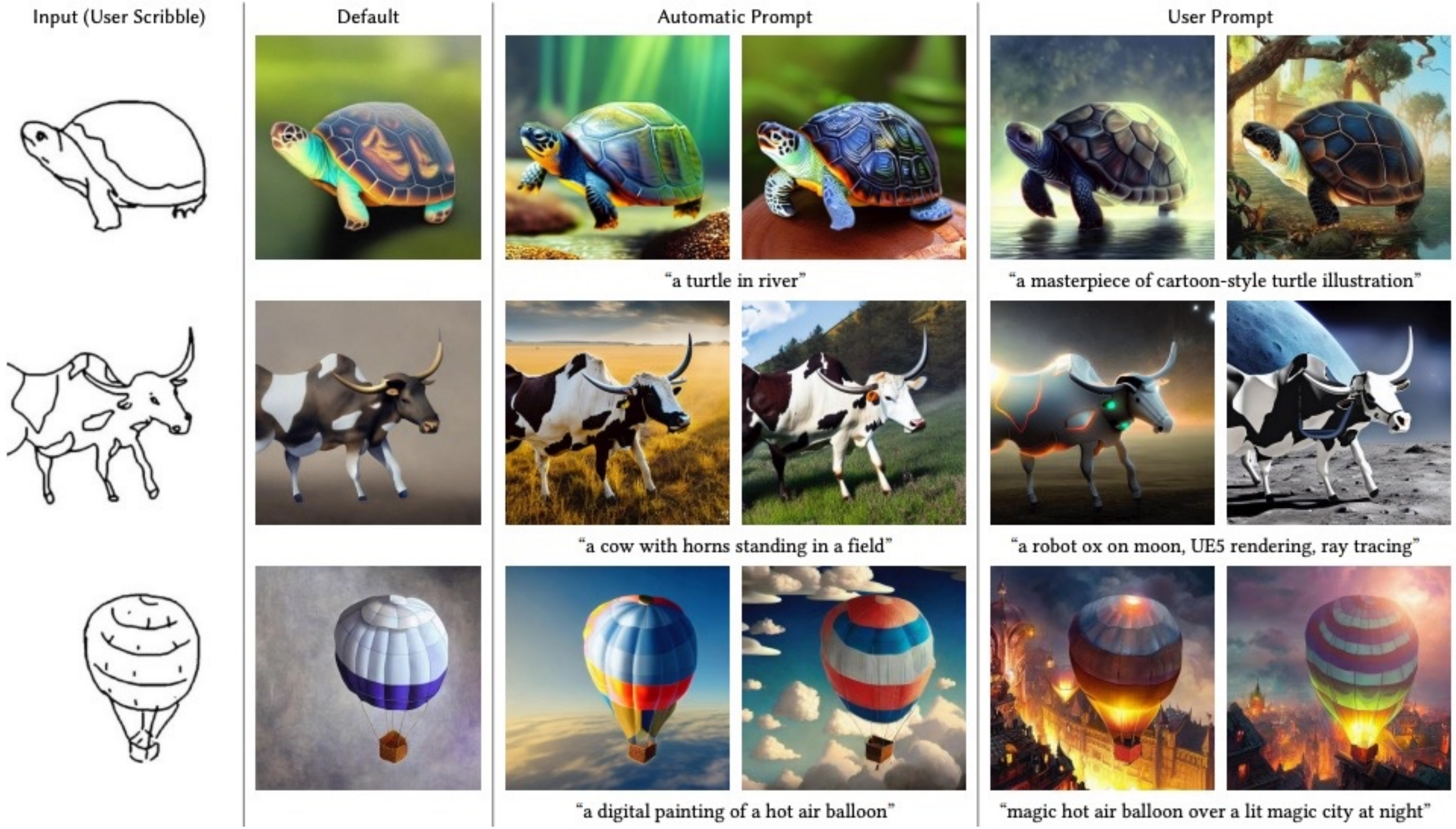


(a) Stable Diffusion (b) ControlNet

ControlNet

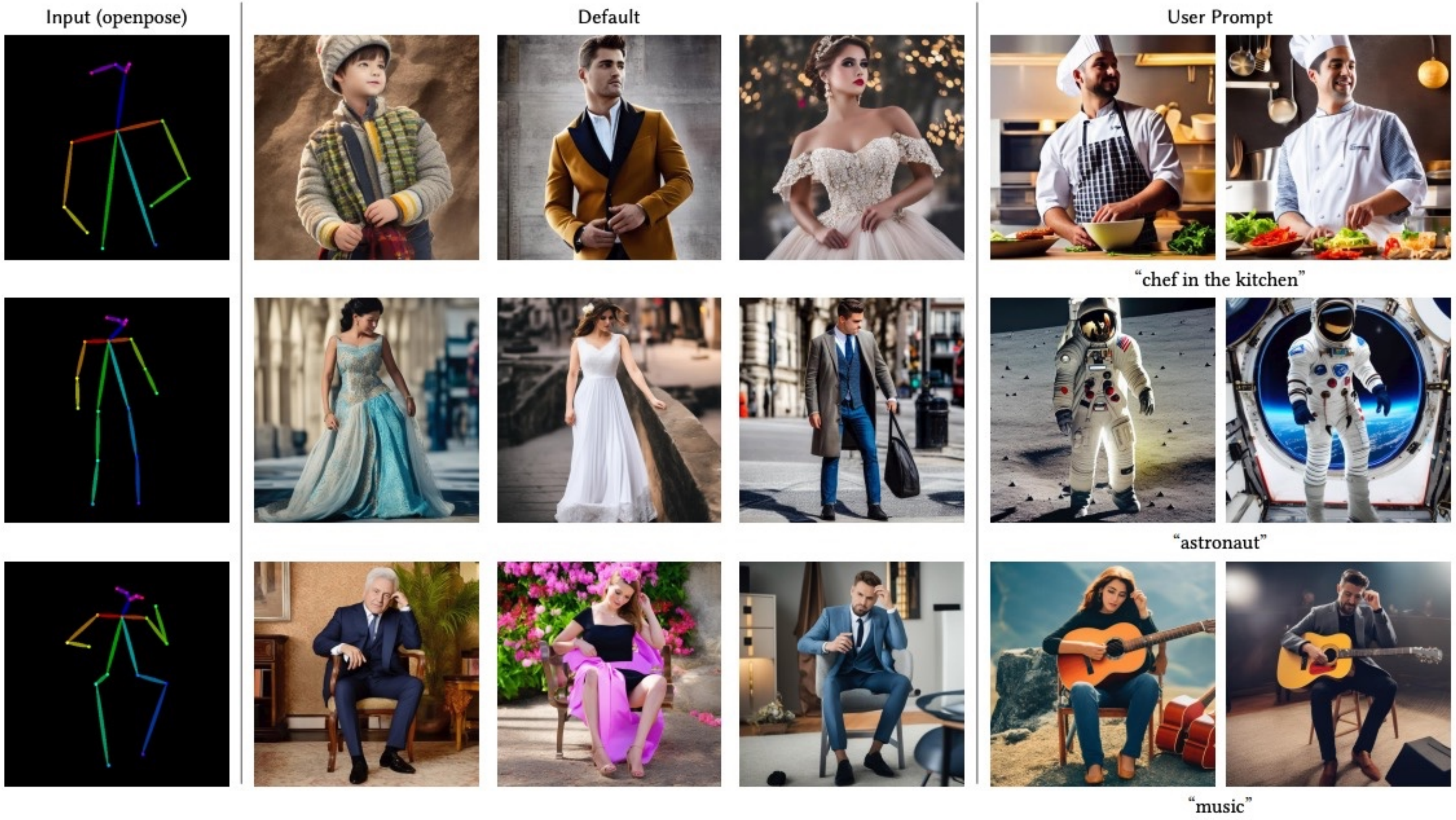
L. Zhang and M. Agrawala. [Adding Conditional Control to Text-to-Image Diffusion Models](). arXiv 2023

ControlNet



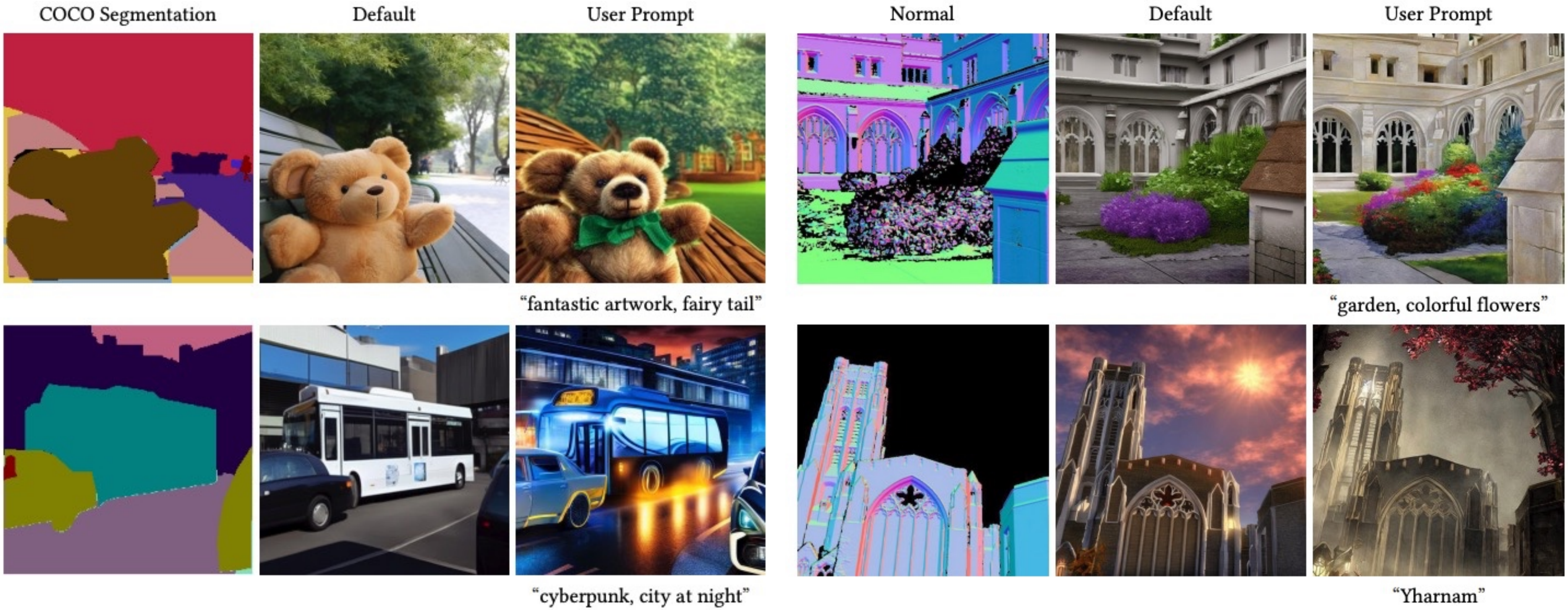
L. Zhang and M. Agrawala. [Adding Conditional Control to Text-to-Image Diffusion Models](). arXiv 2023

ControlNet



L. Zhang and M. Agrawala. [Adding Conditional Control to Text-to-Image Diffusion Models](). arXiv 2023

ControlNet



L. Zhang and M. Agrawala. [Adding Conditional Control to Text-to-Image Diffusion Models](). arXiv 2023

Outline

- **Part 3: Applications and Implementation; Ethical Issues**
- Customizing Diffusion Models
  - Textual Inversion
  - DreamBooth
  - Low Rank Approximation (LoRA)
  - ZipLoRA
- ControlNet
- Prompt-to-Prompt and InstructPix2Pix

Prompt-to-Prompt Image Editing

“The boulevards are crowded today.”

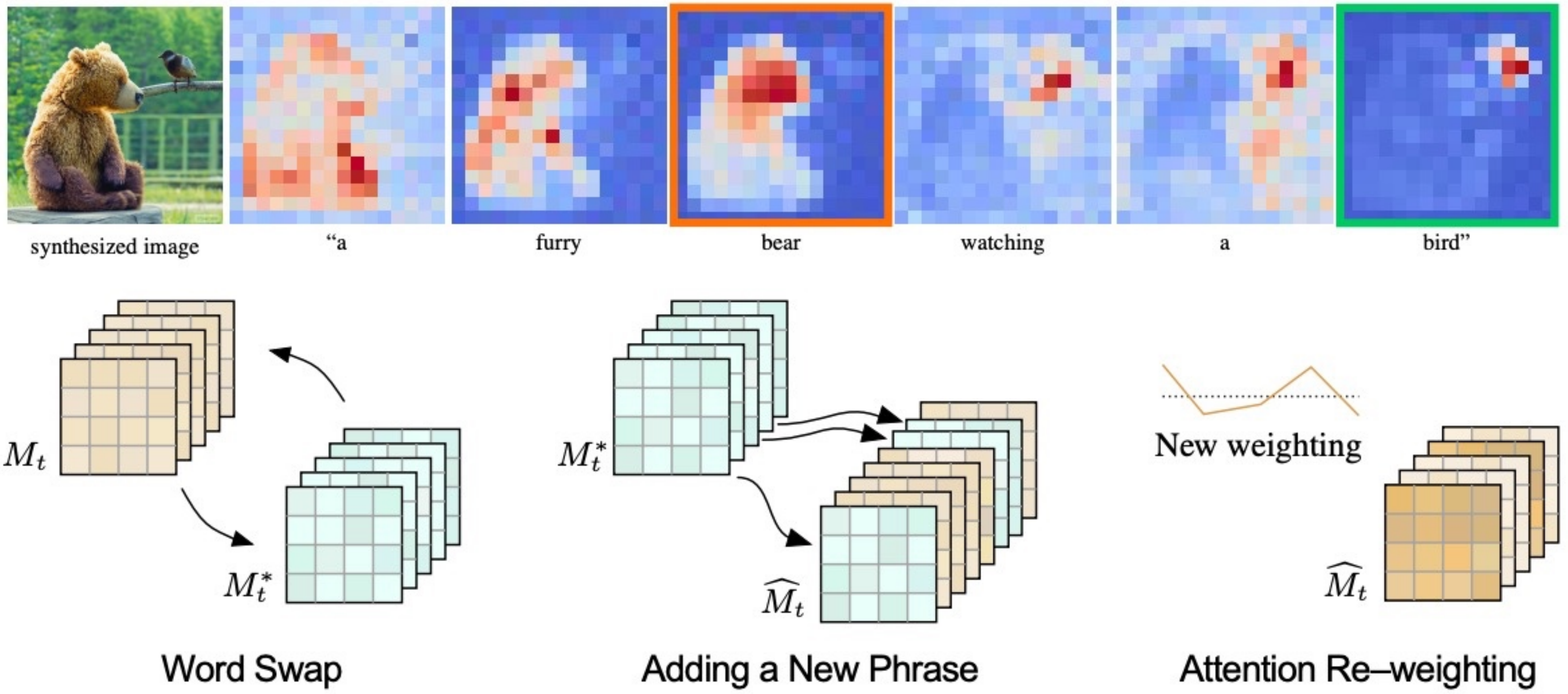
“Photo of a cat riding on a ~~bicycle~~.”
car

“Children drawing of a castle next to a river.”

“a cake with decorations.”
jelly beans

A. Hertz et al. [Prompt-to-Prompt Image Editing with Cross Attention Control](). arXiv 2022

Prompt-to-Prompt Image Editing



Prompt-to-Prompt Image Editing

InstructPix2Pix

Figure 1. Given **an image** and **an instruction** for how to edit that image, our model performs the appropriate edit. Our model does not require full descriptions for the input or output image, and edits images in the forward pass without per-example inversion or fine-tuning.

T. Brooks, A. Holynski, and A. Efros. [InstructPix2Pix: Learning to Follow Image Editing Instructions]. CVPR 2023

InstructPix2Pix

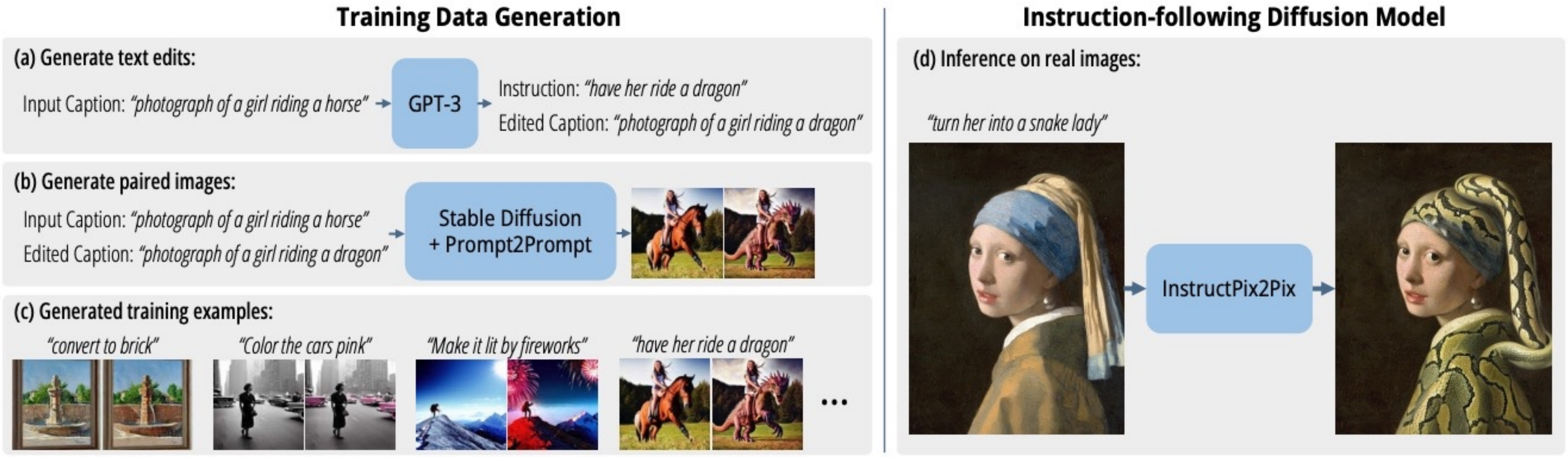


Figure 2. Our method consists of two parts: generating an image editing dataset, and training a diffusion model on that dataset. (a) We first use a finetuned GPT-3 to generate instructions and edited captions. (b) We then use StableDiffusion [52] in combination with Prompt-to-Prompt [17] to generate pairs of images from pairs of captions. We use this procedure to create a dataset (c) of over 450,000 training examples. (d) Finally, our InstructPix2Pix diffusion model is trained on our generated data to edit images from instructions. At inference time, our model generalizes to edit real images from human-written instructions.

T. Brooks, A. Holynski, and A. Efros. [InstructPix2Pix: Learning to Follow Image Editing Instructions](). CVPR 2023

InstructPix2Pix

- Fine-tuning GPT-3:

	Input LAION caption	Edit instruction	Edited caption
Human-written (700 edits)	*Yefim Volkov, Misty Morning*	*make it afternoon*	*Yefim Volkov, Misty Afternoon*
	girl with horse at sunset	*change the background to a city*	*girl with horse at sunset in front of city*
	painting-of-forest-and-pond	*Without the water.*	*painting-of-forest*
	...	...	...
GPT-3 generated (>450,000 edits)	*Alex Hill, Original oil painting on canvas, Moonlight Bay*	*in the style of a coloring book*	*Alex Hill, Original coloring book illustration, Moonlight Bay*
	The great elf city of Rivendell, sitting atop a waterfall as cascades of water spill around it	*Add a giant red dragon*	*The great elf city of Rivendell, sitting atop a waterfall as cascades of water spill around it with a giant red dragon flying overhead*
	Kate Hudson arriving at the Golden Globes 2015	*make her look like a zombie*	*Zombie Kate Hudson arriving at the Golden Globes 2015*
	...	...	...

Table 1. We label a small text dataset, finetune GPT-3, and use that finetuned model to generate a large dataset of text triplets. As the input caption for both the labeled and generated examples, we use real image captions from LAION. Highlighted text is generated by GPT-3.

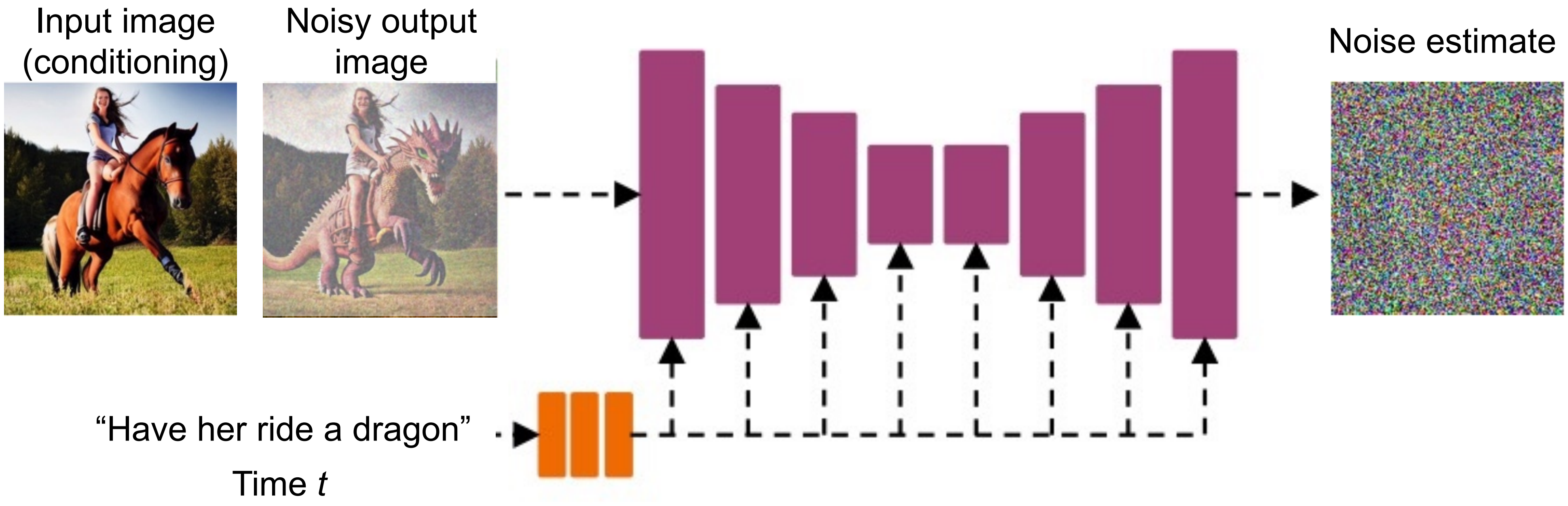
InstructPix2Pix

(a) Without Prompt-to-Prompt. (b) With Prompt-to-Prompt.

Figure 3. Pair of images generated using StableDiffusion [52] with and without Prompt-to-Prompt [17]. For both, the corresponding captions are *"photograph of a girl riding a horse"* and *"photograph of a girl riding a dragon"*.

InstructPix2Pix

- Fine-tuning a DM for image-to-image translation:



InstructPix2Pix: Results

Input | *"Make it a Modigliani painting"* | *"Make it a Miro painting"* | *"Make it an Egyptian sculpture"* | *"Make it a marble roman sculpture"*

Figure 5. *Mona Lisa* transformed into various artistic mediums.

Input | *"Put them in outer space"* | *"Turn the humans into robots"*

Figure 6. *The Creation of Adam* with new context and subjects (generated at 768 resolution).

InstructPix2Pix: Results

Input | "Apply face paint" | "What would she look like as a bearded man?" | "Put on a pair of sunglasses" | "She should look 100 years old"

"What if she were in an anime?" | "Make her terrifying" | "Make her more sad" | "Make her James Bond" | "Turn her into Dwayne The Rock Johnson"

InstructPix2Pix: Results

"Make it Paris" "Make it Hong Kong" "Make it Manhattan" "Make it Prague"

"Make it evening" "Put them on roller skates" "Turn this into 1900s" "Make it underwater"

"Make it Minecraft" "Turn this into the space age" "Make them into Alexander Calder sculptures" "Make it a Claymation"

InstructPix2Pix: Failure cases

“Zoom into the image” “Move it to Mars” “Color the tie blue” “Have the people swap places”

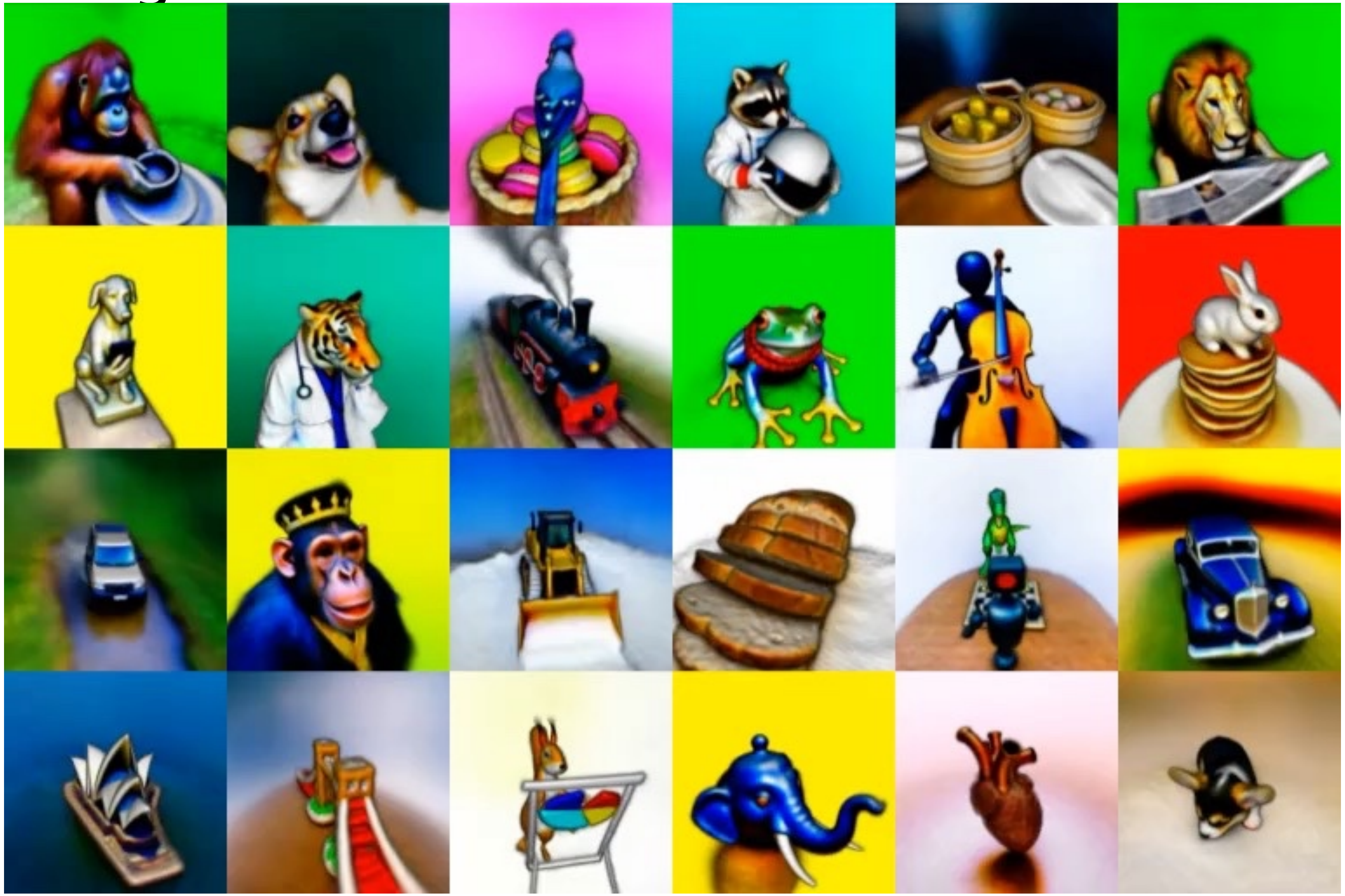
Figure 13. Failure cases. Left to right: our model is not capable of performing viewpoint changes, can make undesired excessive changes to the image, can sometimes fail to isolate the specified object, and has difficulty reorganizing or swapping objects with each other.

Outline

Part 3: Applications and Implementation

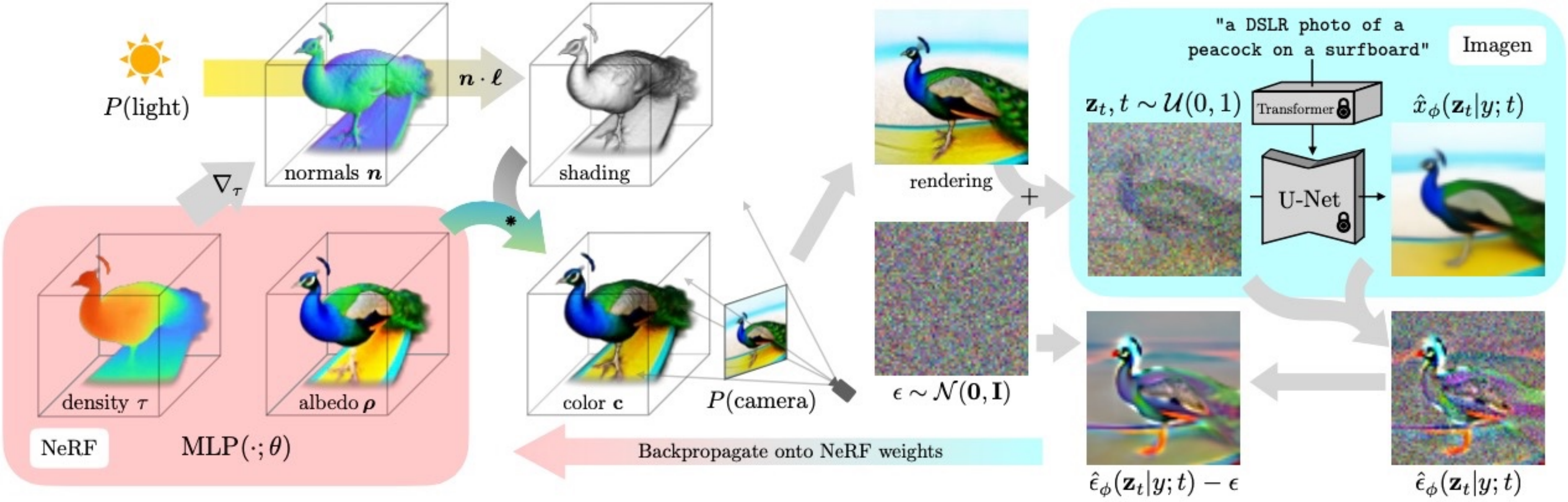
- Customizing Diffusion Models
    - Textual Inversion
    - DreamBooth
    - Low Rank Approximation (LoRA)
    - ZipLoRA
- ControlNet
- Prompt-to-Prompt
- InstructPix2Pix
- DreamFusion

Connecting 2D to 3D: DreamFusion



B. Poole, A. Jain, J. Barron, B. Mildenhall. [DreamFusion: Text-to-3D using 2D Diffusion](). arXiv 2022

Connecting 2D to 3D: DreamFusion



B. Poole, A. Jain, J. Barron, B. Mildenhall. DreamFusion: Text-to-3D using 2D Diffusion. arXiv 2022

Working with Diffusion Models

- Fast moving area with new research coming in everyday
- Variety of pre-trained, open-source models available
- Working directly with the implementations and codebases helps!
- Starting point: diffusers library by huggingface
- Popular open-source models:
    - Stable Diffusion
    - SDXL
    - SD3 (recently announced)
    - Deep-Floyd IF (open-source alternative of Imagen)
- Some useful websites
    - Huggingface, Reddit threads on Stable Diffusion; Civit.ai; publicprompts.art

SC395:

Image Generative Models in Computer Vision

Viraj Shah

Lecture S2

Jan 19th, 2026

sc395.virajshah.com

- stochiastic
- Markov process
- with no jumps

Complex Distribution

↓

Simple Distribution

$$dX_t = a(X_t, t)dt + (random\ mean\ zero\ variance\ \mu(X_t, t)dt)$$

$$q(x_t|x_{t-1}) \sim \mathcal{N}(x_t; \sqrt{1-\beta_t}x_{t-1}, \beta_t\mathbb{I})$$

Forward

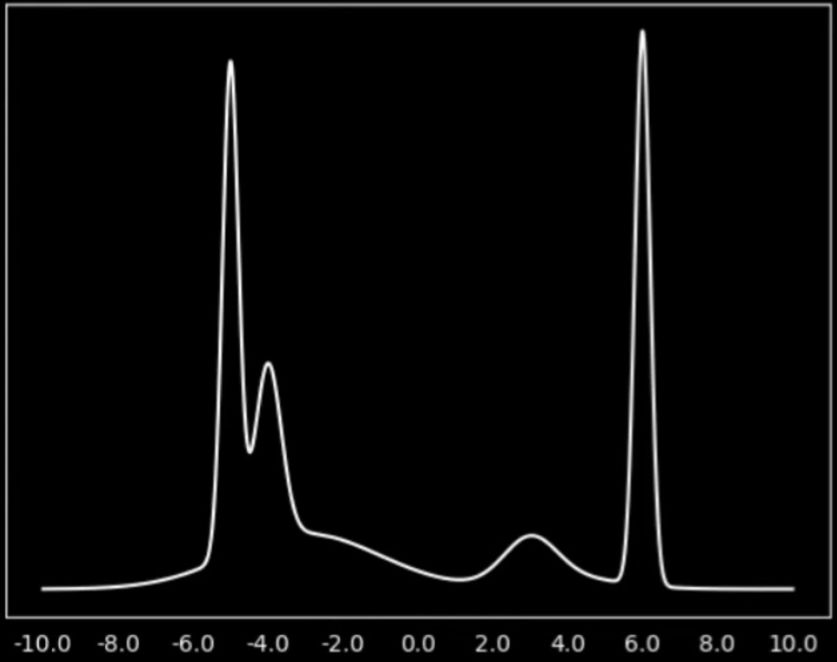
x_0

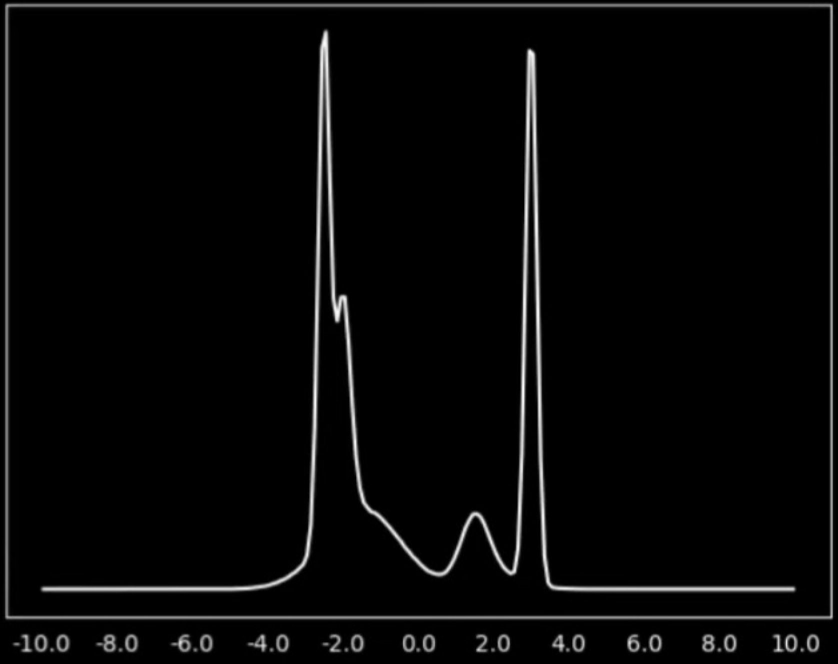
x_T

Reverse

$$x_t = \sqrt{1-\beta}\,x_{t-1} + \sqrt{\beta}\,\mathcal{N}(0, \mathbb{I})$$

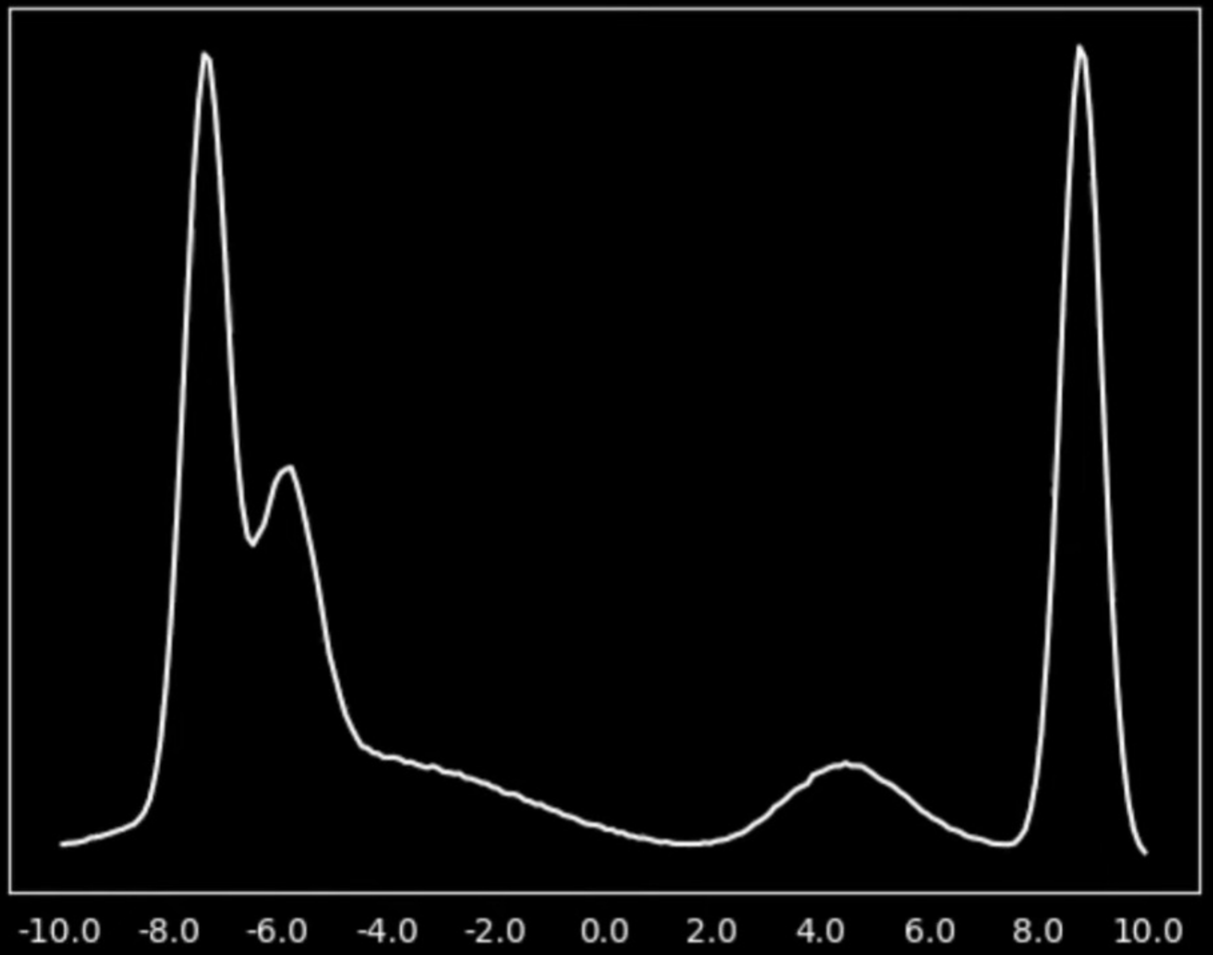
$$x_t = \alpha x_{t-1} + \beta \mathcal{N}(0, \mathbb{I})$$ α = 0.5, β = 0.1



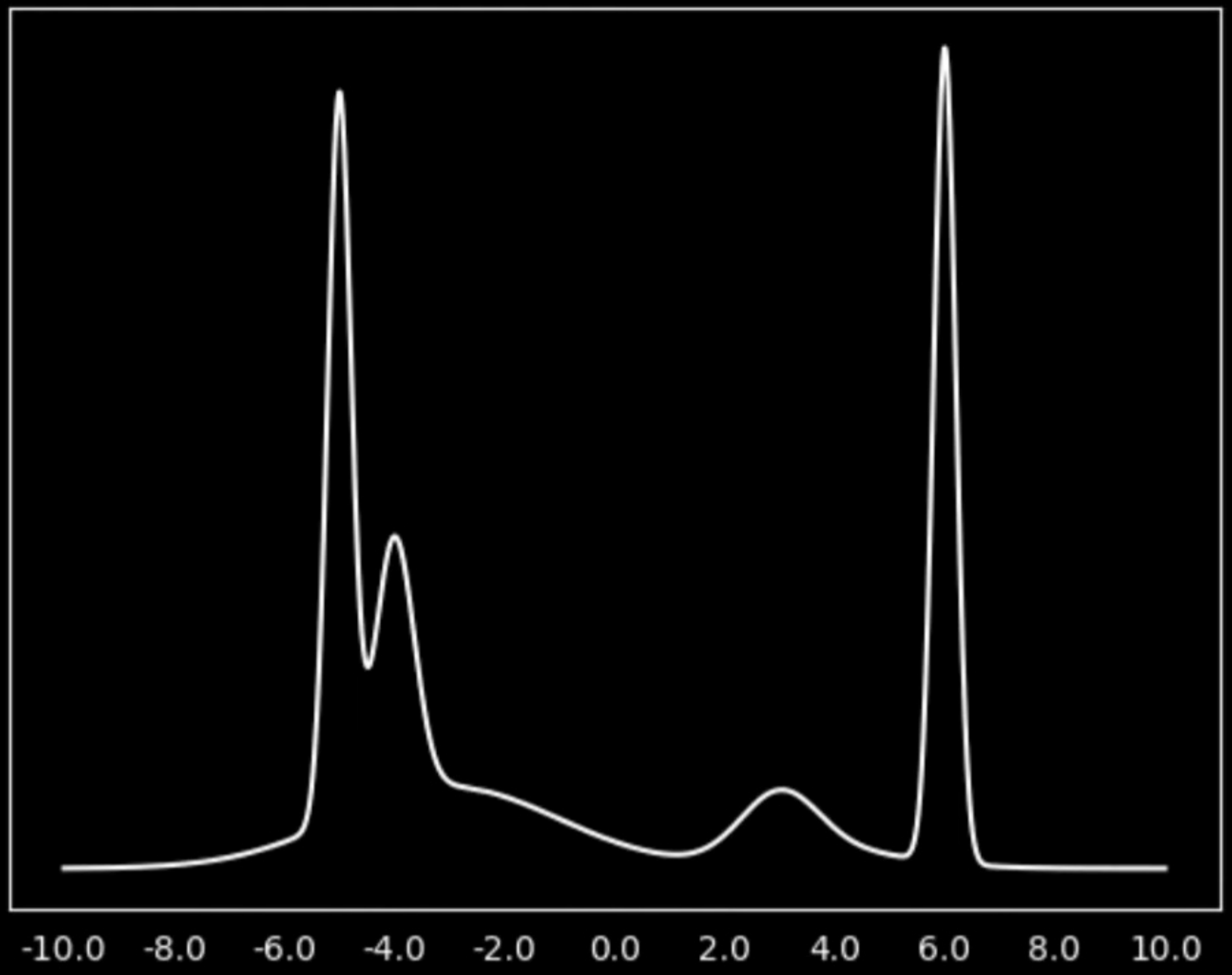


$$x_t = \alpha x_{t-1} + \beta \mathcal{N}(0, \mathbb{I})$$

α = 1.1, β = 0.1

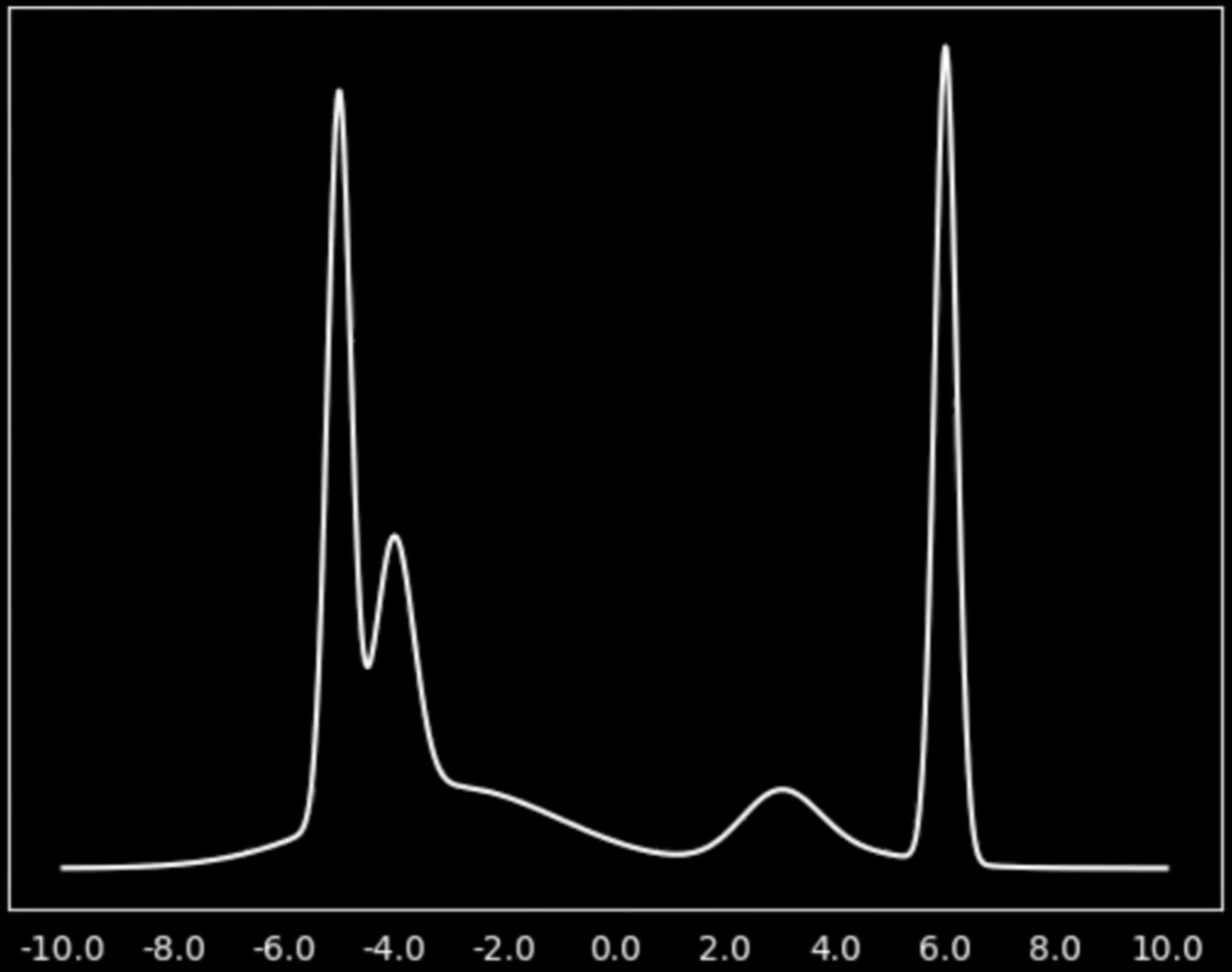


$$x_t = \alpha x_{t-1} + \beta \mathcal{N}(0, \mathbb{I})$$



$$x_t = \alpha x_{t-1} + \beta \mathcal{N}(0, \mathbb{I})$$

α = 0.999, β = 1



$$x_t = \sqrt{\alpha}x_{t-1} + \sqrt{\beta}\mathcal{N}(0,\mathbb{I})$$

$$x_T = \sqrt{1-\beta}^T x_0 + \ldots + \underbrace{\sqrt{1-\beta}\sqrt{1-\beta}\sqrt{\beta}\mathcal{N}(0,\mathbb{I})}_{\beta(1-\beta)(1-\beta)} + \underbrace{\sqrt{1-\beta}\sqrt{\beta}\mathcal{N}(0,\mathbb{I})}_{\beta(1-\beta)} + \underbrace{\sqrt{\beta}\mathcal{N}(0,\mathbb{I})}_{\beta}$$

$$0$$

$$\beta\frac{1-(1-\beta)^T}{1-(1-\beta)} \sim \frac{\beta}{\beta} = 1$$

$$x_T = \sqrt{\alpha}^T x_0 + \ldots \underbrace{\sqrt{\alpha}\sqrt{\alpha}\sqrt{\beta}\mathcal{N}(0,\mathbb{I})}_{\beta\alpha^2} + \underbrace{\sqrt{\alpha}\sqrt{\beta}\mathcal{N}(0,\mathbb{I})}_{\beta\alpha} + \underbrace{\sqrt{\beta}\mathcal{N}(0,\mathbb{I})}_{\beta}$$

$$\beta\frac{1}{1-\alpha}$$

4. So to get the value of x at t=1000, I need to apply transtion 1000 times ?

$$x_t = \sqrt{1-\beta_t}x_{t-1} + \sqrt{\beta_t}\mathcal{N}(0, \mathbb{I}) \qquad \alpha_t = 1 - \beta_t$$

$$x_t = \sqrt{\alpha_t} x_{t-1} + \sqrt{1-\alpha_t}\epsilon_t$$

$$x_t = \sqrt{\alpha_t}(\sqrt{\alpha_{t-1}} x_{t-2} + \sqrt{1-\alpha_{t-1}}\epsilon_{t-1}) + \sqrt{1-\alpha_t}\epsilon_t$$

$$x_t = \sqrt{\alpha_t}\sqrt{\alpha_{t-1}} x_{t-2} + \underbrace{\sqrt{\alpha_t}\sqrt{1-\alpha_{t-1}}\epsilon_{t-1}}_{\alpha_t - \alpha_t\alpha_{t-1}} + \underline{\sqrt{1-\alpha_t}\epsilon_t}\ 1-\alpha_t$$

$$x_t = \sqrt{\alpha_t\alpha_{t-1}} \underline{x_{t-2}} + \sqrt{1-\alpha_t + \alpha_t - \alpha_t\alpha_{t-1}}\epsilon$$

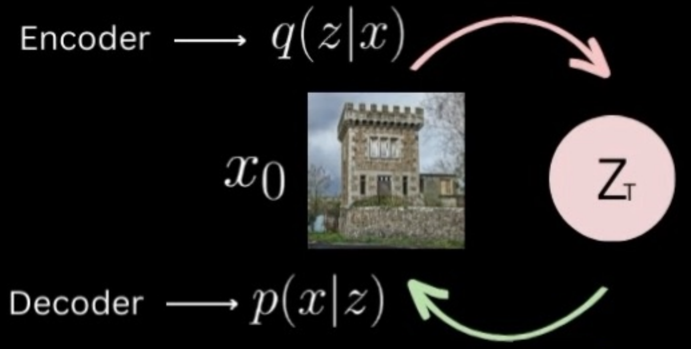
$$x_t = \sqrt{\alpha_t\alpha_{t-1}}(\sqrt{\alpha_{t-2}} x_{t-3} + \sqrt{1-\alpha_{t-2}}\epsilon_{t-2}) + \sqrt{1-\alpha_t\alpha_{t-1}}\epsilon$$

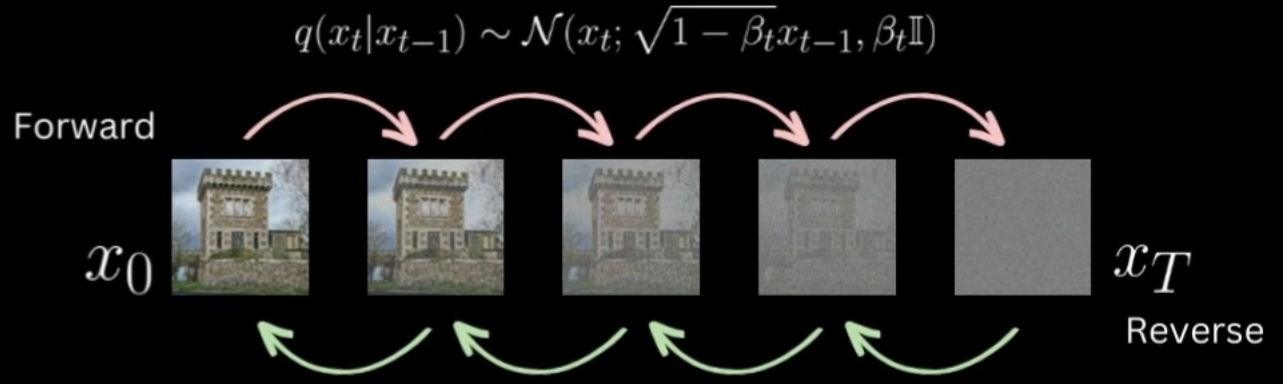
$$x_t = \sqrt{\alpha_t\alpha_{t-1}\alpha_{t-2}} x_{t-3} + \sqrt{\alpha_t\alpha_{t-1}}\sqrt{1-\alpha_{t-2}}\epsilon_{t-2} + \sqrt{1-\alpha_t\alpha_{t-1}}\epsilon$$

$$x_t = \sqrt{\alpha_t\alpha_{t-1}\alpha_{t-2}} x_{t-3} + \sqrt{1-\alpha_t\alpha_{t-1}\alpha_{t-2}}\epsilon$$

$$x_t = \underline{\sqrt{\alpha_t\alpha_{t-1}\alpha_{t-2}\ldots\alpha_2\alpha_1} x_0} + \sqrt{1-\underline{\alpha_t\alpha_{t-1}\alpha_{t-2}\ldots\alpha_2\alpha_1}}\epsilon \qquad \bar{\alpha}_t = \prod_{i=1}^{t} \alpha_i$$

$$x_t = \sqrt{\bar{\alpha}_t} x_0 + \sqrt{1-\bar{\alpha}_t}\epsilon$$





$$\log p(x) = \log \int p(x, z) dz$$

$$\log p(x) = \log \int p(x, z) \frac{q(z|x)}{q(z|x)} dz$$

$$\log p(x) = \log \mathbf{E}_{q(z|x)} \left[\frac{p(x, z)}{q(z|x)} \right]$$

$$\log p(x) \geq \mathbf{E}_{q(z|x)} \left[\log \frac{p(x, z)}{q(z|x)} \right]$$

$$\log p(x) \geq \mathbf{E}_{q(z|x)} \left[\log p(x|z) \right] - \mathbb{D}_{KL}(q(z|x) || p(z))$$

$$\log p(x_0) = \log \int p(x_{0:T}) dx_{1:T}$$

$$\log p(x_0) = \log \int p(x_{0:T}) \frac{q(x_{1:T}|x_0)}{q(x_{1:T}|x_0)} dx_{1:T}$$

$$\log p(x_0) = \log \mathbf{E}_{q(x_{1:T}|x_0)} \left[\frac{p(x_{0:T})}{q(x_{1:T}|x_0)} \right]$$

$$\log p(x_0) \geq \mathbf{E}_{q(x_{1:T}|x_0)} \left[\log \frac{p(x_{0:T})}{q(x_{1:T}|x_0)} \right]$$

$$\log \frac{p(x_{0:T})}{q(x_{1:T}|x_0)}$$

$$\log \frac{p(x_T)\prod_{t=1}^{T} p_\theta(x_{t-1}|x_t)}{\underline{q(x_1|x_0)\prod_{t=2}^{t=T} q(x_t|x_{t-1},x_0)}}$$

$$q(x_1|x_0)\prod_{t=2}^{t=T} q(x_t|x_{t-1},x_0)$$

$$q(x_1|x_0)\prod_{t=2}^{t=T} \frac{q(x_{t-1}|x_t,x_0)q(x_t|x_0)}{q(x_{t-1}|x_0)}$$

$$\cancel{q(x_1|x_0)}\frac{q(x_{T-1}|x_T,x_0)q(x_T|x_0)q(x_{T-2}|x_{T-1},x_0)\cancel{q(x_{T-1}|x_0)}\ldots q(x_2|x_3,x_0)q(x_3|x_0)q(x_1|x_2,x_0)\cancel{q(x_2|x_0)}}{\cancel{q(x_{T-1}|x_0)}q(x_{T-2}|x_0)\ldots\cancel{q(x_2|x_0)}\cancel{q(x_1|x_0)}}$$

$$\log \frac{p(x_{0:T})}{q(x_{1:T}|x_0)}$$

$$\log \frac{p(x_T) \prod_{t=1}^{T} p_\theta(x_{t-1}|x_t)}{q(x_1|x_0) \prod_{t=2}^{t=T} q(x_t|x_{t-1}, x_0)} \longrightarrow q(x_T|x_0) \prod_{t=2}^{t=T} q(x_{t-1}|x_t, x_0)$$

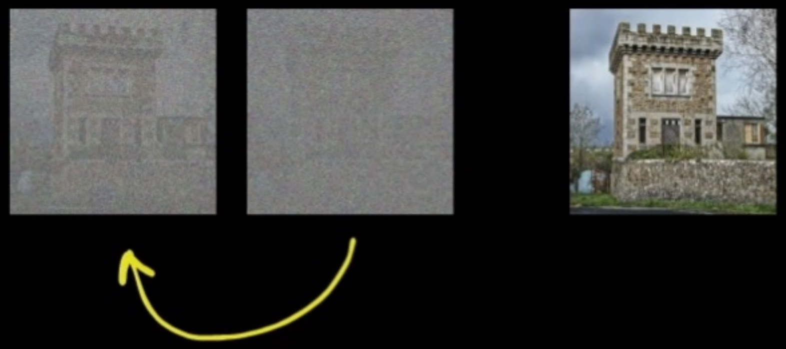
$$\log \frac{p(x_T) \prod_{t=1}^{t=T} p_\theta(x_{t-1}|x_t)}{q(x_T|x_0) \prod_{t=2}^{t=T} q(x_{t-1}|x_t, x_0)}$$

$$\log \frac{p(x_T) p_\theta(x_0|x_1) \prod_{t=2}^{t=T} p_\theta(x_{t-1}|x_t)}{q(x_T|x_0) \prod_{t=2}^{t=T} q(x_{t-1}|x_t, x_0)}$$

$$\log \frac{p(x_T)}{q(x_T|x_0)} + \log p_\theta(x_0|x_1) + \sum_{t=2}^{T} \log \frac{p_\theta(x_{t-1}|x_t)}{q(x_{t-1}|x_t, x_0)}$$

6. We are back at same problem as we don’t know reverse distribution.
So how do we move forward ?

$$q(x_{t-1}|xt\ , x_0)$$



$$q(x_{t-1}|x_t,x_0) = \frac{q(x_t|x_{t-1},x_0)q(x_{t-1}|x_0)}{q(x_t|x_0)}$$

$$q(x_{t-1}|x_t,x_0) = \frac{\mathcal{N}(x_t;\sqrt{\alpha_t}x_{t-1},(1-\alpha_t)\mathbb{I})\mathcal{N}(x_{t-1};\sqrt{\bar\alpha_{t-1}}x_0,(1-\bar\alpha_{t-1})\mathbb{I})}{\mathcal{N}(x_t;\sqrt{\bar\alpha_t}x_0,(1-\bar\alpha_{t-1})\mathbb{I})}$$

$$\exp\frac{-1}{2}\left[\frac{(x_t-\sqrt{\alpha_t}x_{t-1})^2}{1-\alpha_t}+\frac{(x_{t-1}-\sqrt{\bar\alpha_{t-1}}x_0)^2}{1-\bar\alpha_{t-1}}-\frac{(x_t-\sqrt{\bar\alpha_t}x_0)^2}{1-\bar\alpha_t}\right]$$

$$\exp\frac{-1}{2}\left[x_{t-1}^2(\frac{\alpha_t}{1-\alpha_t}+\frac{1}{1-\bar\alpha_{t-1}})-2x_{t-1}(\frac{\sqrt{\alpha_t}x_t}{1-\alpha_t}+\frac{\sqrt{\bar\alpha_{t-1}}x_0}{1-\bar\alpha_{t-1}})+(\dots)\right]$$

$$\exp\frac{-1}{2}\left[x_{t-1}^2(\frac{\alpha_t-\bar\alpha_t+1-\alpha_t}{(1-\alpha_t)(1-\bar\alpha_{t-1})})-2x_{t-1}(\frac{\sqrt{\alpha_t}x_t}{1-\alpha_t}+\frac{\sqrt{\bar\alpha_{t-1}}x_0}{1-\bar\alpha_{t-1}})+(\dots)\right]$$

$$\exp\frac{-1}{2}\left[x_{t-1}^2(\frac{\alpha_t-\bar\alpha_t+1-\alpha_t}{(1-\alpha_t)(1-\bar\alpha_{t-1})})-2x_{t-1}(\frac{(1-\bar\alpha_{t-1})\sqrt{\alpha_t}x_t+(1-\alpha_t)\sqrt{\bar\alpha_{t-1}}x_0}{(1-\alpha_t)(1-\bar\alpha_{t-1})})+(\dots)\right]$$

$$\exp\frac{-1}{2}\left[\frac{1}{\frac{(1-\alpha_t)(1-\bar\alpha_{t-1})}{1-\bar\alpha_t}}\left\{x_{t-1}^2-2x_{t-1}(\frac{(1-\bar\alpha_{t-1})\sqrt{\alpha_t}x_t+(1-\alpha_t)\sqrt{\bar\alpha_{t-1}}x_0}{(1-\bar\alpha_t)})+\frac{(1-\alpha_t)(1-\bar\alpha_{t-1})}{1-\bar\alpha_t}(\dots)\right\}\right]$$