Personalization of Diffusion Models

Using 🧨 Diffusers

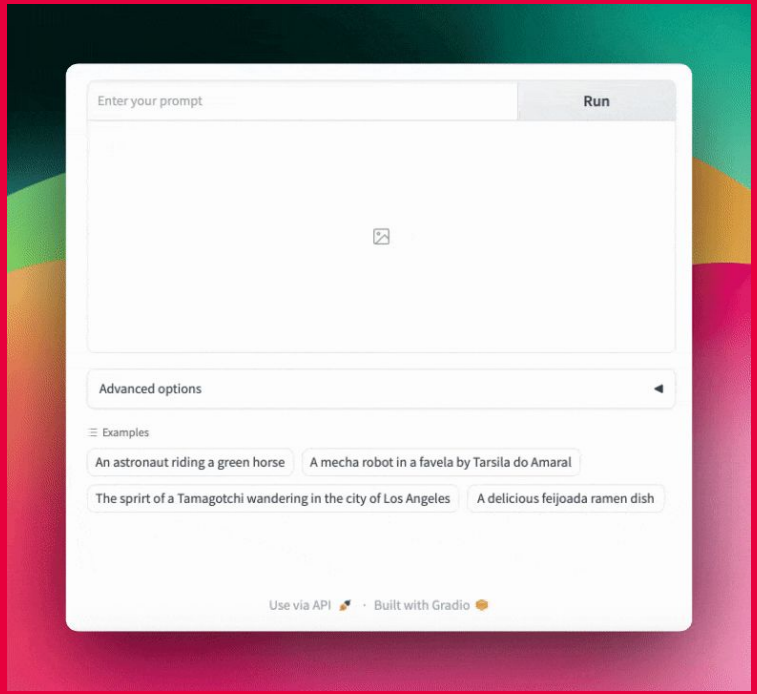
pyladies amsterdam

X

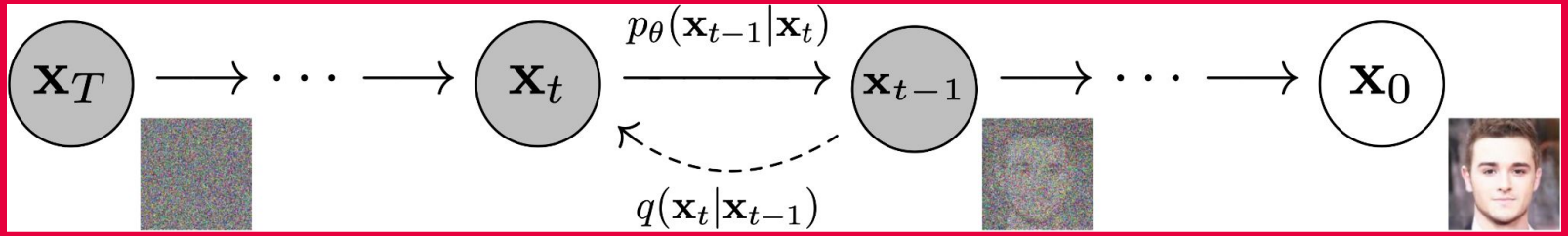
Linoy Tsaban
ML Engineer
Hugging Face

February 29th, 2024

Diffusion Models



Diffusion Models

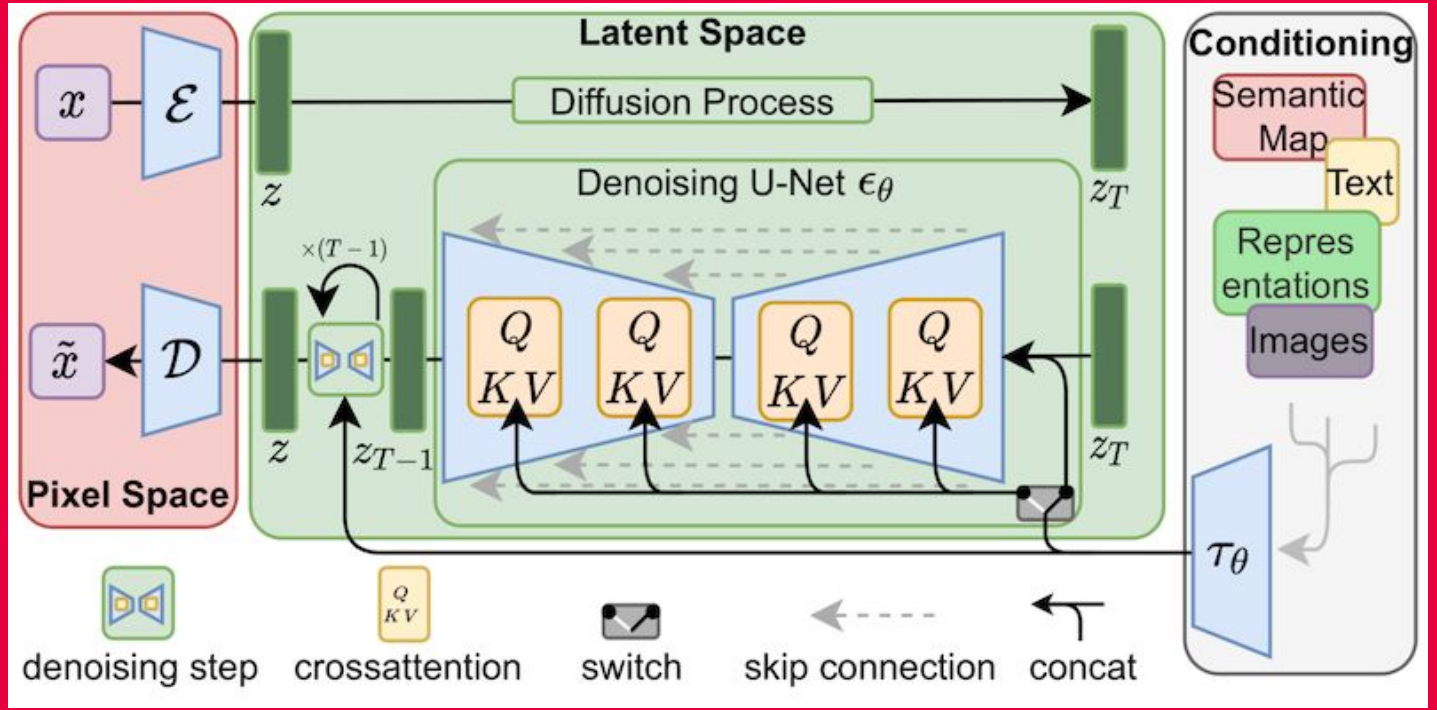


Stable Diffusion

Pixel space

Latent space

Stable Diffusion



Latest Release - Stable Cascade

Click **here** to try it out!

What is **Personalization**?

*Usually**, refers to the process of teaching new concepts to a diffusion model by using specialized forms of fine-tuning.

E.g. Dreambooth, Textual Inversion, LoRA

*

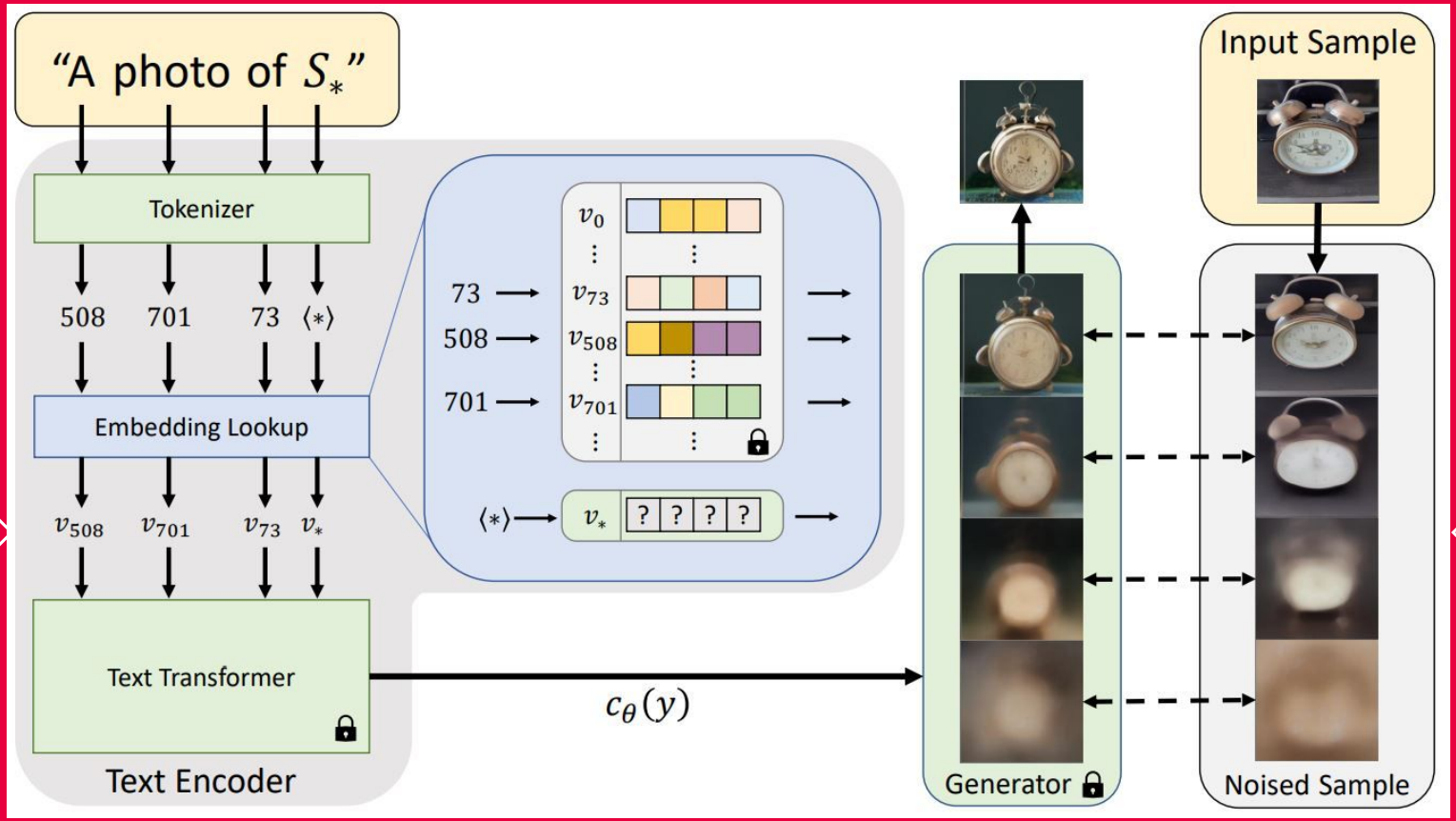
Textual Inversion

Textual Inversion

(Fine Tuning Technique)
When we teach a new concept to Stable Diffusion by finding new text embeddings that represent the concept. These embeddings are then linked to new pseudo-words, which can be incorporated into new sentences like any other word.

Gal, Rinon, et al. "An image is worth one word: Personalizing text-to-image generation using textual inversion."

Textual Inversion



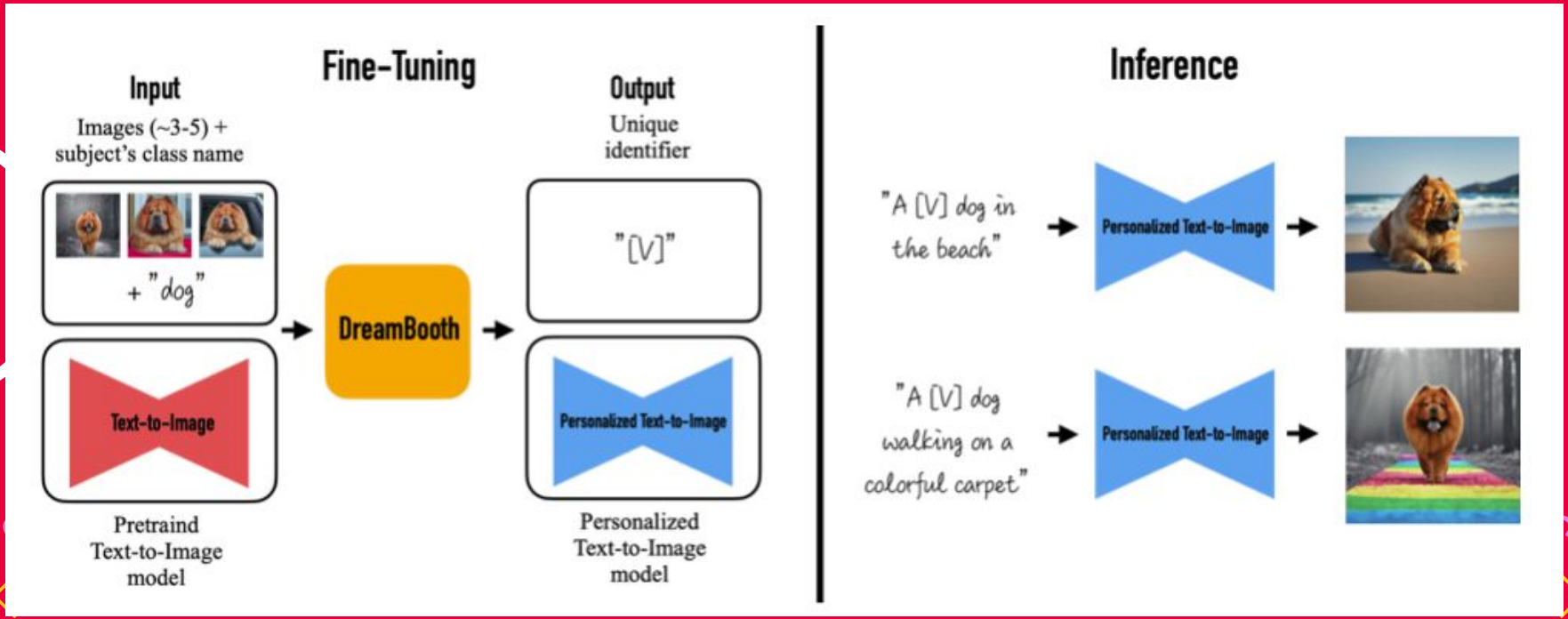
Dreambooth

Dreambooth

(Fine Tuning Technique)
When we teach a new concept to Stable Diffusion by fine tuning on 3-5 input images paired with a text prompt containing a unique identifier and the name of the class the subject belongs to (e.g., "A photo of a [T] dog”)

Ruiz, Nataniel, et al. "Dreambooth: Fine tuning text-to-image diffusion models for subject-driven generation."

Dreambooth



Ruiz, Nataniel, et al. "Dreambooth: Fine tuning text-to-image diffusion models for subject-driven generation."

LoRA

Hu, Edward J., et al. "Lora: Low-rank adaptation of large language models."

1. First introduced for LLMs, and later proposed by Simo Ryu for diffusion models
2. Proposes to freeze pre-trained model weights and inject trainable layers in transformer blocks
3. Quality on par with full fine-tuning while being faster & needing less compute

https://github.com/cloneofsimo/lora

LoRA

LoRA: Low-Rank Adaptation of Large Language Models

For a pre-trained weight matrix *W*0, LoRA decomposes its update into **W0 + ΔW** and represents ΔW as a product of **two low-rank matrices, B and A.**

Both B and A have ranks significantly smaller than the dimensions of W0, dramatically reducing the number of trainable parameters.

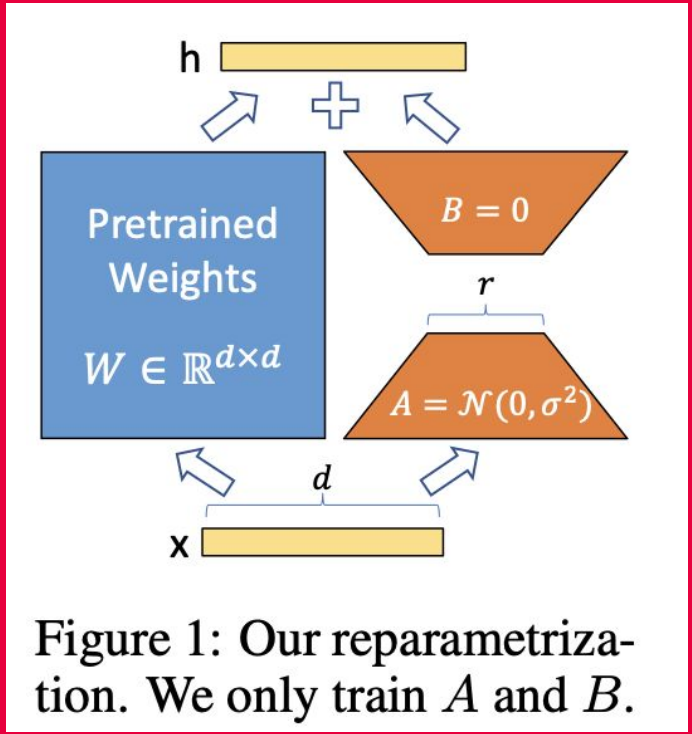


Figure 1: Our reparametrization. We only train A and B.

Hu, Edward J., et al. "Lora: Low-rank adaptation of large language models."

Recap

	Output
Dreambooth	Fine tuned model
LoRA	Injectable layers
Textual Inversion	Text embeddings

Let’s Train 

What if
we combine
Dreambooth
with **LoRA?**

Recap

	Output
Dreambooth	Fine tuned model
LoRA	Injectable layers
Textual Inversion	Text embeddings
Dreambooth LoRA	**Injectable layers**

What if **#2:**
We add in
Textual
inversion?

Share your models with the community!

linoyts / **2000_ads** ♡ like 2

Text-to-Image | Diffusers | stable-diffusion-xl | stable-diffusion-xl-diffusers | lora | template:sd-lora | License: openrail++

Model card | Files and versions | Community | Settings | Deploy | Use in Diffusers

Edit model card

SDXL LoRA DreamBooth - LinoyTsaban/2000_ads

Prompt
<s0><s1> ad of a llama wearing headphones

Prompt
<s0><s1> ad for banana flavored toothpaste

Prompt
<s0><s1> ad of the Mona Lisa eating ramen

Prompt
<s0><s1> ad for speakers

Prompt
<s0><s1> ad of Santa clause

Prompt
<s0><s1> ad of the an astronaut

Downloads last month
232

Inference API

Text-to-Image | Examples

<s0><s1> ad for speakers | Compute

This model can be loaded on the Inference API on-demand.

Pivotal Tuning

Pivotal Tuning

(Fine Tuning Technique)
combines Textual Inversion with regular diffusion fine-tuning by inserting new tokens into the text encoders of the model, instead of reusing existing ones and optimize the associated embeddings. We continue to fine-tune the LoRA layers of the unet with the new textual embeddings.

Recap

	Output
Dreambooth	Fine tuned model
LoRA	Injectable layers
Textual Inversion	Text embeddings
Dreambooth LoRA	**Injectable layers**
Dreambooth LoRA with Pivotal Tuning	**Injectable layers & text embeddings**

Learn More 🤗

🎬**Introduction to Diffusion Models:**
Hugging Face Diffusion Models Course

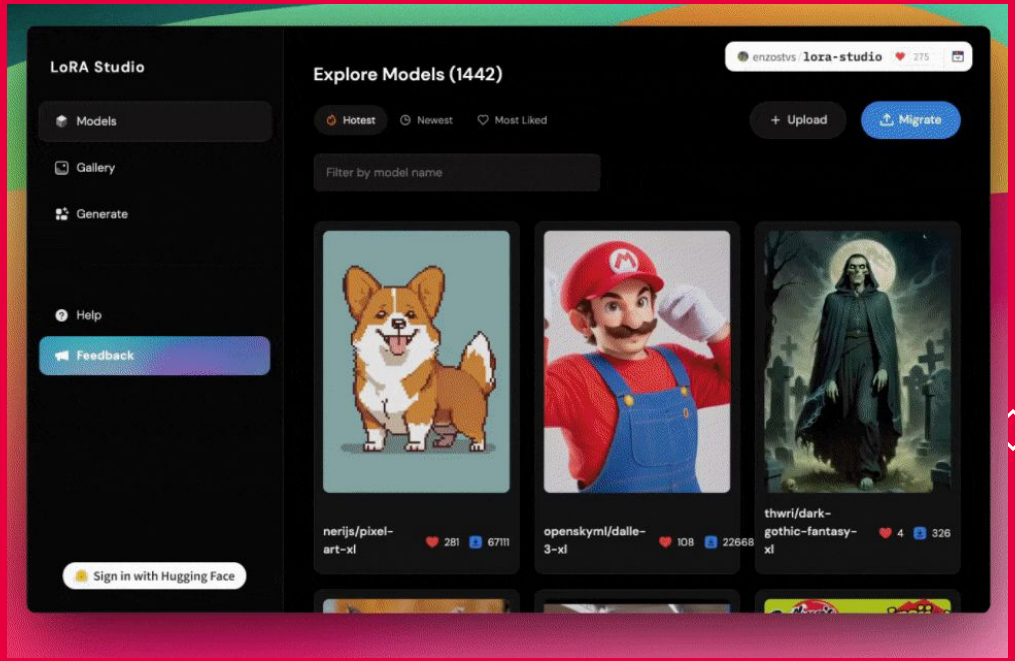
💌**Blog Posts:**
Blog post about Dreambooth
Blog post about LoRAs
Blog post about current SOTA practices for Dreambooth LoRA

📚**Useful Docs:**
Dreambooth with Diffusers
Textual Inversion with Diffusers
LoRA with Diffusers

Personalization Playground

Demos worth trying 🔥

- ▸ LoRA Studio
- ▸ LoRA the explorer
- ▸ LoRA roulette
- ▸ IP Adapter
- ▸ Instant ID



Thank You!