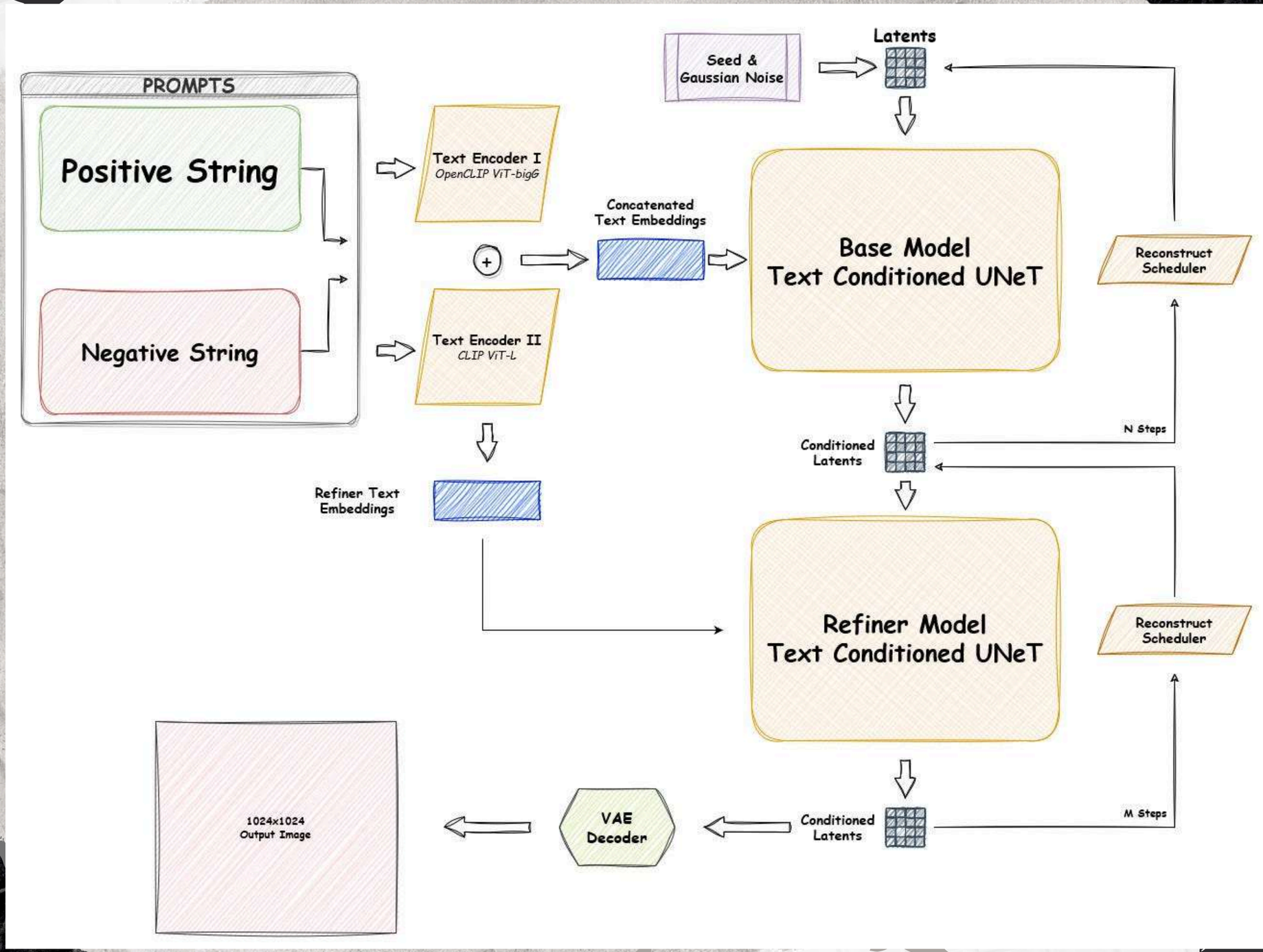
SDXL

1. Gets text input and tokenizes it (**CLIP**)
2. Generates latent space
3. Denoising (**U-Net**)
4. Decoding (**VAE**)
5. Adding details (**Refiner**)


PROMPTS
Positive String
Negative String
Text Encoder I
OpenCLIP ViT-bigG
Text Encoder II
CLIP ViT-L
Concatenated
Text Embeddings
Refiner Text
Embeddings
Seed &
Gaussian Noise
Latents
Base Model
Text Conditioned UNeT
Reconstruct
Scheduler
N Steps
Conditioned
Latents
Refiner Model
Text Conditioned UNeT
Reconstruct
Scheduler
M Steps
Conditioned
Latents
VAE
Decoder
1024x1024
Output Image

LORA

1. LoRA adds two low-rank matrix multiplications (A and B) (rank=4 f.i) to specific SDXL blocks.
2. These multiplications are trained on your custom data (e.g., tattoo images).
3. During the inference, the SDXL model uses the basic weights + LoRA add-ons.

You can mix styles (multiple LoRAs), turn stylization on/off **without retraining the entire model**.

DATASET

Split (1)

train · 2.08k rows

Search this dataset

image image · *width (px)* 512 512	text string · *lengths* 6 70
	an anchor with flowers on it
	a rose
	a skull with roses in the background
	a flower

MODEL AND TRAIN

- Diffusers + PEFT
- resolution = 512
- rank = 8
- 8-bit Adam
- learning_rate = 1e-4
- epochs = 5
- training text encoder

INFERENCE

Euler Ancestral Scheduler

Mixture-of-Experts

Base (80%): LoRA → Applies concept/style, building the rough structure and layout of the image.

Refiner (20%): SDXL → Adds fine details and improving visual quality.

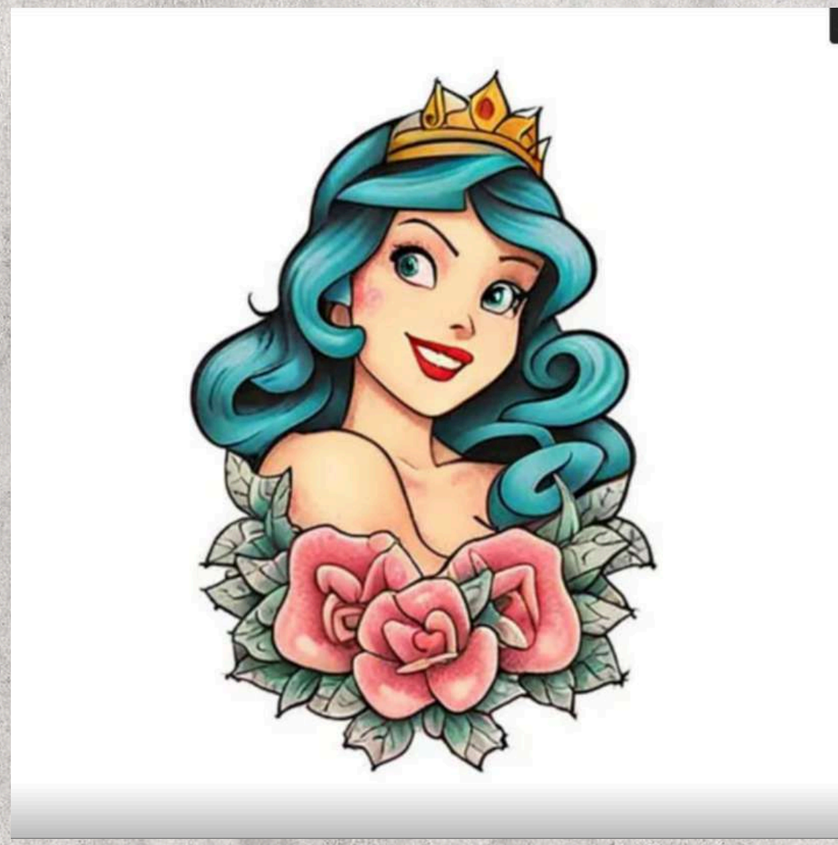
COMPARISON

MODELS

"A disney princess tattoo"

SD 1.5 + LoRA

SDXL



SDXL + LoRA

MODELS

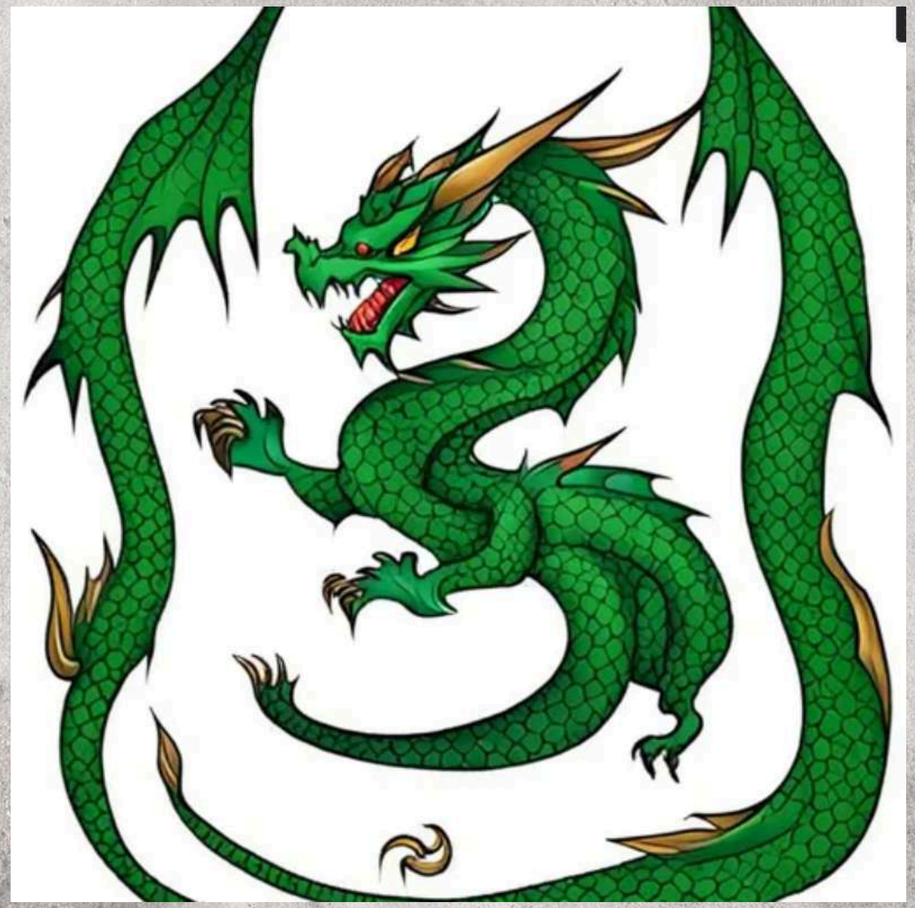
“A crazy clown tattoo”

SDXL

SDXL + LoRA

NEGATIVE PROMPT

“A green dragon tattoo”



GRAYSCALE

SDXL

SDXL
+
LoRA

COMPARISON

Service

Our