Navigating Text-To-Image Customization: From LyCORIS Fine-Tuning to Model Evaluation

Shih-Ying Yeh[1], Yu-Guan Hsieh[2], Zhidong Gao[3], Bernard B W Yang[4], Giyeong Oh[5], Yanmin Gong[3]

[1]National Tsing Hua University [2]Apple [3]University of Texas at San Antonio [4]University of Toronto [5]Yonsei University

Q & A

Q: What does LyCORIS stands for?
A: LyCORIS is the abbreviation of *"**L**ora be**Y**ond **C**onventional methods, **O**ther **R**ank adaptation **I**mplementations for **S**table diffusion"*

Q: Where does the name LyCORIS comes from?
A: LyCORIS refers to the flower Lycoris radiata, aka red spider lily, or "higanbana". In particular, the name does **NOT** come from the anime Lycoris Recoil.

Q: Why does someone claim LyCORIS leads to better results whereas the others claim the opposite?
A: There can be many reasons. To name a few, they may use different algorithms from LyCORIS, train on very different datasets, or run with different hyperparameters / optimization algorithms ...

Related Projects

- Automatic1111/stable-diffusion-webui
- kohya-ss/sd-scripts
- cloneofsimo/lora

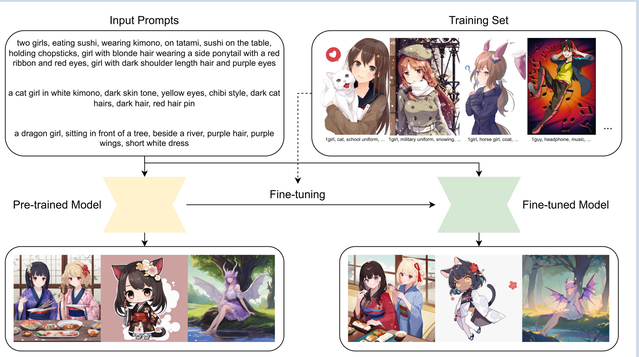
Algorithm References

- (LoRA) LoRA: Low-Rank Adaptation of Large Language Models. Hu et al., ICLR 2022.
- (IA^3) Few-shot parameter-efficient fine-tuning is better and cheaper than in-context learning. Liu et al, NeurIPS 2022.
- (GLoRA) One-for-all: Generalized lora for parameter-efficient fine-tuning. Chavan et al, 2023.
- (DoRA) Dora: Weight-decomposed low-rank adaptation. Liu et al, 2024.
- (OFT) Controlling text-to-image diffusion by orthogonal finetuning. Qiu et al., NeurIPS 2023.
- (BOFT) Parameter-efficient orthogonal finetuning via butterfly factorization. Liu et al., ICLR 2024.

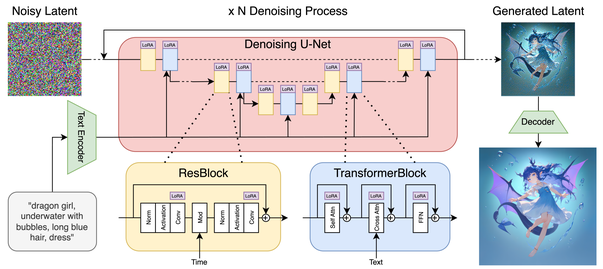
Image Credits

- 米望ゆめ / bogdan / Animenbo / ke-ta
- Spider Lily / アニプレックス・ABCアニメーション・BS11

Customization of Text-To-Image Models



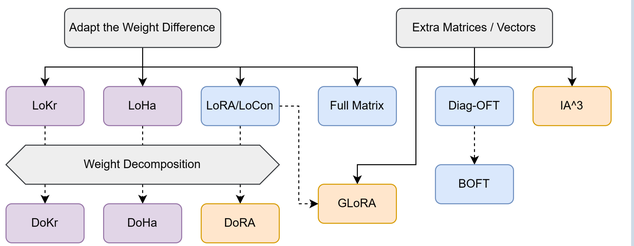
Stable Diffusion and LoRA Fine-Tuning



LoRA : $h' = W_0 h + b + \gamma\, \Delta W\, h = W_0 h + b + \gamma\, BA\, h$

where $B \in \mathbb{R}^{p \times r}$, $A \in \mathbb{R}^{r \times q}$, and $\gamma = \alpha / r$

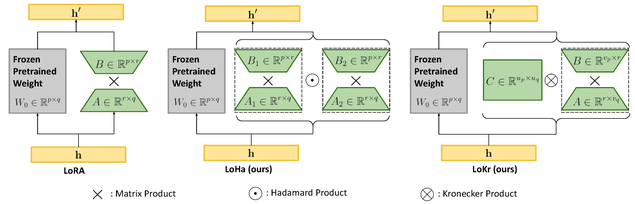
Algorithms Implemented in LyCORIS



TL;DR

1. LyCORIS is an open-source library that promotes easy adaption of various fine-tuning algorithms for Stable Diffusion
2. To enable fair comparison, we propose an extensive framework for model evaluation and conduct it for a subset of algorithms in LyCORIS

LoCon LoHa LoKr



LoHa : $h' = W_0 h + b + \gamma\, \Delta W\, h = W_0 h + b + \gamma\, [(B_1 A_1) \odot (B_2 A_2)]\, h$

LoKr : $h' = W_0 h + b + \gamma\, \Delta W\, h = W_0 h + b + \gamma\, [C \otimes (BA)]\, h$

Theorem (Effect of γ)

Assume that forward pass is modified to

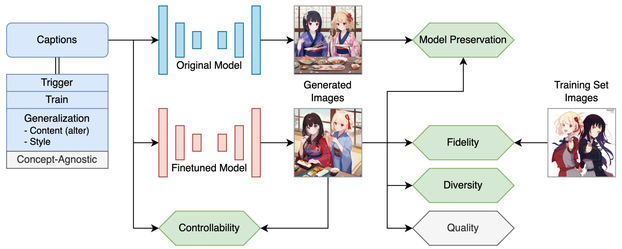
$$h' = W_0 h + b + \gamma\, \Delta W\, h = W_0 h + b + \gamma\, \mathcal{T}(A_1, \ldots, A_m)\, h$$

where T is a homogeneous operator

Mathematically, it is equivalent to training without γ and

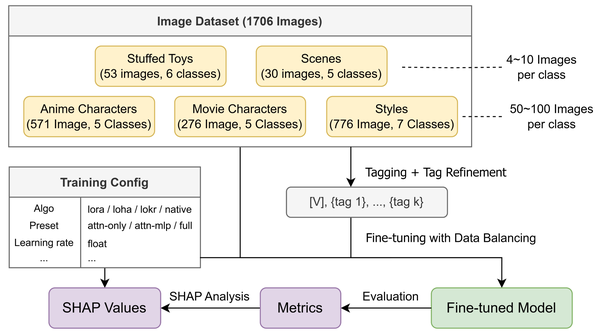
- scaling the initialization parameters by $\gamma^{1/m}$
- scaling the learning rate by $\gamma^{c/m}$, where c=2 for SGD and c=1 for Adam, etc

Evaluation Criteria



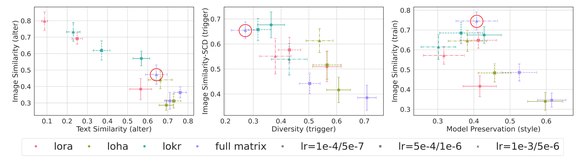
Experiments

Setup



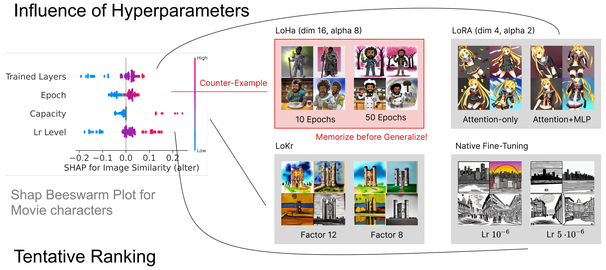
Results

- High-Level Observations
  - Image Fidelity is mostly negative correlated with other metrics
  - Choice of prompt can affect significantly the evaluation



(Movie characters, 30 epochs)

- Influence of Hyperparameters



Shap Beeswarm Plot for Movie characters

- Tentative Ranking

	LoRA	LoHa	LoKr	Native lr 5e-6	Native lr 1e-6
Fidelity	★★★	★★	★★★★	★★★★★	★
Controllability	★★	★★★★	★	★★★	★★★★★
Diversity	★★★	★★★★	★★	★	★★★★★