Adding Conditional Control to Diffusion Models with Reinforcement Learning

Yulai Zhao*[1,2], Masatoshi Uehara[2]*, Gabriele Scalia[2], Tommaso Biancalani[2], Sergey Levine[3], Ehsan Hajiramezanali[2]
[1]Princeton University, [2]Genentech, [3]University of California, Berkeley

Genentech
A Member of the Roche Group

Biology Research | AI Development

Background

- A **pre-trained conditional diffusion** model excels at modeling **p(x|c)**.
  - For example, in Stable Diffusion, c∈C is a prompt, and x∈X is the image generated according to this prompt.
  - Many tailored DMs are able to generate biological sequences (e.g., DDSM).
- In practice, **we often want to add additional controls into pre-trained diffusion models**, e.g.
  - **Stable Diffusion**.
    - existing condition: prompts
    - new condition: certain layouts or backgrounds.
  - **DDSM tailored for generating DNA enhancers**.
    - existing condition: activity level in HepG2
    - new condition: activity level in other cell lines such as K562.

Settings

- Given the pre-trained model, which enables us to sample from $p^{pre}(x|c): C \mapsto \Delta(X)$.
- **Goal**: add new conditional controls **y∈Y** such that we can sample from **p(x|c,y)**.
- Assume we can access to offline data:

$$D = \{(c^{(i)}, x^{(i)}, y^{(i)})\}_{i=1}^{n}$$

where conditional distribution is denoted by $p^{\diamond}(y|x,c)$.

Target Distribution

our goal is to obtain a diffusion model such that we can sample from

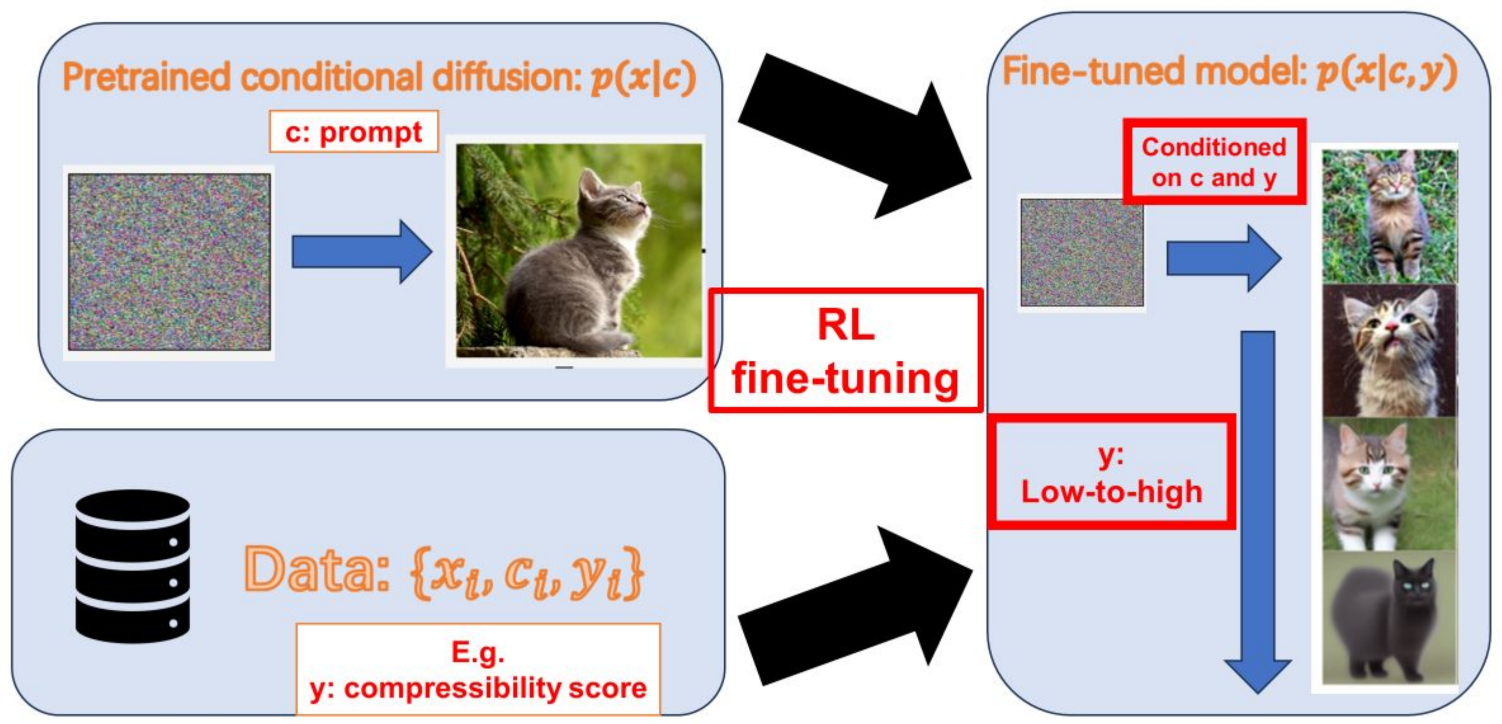
$$p_{\gamma}(\cdot|c,y) \propto (p^{\diamond}(y|\cdot,c))^{\gamma}\, p^{pre}(\cdot|c)$$

where γ represents the strength of the additional guidance.

Methodology & Results

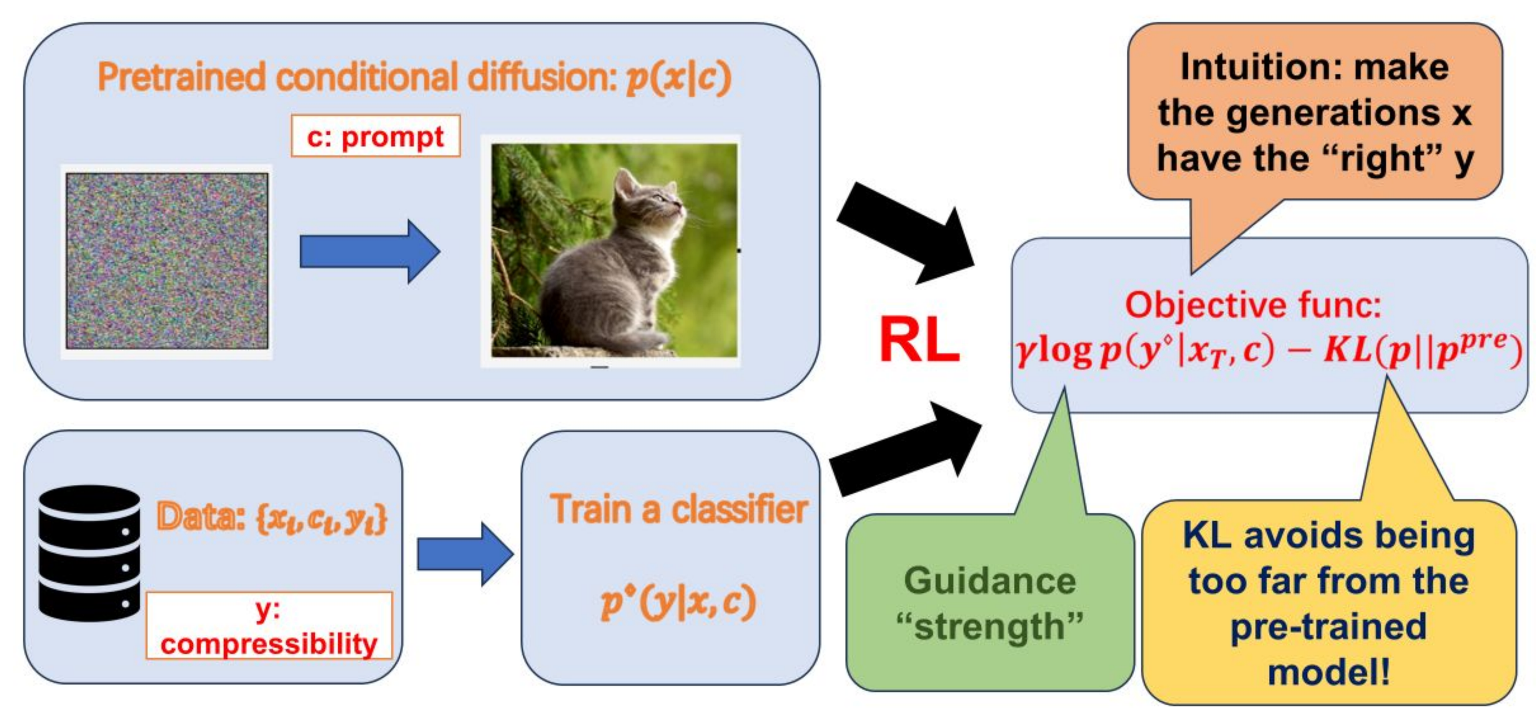
Roadmap

Our goal: adding control via fine-tuning



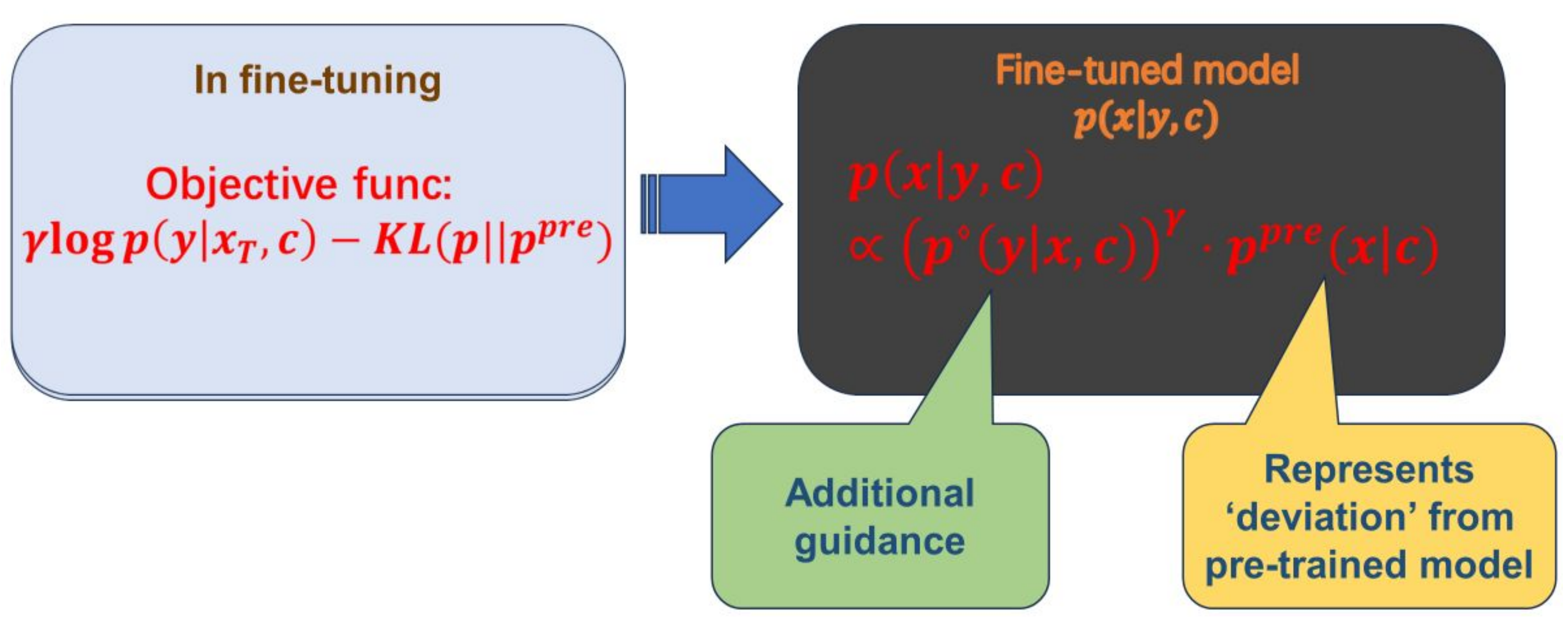
RL-based Fine-tuning

Methodology



Advantages of Our Approach

Theoretical justification (incomplete)



★ **Compared to classifier-free guidance**

Our method demonstrates superiority by **leveraging the conditional independence**.

- **Example 1: Y ⊥ C | X.** Then

$$p^{\diamond}(y|x,c) = p^{\diamond}(y|x)$$

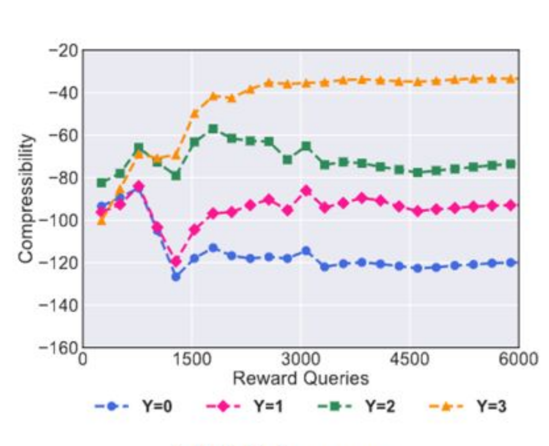
  - This means we only need **(x,y)** sample pairs to train the classifier, rather than triplets **(c,x,y)**.
- **Example 2 (multi-task): $Y_1 \perp Y_2$ | C, X, $Y_1 \perp C$ | X and $Y_2 \perp C$ | X.** Then

$$\log p^{\diamond}(y_1,y_2|x,c) = \log p^{\diamond}(y_1|x) + \log p^{\diamond}(y_2|x)$$

  - This means we only need (x,y_1) and (x,y_2) pairs.
  - Significantly simplifying dataset construction: classifier free guidance must require quadruples (c,x,y_1,y_2) !

Results

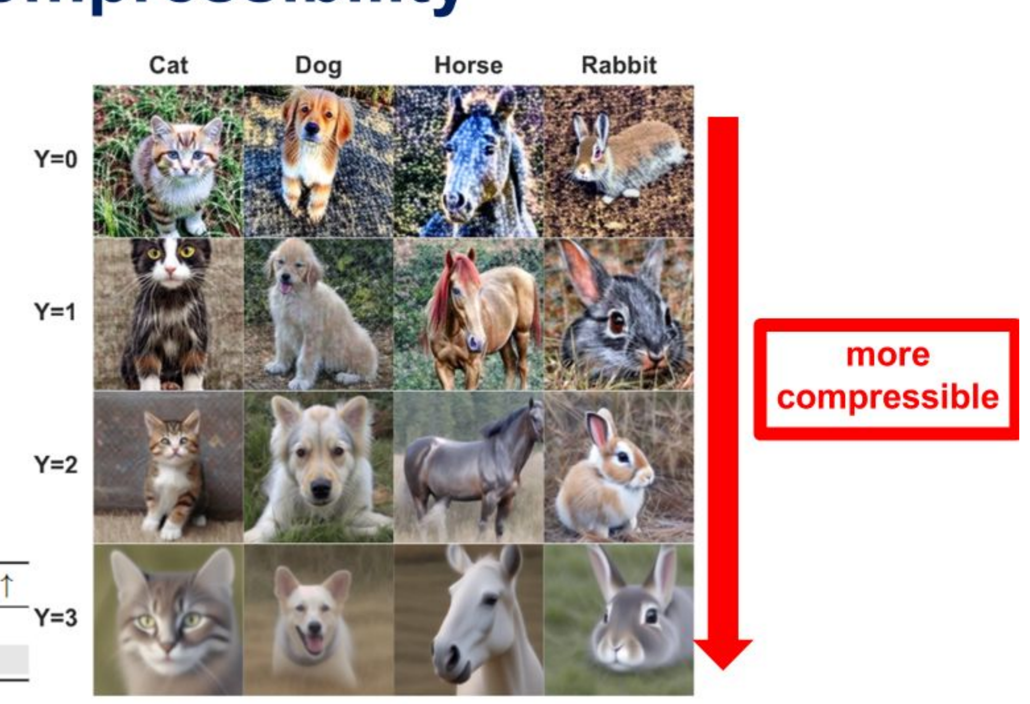
Example 1: Compressibility



(a) Training curve

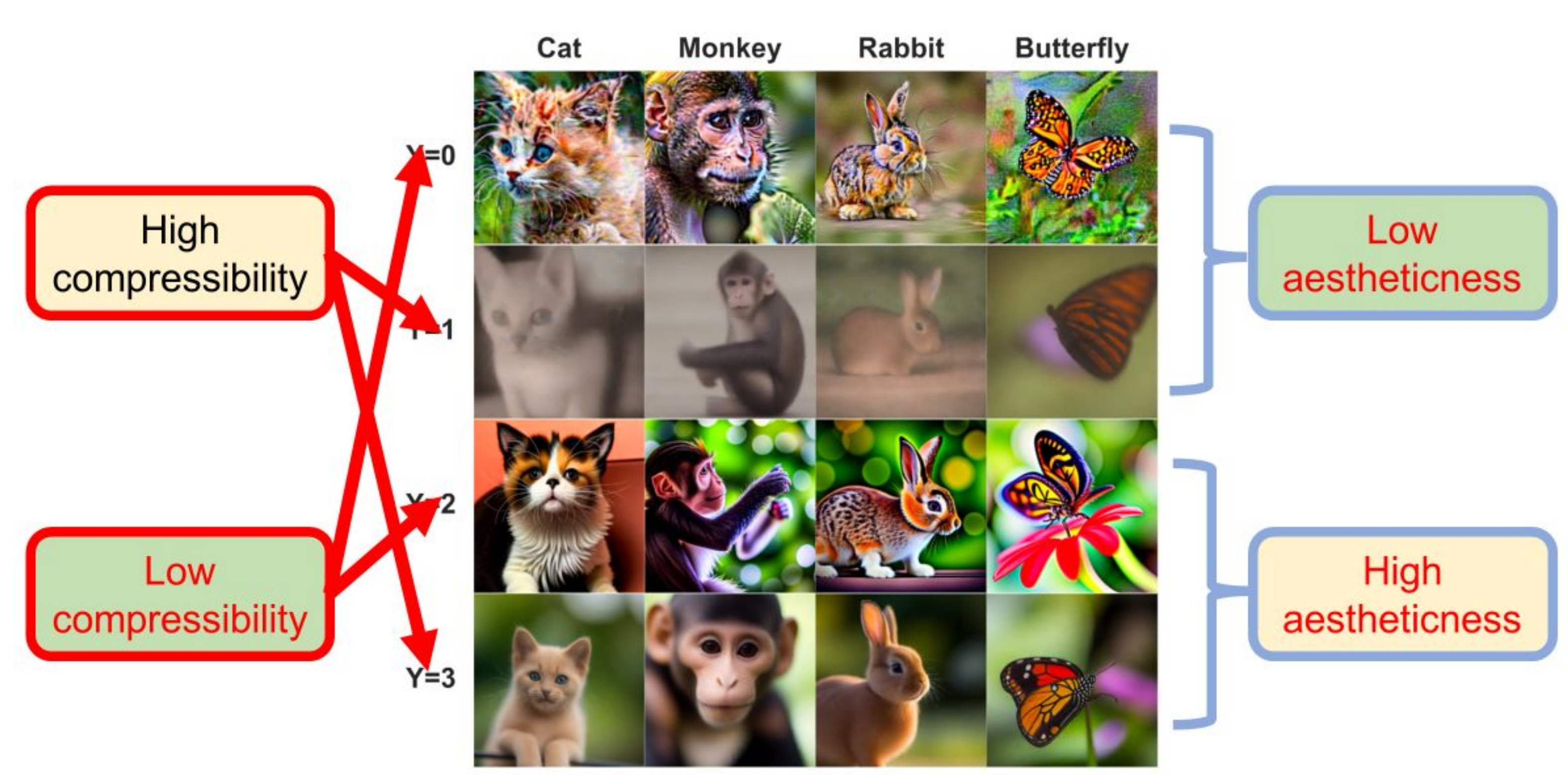
	Accuracy ↑	Macro F1 score ↑
DPS	0.45	0.44
CTRL (Ours)	**1.0**	**1.0**

(b) Evaluation of conditional generations



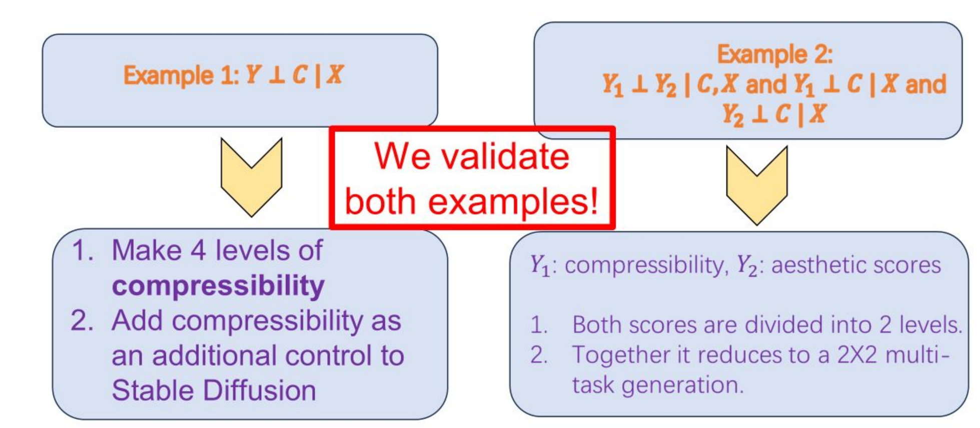
(c) Generated images

Example 2: Compressibility & Aestheticness

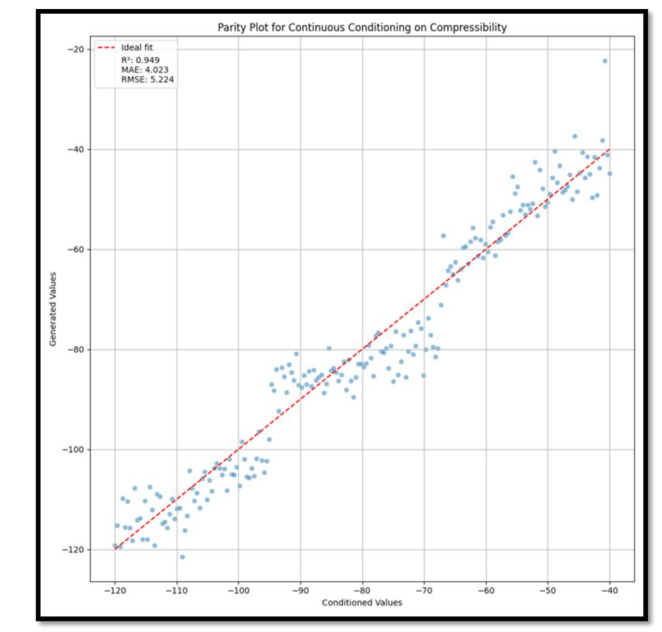


Experimental Details

Settings



Extension: conditioned on continuous y



- We can use the **model fine-tuned on discretized y** and **interpolate** discrete class embeddings for **continuous y**.
- This naive approach can achieve R^2**=0.95**.

Conclusions

- We introduce an **RL-based fine-tuning** approach for conditioning pre-trained diffusion models on new additional labels.
- Compared to **classifier-free guidance**, our proposed method allows for **leveraging the conditional independence**, thereby greatly simplifying the construction of the offline dataset.
- We also **theoretically justify** our approach and build the connection with **classifier-based guidance**.
- Future work includes applying this method to design **DNA enhancers** and **5'UTR**.
- The goal is achieve **cell-specific promoters design**!

arXiv