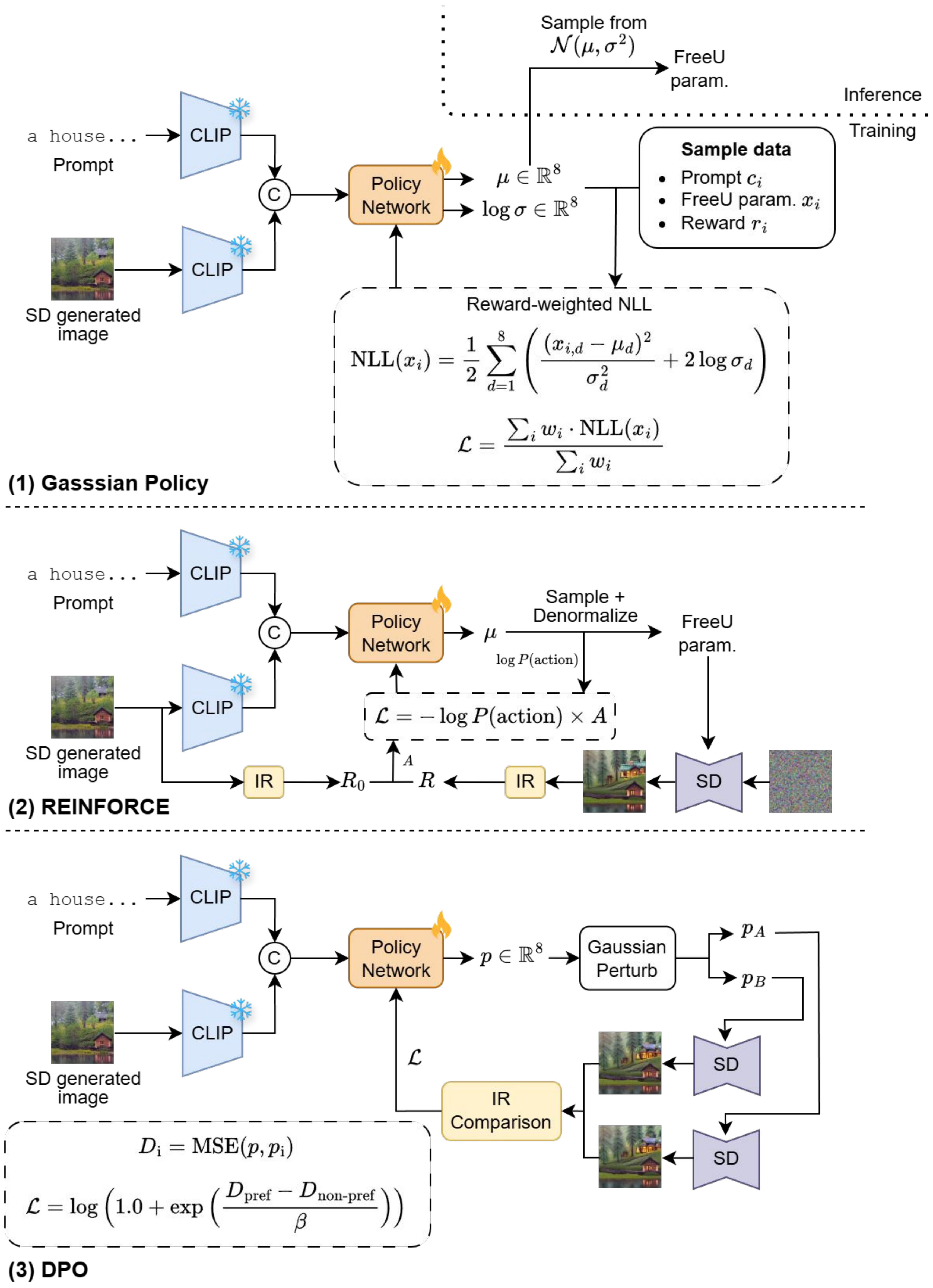
AdaFreeU: Adaptive FreeU Parameter Prediction for Text-to-Image Diffusion Models

Yi-Hsiang Ho, Ting-Wei Chou, Yi-Cheng Lai

Abstract

We propose **AdaFreeU**, an adaptive extension of FreeU for text-to-image diffusion models such as Stable Diffusion (SD) [1,4]. FreeU improves generation by re-weighting U-Net backbone and skip features during inference through two scaling factors, but its fixed parameters may not generalize well across different prompts, seeds, and visual styles. We aim to solve this limitation by predicting FreeU parameters adaptively using strategies such as Gaussian policy prediction, REINFORCE, and DPO [5]. Experimental results show that AdaFreeU improves over both standard SD and default FreeU, with DPO achieving the highest mean ImageReward [2]. Compared with default FreeU, DPO consistently improves the mean ImageReward gain over SD, with 2.1× larger gains in constant mode and 7.0× larger gains in spatial mode.

Methodology



Experimental Setup & Dataset

Our experiments are conducted on Stable Diffusion 1.5, which is based on the latent diffusion framework [4], with a custom FreeU implementation that enables per-layer control over the four U-Net upsampling layers [1]. We compare standard Stable Diffusion, default FreeU, and adaptive FreeU variants that predict FreeU parameters from the prompt and baseline SD image. Large-scale evaluation is performed on the MJHQ-30K dataset [3], using prompts and category metadata to assess performance across diverse image types. ImageReward-v1.0 [2] is used as the primary evaluation metric, measuring prompt-image quality according to learned human preference.

Experiment Result

We evaluate each method on 5,000 MJHQ-30K prompts sampled uniformly across categories, using ImageReward as the main evaluation metric. Adaptive FreeU methods generally outperform standard SD and default FreeU, with DPO achieving the highest mean ImageReward of 0.076 under constant prediction.

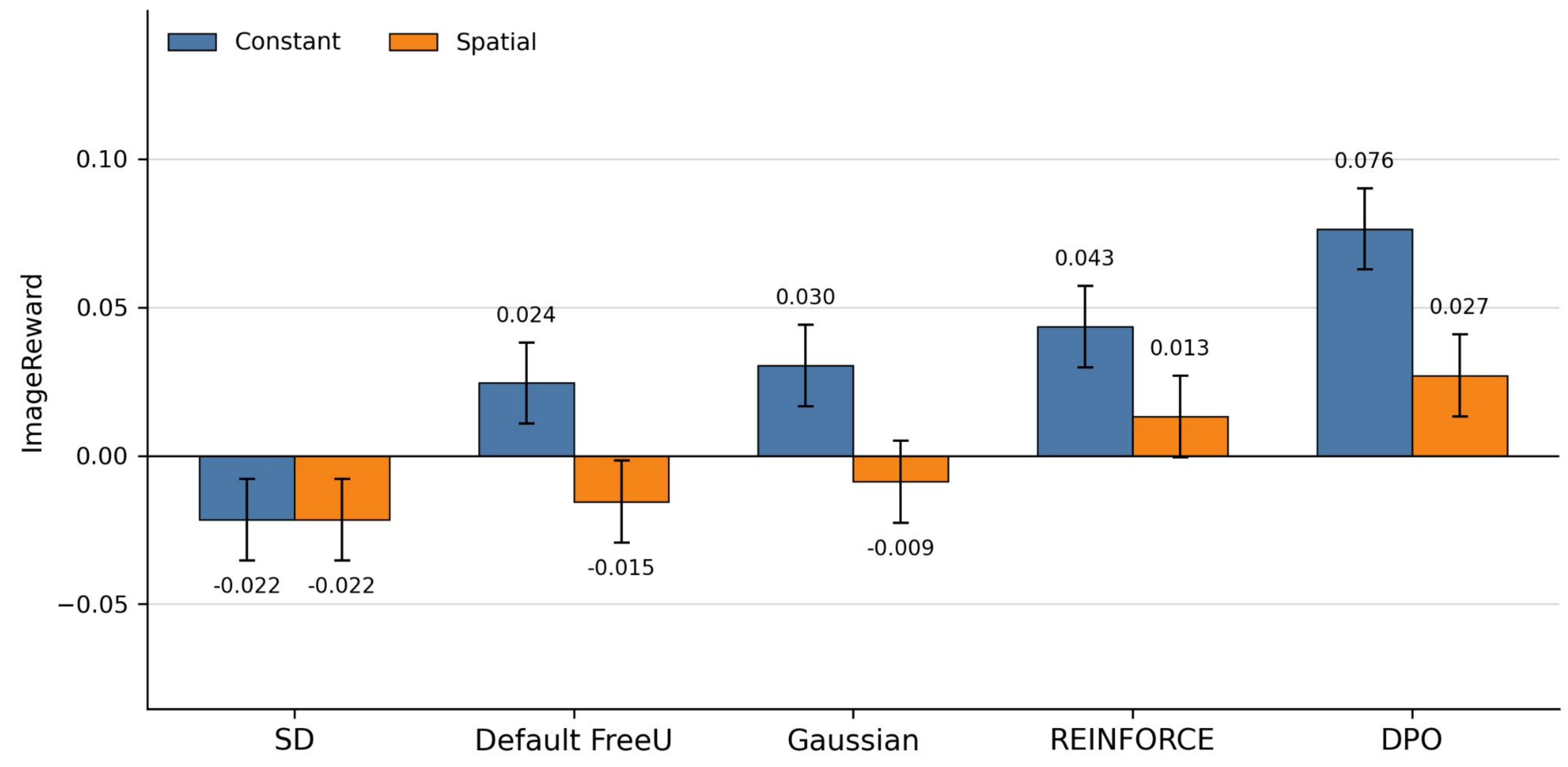


Figure 1. ImageReward comparison across FreeU adaptation strategies

The user study further supports that the Gaussian-based adaptive prediction method is preferred over both SD and default FreeU.

Mode	SD	Default FreeU	Gaussian
Constant	30.0% (189/630)	23.7% (149/630)	**46.3% (292/630)**
Spatial	14.8% (93/630)	23.5% (148/630)	**61.7% (389/630)**

Table 1. User preference comparison

Qualitative examples show that different FreeU strategies affect object structure, composition, and visual detail, suggesting that optimal FreeU parameters are prompt-dependent.

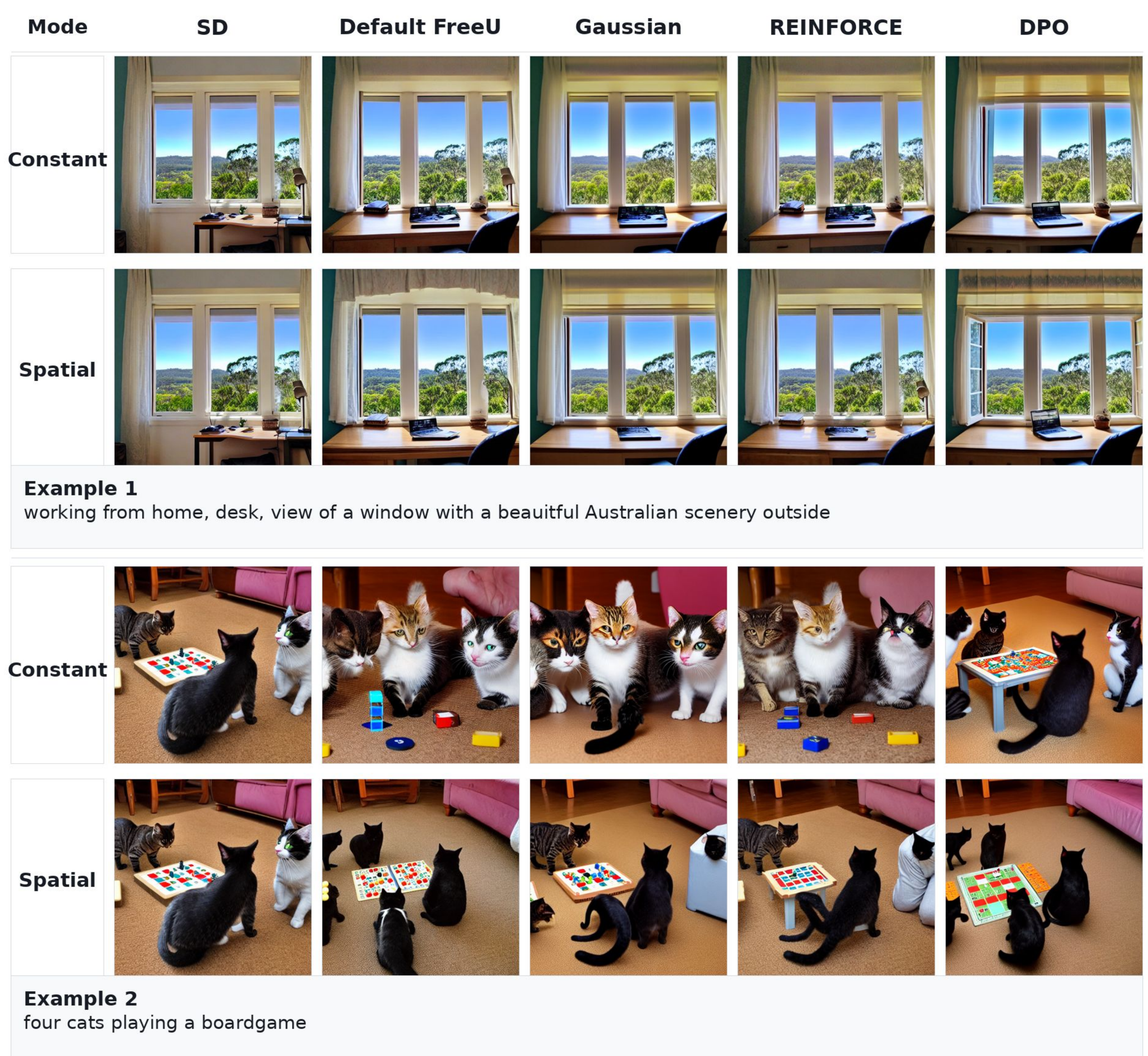


Figure 2. Visual comparison of SD, default FreeU, and adaptive FreeU methods

Reference

[1] C. Si, Z. Huang, Y. Jiang, and Z. Liu, "FreeU: Free Lunch in Diffusion U-Net," CVPR, 2024.
[2] J. Xu, X. Liu, Y. Wu, Y. Tong, Q. Li, M. Ding, J. Tang, and Y. Dong, "ImageReward: Learning and Evaluating Human Preferences for Text-to-Image Generation," NeurIPS, 2023.
[3] Playground AI, "MJHQ-30K Benchmark," Hugging Face Dataset, 2024.
[4] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-Resolution Image Synthesis with Latent Diffusion Models," CVPR, 2022.
[5] R. Rafailov, A. Sharma, E. Mitchell, C. D Manning, S. Ermon, and C. Finn, "Direct preference optimization: Your language model is secretly a reward model," NeurIPS, 2023.