Diffusion-NPO: Negative Preference Optimization for Better Preference Aligned Generation of Diffusion Models

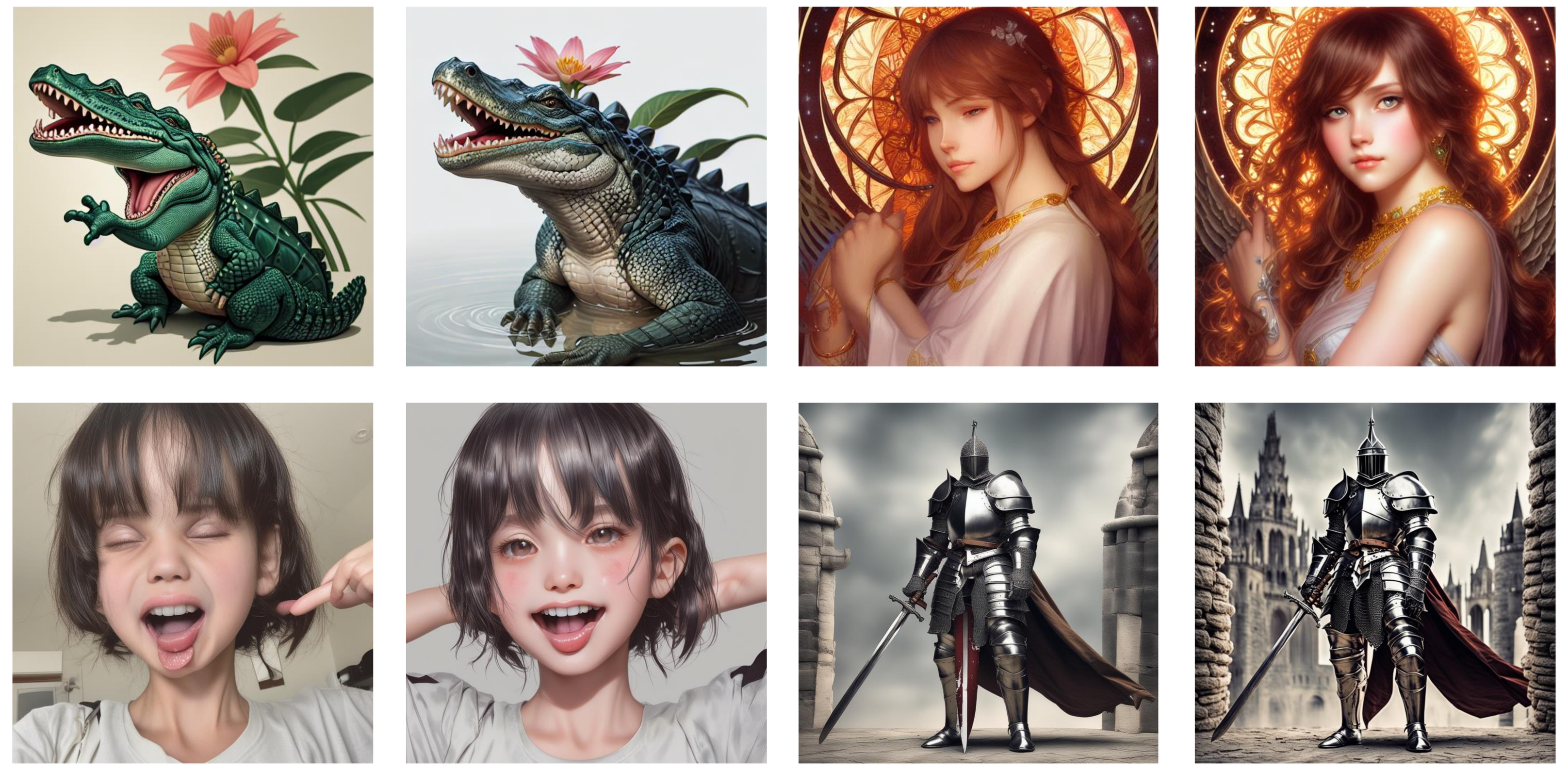
Fu-Yun Wang[1] Yunhao Shui[2] Jingtan Piao[1] Keqiang Sun[1] Hongsheng Li[1]

[1]CUHK MMLab [2]Shanghai Jiao Tong University

Motivation

- Diffusion models excel in image generation, but those trained on vast, uncurated datasets often produce results that diverge from human preferences. Various fine-tuning techniques have improved alignment with human expectations.
- We contend that current alignment methods overlook the importance of managing negative-conditional outputs, reducing their ability to prevent unwanted results. To address this, we introduce Diffusion-NPO, a simple yet highly effective solution.

Effectiveness of Diffusion-NPO



w/o NPO w/ NPO w/o NPO w/ NPO

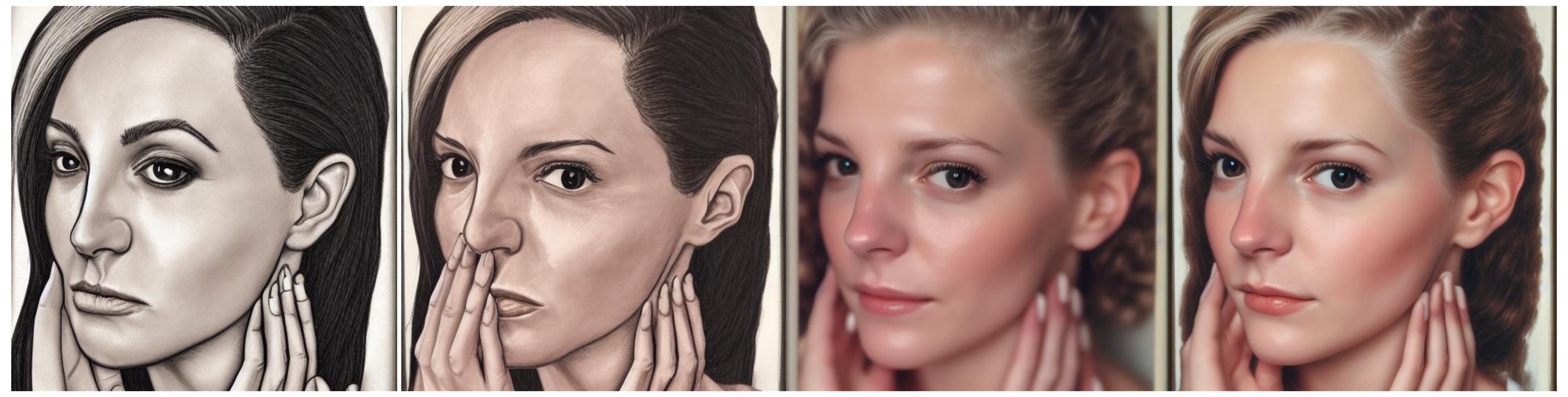
Diffusion-NPO enhances high-frequency details, color and lighting, and low-frequency structures in images by aligning human's negative preference.

Effectiveness on Video Generation.

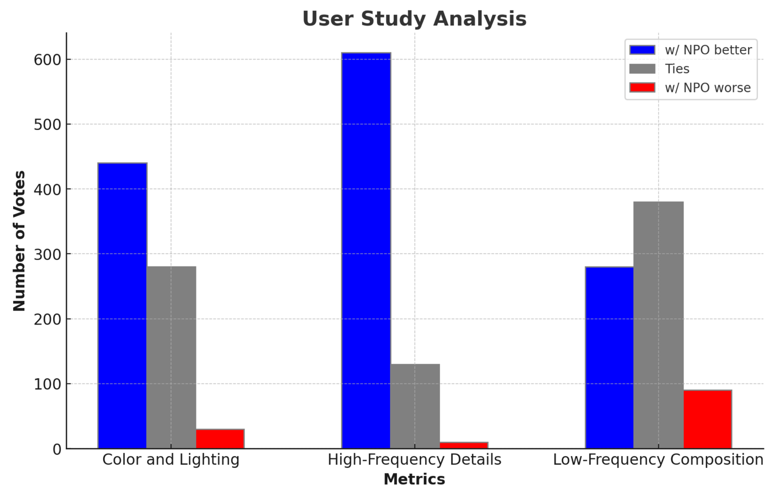
Prompt: "A person playing a guitar by a campfire under a starry sky."

VideoCrafter2 VADER VADER + NPO (HPSv2) VADER + NPO (PickScore)

Compatibility of Diffusion-NPO and User Study



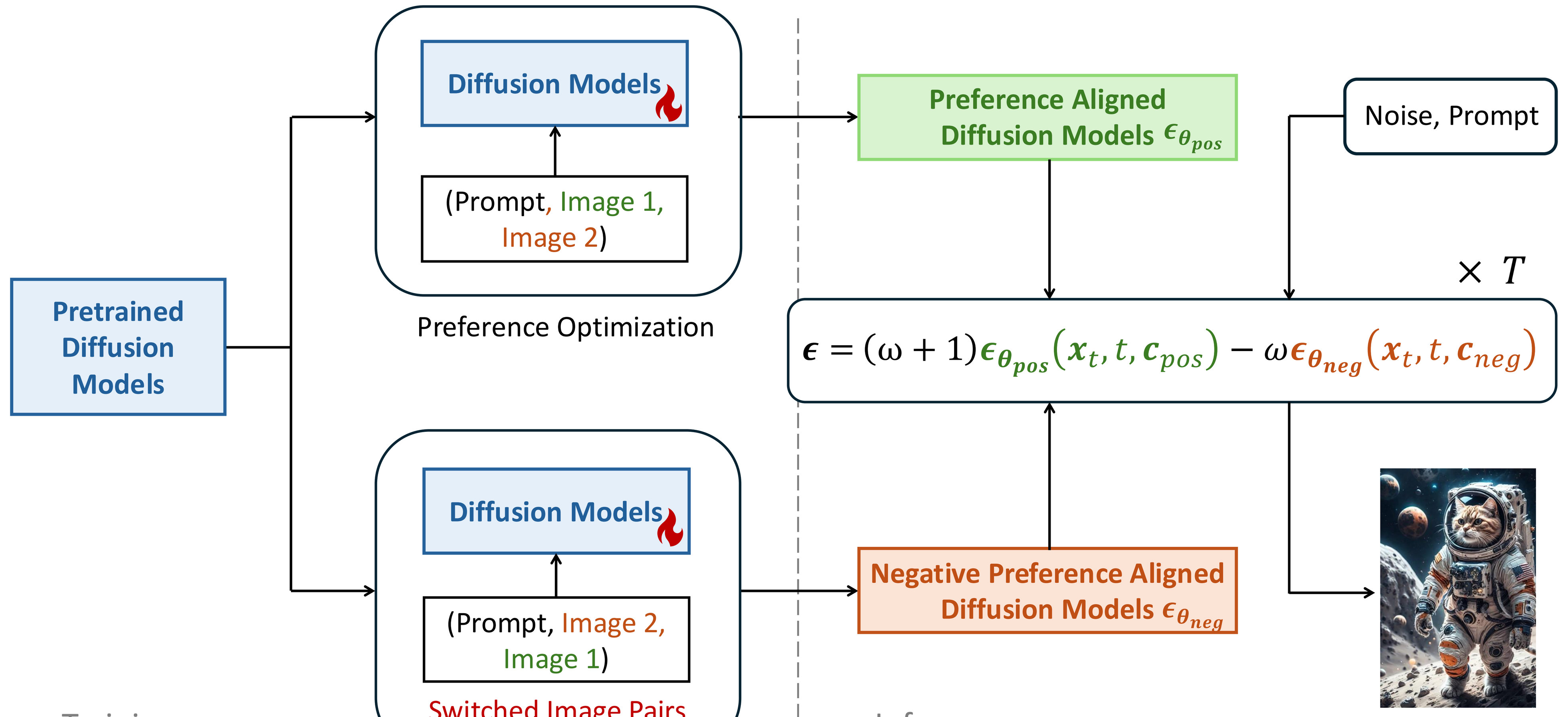
Stable Diffusion v1-5 Diffusion-DPO Diffusion-DPO+NPO w/o WM Diffusion-DPO+NPO



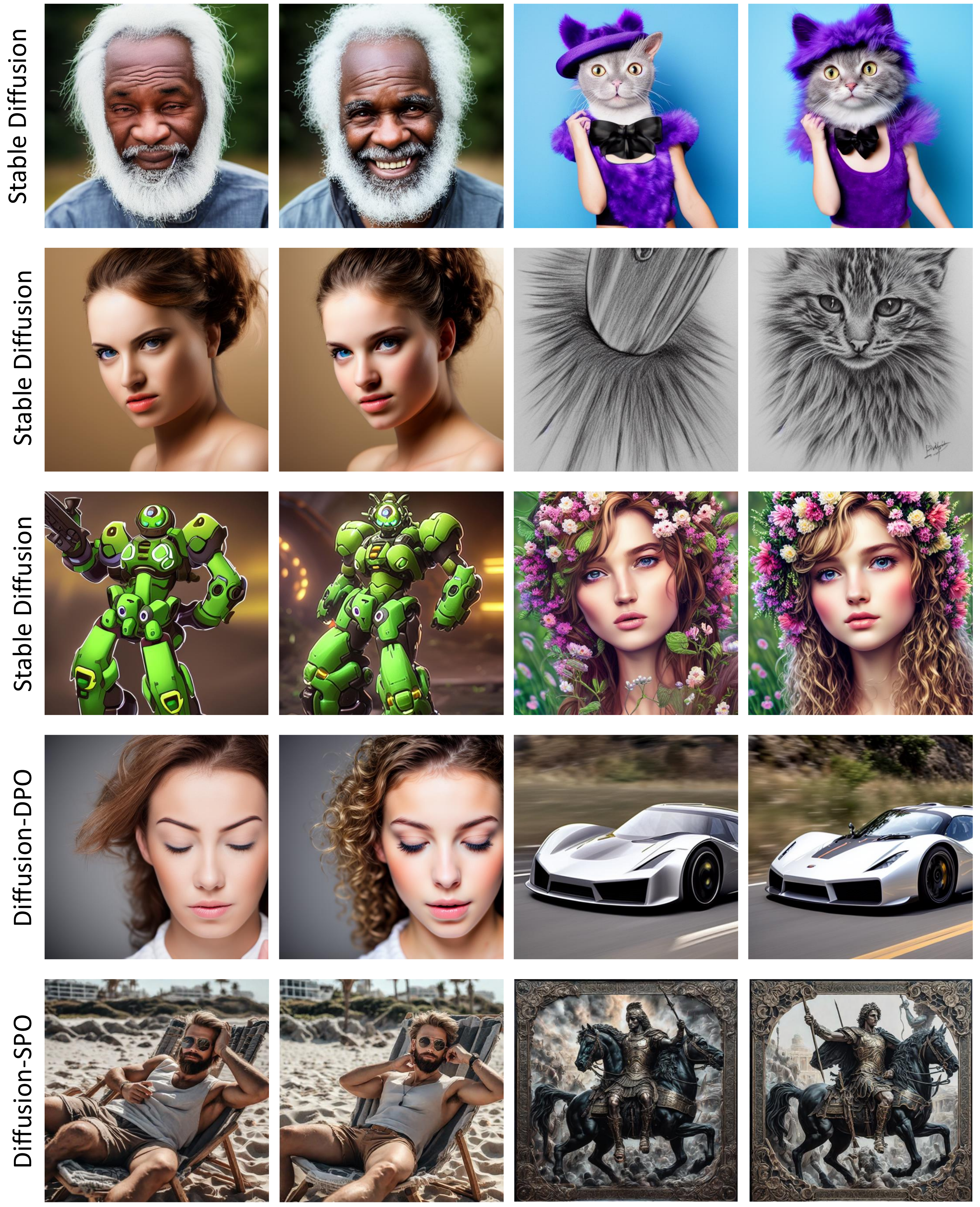
NPO augments the resolution of high-frequency details in generated outputs, while optimizing color and lighting to better correspond with human perceptual preferences. Furthermore, NPO moderately enhances the compositional integrity of the resulting images.

Methodology

- Our crucial insight is that training such a negative preference aligned model requires no new training strategies or datasets, only minor modifications to existing methods.
- **Training of Diffusion-NPO.** In essence, all strategies can be perceived as reversing the order of image pairs in the collected preference data by adapting the same training procedure.
- **Inference of Diffusion-NPO**: Leveraging classifier-free guidance, we apply a preference-aligned model for conditional outputs and a negatively aligned model for negative-conditional outputs to maximize preference alignment.



More Generation Results



w/o NPO w/ NPO w/o NPO w/ NPO