

FreeEnhance: Tuning-Free Image Enhancement via Content-Consistent Noising-and-Denoising Process -ACMMM 2024 Supplementary Materials

Anonymous Authors

Table 1: Experimental Settings of FreeEnhance.

Configuration	Value
Num inference steps	100
t_0	500
Scheduler	DDIM
Eta of DDIM	0
Guidance scale	1
\mathcal{F}_{blur} : Kernel size	3
\mathcal{F}_{blur} : Sigma	1

Table 2: Text-to-image Generation Settings.

Configuration	Value
Num inference steps	100
Scheduler	DDIM
Eta of DDIM	0
Guidance scale	7.5
\mathcal{F}_{blur} : Kernel size	3
\mathcal{F}_{blur} : Sigma	1

The supplementary material contains: 1) implementation details of FreeEnhance; 2) implementation details of applying FreeEnhance in the text-to-image task; 3) empirical analyses exploring the trade-off parameters in the noising and denoising stages of FreeEnhance; 4) more qualitative results of FreeEnhance and the comparisons with Magnific AI; 5) more qualitative results of FreeEnhance in the text-to-image application; 6) a sample video demonstrating qualitative comparisons of images before and after applying FreeEnhance.

1 IMPLEMENTATION OF FREEENHANCE

In addition to the description of FreeEnhance in Section 4.1 of the main paper, Table 1 outlines the specific configurations applied in our FreeEnhance for image enhancement tasks. We leverage the Image-to-Image pipeline of Stable Diffusion XL from HuggingFace Diffusers as a foundation. We then implement the noising and denoising stages described in Sections 3.2 and 3.3 of the main paper, respectively. For the adversarial regularization in the denoising stage, the \mathcal{F}_{blur} utilizes the Gaussian blur with a kernel size of 3 and a sigma value of 1. For the assessment of NR-IQA metrics, we engage the IQA-Pytorch library [2]. Additionally, the HPSv2 scores are evaluated using the official HPSv2 implementation, with the compressed v2.1 version of the HPS model.

Table 3: Performances of FreeEnhance with different noise strength (determined by t_0) attached to input images on HPDv2 Benchmark.

t_0	300	400	500	600	700
HPSv2 \uparrow	20.30	21.63	29.32	28.80	28.20

Table 4: Effect of the trade-off parameter τ in FreeEnhance on HPDv2 Benchmark.

τ	1.0	0.7	0.6	0.5	0.4	0.0
HPSv2 \uparrow	28.94	29.30	29.32	29.23	29.18	28.07

2 FREEENHANCE FOR TEXT-TO-IMAGE

In addition to the description in Section 4.5 of the main paper, here we elaborate the way to apply FreeEnhance to the task of Text-to-Image image generation. Given the public availability of only the Stable Diffusion 1.5 based PAG [1], we utilize the Text-to-Image pipeline from HuggingFace Diffusers, specifically implemented for Stable Diffusion 1.5, as our baseline for fair comparisons. Following the methodologies outlined in Section 3.3 of the main paper, we remold the denoising stage to integrate our FreeEnhance approach. Table 2 shows the detailed configuration of the denoising process for Text-to-Image generation.

3 ANALYSIS OF TRADE-OFF PARAMETERS

Effect of Noise Strength Parameter t_0 . We investigate the influence of the hyper-parameter t_0 , which determines the intensity of noise applied to the input image. Table 3 details HPSv2 scores achieved with different t_0 values. Overall, the highest human preference score is observed with $t_0 = 500 = 0.57$, indicating that moderately adding noise to input images benefits the image enhancement process. Lower t_0 values lead to diminished image quality. We hypothesize that this deterioration may stem from the excessive artifacts introduced into the low-frequency image regions where the input image does not provide guidance during the internal denoising process ($x_T^c \rightarrow x_0^c$ within the Creative Stream). The HPSv2 experiences a slight decline as we increase t_0 beyond 500. A higher t_0 reduces the number of denoising steps performed by the Creative Stream, resulting in a limited generation of image details, which are crucial factors in human assessments of image quality.

Effect of the trade-off parameter τ . To verify the effect of the trade-off parameter τ in Eq.(5) of the main paper, we detail the enhancement performances (i.e., HPSv2) with different values of the trade-off parameter τ in Table 4. In the extreme case of $\tau = 1.0$, the low-frequency regions of the blended noisy image are solely influenced by the Stable Stream, resulting in the inability to incorporate additional details into the input image and sub-optimal

Table 5: Comparison of FreeEnhance with and without Distribution Calibration on HPDv2 Benchmark.

	without calibration	with calibration
HPSv2 \uparrow	27.50	29.32

Table 6: Comparison of FreeEnhance employing different values of ρ_{acu} , ρ_{dist} , and ρ_{adv} on HPDv2 Benchmark.

ρ_{acu}	0	2	4	6	8
HPSv2 \uparrow	29.15	29.25	29.32	29.21	29.01
ρ_{dist}	0	20	40	60	80
HPSv2 \uparrow	29.24	29.32	29.29	29.28	29.27
ρ_{adv}	0	0.3	0.5	0.7	0.9
HPSv2 \uparrow	28.92	29.32	29.12	28.79	28.43

results. When τ is reduced to 0.7, a improvement in image quality is observed, indicating that the two-stream noising scheme facilitates the generation of image details. The performance change is relatively smooth as τ varies within the range of 0.7 to 0.4, with the optimal result achieved at $\tau = 0.6$. Moreover, the HPSv2 experiences a significant decline when $\tau = 0$. We postulate that this decrease may stem from artifacts in the low-frequency regions of the image where the noise is exclusively provided by the Creative Stream.

Effect of the Distribution Calibration. In addition to the qualitative demonstration in Figure 4 of the main paper, Table 5 here further quantitatively elaborates the effectiveness of the distribution calibration employed in the noising stage of FreeEnhance.

Effect of the Trade-off Parameters ρ_{acu} , ρ_{dist} , and ρ_{adv} . To examine the impact of the weights of three regularizations utilized by FreeEnhance on the quality of enhanced images, we assess FreeEnhance configured with various values of ρ_{acu} , ρ_{dist} , and ρ_{adv} , and summarize the corresponding HPSv2 scores in Table 6. Our results indicate that FreeEnhance experiences performance degradation when any one of the regularizations is disabled (i.e., when the weight of regularization is set to 0). Regarding ρ_{acu} , FreeEnhance achieves optimal results when $\rho_{acu} = 4$, but experiences performance degradation when $\rho_{acu} \geq 8$ due to an overemphasis on image acutance. Similarly, the highest HPSv2 score is attained when $\rho_{adv} = 0.3$, whereas larger weights of the adversarial regularization negatively impact human preference for the images. The HPSv2 fluctuates within a range of 0.05 when the values of ρ_{dist} are adjusted from 20 to 80. In the experiments detailed in the main paper, ρ_{acu} , ρ_{dist} , and ρ_{adv} are fixed at 4, 20, and 0.3, respectively.

4 QUALITATIVE RESULTS OF IMAGE ENHANCEMENT

We showcase more qualitative results in Figure 1, Figure 2, and Figure 3. Similar to the observation in the main paper, the visual quality of input images is consistently improved and image details are properly added in a content-consistent manner by FreeEnhance.

5 QUALITATIVE RESULTS OF TEXT-TO-IMAGE GENERATION

In addition to Figure 9 of the main paper, Figure 4 provided here offers additional qualitative comparisons among Stable Diffusion 1.5 (SD 1.5), SAG [3], PAG [1], and FreeEnhance about text-to-image generation. Images generated by FreeEnhance consist of more detailed structures and faithfully reflect the text prompts, surpassing both SD 1.5, SAG and PAG.

6 SAMPLE VIDEO

We summarize several qualitative results of FreeEnhance into a video (“FreeEnhance.mp4”) and demonstrate the effectiveness of image enhancement by placing images before and after enhancement on top of each other and displaying them alternately.

REFERENCES

- [1] Donghoon Ahn, Hyoungwon Cho, Jaewon Min, Woosok Jang, Jungwoo Kim, SeonHwa Kim, Hyun Hee Park, Kyong Hwan Jin, and Seungryong Kim. 2024. Self-Rectifying Diffusion Sampling with Perturbed-Attention Guidance. *arXiv preprint arXiv:2403.17377* (2024).
- [2] Chaofeng Chen and Jiadi Mo. 2022. IQA-PyTorch: PyTorch Toolbox for Image Quality Assessment. [Online]. Available: <https://github.com/chaofengc/IQA-PyTorch>.
- [3] Susung Hong, Gyuseong Lee, Woosok Jang, and Seungryong Kim. 2023. Improving sample quality of diffusion models using self-attention guidance. In *ICCV*.

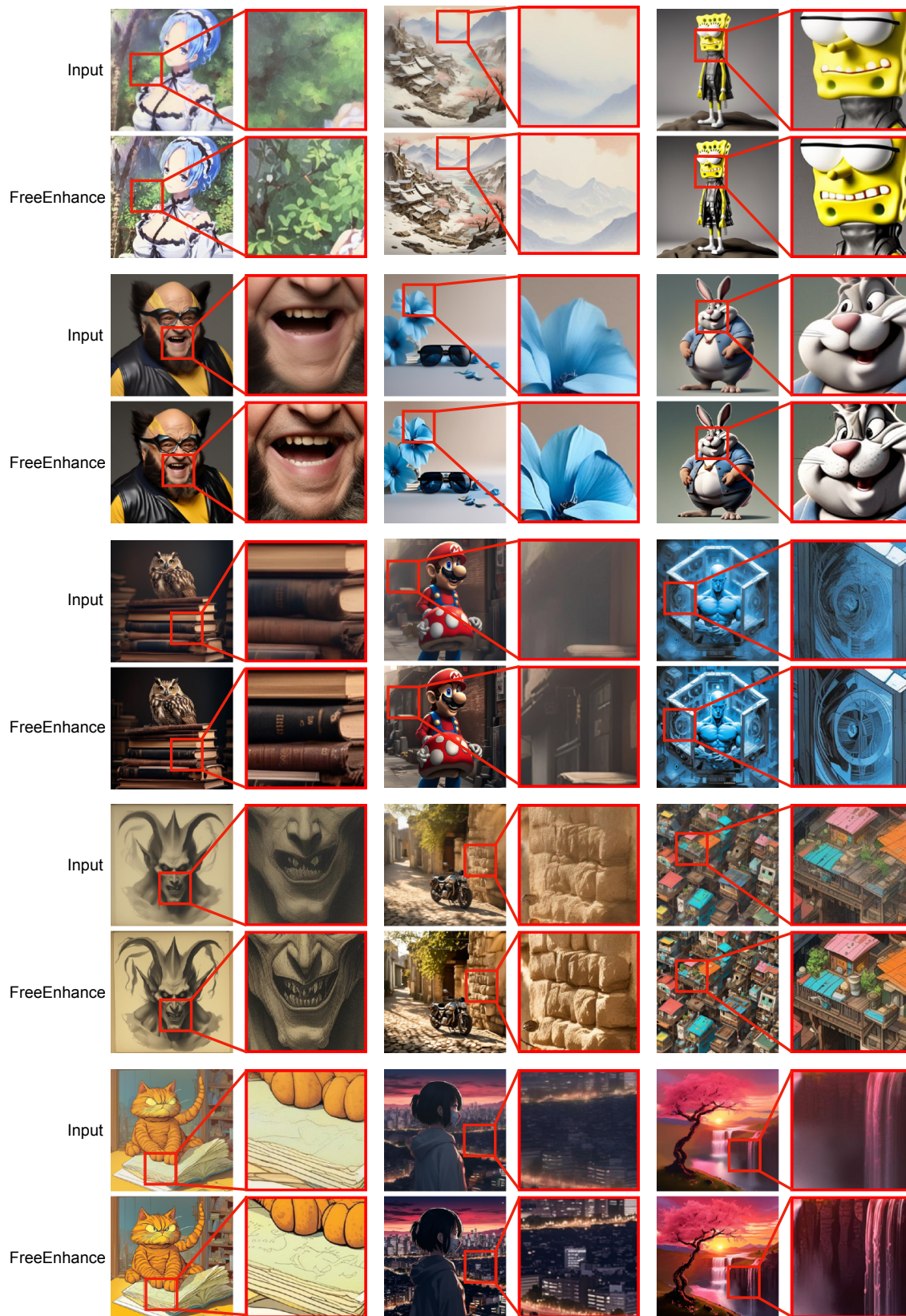


Figure 1: Examples of image enhancement results of FreeEnhance on HPSv2 benchmark.

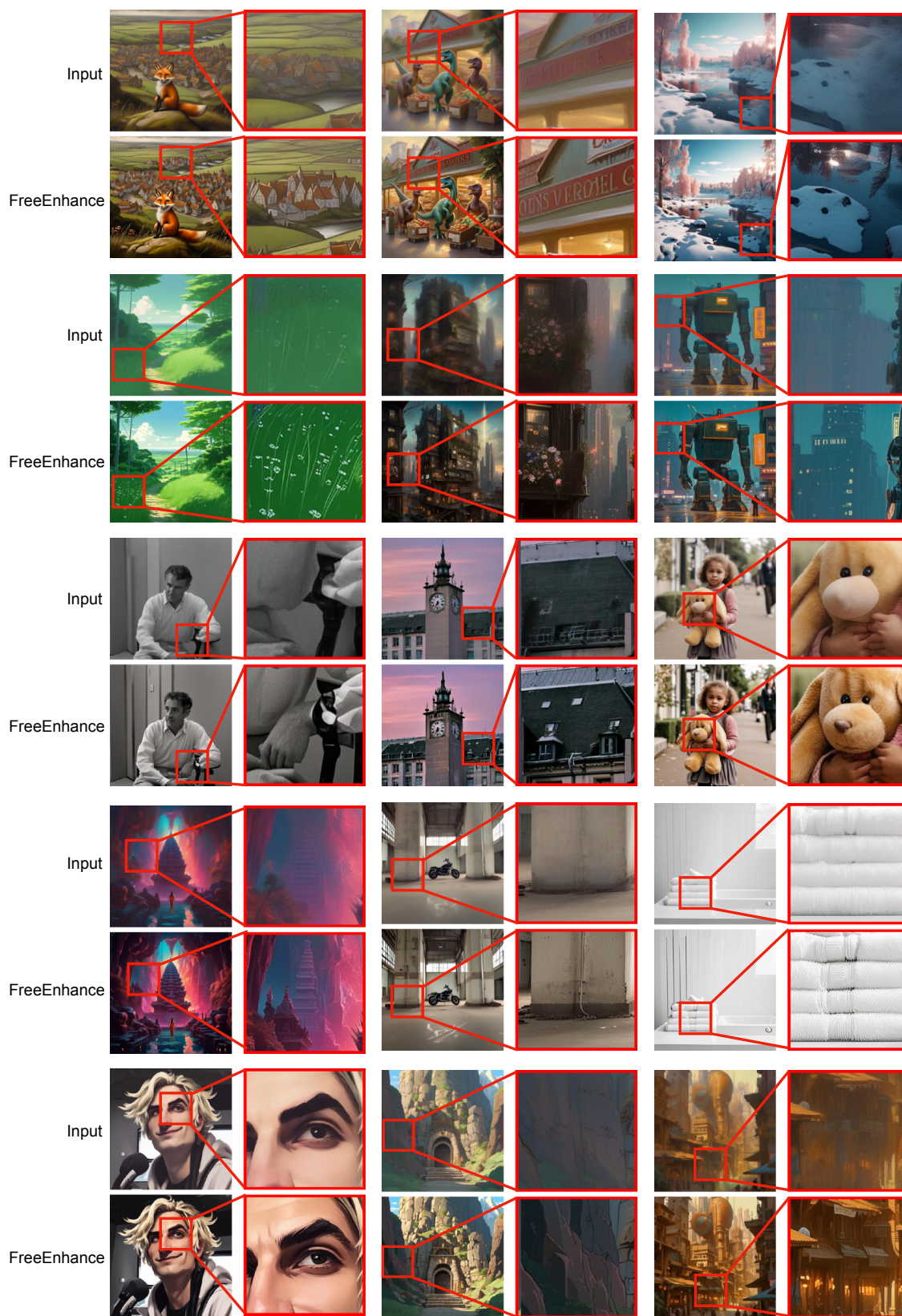


Figure 2: Examples of image enhancement results of FreeEnhance on HPSv2 benchmark.



Figure 3: More comparisons of images enhanced by Magnific AI and our FreeEnhance.

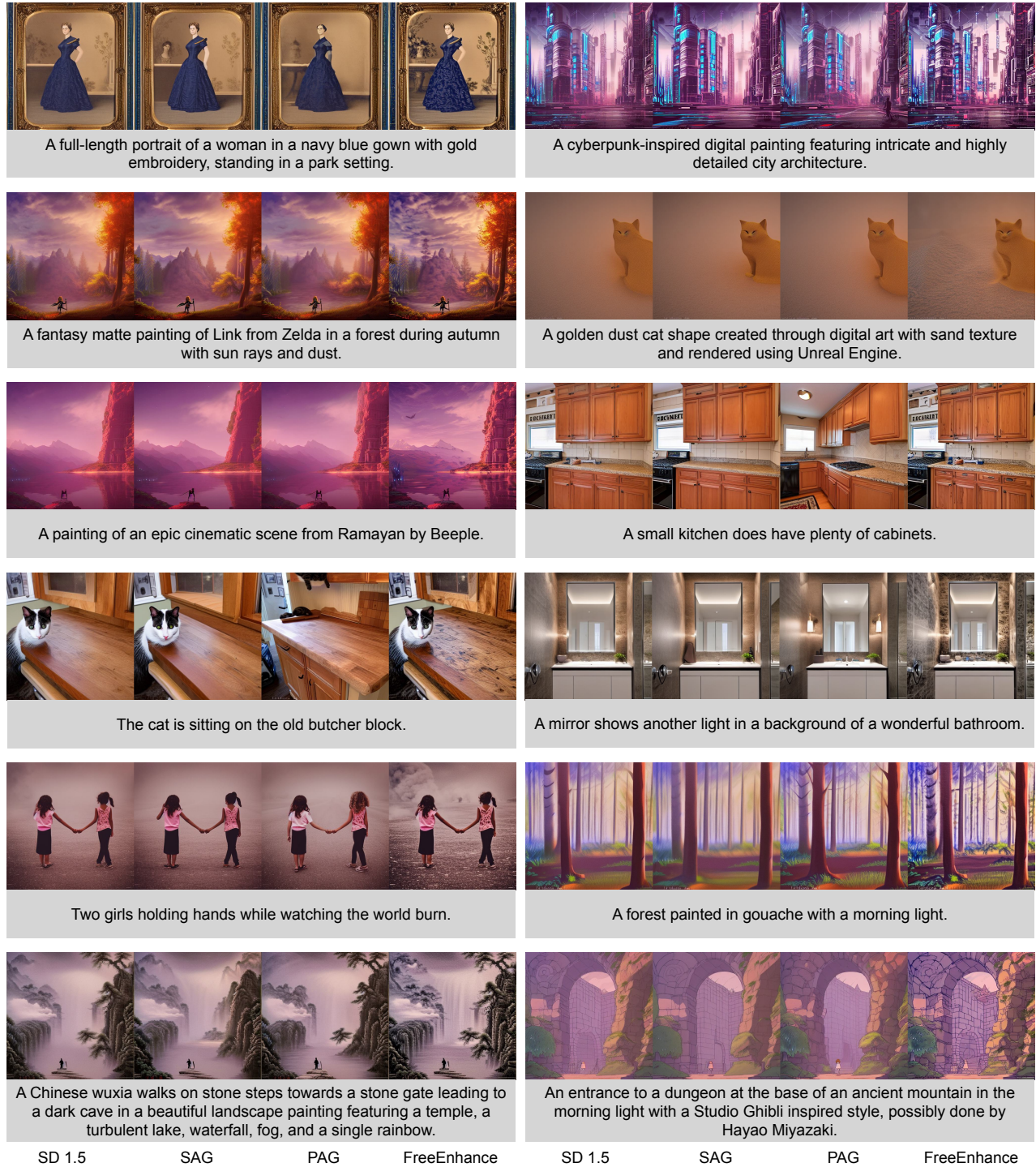


Figure 4: Qualitative comparisons of images synthesized using various denoising approaches in the text-to-image scenario, using prompts from HPDv2 benchmark.