| Image-text alignment | | | Image quality | | |
|---|---|---|---|---|---|
| Win | Lose | Tie | Win | Lose | Tie |
| 67.8 $_{\pm 12.1}$ | 5.6 $_{\pm 2.7}$ | 26.6 $_{\pm 11.3}$ | 33.4 $_{\pm 18.8}$ | 9.7 $_{\pm 8.2}$ | 56.9 $_{\pm 22.6}$ |

Table 1: Percentage of generated images from our RL model that are better than (win), tied with, or worse than (lose) compared to the original stable diffusion model in terms of image-text alignment and image quality (fidelity).

| Image-text alignment | | | Image quality | | |
|---|---|---|---|---|---|
| Win | Lose | Tie | Win | Lose | Tie |
| 52.8 $_{\pm 12.8}$ | 7.8 $_{\pm 3.4}$ | 40.0 $_{\pm 13.8}$ | 79.7 $_{\pm 10.6}$ | 5.3 $_{\pm 5.5}$ | 15.0 $_{\pm 12.3}$ |

Table 2: Percentage of generated images from RL model that are better than (win), tied with, or worse than (lose) compared to the SFT model in terms of image-text alignment and image quality (fidelity).

| | MS-CoCo | | Drawbench | |
|---|---|---|---|---|
| | Original model | RL model | Original model | RL model |
| ImageReward score | 0.22 | **0.55** | 0.13 | **0.58** |
| Aesthetic score | 5.39 | 5.43 | 5.31 | 5.35 |

Table 3: ImageReward scores and Aesthetic scores from the original model, and RL fine-tuned model, all are trained on multiple prompts from MS-CoCo (104 prompts) and Drawbench (183 prompts). We report the average ImageReward and Aesthetic scores across 3120 and 5490 images on MS-CoCo and Drawbench, respectively (30 images per each prompt).
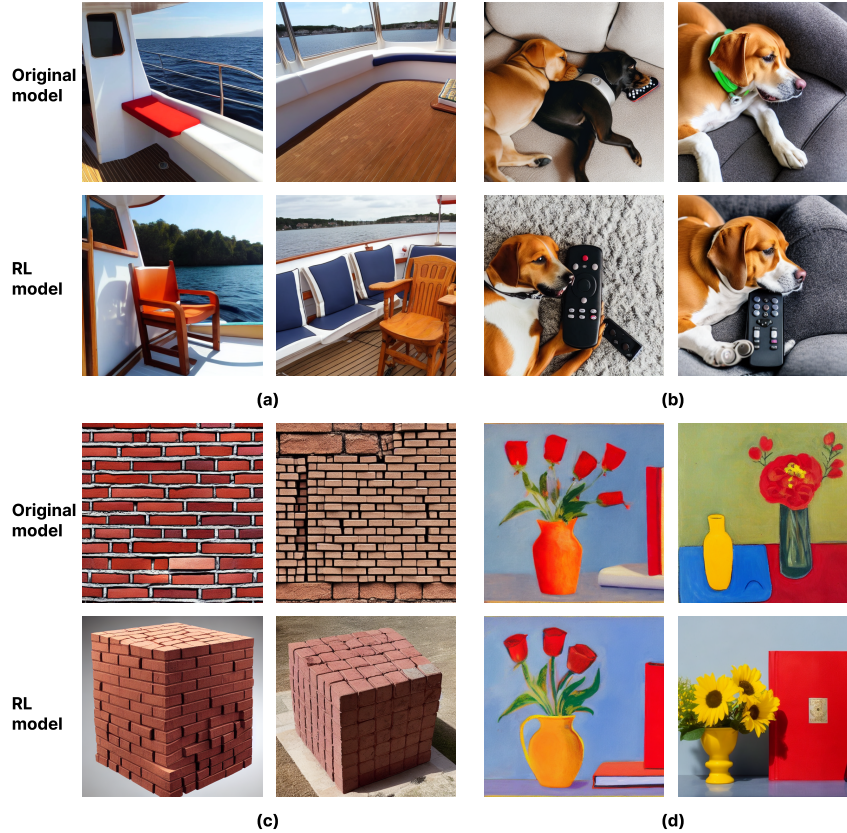


Figure 1: Sample images generated from prompts: (a) "A chair in the corner on a boat"; (b) "A dog is laying with a remote controller"; (c) "A cube made of brick"; (d) "A red book and a yellow vase", from the original model and RL model respectively. Images in the same column are generated with the same random seed. Prompts (a) (b) are from MS-CoCo and (c) (d) are from Drawbench.