

DiverseGRPO: Mitigating Mode Collapse in Image Generation via Diversity-Aware GRPO

Supplementary Material

1. Human Evaluation

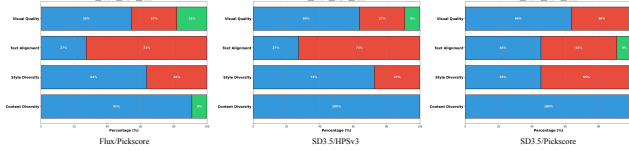


Figure 1. Human evaluation results of DiverseGRPO across different model architectures (Flux/SD3.5-M) and different reward models (PickScore/HPSv3) (image quality includes Visual Quality and Text Alignment, image diversity includes Content Diversity and Style Diversity).

To comprehensively evaluate the effectiveness of the proposed method, we conducted a user study involving a systematic comparison across two generative architectures (Flux [1] and SD3.5-M [2]) and two reward models (PickScore [3] and HPSv3 [5]). The study invited 15 experts in visual computing and graphics to complete a questionnaire comprising 20 sets of image comparisons. For each comparison, based on a given text prompt (e.g., "Marilyn Monroe wearing a shirt that reads Worship"), participants selected their preference between the results generated by our method and the baseline method. The evaluation covered two primary dimensions: image quality (including visual quality and text alignment) and image diversity (including style diversity and content diversity). A total of 1,200 valid votes were collected, spanning various generation themes such as portraits, still lifes, realistic imagery, and abstract art, thereby ensuring broad and representative assessment coverage.

The statistical results presented in Fig. 1 indicate that our method demonstrates significant advantages across all tested configurations, achieving consistently superior win rates. In terms of image quality, our method attained a maximum win rate of 64% in visual quality (SD3.5-M/HPSv3) while maintaining a steady lead in text alignment. Regarding image diversity, our method exhibited particularly outstanding performance, achieving a 100% win rate in content diversity and significantly outperforming the baseline in style diversity, as evidenced by a 45% win rate in the SD3.5-M/PickScore combination. Furthermore, the method demonstrated robust generalization and stability, sustaining high win rates across all four evaluation metrics regardless of the generative architecture (Flux or SD3.5-M) or the reward model (PickScore or HPSv3). It is noteworthy that the experimental setup encompassed a wide spectrum of generated content, ranging from portraits (including males, fe-

males, and celebrities) to still lifes, and from realistic to abstract styles.

2. More Visualization

In this section, we present a visual comparison of the generated results from different backbone networks combined with different preference reward models. In each set of images (Fig. 2, 3, 4), the top row represents Flow-GRPO [4], while the bottom row corresponds to our method. It shows that although the images generated by different backbones exhibit stylistic differences, significant mode collapse occurs after Flow-GRPO training.

Specifically, in the first group (top) of Figure 2, the generated human appearances are highly consistent, even though the corresponding captions do not specify particular individuals. This reflects an unintended correlation between reward scores and appearance features. In contrast, the images in the bottom row of the same group maintain high quality and semantic consistency while displaying diverse human appearances and shooting poses (e.g., front/side views, close-ups/extreme close-ups). Further observation of the fourth group (top) in Figure 3 reveals that the generated images of Marilyn Monroe show high similarity in background and style, consistently featuring bokeh backgrounds and blurred crowds. In comparison, the proposed DiverseGRPO method generates more diverse backgrounds (such as solid gray-white, sky with clouds, and indoor scenes) and styles (e.g., retro or modern). Additionally, in the fourth group (top) of Figure 4, the teddy bear consistently appear on the left side of the image, and the cars are mostly red, even though the captions do not impose constraints on object positions or colors. In contrast, the images generated by our proposed method exhibit greater flexibility in object layout, with car colors and teddy bear poses showing richer diversity.

References

- [1] Black Forest Labs. Flux: Official inference code for flux.1 models. <https://github.com/black-forest-labs/flux>, 2024. Version: commit hash or version number if available. ¹
- [2] Patrick Esser, Sumith Kulal, Andreas Blattmann, Rahim Entezari, Jonas Müller, Harry Saini, Yam Levi, Dominik Lorenz, Axel Sauer, Frederic Boesel, et al. Scaling rectified flow transformers for high-resolution image synthesis. In *Forty-first international conference on machine learning*, 2024. ¹
- [3] Yuval Kirstain, Adam Polyak, Uriel Singer, Shahbuland Matiana, Joe Penna, and Omer Levy. Pick-a-pic: An open dataset of user preferences for text-to-image generation. *Advances in neural information processing systems*, 36:36652–36663, 2023. ¹
- [4] Jie Liu, Gongye Liu, Jiajun Liang, Yangguang Li, Jiaheng Liu, Xintao Wang, Pengfei Wan, Di ZHANG, and Wanli Ouyang. Flow-grpo: Training flow matching models via online rl. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*. ¹
- [5] Yuhang Ma, Xiaoshi Wu, Keqiang Sun, and Hongsheng Li. Hpsv3: Towards wide-spectrum human preference score. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 15086–15095, 2025. ¹

Raw photo, pale albino alien girl with big fish eyes and white hair, horror, headshot photo, nikon, dslr, wildlife photography, 8k uhd, highly detailed skin



Elephant in the style of Jeff Koonz



A man submerged in a tar pit, caught in the process of transforming into a black gooey latex lioness, emphasizing the slick and shiny latex effect. Award-winning wildlife photography capturing the essence of the metamorphosis, featuring slime, goo, and the solo subject mid-transformation. Wildlife Photography, DSLR, transformation-focused, dynamic scene



SNK The King Of Fighters artwork dark medieval gothic trio



Figure 2. Visualization trained with Flux backbone rewarded by Pickscore.

Steampunk cat in a steampunk city



Tiger in suit wearing glasses, anthropomorphic tiger



Frontal portrait of queen elizabeth, by Van Gogh

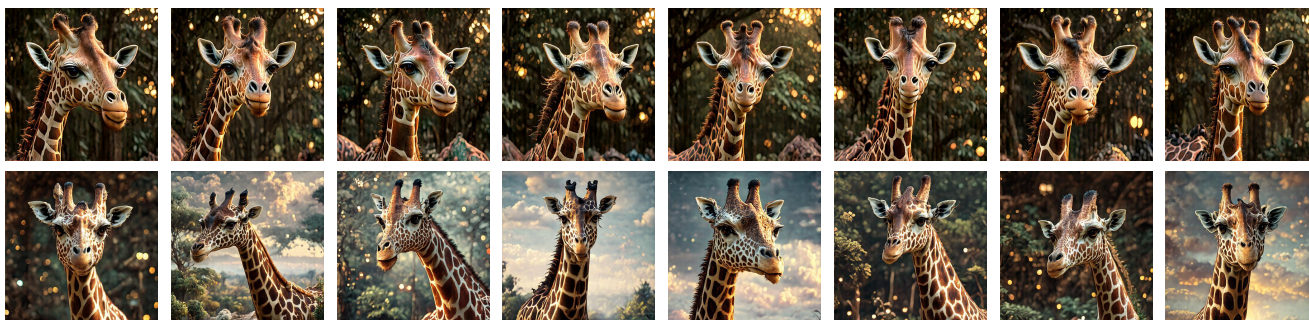


Marilyn Monroe wearing a shirt that reads Worship

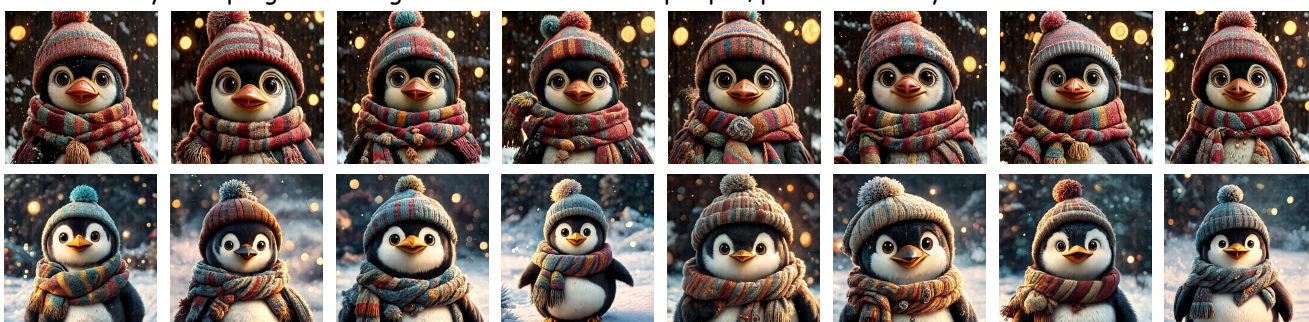


Figure 3. Visualization trained with SD3.5-M backbone rewarded by Pickscore.

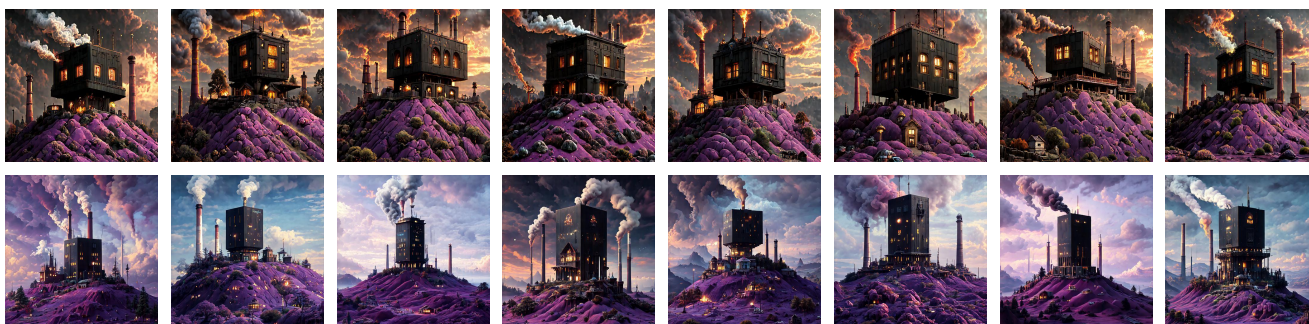
A giraffe by escher, insanely detailed, photorealistic, 8k, ultra high resolution, volumetric lighting, taken with canon eos.



An incredibly cute penguin wearing a scarf and a hat with pompon, pixar movies style



Enigmatic black square building on top of a purple hill, smoke stacks



Teddy bear and a Morris Mini-Minor, big dancing teddy bear



Figure 4. Visualization trained with SD3.5-M backbone rewarded by HPSv3.