Correlation-Weighted Multi-Reward Optimization for Compositional Generation

Jungmyung Wi[1], Hyunsoo Kim[1,2], and Donghyun Kim[†1]

[1] Korea University
[2] Korea Institute of Science and Technology
{wjm333, climba, d_kim}@korea.ac.kr

Abstract. Text-to-image models produce images that align well with natural language prompts, but compositional generation has long been a central challenge. Models often struggle to satisfy multiple concepts within a single prompt, frequently omitting some concepts and resulting in partial success. Such failures highlight the difficulty of jointly optimizing multiple concepts during reward optimization, where competing concepts can interfere with one another. To address this limitation, we propose Correlation-Weighted Multi-Reward Optimization (CMO), a framework that leverages the correlation structure among concept rewards to adaptively weight each attribute concept in optimization. By accounting for interactions among concepts, CMO balances competing reward signals and emphasizes concepts that are partially satisfied yet inconsistently generated across samples, improving compositional generation. Specifically, we decompose multi-concept prompts into pre-defined concept groups (*e.g.*, objects, attributes, and relations) and obtain reward signals from dedicated reward models for each concept. We then adaptively reweight these rewards, assigning higher weights to conflicting or hard-to-satisfy concepts using correlation-based difficulty estimation. By focusing optimization on the most challenging concepts within each group, CMO encourages the model to consistently satisfy all requested attributes simultaneously. We apply our approach to train state-of-the-art diffusion models, SD3.5 and FLUX.1-dev, and demonstrate consistent improvements on challenging multi-concept benchmarks, including ConceptMix, GenEval 2, and T2I-CompBench.

Keywords: Compositional Image Generation · Diffusion Model · Multi-Reward for Reinforcement Learning

1 Introduction

Text-to-image diffusion models [6, 13, 33, 35, 36, 43, 44, 46] have become the standard paradigm for text-conditioned image generation, producing images that align well with natural language prompts. Early breakthroughs in this area were primarily driven by UNet-based diffusion models [33, 35]. Recently, the

[†]Corresponding author.

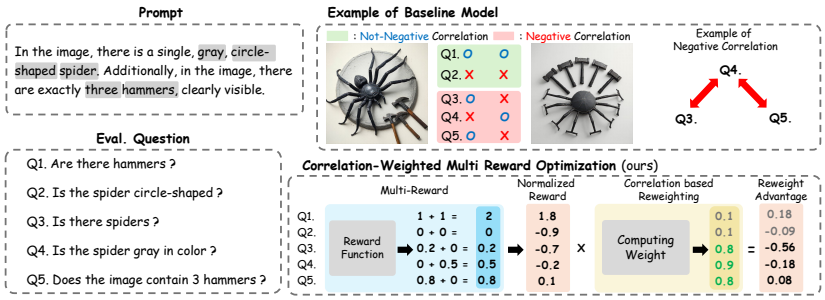

Fig. 1: Motivation of Correlation-based Reweighting. Multi-concept compositional generation often yields different partial successes distributed across generated images. Thus, naively aggregating rewards from each concept tends to focus on concepts that are consistently satisfied across images, while failing to highlight more difficult concepts. Correlation computes concept interactions among generated instances to identify difficult concepts, which are then assigned higher weights during normalized advantage reweighting (*i.e.*, Q3, Q4, and Q5).

field has rapidly evolved with the introduction of scalable flow-matching based models [6, 13, 36, 43] and autoregressive generative models [44, 46]. As model architectures and training corpora continue to scale, modern generative models are increasingly capable of generating diverse visual concepts described in text with high fidelity.

Despite these advances in generation quality, compositional generation has long been a central challenge. Real-world prompts often require models to align multiple concepts simultaneously while correctly binding each attribute to its intended object and relation. Early generative models struggled with token entanglement, frequently failing when multiple concepts are expressed within a single prompt [16, 20]. Prior works attempted to mitigate this issue through inference-time attention control [5, 15, 19, 29] or attention map manipulation using additional conditioning signals [10, 11, 50]. More recent architectures, such as DiT and flow-matching models [6, 13, 23, 32, 43], improve token separation and partially alleviate this problem.

To further improve compositional generation, recent diffusion reinforcement learning (RL) approaches [17, 26, 28, 41, 49] have successfully optimized a simple concept generation using evaluation metrics (*e.g.*, GenEval [16]) as rewards. However, applying this RL paradigm to multi-concept prompts reveals a critical bottleneck. Models frequently fail to generate all concepts together, resulting in partial success (*e.g.*, correct objects but wrong numeracy) [22, 24, 42, 45]. Existing approaches [17,26,28,41,49] attempt to solve this by naively aggregating individual rewards for each concept. However, as illustrated in Figure 1, such naive aggregation tends to overlook difficult or negatively correlated concepts (*i.e.*, Q3, Q4, and Q5). For example, it may overemphasize easier concepts (Q1), while optimizing for the gray color (Q4) often leads to the omission of the spider

(Q3) or the hammers (Q5). Consequently, the model prioritizes easier concepts and fails to generate all requested attributes simultaneously.

To address this, we propose **C**orrelation-Weighted **M**ulti-Reward **O**ptimization (**CMO**), which adaptively reweights concept-wise rewards by leveraging the correlation structure among reward signals across generated samples. Since difficulty varies depending on the specific combination of concepts in each prompt, manual weight adjustment becomes impractical. To address this, CMO automatically estimates concept difficulty within each sampled group. By assigning higher weights to these hard-to-align concepts (*e.g.*, 'three hammers' in Figure 1), CMO encourages the model to generate all requested concepts together. To enable this optimization, we first design a *multi-concept reward* that decomposes multi-concept prompts into fine-grained objectives covering objects, attributes, numeracy, and spatial relations. Each concept is evaluated with dedicated reward functions, producing a set of concept-wise rewards that reflect different aspects of compositional correctness. These rewards are then used by CMO to estimate concept interactions and dynamically reweight difficult concepts. For optimization, we build on recent reward optimization recipes that employ group-wise comparisons for stable learning. In particular, we adopt MixGRPO [26] for efficient mixed ODE–SDE sampling and GDPO [30] for reward-decoupled normalization across multiple rewards.

We apply CMO to SD3.5 [13] and FLUX.1-dev [23]. Extensive experiments demonstrate consistent improvements on multi-concept benchmarks such as ConceptMix [45] and Geneval 2 [22], as well as on T2I-CompBench [20], which evaluates fine-grained attribute binding. Our contributions can be summarized as follows:

- We identify a key limitation of existing RL approaches for multi-concept compositional generation: naive aggregation of concept-wise rewards overemphasizes easy concepts while overlooking difficult or negatively correlated ones.
- We propose Correlation-Weighted Multi-Reward Optimization (CMO), which estimates concept difficulty from correlations among reward signals across generated samples and adaptively reweights concept-wise rewards to emphasize hard-to-satisfy concepts.
- We apply CMO to train state-of-the-art diffusion models (SD3.5 and FLUX.1-dev), demonstrating consistent improvements on challenging multi-concept benchmarks (*e.g.*, ConceptMix, GenEval 2, and T2I-CompBench).

2 Related Work

2.1 Compositional Generation in T2I Models

Compositional generation has long been recognized as a central challenge in text-to-image (T2I) diffusion models, where early models struggle to render multiple concepts such as objects, attributes, and spatial relations expressed in a single prompt [16, 20]. To address these issues, prior works have focused on improving

attribute-level disentanglement through attention manipulation, token interaction control, or additional conditioning mechanisms [5, 11, 15, 19, 29, 50]. While architectural advancements such as DiT and flow-matching models [13, 23, 32] have largely mitigated initial entanglement problems, generating multiple concept together remains challenging [22, 42, 45]. Recent efforts have explored diffusion RL to directly optimize reward signals that reflect multi-constraint satisfaction [41, 51]. However, prior works fail to generate images that satisfy all attributes in a multi-concept prompt simultaneously because they do not consider the relationships between individual attributes within multiple different concepts. To prevent such failures, we propose CMO, a multi-reward optimization framework specifically designed to satisfy multiple concepts simultaneously. By automatically focusing on the hardest concepts, our method ensures that every requested concept is generated at the same time.

2.2 Reinforcement Learning for Diffusion Model

Reinforcement learning is widely used to align diffusion/flow-based T2I models [13, 23, 43] by optimizing rewards along the denoising process under distributional constraints. Prior work includes PPO-style policy gradient methods with KL regularization (e.g., DDPO [3] and DPOK [14]), diffusion-adapted preference learning from win–lose pairs without explicit reward-model training (e.g., Diffusion-DPO [39]), sampler-based optimization with differentiable rewards (e.g., DRaFT [8] and AlignProp [34]), and scalable black-box reward fine-tuning over large prompt sets (e.g., PRDP [12]). Recent work extends RL alignment to flow-matching and rectified-flow backbones using group-relative optimization and stochastic exploration to improve compositional correctness and preference alignment, as exemplified by Flow-GRPO [28], DanceGRPO [47], MixGRPO [26] and TempGRPO [17]. However, these methods simply average multiple reward signals without considering the relationships between them, which can cause optimization to be dominated by easily satisfied concepts while neglecting more difficult or conflicting ones. To address this, we propose CMO, which dynamically reweights multiple rewards to encourage the model to satisfy them simultaneously.

3 Method

In this work, we propose Correlation-Weighted Multi-Reward Optimization (CMO), a new reward optimization framework tailored for multi-concept compositional generation. The overall pipeline is illustrated in Figure 2. First, we present the preliminaries in Sec. 3.1. In Sec. 3.2, we design a multi-concept reward, which decomposes the multi-concept prompt into fine-grained objectives covering objects, attributes, numeracy, and spatial relations, and evaluates each concept with dedicated reward functions. And then in Sec. 3.3, we introduce a correlation-based reweighting mechanism that dynamically assigns higher weights to hard-to-align, conflicting concepts to ensure balanced optimization and encourage the model to satisfy all requested concepts simultaneously.

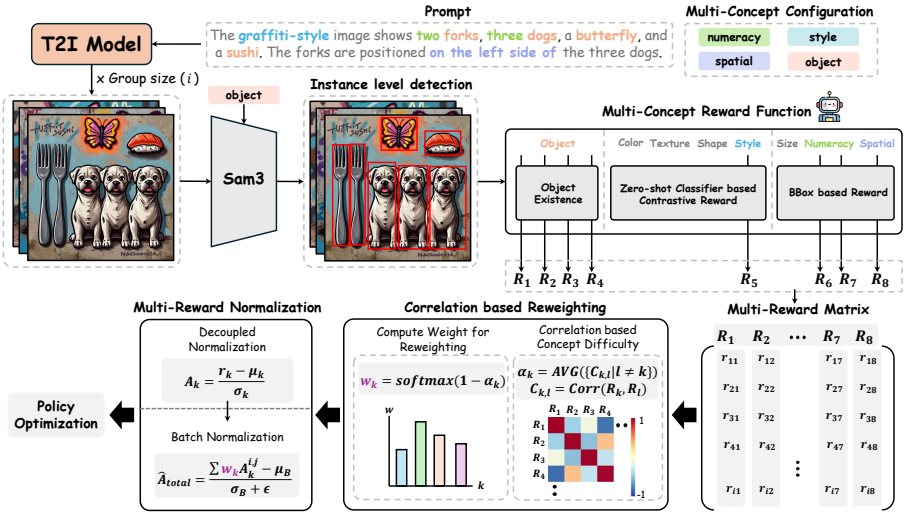

Fig. 2: Overview of Correlation-Weighted Multi-Reward Optimization (CMO). Given a multi-concept prompt, the text-to-image (T2I) model generates a group of images, which are evaluated using a set of concept-specific reward functions covering objects, attributes, and spatial relations. The resulting concept rewards form a Multi-Reward Matrix across generated instances. CMO then computes correlations among reward signals to estimate concept difficulty based on their interactions, assigning higher weights to concepts that are harder to consistently satisfy. The reweighted rewards are normalized and aggregated to guide policy optimization, encouraging the model to jointly satisfy all requested concepts and improving compositional generation.

3.1 Preliminaries

Mixed ODE-SDE Sampling. To mitigate the computational overhead of full-step SDE sampling, MixGRPO [26] restricts stochastic exploration to a sliding window $S \subseteq [0, T]$. It employs SDE sampling within S for RL exploration, and deterministic ODE sampling elsewhere:

$$dx_t = \begin{cases} [v_t - \frac{1}{2} g^2(t) s_t]dt + g(t)dw, & t \in S \\ v_t dt, & t \notin S \end{cases} \tag{1}$$

where v_t and s_t are the velocity field and score function. This concentrates gradients within S, significantly shortening the optimization timestep.

Multi-Reward Normalization. GRPO [38] often suffers from reward collapse in multi-reward settings, mapping distinct reward combinations to identical advantages due to pre-normalization aggregation. GDPO [30] resolves this by decoupling the normalization process. For K rewards, the group-relative advantage $A_k^{(i,j)}$ for the k-th reward is computed independently, then weighted, summed,

and batch-normalized to obtain the final advantage $\hat{A}_{total}^{(i,j)}$:

$$A_k^{(i,j)} = \frac{r_k^{(i,j)} - \mu_k^{(i)}}{\sigma_k^{(i)}}, \quad \hat{A}_{total}^{(i,j)} = \frac{\sum A_k^{(i,j)} - \mu_B}{\sigma_B + \epsilon} \tag{2}$$

where $\mu_k^{(i)}, \sigma_k^{(i)}$ are the group-level mean and standard deviation for the k-th reward, and μ_B, σ_B are the batch-level statistics of the weighted sum.

3.2 Multi-Concept Reward Decomposition

To evaluate compositional generation, we decompose a multi-concept prompt into a set of fine-grained concept objectives. Instead of computing a single scalar reward, our framework evaluates multiple concept-level rewards corresponding to objects, attributes, numeracy, and spatial relations. This decomposition allows each concept in the prompt to be evaluated independently, providing more informative supervision for reward optimization. Formally, given a generated image x and a multi-concept configuration $\mathcal{C}$ (*e.g.*,'style':'graffti', 'object':'dog', 'numeracy':'two' in Figure 2) extracted from the prompt, we compute a set of concept-wise rewards. Each reward corresponds to a specific aspect of compositional correctness. Detailed formulations and implementation details for each reward component are provided in the Appendix.

Object Existence Reward. For each object specification in $\mathcal{C}$, we extract the specific text corresponding to the 'object' key in $\mathcal{C}$ and feed it as a prompt into a segmentation model (*e.g.*, SAM 3 [4]) to perform text-guided instance object detection. Let $\mathcal{O}_i$ denote the set of detected instances for object i. For each valid instance, we extract its bounding box b and segmentation mask m. Based on these detections, we assign a binary existence reward:

$$r_i^{\text{exist}} = \begin{cases} 1, & \text{if } |\mathcal{O}_i| > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

Crucially, the extracted bounding boxes and masks serve as the precise foundation for subsequent localized evaluations, effectively isolating entities from background noise.

Zero-shot Classifier based Contrastive Reward for Attribute Reward. For visual attributes such as color, texture, shape, and style, we measure semantic alignment using a zero-shot classifier (*i.e.*, OpenCLIP-H [7]). To isolate the target object from background noise, we extract the image feature $f_{\text{img}}(x \odot b_i)$ from the cropped region defined by the bounding box b_i, where x is the generated image and b_i is the bounding box of the target object. For each attribute category k (e.g., color), we maintain a predefined candidate set $\mathcal{V}_k$ (e.g., red, green, blue). Let f_v denote the normalized text feature corresponding to candidate attribute $v \in \mathcal{V}_k$. We compute similarities between the image feature and all candidate attributes, and use a softmax classifier to estimate the probability of

the target attribute v^* specified in the prompt. The attribute reward is defined as:

$$r_i^{\text{attr}} = \frac{\exp\left(\langle f_{\text{img}}, f_{v^*} \rangle\right)}{\sum_{v \in \mathcal{V}_k} \exp\left(\langle f_{\text{img}}, f_v \rangle\right)}. \tag{4}$$

Additional details on the construction and normalization of the candidate text features are provided in the Appendix.

BBox based Reward. For numerical reward, instead of a binary penalty such as object existence, we propose a differentiable heuristic based on count deviation. Given the detected count $c_i = |\mathcal{O}_i|$ and the target count c_i^*, the reward is formulated as:

$$r_i^{\text{num}} = \frac{1}{(|c_i - c_i^*| + 1)^2}. \tag{5}$$

Similarly, size attributes (*e.g.*, huge, tiny) are evaluated using a rank-based inverse squared decay over the areas of all detected objects in the scene. For specified spatial relations between objects i and j, we formulate geometric heuristics to assign partial scores (1.0, 0.5, or 0.0) based on alignment constraints.

Notably, for 3D depth relations (*e.g.*, in front of, behind), we leverage the depth estimation from Depth Anything 3 [27] to extract a normalized depth map $\mathcal{D}(x)$. We compute the average depth $\bar{d}_i$ within the masked region m_i:

$$\bar{d}_i = \frac{1}{|m_i|} \sum_{p \in m_i} \mathcal{D}(x)[p]. \tag{6}$$

The depth-based relation reward seamlessly incorporates a tolerance margin ϵ to account for estimation noise. For instance, the reward for i being *in front of* j is formulated as:

$$r_{i,j}^{\text{rel}} = \begin{cases} 1.0, & \text{if } \bar{d}_i < \bar{d}_j - \epsilon, \\ 0.5, & \text{if } \bar{d}_i < \bar{d}_j, \\ 0.0, & \text{otherwise.} \end{cases} \tag{7}$$

Conversely, the reward for the *behind* relation is evaluated using the exact opposite condition. Detailed formulations for 2D spatial relations (*e.g.*, above, below, to the left of, to the right of) are provided in the Appendix.

3.3 Correlation-based Reweighting for Multi-Reward Optimization

Motivation. While the multi-concept reward provides fine-grained supervision for each concept, naively aggregating these rewards can lead to suboptimal optimization. GRPO [31] aggregates multiple rewards through naive summation, and GDPO [30] decouples with reward normalization, and it aggregates advantages. As a result, concepts that are easy to satisfy tend to dominate the optimization process, while difficult or conflicting concepts receive insufficient training signals. To verify this, we analyze correlations of each concept reward across generated instances. Using the ConceptMix set, we define a generation group as 10 images synthesized from a single prompt with different random seeds to analyze

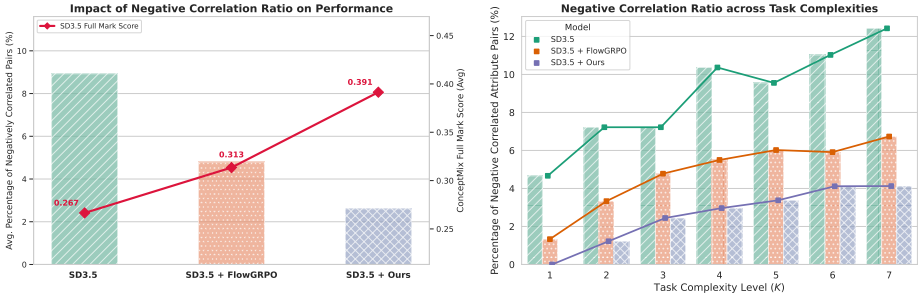

Fig. 3: Quantitative Analysis on Multi-Concept Generation Conflicts. Left: The graph compares the average ratio of negative correlations and the overall Full Mark Score on Conceptmix [45], showing that reducing negative correlations between concepts directly improves multi-concept generation. **Right:** The graph tracks the percentage of concept pairs exhibiting negative correlation as the number of concepts (task complexity level K) increases. Baselines exhibit increasing negative correlations as complexity grows, whereas our method maintains a low negative correlation ratio with correlation-based reweighting.

the relationship between concepts. Figure 3 (left) shows a negative correlation between concepts and the full-mark score on the ConceptMix [45] benchmark across baselines (SD 3.5, SD 3.5 + FlowGRPO, and SD 3.5 + Ours). There is a clear tendency that higher negative correlation leads to lower full-mark scores, indicating a generation conflict where successfully generating one concept limits the ability to generate another. Crucially, Figure 3 (right) shows that baseline models exhibit a sharp increase in the ratio of negatively correlated concept pairs as the number of concepts (prompt complexity, K) grows. This observation suggests that naive aggregation exacerbates trade-offs between concepts. However, our correlation-based reweighting shows that prioritizing difficult and negatively correlated concepts can mitigate generation conflicts and significantly improve multi-concept compositional generation even at high levels of complexity.

Correlation-Based Concept Difficulty Estimation. For a given prompt i, we generate a group of G images. In reward optimization for diffusion models, concept-wise rewards are collected to evaluate generation quality, yielding a reward matrix $\mathbf{R}^{(i)} \in \mathbb{R}^{G\times K}$, where K is the number of concepts. To stably estimate the difficulty among the concepts, we calculate the Pearson correlation [37] matrix $\mathbf{C} \in \mathbb{R}^{K\times K}$ across the K concepts over the G samples:

$$C_{k,l} = \mathrm{Corr}(\mathbf{R}_k^{(i)}, \mathbf{R}_l^{(i)}). \tag{8}$$

During fine-tuning, certain concepts may be perfectly satisfied across all samples in a group, resulting in zero variance ($\mathrm{std}(\mathbf{R}_{:,k}^{(i)}) \approx 0$). In such cases, the correlation is mathematically undefined. To ensure training stability, we simply bypass the correlation estimation for these static concepts and assign a default maximum correlation score to prevent unstable gradient. For each concept k, we

define a score α_k as the average correlation with other attributes:

$$\alpha_k = \begin{cases} 1.0, & \text{if } \mathrm{std}(\mathbf{R}_k^{(i)}) < \epsilon, \\ \frac{1}{K-1} \sum_{l \neq k} C_{k,l}, & \text{otherwise.} \end{cases} \tag{9}$$

Since $C_{k,l} \in [-1, 1]$, negatively correlated (i.e., difficult to generate simultaneously) concepts yield lower scores. To ensure training stability, if the generated group contains incomplete padding values, we safely bypass this correlation estimation.

Correlation-based Reweighting. Concepts with high scores are easier to satisfy jointly and therefore receive lower gradient emphasis. Conversely, weakly correlated concepts are considered more difficult and receive higher weights. We define an correlation-based weighting using a softmax function:

$$w_k = \frac{\exp\big((1 - \alpha_k)\big)}{\sum_{n=1}^{K} \exp\big((1 - \alpha_n)\big)}, \tag{10}$$

To compute the final advantage for the j-th sample in group i, we substitute the aggregation in Eq. 2 with our correlation-based weights. The weighted sum is then batch-normalized to obtain the final advantage :

$$\hat{A}_{\text{total}}^{(i,j)} = \frac{\sum w_k A_k^{(i,j)} - \mu_B}{\sigma_B + \epsilon} \tag{11}$$

This strategy dynamically prioritizes concepts difficult to generate simultaneously via weighted averaging, encouraging balanced reward optimization. a

4 Experiment

In this section, we evaluate the effectiveness of Correlation-Weighted Multi-Reward Optimization (CMO) for improving compositional generation. We first describe the experimental setup in Sec. 4.1, including the training configuration and evaluation benchmarks. We then compare CMO with existing diffusion RL baselines across multiple compositional generation benchmarks in Sec. 4.2. To further analyze the behavior of CMO, we provide detailed ablations on the correlation-based reweighting mechanism and study its impact on resolving concept conflicts in Sec. 4.3. Finally, we present qualitative examples demonstrating improved multi-concept generation under increasing compositional complexity in Sec. 4.4.

4.1 Experimental Setup

Training Prompt Dataset. Our training dataset comprises 5k prompts: 2.5k complex multi-concept and 2.5k single-attribute prompts. For the complex multi-concept set, we synthesize prompts binding up to eight concepts per entity

($K + 1 = 8$) using the ConceptMix [45] pipeline and Qwen3-30B [48]. Motivated by recent findings [17,28] that a single-attribute prompt set(*e.g.*, '{color} {object}', '{color} {object} and {color} {object}) enhances overall compositional generation, we include the single-attribute set. To focus the policy on the most challenging attributes, we heavily skew their sampling distribution (Color : Texture : Shape : Size : Spatial : Numeracy) to a 3:3:2:1:10:7 ratio.

Base Models and Training Setup. We fine-tune Stable Diffusion 3.5 [13] and FLUX.1-dev [23] using our based multi-concept reward optimization with LoRA [18]. Each training iteration samples 8 prompts, generating 16 images per prompt across 4 GPUs, yielding a global batch size of 512. The learning rate is set to 3×10^{-4}. During training, we use 10 sampling steps for SD3.5 and 6 for FLUX.1-dev. To balance exploration and computational efficiency, we employ a mixed ODE-SDE sampling strategy with Coefficient Preserving Sampling (CPS) [40]. Following MixGRPO [26], stochastic SDE updates are confined strictly to the first 3 timesteps, while the remaining steps rely on deterministic ODE sampling.

Evaluation Protocol. We evaluate our method on three compositional text-to-image benchmarks: ConceptMix [45], GenEval 2 [22], and T2I-CompBench [20]. To account for generative stochasticity, we synthesize 10 images per prompt using different random seeds across all benchmarks and report the averaged metrics. While following official protocols, we adapt ConceptMix: originally, it relies on GPT-4o [21] as its primary evaluator and provides DeepSeek-VL-7B-chat [31] as an open-source alternative. To ensure stable reproducibility, we instead employ the more capable Qwen3-VL-8B [1]. All baselines are strictly re-evaluated using this exact setup for fair comparison.

Baselines. We compare our method against a variety of representative compositional T2I baselines. These include Qwen-Image [43] as a strong open-weight generator, and recent RL-based alignment methods such as IterComp [51], Flow-GRPO [28] on SD3.5, and Pref-GRPO [41] on FLUX.1-dev. Due to evaluation costs, NanoBanana [9] is evaluated exclusively on T2I-CompBench. To demonstrate backbone-agnostic scalability, we report results for both SD3.5 and FLUX.1-dev in T2I-CompBench. For other benchmarks, we focus on the high-capacity FLUX.1-dev model to evaluate our method against large-scale baselines such as Qwen-Image.

4.2 Experimental Results

Results on ConceptMix. We first evaluate our proposed method on the ConceptMix benchmark, which rigorously tests a model's ability to handle mixed-concept prompts without attribute leakage. We report the Full Mark Score (Table 1) and Concept Fraction Score (Table 2) across varying difficulty levels ($k \in [1, 7]$), where k represents the number of distinct visual concepts required.

Full Mark Score of T2I Models. As shown in Table 1, CMO consistently outperforms all evaluated baselines across every difficulty level. While the performance of existing models drops sharply as k increases due to attribute leakage and object omission, CMO demonstrates remarkable robustness, achieving

Table 1: Full Mark Score of T2I Models on ConceptMix. We report full mark scores across difficulty levels $k = 1$ to $k = 7$, where k denotes the number of required concepts (thus each prompt contains $k+1$ concepts). As k increases, performance degrades for all models, while our method consistently mitigates degradation at higher complexity. **Bold** and underline indicate the best and second-best results, respectively.

Models	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$
Autoregressive Models							
Janus [44]	0.5340	0.3063	0.1710	0.0713	0.0383	0.0173	0.0083
Show-o [46]	0.7132	0.4560	0.2747	0.1507	0.1100	0.0583	0.0357
HermesFlow [49]	0.6993	0.4560	0.2690	0.1467	0.1167	0.0540	0.0403
Non-Autoregressive Models							
StableDiffusion v1.4 [35]	0.4780	0.1943	0.0717	0.0203	0.0040	0.0007	0.0000
StableDiffusion v2 [33]	0.5170	0.2683	0.1010	0.0437	0.0120	0.0040	0.0010
StableDiffusion XL Turbo [36]	0.5886	0.3073	0.1343	0.0493	0.0190	0.0103	0.0003
PixArt-α [6]	0.6170	0.3133	0.1353	0.0490	0.0207	0.0097	0.0013
StableDiffusion XL [33]	0.6237	0.3517	0.1670	0.0553	0.0273	0.0127	0.0050
Playground V2.5 [25]	0.6963	0.3830	0.1917	0.0690	0.0373	0.0113	0.0017
StableDiffusion 3.5 [13]	0.7003	0.4467	0.3010	0.1673	0.1487	0.0597	0.0430
FLUX.1-dev [23]	0.6647	0.4127	0.2680	0.1620	0.1173	0.0567	0.0372
IterComp [51]	0.6570	0.3790	0.2143	0.0740	0.0417	0.0213	0.0040
FlowGRPO [28]	0.7600	0.5457	0.3517	0.2103	0.1727	0.0827	0.0713
PrefGRPO [41]	0.7037	0.4790	0.3023	0.2120	0.1603	0.0897	0.0647
Qwen-Image [43]	0.8183	0.6583	0.5060	0.3803	0.3353	0.2213	0.1747
CMO	**0.8410**	**0.7063**	**0.5700**	**0.3947**	**0.3560**	**0.2317**	**0.1883**

Table 2: Concept Fraction Score of T2I Models on ConceptMix. We report the concept fraction score across difficulty levels $k = 1$ to $k = 7$, where k denotes the number of required concepts. As k increases, the scores of all models decrease, but at different rates. **Bold** and underline indicate the best and second-best results, respectively.

Models	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$
Autoregressive Models							
Janus [44]	0.7385	0.6688	0.6318	0.5681	0.5602	0.5585	0.5285
Show-o [46]	0.8435	0.7663	0.7260	0.6741	0.6691	0.6378	0.6331
HermesFlow [49]	0.8362	0.7668	0.7193	0.6775	0.6705	0.6390	0.6348
Non-Autogressive Models							
Stable Diffusion v1.4 [35]	0.7153	0.5897	0.5190	0.4508	0.4105	0.3807	0.3475
Stable Diffusion v2 [35]	0.7397	0.6467	0.5822	0.5086	0.4702	0.4458	0.4228
StableDiffusion XL Turbo [36]	0.7824	0.6782	0.6324	0.5661	0.5330	0.4968	0.4630
PixArt-α [6]	0.7900	0.6858	0.6312	0.5634	0.5262	0.4936	0.4625
StableDiffusion XL [33]	0.8043	0.7169	0.6657	0.5958	0.5727	0.5337	0.5023
Playground V2.5 [25]	0.8399	0.7298	0.6701	0.5962	0.5800	0.5452	0.5000
StableDiffusion 3.5 [13]	0.8367	0.7642	0.7368	0.6937	0.6943	0.6603	0.6565
FLUX.1-dev [23]	0.8162	0.7430	0.7093	0.6706	0.6779	0.6467	0.6576
IterComp [51]	0.8247	0.7293	0.6984	0.6139	0.6038	0.5727	0.5419
FlowGRPO [28]	0.8747	0.8144	0.7726	0.7235	0.7263	0.7010	0.6953
Pref-GRPO [41]	0.8395	0.7819	0.7394	0.7159	0.7129	0.6868	0.6846
Qwen-Image [43]	0.9052	0.8600	0.8403	**0.8193**	0.8171	0.7869	**0.7942**
CMO	**0.9147**	**0.8850**	**0.8635**	0.8188	**0.8214**	**0.7917**	0.7884

the highest score of 0.8410 at $k = 1$. Most notably, at the extreme complexity of $k = 7$, our model successfully maintains a score of 0.1883. This not only yields a substantial improvement over the base FLUX.1-dev model (0.0372) but also decisively surpasses the highly competitive Qwen-Image (0.1747). These results empirically prove that our approach effectively handles dense multi-concept prompts, setting a new state-of-the-art for strict multi-concept generation.

Concept Fraction Score of T2I Models. As shown in Table 2, we also report the Concept Fraction Score to evaluate the average proportion of correctly generated concepts within a prompt, accounting for partial successes. Similar to the strict exact-match evaluation, our proposed method consistently demonstrates superior performance across all difficulty levels. Notably, even under the extreme compositional constraint of $k = 7$, our model successfully preserves a high concept fraction score of 0.7884. This result not only significantly improves upon the base FLUX.1-dev (0.6576) and the RL-aligned Pref-GRPO (0.6846), but also closely rivals or exceeds the massive Qwen-Image model across most complexity levels. These findings confirm that difficulty-aware compositional reward effectively prevents catastrophic object omission, enabling the model to retain strong attribute binding even when densely packed with concepts.

Evaluation on GenEval 2. Table 3 presents the results on GenEval 2 using the Soft-TIFA metric, which evaluates atom-level (TIFA_{AM}) and prompt-level (TIFA_{GM}) correctness. Our method demonstrates superior generalization by establishing a new state-of-the-art, achieving **80.0** in TIFA_{AM} and **34.0** in TIFA_{GM}. This decisively outperforms existing RL-aligned baselines (FlowGRPO, PrefGRPO) and notably surpasses the large-scale Qwen-Image in prompt-level correctness (34.0 vs. 31.4). These results confirm that our fine-grained reward structure successfully enforces complex relational reasoning and holistic composition, generalizing robustly to unseen prompt distributions rather than merely relying on memorized attribute associations.

Table 3: Soft-TIFA Evaluation of T2I Models on GenEval 2.

Model	Soft TIFA_{AM}	Soft TIFA_{GM}
Autoregressive Models		
Janus [44]	52.9	11.0
Show-o [46]	63.4	16.0
HermesFlow [49]	63.5	16.1
Non-Autogressive Models		
SD 2.1 [35]	40.8	5.2
SDXL [33]	50.1	9.1
SD3.5-M [13]	63.7	15.8
FLUX.1-dev [23]	64.1	17.1
IterComp [51]	53.0	8.5
FlowGRPO [28]	70.9	21.8
PrefGRPO [41]	70.3	23.0
Qwen-Image [43]	**80.0**	31.4
CMO	**80.0**	**34.0**

Evaluation on T2I-CompBench. To systematically evaluate fine-grained compositional capabilities, we present the results on T2I-CompBench in Table 4. This benchmark assesses models across multiple dimensions, including attribute binding (color, shape, texture), object relationships (spatial, non-spatial), numeracy, and complex compositions. Our proposed framework achieves state-of-the-art results in the majority of individual categories and consistently yields the highest overall averages. Specifically, our method equipped with Stable Diffusion 3.5 records a state-of-the-art Avg score of **0.6141**, excelling significantly in fundamental attribute binding (Color: **0.8692**, Shape: **0.6427**, Texture: **0.7959**) and spatial relationships (**0.5475**). Furthermore, when applied to the FLUX.1-

Table 4: Evaluation Results on T2I-CompBench [20]. We report attribute binding, object relationships, numeracy, complex compositions, and their overall **Avg** (arithmetic mean over the seven metrics). **Bold** and underline indicate the best and second-best results, respectively.

Model	Attribute Binding			Object Relationship		Numeracy↑	Complex↑	Avg↑
	Color ↑	Shape↑	Texture↑	Spatial↑	Non-Spatial↑			
	BLIP-VQA	BLIP-VQA	BLIP-VQA	UniDet	CLIPscore	UniDet	3-in-1	
Stable Diffusion XL [33]	0.5879	0.4687	0.5299	0.2133	0.3119	0.4991	0.3237	0.4192
PixArt-α [6]	0.4123	0.3809	0.4559	0.2057	0.3083	0.5085	0.3604	0.3760
DALL-E 3 [2]	0.7785	0.6205	0.7036	0.2865	0.3003	0.5926	0.3373	0.5170
Stable Diffusion 3.5 [13]	0.7926	0.5617	0.7338	0.2914	0.3139	0.5898	0.3842	0.5239
FLUX.1-dev [23]	0.7358	0.4802	0.5989	0.2461	0.3067	0.6107	0.4281	0.4866
IterComp [51]	0.7295	0.5356	0.6526	0.2471	**0.3227**	0.5430	0.3883	0.4884
FlowGRPO [28]	0.8263	0.6177	0.7241	0.5237	0.3184	0.6987	0.3905	0.5856
PrefGRPO [41]	0.7895	0.5595	0.6688	0.2846	0.3102	0.6559	0.4661	0.5335
NanoBanana [9]	0.8194	0.6275	0.7159	0.3240	0.3023	0.6321	0.3741	0.5422
Qwen-Image [43]	0.8395	0.5882	0.7407	0.4454	0.3136	**0.7553**	0.4414	0.5892
StableDiffusion 3.5 + CMO	**0.8692**	**0.6427**	**0.7959**	**0.5475**	0.3201	0.7200	0.4036	**0.6141**
FLUX.1-dev + CMO	0.8607	0.6188	0.7206	0.4577	0.3171	0.7066	**0.4909**	0.5961

Table 5: Ablation on Design Choices. We explicitly isolate the impact of our baseline optimization, training data, and correlation mechanism. **DRO** (Decoupled Reward Optimization) denotes our base multi-reward setup using MixGRPO [26] and GDPO [30]. **Single-Attr.** and **Multi-Concept** represent the inclusion of training sets consisting of single-attribute binding and multi concepts, respectively. **CR** denotes Correlation-based Reweighting.

Method	DRO	Single-Attr.	Multi-Concept	CR	Avg. Full Mark	Avg. Concept Frac.
SD3.5-M (Base)	-	-	-	-	0.2667	0.7203
Ours	✓	×	×	×	0.2713	0.7411
	✓	✓	×	×	0.3305	0.7752
	✓	✓	✓	×	0.3531	0.7993
	✓	✓	✓	✓	**0.3913**	**0.8126**

dev backbone, our method demonstrates exceptional proficiency in higher-order reasoning, achieving the highest score in the Complex category (**0.4909**) and an impressive Avg of 0.5961. These comprehensive improvements across diverse and granular criteria confirm that our difficulty-aware reward optimization avoids overfitting to specific attributes, ensuring a well-balanced and robust compositional generation capability.

4.3 Ablation Study

To provide a more in-depth discussion of the findings from Table 5 in the main manuscript, Table 5 delves deeper by explicitly isolating the contributions of our reward optimization baseline (**DRO**), training datasets (**Single-Attr.** and **Multi-Concept**), and our core contribution (**CR**). To comprehensively analyze the optimization, we evaluate both Avg. Full Mark score and Avg. Concept Fraction score.

The results clearly demonstrate the necessity of each proposed component. Applying our advanced optimization (**DRO**) solely on the standard baseline

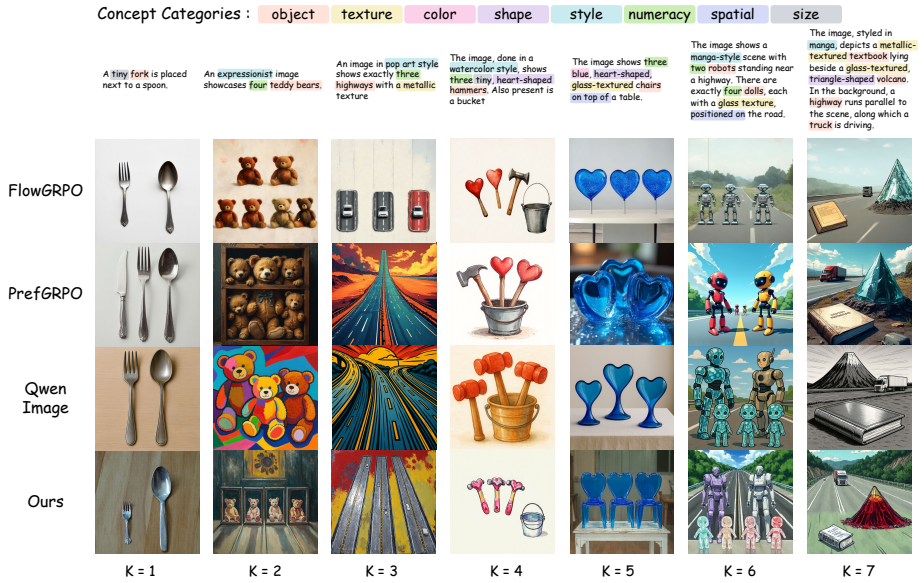

Fig. 4: Qualitative comparison across varying prompt complexity ($K = 1 \sim 7$). Baseline models frequently exhibit concept omission or attribute leakage as complexity increases. In contrast, our method consistently maintains high faithfulness and accurate attribute binding even under extreme constraints ($K = 7$).

dataset severely degrades performance (Avg. Full Mark score: 0.2713), proving that standard datasets lack the structural density required to ground fine-grained multi-concept rewards. Swapping to our proposed **Single-Attr.** and **Multi-Concept** datasets provides the essential training signals, stabilizing the optimization and recovering the performance to 0.3531. Ultimately, even with optimal data and baseline optimization, the breakthrough is strictly driven by **CR**. By explicitly resolving the remaining negative correlations among conflicting concepts, **CR** achieves the state-of-the-art performance (Avg. Full Mark score: 0.3913, Avg. Concept Fraction: 0.8126).

4.4 Qualitative Result on ConceptMix

Figure 4 illustrates the qualitative performance across complexity levels $K = 1$ to $K = 7$. While baselines often suffer from object omission or misaligned attributes as K increases, our method consistently generates faithful images by accurately aligning a diverse categories(*e.g.*, object, color, numeracy, spatial, size).

5 Discussions

The current design of CMO also suggests several directions for future work. First, future studies may develop more adaptive training data construction strategies that adjust the sampling frequency of each concept group based on observed failure patterns or compositional conflicts. In particular, the sampling ratio could be

updated according to which reward groups are repeatedly involved in failed generations, rather than being fixed a priori. This would make the training prompt distribution itself part of the difficulty-aware optimization loop. Such data-centric refinement could help determine whether failures in specific attributes can be mitigated by exposing the model to more targeted concept combinations. Second, CMO may be extended to preference-aware settings by jointly considering compositional prompts and perceptual-quality constraints during data construction and reward design. Finally, studying the relationship between pre-training data scale and post-training multi-reward optimization would further clarify when structured reward optimization is most effective for compositional generation.

6 Conclusion

In this work, we address the persistent challenge of multi-concept compositional generation in text-to-image models,where models often struggle to align multiple concepts within a prompt simultaneously. To overcome this, we propose Correlation-Weighted Multi-Reward Optimization (CMO). Crucially, by computing the correlation among concept-wise reward signals across generated samples, our framework explicitly estimates concept generation difficulty. This allows the model to dynamically assign higher optimization weights to negatively correlated, hard-to-align concepts. Extensive experiments on state-of-the-art architectures, including Stable Diffusion 3.5 and FLUX.1-dev, demonstrate that CMO significantly mitigates negative correlations among multi-concepts even under extreme prompt complexities. By establishing new state-of-the-art performance on challenging benchmarks like ConceptMix, Geneval 2, and T2I-CompBench, our findings demonstrate that designing a structured, difficulty-aware multi-reward landscape is the definitive key to mastering true compositional generation.

Acknowledgements

This research was supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No. RS-2019-II190079, Artificial Intelligence Graduate School Program(Korea University), 1%; High-Performance Research AI Computing Infrastructure Support at the 2 PFLOPS Scale, 1%), Culture, Sports and Tourism R&D Program through the Korea Creative Content Agency grant funded by the Ministry of Culture, Sports and Tourism in 2024(Project Name: International Collaborative Research and Global Talent Development for the Development of Copyright Management and Protection Technologies for Generative AI, Project Number: RS-2024-00345025, 10%), the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. RS-2024-00341514, 28%; No. RS-2025-02263628, 60%)

References

1. Bai, S., Cai, Y., Chen, R., Chen, K., Chen, X., Cheng, Z., Deng, L., Ding, W., Gao, C., Ge, C., et al.: Qwen3-vl technical report. arXiv preprint arXiv:2511.21631 (2025)
2. Betker, J., Goh, G., Jing, L., Brooks, T., Wang, J., Li, L., Ouyang, L., Zhuang, J., Lee, J., Guo, Y., et al.: Improving image generation with better captions. Computer Science. https://cdn. openai. com/papers/dall-e-3. pdf **2**(3), 8 (2023)
3. Black, K., Janner, M., Du, Y., Kostrikov, I., Levine, S.: Training diffusion models with reinforcement learning. arXiv preprint arXiv:2305.13301 (2023)
4. Carion, N., Gustafson, L., Hu, Y.T., Debnath, S., Hu, R., Suris, D., Ryali, C., Alwala, K.V., Khedr, H., Huang, A., et al.: Sam 3: Segment anything with concepts. arXiv preprint arXiv:2511.16719 (2025)
5. Chefer, H., Alaluf, Y., Vinker, Y., Wolf, L., Cohen-Or, D.: Attend-and-excite: Attention-based semantic guidance for text-to-image diffusion models. ACM transactions on Graphics (TOG) **42**(4), 1–10 (2023)
6. Chen, J., Yu, J., Ge, C., Yao, L., Xie, E., Wu, Y., Wang, Z., Kwok, J., Luo, P., Lu, H., et al.: Pixart-*alpha*: Fast training of diffusion transformer for photorealistic text-to-image synthesis. arXiv preprint arXiv:2310.00426 (2023)
7. Cherti, M., Beaumont, R., Wightman, R., Wortsman, M., Ilharco, G., Gordon, C., Schuhmann, C., Schmidt, L., Jitsev, J.: Reproducible scaling laws for contrastive language-image learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 2818–2829 (2023)
8. Clark, K., Vicol, P., Swersky, K., Fleet, D.J.: Directly fine-tuning diffusion models on differentiable rewards. arXiv preprint arXiv:2309.17400 (2023)
9. Comanici, G., Bieber, E., Schaekermann, M., Pasupat, I., Sachdeva, N., Dhillon, I., Blistein, M., Ram, O., Zhang, D., Rosen, E., et al.: Gemini 2.5: Pushing the frontier with advanced reasoning, multimodality, long context, and next generation agentic capabilities. arXiv preprint arXiv:2507.06261 (2025)
10. Dahary, O., Patashnik, O., Aberman, K., Cohen-Or, D.: Be yourself: Bounded attention for multi-subject text-to-image generation. In: European Conference on Computer Vision. pp. 432–448. Springer (2024)
11. Dat, D.H., Hyeon-Woo, N., Mao, P.Y., Oh, T.H.: Vsc: Visual search compositional text-to-image diffusion model. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 19153–19162 (2025)
12. Deng, F., Wang, Q., Wei, W., Hou, T., Grundmann, M.: Prdp: Proximal reward difference prediction for large-scale reward finetuning of diffusion models. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 7423–7433 (2024)
13. Esser, P., Kulal, S., Blattmann, A., Entezari, R., Müller, J., Saini, H., Levi, Y., Lorenz, D., Sauer, A., Boesel, F., et al.: Scaling rectified flow transformers for high-resolution image synthesis. In: Forty-first international conference on machine learning (2024)
14. Fan, Y., Watkins, O., Du, Y., Liu, H., Ryu, M., Boutilier, C., Abbeel, P., Ghavamzadeh, M., Lee, K., Lee, K.: Dpok: Reinforcement learning for fine-tuning text-to-image diffusion models. Advances in Neural Information Processing Systems **36**, 79858–79885 (2023)
15. Feng, W., He, X., Fu, T.J., Jampani, V., Akula, A., Narayana, P., Basu, S., Wang, X.E., Wang, W.Y.: Training-free structured diffusion guidance for compositional text-to-image synthesis. arXiv preprint arXiv:2212.05032 (2022)

16. Ghosh, D., Hajishirzi, H., Schmidt, L.: Geneval: An object-focused framework for evaluating text-to-image alignment. Advances in Neural Information Processing Systems **36**, 52132–52152 (2023)
17. He, X., Fu, S., Zhao, Y., Li, W., Yang, J., Yin, D., Rao, F., Zhang, B.: Tempflow-grpo: When timing matters for grpo in flow models, 2025. URL https://arxiv.org/abs/2508.04324 (2025)
18. Hu, E.J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., Chen, W., et al.: Lora: Low-rank adaptation of large language models. Iclr **1**(2), 3 (2022)
19. Hu, T., Li, L., van de Weijer, J., Gao, H., Shahbaz Khan, F., Yang, J., Cheng, M.M., Wang, K., Wang, Y.: Token merging for training-free semantic binding in text-to-image synthesis. Advances in Neural Information Processing Systems **37**, 137646–137672 (2024)
20. Huang, K., Sun, K., Xie, E., Li, Z., Liu, X.: T2i-compbench: A comprehensive benchmark for open-world compositional text-to-image generation. Advances in Neural Information Processing Systems **36**, 78723–78747 (2023)
21. Hurst, A., Lerer, A., Goucher, A.P., Perelman, A., Ramesh, A., Clark, A., Ostrow, A., Welihinda, A., Hayes, A., Radford, A., et al.: Gpt-4o system card. arXiv preprint arXiv:2410.21276 (2024)
22. Kamath, A., Chang, K.W., Krishna, R., Zettlemoyer, L., Hu, Y., Ghazvininejad, M.: Geneval 2: Addressing benchmark drift in text-to-image evaluation. arXiv preprint arXiv:2512.16853 (2025)
23. Labs, B.F.: Flux. `https://github.com/black-forest-labs/flux` (2024)
24. Li, B., Lin, Z., Pathak, D., Li, J.E., Xia, X., Neubig, G., Zhang, P., Ramanan, D.: Genai-bench: A holistic benchmark for compositional text-to-visual generation. In: Synthetic Data for Computer Vision Workshop@ CVPR 2024 (2024)
25. Li, D., Kamko, A., Akhgari, E., Sabet, A., Xu, L., Doshi, S.: Playground v2. 5: Three insights towards enhancing aesthetic quality in text-to-image generation. arXiv preprint arXiv:2402.17245 (2024)
26. Li, J., Cui, Y., Huang, T., Ma, Y., Fan, C., Yang, M., Zhong, Z.: Mixgrpo: Unlocking flow-based grpo efficiency with mixed ode-sde. arXiv preprint arXiv:2507.21802 (2025)
27. Lin, H., Chen, S., Liew, J., Chen, D.Y., Li, Z., Shi, G., Feng, J., Kang, B.: Depth anything 3: Recovering the visual space from any views. arXiv preprint arXiv:2511.10647 (2025)
28. Liu, J., Liu, G., Liang, J., Li, Y., Liu, J., Wang, X., Wan, P., Zhang, D., Ouyang, W.: Flow-grpo: Training flow matching models via online rl. arXiv preprint arXiv:2505.05470 (2025)
29. Liu, N., Li, S., Du, Y., Torralba, A., Tenenbaum, J.B.: Compositional visual generation with composable diffusion models. In: European conference on computer vision. pp. 423–439. Springer (2022)
30. Liu, S.Y., Dong, X., Lu, X., Diao, S., Belcak, P., Liu, M., Chen, M.H., Yin, H., Wang, Y.C.F., Cheng, K.T., et al.: Gdpo: Group reward-decoupled normalization policy optimization for multi-reward rl optimization. arXiv preprint arXiv:2601.05242 (2026)
31. Lu, H., Liu, W., Zhang, B., Wang, B., Dong, K., Liu, B., Sun, J., Ren, T., Li, Z., Yang, H., et al.: Deepseek-vl: towards real-world vision-language understanding. arXiv preprint arXiv:2403.05525 (2024)
32. Peebles, W., Xie, S.: Scalable diffusion models with transformers. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 4195–4205 (2023)

33. Podell, D., English, Z., Lacey, K., Blattmann, A., Dockhorn, T., Müller, J., Penna, J., Rombach, R.: Sdxl: Improving latent diffusion models for high-resolution image synthesis. arXiv preprint arXiv:2307.01952 (2023)
34. Prabhudesai, M., Goyal, A., Pathak, D., Fragkiadaki, K.: Aligning text-to-image diffusion models with reward backpropagation (2023)
35. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B.: High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 10684–10695 (2022)
36. Sauer, A., Lorenz, D., Blattmann, A., Rombach, R.: Adversarial diffusion distillation. In: European Conference on Computer Vision. pp. 87–103. Springer (2024)
37. Sedgwick, P.: Pearson's correlation coefficient. Bmj **345** (2012)
38. Shao, Z., Wang, P., Zhu, Q., Xu, R., Song, J., Bi, X., Zhang, H., Zhang, M., Li, Y., Wu, Y., et al.: Deepseekmath: Pushing the limits of mathematical reasoning in open language models. arXiv preprint arXiv:2402.03300 (2024)
39. Wallace, B., Dang, M., Rafailov, R., Zhou, L., Lou, A., Purushwalkam, S., Ermon, S., Xiong, C., Joty, S., Naik, N.: Diffusion model alignment using direct preference optimization. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 8228–8238 (2024)
40. Wang, F., Yu, Z.: Coefficients-preserving sampling for reinforcement learning with flow matching. arXiv preprint arXiv:2509.05952 (2025)
41. Wang, Y., Li, Z., Zang, Y., Zhou, Y., Bu, J., Wang, C., Lu, Q., Jin, C., Wang, J.: Pref-grpo: Pairwise preference reward-based grpo for stable text-to-image reinforcement learning. arXiv preprint arXiv:2508.20751 (2025)
42. Wei, X., Zhang, J., Wang, Z., Wei, H., Guo, Z., Zhang, L.: Tiif-bench: How does your t2i model follow your instructions? arXiv preprint arXiv:2506.02161 (2025)
43. Wu, C., Li, J., Zhou, J., Lin, J., Gao, K., Yan, K., Yin, S.m., Bai, S., Xu, X., Chen, Y., et al.: Qwen-image technical report. arXiv preprint arXiv:2508.02324 (2025)
44. Wu, C., Chen, X., Wu, Z., Ma, Y., Liu, X., Pan, Z., Liu, W., Xie, Z., Yu, X., Ruan, C., et al.: Janus: Decoupling visual encoding for unified multimodal understanding and generation. In: Proceedings of the Computer Vision and Pattern Recognition Conference. pp. 12966–12977 (2025)
45. Wu, X., Yu, D., Huang, Y., Russakovsky, O., Arora, S.: Conceptmix: A compositional image generation benchmark with controllable difficulty. Advances in Neural Information Processing Systems **37**, 86004–86047 (2024)
46. Xie, J., Mao, W., Bai, Z., Zhang, D.J., Wang, W., Lin, K.Q., Gu, Y., Chen, Z., Yang, Z., Shou, M.Z.: Show-o: One single transformer to unify multimodal understanding and generation. arXiv preprint arXiv:2408.12528 (2024)
47. Xue, Z., Wu, J., Gao, Y., Kong, F., Zhu, L., Chen, M., Liu, Z., Liu, W., Guo, Q., Huang, W., et al.: Dancegrpo: Unleashing grpo on visual generation. arXiv preprint arXiv:2505.07818 (2025)
48. Yang, A., Li, A., Yang, B., Zhang, B., Hui, B., Zheng, B., Yu, B., Gao, C., Huang, C., Lv, C., et al.: Qwen3 technical report. arXiv preprint arXiv:2505.09388 (2025)
49. Yang, L., Zhang, X., Tian, Y., Shang, C., Xu, M., Zhang, W., Cui, B.: Hermesflow: Seamlessly closing the gap in multimodal understanding and generation. arXiv preprint arXiv:2502.12148 (2025)
50. Zhang, X., Yang, L., Cai, Y., Yu, Z., Xie, J., Tian, Y., Xu, M., Tang, Y., Yang, Y., Cui, B.: Realcompo: Dynamic equilibrium between realism and compositionality improves text-to-image diffusion models. CoRR (2024)
51. Zhang, X., Yang, L., Li, G., Cai, Y., Xie, J., Tang, Y., Yang, Y., Wang, M., Cui, B.: Itercomp: Iterative composition-aware feedback learning from model gallery for text-to-image generation. arXiv preprint arXiv:2410.07171 (2024)