REPAIR: MINING MULTI-DIRECTIONAL REGIONAL PREFERENCES FROM IMPERFECT PAIRS

Anonymous authors
Paper under double-blind review



Figure 1: **RePair improves fine-grained text-to-image alignment on Ideogram 4.** Compared with the base model, RePair leverages paired imperfect generations to correct diverse localized failures.

ABSTRACT

Post-training plays a crucial role in advancing generative models, yet acquiring reliable reward signals is costly and sparse. Conventional approaches evaluate an entire image with a global reward, implicitly assuming that artifacts are evenly distributed. However, modern text-to-image models usually produce localized failures rather than global degradation. To supervise local errors, recent preference construction heavily rely on manually designed contrasts, which inherently assume that all regional preferences point toward a target image, imposing a restriction we term the *unidirectional preference constraint*. Instead, we explore a different source of supervision by leveraging the intrinsic stochasticity of text-to-image models, where local successes and failures are often mixed, formalizing this rich signal as *multi-directional regional preferences*. Building on this insight, we introduce RePair, which transforms cross-generation variance into reliable regional supervision. First, correlated sampling creates candidate images where global structures remain aligned while local differences are preserved. Second, visual grounding directly mines verified regional preferences without requiring artificial editing. Finally, we optimize each regional relation independently using a signed, region-normalized objective, allowing a single image pair to simultaneously provide multiple preference directions. Extensive experiments on FLUX.2 and Ideogram 4 demonstrate that RePair, trained on only 5K automatically mined pairs, achieves substantial improvements across compositional benchmarks.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

1 INTRODUCTION

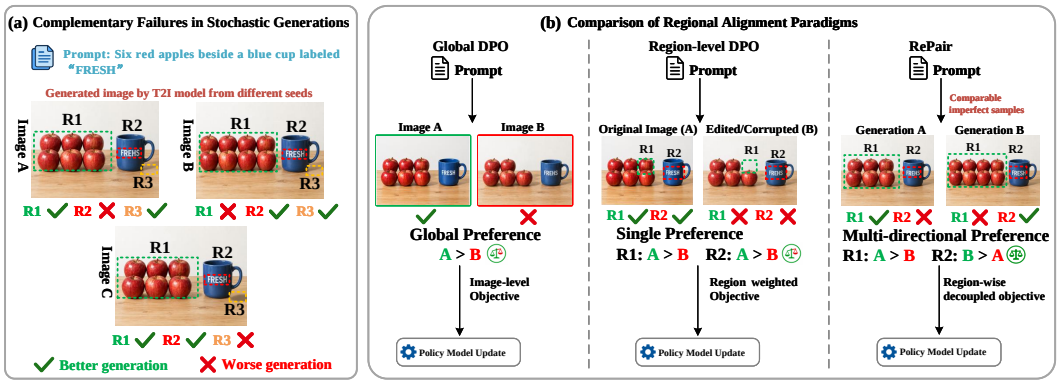

Figure 2: **Multi-directional regional preferences from stochastic generations.** (a) Stochastic generations naturally contain complementary local successes and failures. (b) Comparison between existing regional alignment paradigms and RePair, which preserves multiple regional preference directions within a single image pair.

Modern text-to-image (T2I) models Wu et al. (2025a); Labs (2025); Wu et al. (2025b); Team (2025) exhibit remarkable generative capabilities, yet frequently struggle with localized requirements within complex compositional prompts. To bridge this gap, feedback-driven post-training Xu et al. (2023); Liang et al. (2024); Clark et al. (2024) has emerged as an indispensable paradigm. However, conventional alignment approaches Wallace et al. (2024); Liu et al. (2025) rely on monolithic feedback, evaluating an entire image with a single global reward. This implicitly assumes that generation artifacts are evenly distributed, directly contradicting the reality that modern T2I failures are highly localized, often manifesting as specific attribute mismatches or missing objects confined to isolated regions.

Since prevailing post-training paradigms Wallace et al. (2024); Liu et al. (2025) propagate a global preference to all semantic components. An excellently rendered region within an otherwise flawed generation is inadvertently penalized, while a defective region in a successful generation is mistakenly encouraged. Framing this within the context of preference optimization, we formalize this structural limitation as the *unidirectional preference constraint*, the flawed assumption that a single preference label applies uniformly across all local regions. Recent region-aware preference methods Liu et al. (2026); Huang et al. (2026) partially relax this assumption by introducing spatially localized supervision, but they still inherit this constraint by sharing one identical preference label per image pair. By forcing a uniform optimization polarity, this constraint artificially bottlenecks the learning signal and completely fails to capture the stochastic reality where local successes and failures frequently coexist within the same generation.

To break free from this structural constraint, we explore the intrinsic stochasticity of T2I models as a scalable, natural source of supervision. Because different samples generated from the same prompt realize semantic requirements with varying degrees of fidelity, natural cross-generation variance already contains rich semantic-spatial contrasts, as illustrated in Fig. 2 (a). A region that fails in one sample may be rendered perfectly in another, and vice versa. Guided by this perspective, we formalize this naturally occurring, mixed signal as *multi-directional regional preferences* (Fig. 2 (b)). In this paradigm, a single pair of stochastic generations can simultaneously provide multiple valid local comparisons that favor different images, effectively liberating preference learning from the unidirectional assumption. The primary challenge in mining these natural signals lies in the *local attribution problem*: independent samples often differ drastically in overall layout, making it highly unreliable to isolate local rendering variations from irrelevant global composition changes.

To address this challenge, we introduce RePair, a novel framework that first generates *comparable imperfections* via correlated sampling. This creates controlled variations where semantic structures remain aligned while local generation quality naturally varies. Our pipeline subsequently grounds candidate failures, aligns semantically corresponding regions across generations, and filters out ambiguous pairs. Ultimately, this verification process extracts multiple verified regional preferences

from the same imperfect pair, successfully capturing these multi-directional signals without forcing them to align with a single image. Instead of aggregating these localized signals into a single pairwise preference, our framework optimizes each regional relation independently using a signed, regional objective. This crucial design preserves these multi-directional local comparisons within the same image pair, preventing distinct regional signals from canceling each other out. Extensive evaluations on both JSON-conditioned and natural-language-conditioned image generators demonstrate that RePair can effectively exploit this automatically mined supervision, achieving consistent improvements on compositional benchmarks using only 5K automatically mined preference pairs.

Our contributions are summarized as: (1) We identify the *unidirectional preference constraint* as a shared bottleneck of prior preference learning for text-to-image generation, and formalize *multi-directional regional preferences* to capture the coexisting local successes and failures within a single generation pair. (2) We introduce RePair, a fully automated framework that mines verified regional preferences from the intrinsic stochasticity of text-to-image generation. (3) We propose a signed, region-normalized objective that optimizes each regional relation independently, avoiding the gradient cancellation of image-level aggregation. (4) Extensive experiments on Ideogram 4 and FLUX.2 show that RePair, trained on only 5K automatically mined pairs, achieves consistent gains across compositional benchmarks.

2 Related Work

2.1 Post-Training for Text-to-Image Generation

Inspired by the success of reinforcement learning from human feedback in language models Christiano et al. (2017); Ouyang et al. (2022); Rafailov et al. (2023), post-training has been widely adopted for text-to-image generation. Existing approaches learn reward models from human or model-generated preferences and optimize generative models through reward backpropagation or reinforcement learning Kirstain et al. (2023); Wu et al. (2023); Xu et al. (2023); Liang et al. (2024); Prabhudesai et al. (2024); Clark et al. (2024); Black et al. (2024). Recent studies further introduce semantic decomposition and spatially grounded feedback for more targeted compositional alignment Xu et al. (2026); Feng et al. (2026); Hu et al. (2023); Lin et al. (2024); Xing et al. (2025); Guo et al. (2025); Yang et al. (2026); Zhang et al. (2026). Another line directly optimizes models from preference supervision and improves preference construction, data selection, and optimization objectives Wallace et al. (2024); Borso et al. (2025); Zhang et al. (2025); Luo et al. (2026); Li et al. (2025); Pan et al. (2025); Oh et al. (2026); Hu et al. (2025); Fu et al. (2025).

2.2 Region-Level Preference Learning

Recent studies introduce localized structures into preference optimization. Existing approaches construct regional preference supervision through controlled editing, corruption, fusion, or object-centric perturbation Liu et al. (2026); Huang et al. (2026); Hu et al. (2025); Oh et al. (2026), while others define preference units over predefined semantic, glyph-level, or temporal structures Sun et al. (2026); Shuai et al. (2026); Wu et al. (2025c). Beyond preference construction, prior work has also improved preference reliability through inversion-based alignment, group-wise selection, and stabilized preference objectives Li et al. (2025); Luo et al. (2026); Fu et al. (2025). Unlike these approaches, which construct localized contrasts or define their correspondence structures before optimization, RePair discovers semantic-spatial preference relations directly from naturally occurring variations among imperfect stochastic generations. By leveraging these intrinsic differences, RePair extracts multiple valid regional comparisons from the same generation pair without requiring manually designed perturbations or predefined local correspondences.

3 RePair: Regional Preference Learning from Paired Imperfect Generations

This section presents the proposed **RePair** method. We begin by reviewing the standard global preference formulation (Sec. 3.1). We then detail the two stages of our approach: (1) discovering hidden regional preference relations from paired imperfect generations (Sec. 3.2), and (2) optimizing these decomposed relations without enforcing a global winner (Sec. 3.3).

3.1 Preliminaries: Preference Optimization in Flow Models

Modern text-to-image architectures Labs et al. (2025); Wu et al. (2025a) widely adopt rectified flow Liu et al. (2023); Lipman et al. (2023), which transports noise $z_0 \sim \mathcal{N}(0, I)$ to data x via the trajectory $z_t = (1-t)z_0 + tx$. The model learns a velocity field v_θ to match the target velocity $(x - z_0)$. For brevity, we denote the squared velocity matching error for a sample x conditioned on prompt c as

$$e_\theta(x, c, t, z_0) = \|v_\theta(z_t, t, c) - (x - z_0)\|_2^2. \tag{1}$$

Originally introduced for diffusion models Wallace et al. (2024), direct preference optimization Rafailov et al. (2023) (DPO) is naturally extended to rectified flow models to align them using pairwise preference data $x_w \succ x_l$. The implicit relative penalty for a sample x is defined as the difference in matching errors between the policy v_θ and a reference model v_{ref}:

$$s(x, c, t, \theta) = e_\theta(x, c, t, z_0) - e_{\text{ref}}(x, c, t, z_0). \tag{2}$$

The policy is then optimized to reduce the matching error of the preferred sample relative to the dispreferred one:

$$\mathcal{L}_{\text{Flow-DPO}} = -\mathbb{E}_{t, z_0} \left[\log \sigma \left(-\beta \left(s(x_w, c, t, \theta) - s(x_l, c, t, \theta) \right) \right) \right]. \tag{3}$$

This standard formulation fundamentally assumes a single, global preference relation $x_w \succ x_l$. In contrast, RePair replaces this monolithic assumption by considering a pair of imperfect stochastic generations (I_A, I_B), decomposing their supervision into a set of localized relations $\mathcal{P}(I_A, I_B) = \{(R_k, y_k)\}_{k=1}^K$. Here, R_k denotes a grounded region and $y_k \in \{+1, -1\}$ indicates its specific preference direction.

3.2 Discovering Hidden Regional Preferences in Imperfect Generations.

As illustrated in Fig. 3, our data construction consists of two stages: we first generate a pool of structurally comparable yet locally diverse imperfect samples, and then mine verified regional preferences from their cross-generation differences. The resulting data directly capture which generation better realizes each semantic requirement, without requiring a synthetically corrected counterpart.

Algorithm 1 Independent Candidate Generation versus Correlated Candidate Generation.

(a) Independent Candidate Generation

Require: Generator $\mathcal{G}$, condition c, pool size M
Initialize candidate pool $\mathcal{C} \leftarrow \varnothing$
for $i = 1, \ldots, M$ **do**
  Sample $\xi_i \sim \mathcal{N}(0, \mathbf{I})$
  $I_i \leftarrow \mathcal{G}(c; \xi_i)$
  $\mathcal{C} \leftarrow \mathcal{C} \cup \{I_i\}$
end for
return $\mathcal{C}$

(b) Correlated Candidate Generation

Require: Generator $\mathcal{G}$, condition c, pool size M, coefficient ρ
Initialize candidate pool $\mathcal{C} \leftarrow \varnothing$
Sample shared noise $\xi_0 \sim \mathcal{N}(0, \mathbf{I})$
for $i = 1, \ldots, M$ **do**
  Sample residual $\eta_i \sim \mathcal{N}(0, \mathbf{I})$
  $\xi_i \leftarrow \rho\xi_0 + \sqrt{1-\rho^2}\eta_i$
  $I_i \leftarrow \mathcal{G}(c; \xi_i)$
  $\mathcal{C} \leftarrow \mathcal{C} \cup \{I_i\}$
end for
return $\mathcal{C}$

Correlated Candidate Generation The central challenge in discovering preferences from stochastic generations is the local attribution problem. Independent sampling drastically alters global layout, entangling semantic correctness with irrelevant structural variations and rendering regional attribution unreliable. To restrict this global variance, we introduce *Correlated Candidate Generation* (Algorithm 1 b). Rather than using completely independent noise, each candidate ξ_i blends a shared base Gaussian ξ_0 with an independent residual η_i, governed by a correlation coefficient ρ. This aligns the global layout while preserving natural local diversity, yielding *comparable imperfections* where local successes and failures can be precisely isolated and contrasted.

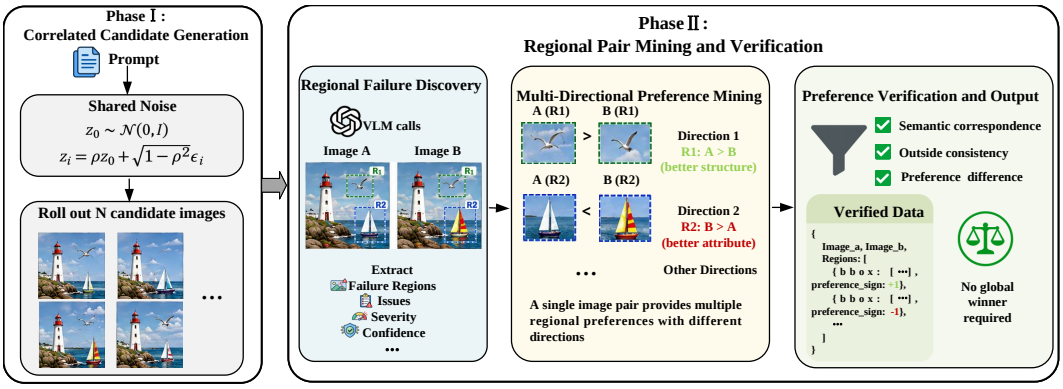

Figure 3: **RePair Data pipeline.** We generate locally alignable candidate images via correlated generation, and leverage VLMs to mine and verify multi-directional regional preferences.

Regional Preference Mining and Verification. Throughout the entire process, we use GPT-5.6 Sol OpenAI (2026) as the visual annotator. Given the candidate pool $\mathcal{C} = \{I_i\}_{i=1}^{M}$, extracting multi-directional preferences proceeds through localization, attribution, and verification. **Localization** grounds semantic failures to specific regions, extracting a corresponding spatial mask R_k. **Attribution** evaluates candidate pairs (I_A, I_B) across matched regions, assigning a preference direction $y_k \in \{+1, -1\}$ to the generation with superior local realization. **Verification** then ensures the overall reliability of the extracted relations. To prevent the comparative signal from being corrupted by unintended changes outside the selected regions, we enforce an *outside-consistency* constraint based on spatially corresponding visual features:

$$S_{\text{out}}(I_A, I_B) = \frac{1}{|\Omega_{\text{out}}|} \sum_{p \in \Omega_{\text{out}}} \cos\left(\Phi_p(I_A), \Phi_p(I_B)\right) \geq \tau, \tag{4}$$

where $\Phi_p(\cdot)$ denotes the DINOv2-Base Oquab et al. (2024) feature at patch p, Ω_{out} denotes the set of spatially corresponding patches outside all selected preference regions, and τ is the consistency threshold. Beyond this structural constraint, candidates undergo a coarse-to-fine filtering pipeline to eliminate ambiguous semantic comparisons. The completely verified output yields the decomposed set of regional relations defined in Sec. 3.1:

$$\mathcal{P}(I_A, I_B) = \{(R_k, y_k)\}_{k=1}^{K}, \qquad y_k = \begin{cases} +1, & I_A \succ_{R_k} I_B, \\ -1, & I_B \succ_{R_k} I_A, \end{cases} \tag{5}$$

This representation preserves all reliable local comparisons and naturally accommodates diverse directional preferences within a single pair.

3.3 Learning from Multi-Directional Regional Preferences

The image-level formulation in Sec. 3.1 assigns a single, global preference relation to each image pair. Our mined supervision is intrinsically decomposed: a single pair provides a set of valid relations, potentially with complementary preference directions. As illustrated in Fig. 4, we localize the underlying flow-matching signal to each verified region and optimize these relations independently.

Preference Scoring over Semantic Regions. Let $e_\theta(p; I_i, c, t, z_0)$ denote the flow-matching error at spatial token p for image I_i, where $i \in \{A, B\}$. Rather than reducing this error over the entire latent representation, we aggregate it within the support of regional relation k. Given a soft region mask $M_i^{(k)}$, the region-normalized preference score is as follows:

$$s_{\theta,i}^{(k)} = -\frac{\sum_p M_i^{(k)}(p)\, e_\theta(p; I_i, c, t, z_0)}{\sum_p M_i^{(k)}(p) + \varepsilon}, \qquad i \in \{A, B\}. \tag{6}$$

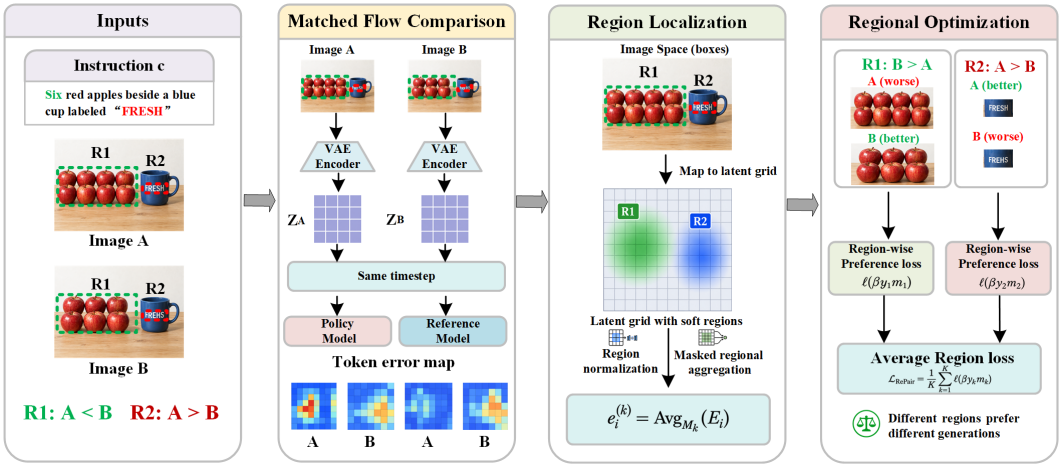

Figure 4: **Overview of RePair optimization method.** Token-level error maps are extracted, spatially localized to specific regions, and normalized. Each regional relation is then optimized independently to preserve multi-directional preferences.

The reference score $s^{(k)}_{\text{ref},i}$ is computed analogously. This spatial normalization ensures the preference signal is invariant to region size, preventing larger objects from dominating the optimization. Both images are evaluated under identical timesteps and noise realizations to eliminate confounding stochasticity. We define the reference-corrected regional score as $a^{(k)}_i = s^{(k)}_{\theta,i} - s^{(k)}_{\text{ref},i}$, and the regional preference margin as $m_k = a^{(k)}_A - a^{(k)}_B$. Since larger regional scores correspond to lower flow-matching errors, $m_k > 0$ favors I_A and $m_k < 0$ favors I_B. Thus, the underlying flow-matching comparison is preserved, while its unit of supervision is refined from the whole generation to an individual semantic-spatial region.

Region-wise Preference Optimization For relation k, we encode its preference direction by $y_k \in \{-1, +1\}$, where $y_k = +1$ denotes $I_A \succ_{R_k} I_B$ and $y_k = -1$ denotes $I_B \succ_{R_k} I_A$. The signed margin $y_k m_k$ therefore provides a unified orientation even when different regions of the same pair favor different generations. Rather than aggregating regional margins into a single image-level comparison, RePair applies the preference objective to each relation independently:

$$\mathcal{L}_{\text{RePair}} = \mathbb{E}_{\mathcal{P},t,z_0}\left[\frac{w(t)}{K}\sum_{k=1}^{K} -\log\sigma(\beta y_k m_k)\right] \tag{7}$$

where $\mathcal{P}$ denotes the regional preference set, K is the number of verified relations, β controls the preference strength, and $w(t)$ denotes the timestep weighting. Summing the regional margins before the loss computation would lead to gradient cancellation, where local successes and failures within the same image pair neutralize each other. By computing the objective for each relation independently, RePair explicitly preserves the exact optimization direction for every semantic requirement. A formal analysis of regional preference decomposition is provided in Appendix B.

4 Experiments

4.1 Experimental Setup

Implementation Details. Following Sec. 3.2, we construct the preference dataset by mining regional relations from imperfect stochastic generations. Detailed statistics of the mined preference structures are provided in Appendix D. For the two evaluated text-to-image models, the resulting preference datasets contain approximately 5K image pairs with verified regional preference relations. Detailed training configurations are provided in Appendix C.1.

Evaluation Protocol. We evaluate RePair on two representative text-to-image models with different conditioning interfaces, including Ideogram 4 AI (2026) with structured JSON prompts and

Table 1: **Main results on Ideogram 4.** We evaluate four capabilities following the official evaluation protocol. Layout control with 7Bench Izzo et al. (2025) mIoU, spatial reasoning with Spatial-GenEval Wang et al. (2026) spatial accuracy, prompt alignment with PRISM Fang et al. (2025), and text rendering with X-Omni Geng et al. (2025) OCR.

Method	Layout Control	Spatial Reasoning	Prompt Alignment	Text Rendering
Base	0.69	0.84	0.92	0.97
SFT	0.68	0.81	0.90	0.95
Vanilla DPO	0.70	0.83	0.89	0.96
RePair	**0.72**	**0.85**	**0.93**	**0.98**

Table 2: **Main results on FLUX.2-klein-base-9B.** RePair consistently improves compositional text-to-image generation across three standard benchmarks.

GenEval

Method	Single Obj.	Two Obj.	Counting	Colors	Position	Color Attr.	Overall
Base	0.99	0.92	0.82	0.90	0.65	0.67	0.83
SFT	0.99	0.93	0.83	0.89	0.63	0.68	0.83
Vanilla DPO	0.99	0.94	0.84	0.92	0.66	0.67	0.84
RePair	**1.00**	**0.96**	**0.85**	**0.93**	**0.78**	**0.72**	**0.87**

DPG-Bench

Method	Global	Entity	Attribute	Relation	Other	Overall
Base	83.51	91.49	89.67	94.33	85.30	85.41
SFT	84.19	91.46	89.58	94.32	86.20	85.47
Vanilla DPO	83.74	91.50	89.65	94.33	86.50	85.55
RePair	83.53	**91.90**	**89.97**	**94.38**	**88.60**	**86.29**

OneIG-Bench EN

Method	Alignment	Text	Reasoning	Style	Diversity	Overall
Base	0.867	0.909	0.302	0.445	0.212	0.547
SFT	0.871	0.921	0.311	0.439	0.181	0.544
Vanilla DPO	0.868	0.921	0.305	0.444	0.210	0.549
RePair	**0.917**	**0.983**	**0.312**	**0.453**	**0.216**	**0.576**

FLUX.2-klein-base-9B Labs (2025) with natural-language prompts. For Ideogram 4, we follow its official structured-generation evaluation protocol. Specifically, we use 7Bench Izzo et al. (2025) mIoU for layout control, SpatialGenEval Wang et al. (2026) spatial accuracy for spatial reasoning, PRISM-Bench Fang et al. (2025) for prompt alignment, and X-Omni OCR Geng et al. (2025) for text rendering. For FLUX.2-klein-base-9B, we use standard natural-language T2I benchmarks, including GenEval Ghosh et al. (2023), DPG-Bench Hu et al. (2024), and OneIG-Bench EN Chang et al. (2025). Ideogram 4 strictly requires structured JSON inputs. We exclude it from standard benchmarks because translating natural language prompts would require an external LLM, confounding the evaluation with external reasoning capabilities and compromising fair comparisons.

4.2 Main Results

We evaluate the effectiveness of RePair on two distinct architectures, comparing it against the Base, SFT, and Vanilla DPO baselines. As shown in Table 2, RePair consistently enhances the FLUX.2-klein-base-9B model across standard text-based benchmarks. Specifically, it increases the overall scores on GenEval from 0.83 to 0.87, on DPG-Bench from 85.41 to 86.29, and on OneIG-Bench EN from 0.547 to 0.576. Similarly, for Ideogram 4 evaluated under structured JSON conditioning (Table 1), RePair comprehensively improves all assessed capabilities, including layout control, spatial reasoning, prompt alignment, and text rendering. Overall, the substantial gains across both architectures demonstrate that **multi-directional regional supervision provides highly effective signals for enforcing localized semantic constraints and resolving complex object relationships.**

Table 3: **Comparison with representative text-to-image models.** All results are evaluated under their corresponding official benchmark protocols.

Model	GenEval ↑	DPG-Bench ↑	OneIG-Bench EN ↑
OmniGen2 Wu et al. (2025b)	**0.95**	83.57	0.475
Qwen-Image Wu et al. (2025a)	0.87	**88.32**	0.539
FLUX.1-dev Labs (2024)	0.66	83.84	0.434
Z-Image Team (2025)	0.84	88.14	0.546
FLUX.2-klein-base-9B (baseline) Labs (2025)	0.83	85.41	0.547
+ RePair	0.87	86.29	**0.576**

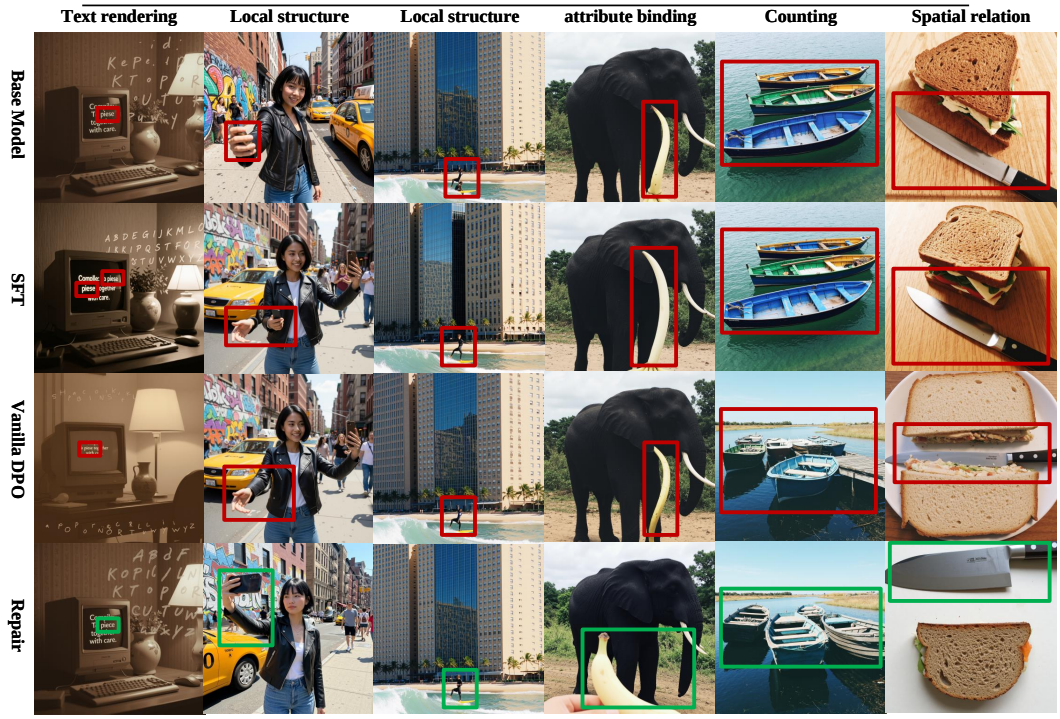

Figure 5: **Qualitative comparison of different methods.** RePair better resolves diverse localized errors, including text rendering, local structure, attribute binding, counting, and spatial relations.

As shown in Table 3, we compare FLUX.2-klein-base-9B aligned via RePair against representative T2I models, including OmniGen2 Wu et al. (2025b), Qwen-Image Wu et al. (2025a), FLUX.1-dev Labs (2024), and Z-Image Team (2025). Beyond consistently outperforming its base model across three benchmarks, RePair achieves the highest score on OneIG-Bench EN among all evaluated models. Furthermore, it remains highly competitive with stronger systems such as Qwen-Image and OmniGen2, underscoring the efficiency of our regional preference optimization. Furthermore, we additionally evaluate RePair on the official HPS v2.1 benchmark. As reported in Appendix A.3, RePair consistently improves HPS v2.1 scores across both FLUX.2 and Ideogram 4, indicating that the proposed regional preference optimization does not compromise general visual quality.

Qualitative Results. Figure 5 presents qualitative comparisons among Base, SFT, Vanilla DPO, and RePair on FLUX.2-klein-base-9B. Compared with the baselines, RePair consistently resolves localized generation failures while preserving the overall scene. **The improvements span a broad range of compositional requirements, including text rendering, structural consistency, attribute binding, counting, and spatial relations.** These observations are consistent with our quantitative results, showing that multi-directional regional supervision provides more precise optimization signals for satisfying complex constraints.

4.3 Ablation Studies

We conduct ablation studies on FLUX.2-klein-base-9B Labs (2025) using a stratified subset of OneIG-Bench EN Chang et al. (2025). Specifically, we sample 20 prompts from each of its five

Table 4: **Ablation of core design components of RePair.** All variants use matched numbers of regional preference relations for downstream training.

Variant	Alignment	Text	Reasoning	Style	Diversity	Overall
w/o Correlated Generation	0.855	0.805	0.294	0.425	**0.208**	0.517
w/o Consistency Verification	0.879	0.862	0.304	0.429	0.199	0.535
Uni-directional Only	0.884	0.850	0.306	0.424	0.199	0.533
Full RePair	**0.880**	**0.879**	**0.311**	**0.438**	0.200	**0.542**

Table 5: **Analysis of regional supervision design choices.** We study the effect of the number of regional preferences per pair and regional loss weighting.

Number of Regional Preferences						
Accumulated Regions ($1 \sim K$)	Alignment	Text	Reasoning	Style	Diversity	Overall
1	**0.886**	0.843	0.302	0.418	0.199	0.530
2	0.880	0.845	0.304	0.423	0.198	0.530
3	0.880	0.848	0.305	0.430	0.200	0.533
4	0.882	0.852	0.304	0.428	**0.201**	0.534
5	0.882	**0.879**	**0.311**	**0.438**	0.200	**0.542**
Region Weighting						
Weighting	Alignment	Text	Reasoning	Style	Diversity	Overall
Preference-Strength	**0.881**	0.856	0.305	0.418	**0.201**	0.536
Uniform	0.880	**0.879**	**0.311**	**0.438**	0.200	**0.542**

evaluation subsets, corresponding to Alignment, Text, Reasoning, Style, and Diversity, resulting in 100 prompts in total. Unless specified, all variants isolate the studied factor while keeping the setup fixed.

Regional Preference Construction. As shown in Table 4, we analyze the key components of our data construction and supervision pipeline. Removing correlated generation causes the severest performance drop by introducing noisy global variations from independent sampling. Similarly, omitting consistency verification degrades the overall score to 0.535 by allowing excessive structural differences outside the target regions to corrupt the local comparative signal. Finally, restricting the training data to *Uni-directional Only* samples, which contain only a single preference direction per image pair, reduces the score to 0.533. These results confirm that **reliable regional preference learning requires strictly isolated local contrasts and the complementary signals inherent in multi-directional supervision.**

Regional Supervision Configuration. Table 5 analyzes the use of regional supervision during training. Using more verified regional relations improves the overall score from 0.530 to 0.542, while uniform weighting outperforms preference-strength weighting (0.542 vs. 0.536). These results indicate that **effective regional supervision benefits from retaining multiple complementary relations and optimizing them with uniform weighting.**

5 Conclusion

In this work, we introduced RePair, a framework that transforms the stochastic diversity of text-to-image generation into fine-grained preference supervision. Unlike existing approaches that rely on manually constructed contrasts with a unidirectional preference structure, RePair discovers naturally occurring semantic-spatial relations from imperfect generations. By combining correlated candidate generation, visual grounding, and a region-wise preference objective, RePair enables a single image pair to provide multiple verified regional preferences with different directions. Extensive experiments on FLUX.2 and Ideogram 4 demonstrate that RePair achieves consistent improvements in compositional alignment using only 5K automatically mined preference pairs. We discuss the limitations and future work of our method in Appendix F.

AI USE STATEMENT

In this work, we used generative AI tools for polishing writing and assisting with data construction, including improving the clarity, readability, and grammatical correctness of the manuscript, as well as supporting the automatic construction and organization of data used in our experiments.

We have not used generative AI tools for generating research ideas, designing the core methodology, deriving theoretical results, conducting experimental analysis, or making scientific claims. All other tasks requiring disclosure that are not explicitly mentioned above are not applicable to this work.

Additionally, we used generative AI tools for recommended disclosure tasks, including improving the presentation quality of the manuscript. All AI-assisted outputs were carefully reviewed and verified by the authors. In particular, AI-assisted writing was manually checked and revised to ensure that it accurately reflects the authors' intended contributions, and automatically constructed data were inspected and validated before being used in experiments. The authors take full responsibility for the final content of this work, including all text, claims, experimental results, and artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

We provide extensive detail to support the reproducibility of RePair. The regional preference mining pipeline, corresponding to Sec. 3.2, is described at the implementation level in Appendix C.2, including correlated candidate generation, regional failure discovery, regional matching and pair construction, preference verification, and outside-region consistency filtering, together with the full region-matching and filtering criteria given in Table 11. The complete training configuration, including the base models, LoRA setup, optimizer, learning-rate schedule, and hardware setup, is provided in Appendix C.1. The evaluation protocol, including all benchmarks and metrics used for FLUX.2-klein-base-9B and Ideogram 4, is described in Sec. 4.1; all evaluation benchmarks (GenEval, DPG-Bench, OneIG-Bench, 7Bench, SpatialGenEval, PRISM-Bench, X-Omni OCR, and HPS v2.1) are publicly available. The configurations of all ablation variants are detailed in Appendix C.3. Statistics of the mined regional preference datasets, including the number of verified preference pairs and the distribution of single- versus multi-directional relations, are reported in Appendix D, as summarized in Table 12. A formal analysis of the regional preference decomposition and its optimization behavior, together with the full proof, is provided in Appendix B. We will publicly release RePair to support reproducibility and future research, including the mined regional preference datasets RePair-FLUX and RePair-Ideogram (approximately 5K pairs each), the full data construction and mining codebase, and the training and evaluation code used to produce all reported results.

REFERENCES

Ideogram AI. Ideogram 4. https://ideogram.ai/blog/ideogram-4.0/, 2026.

Anthropic. Claude sonnet 5 system card. https://www.anthropic.com/system-cards, 2026. Accessed: 2026-09-26.

Kevin Black, Michael Janner, Yilun Du, Ilya Kostrikov, and Sergey Levine. Training diffusion models with reinforcement learning, 2024.

Umberto Borso, Davide Paglieri, Jude Wells, and Tim Rocktäschel. Preference-based alignment of discrete diffusion models, 2025.

Jingjing Chang, Yixiao Fang, Peng Xing, Shuhan Wu, Wei Cheng, Rui Wang, Xianfang Zeng, Gang Yu, and Hai-Bao Chen. Oneig-bench: Omni-dimensional nuanced evaluation for image generation. *arXiv preprint arxiv:2506.07977*, 2025.

Paul F Christiano, Jan Leike, Tom Brown, Miljan Martic, Shane Legg, and Dario Amodei. Deep reinforcement learning from human preferences. In *Advances in Neural Information Processing Systems*, volume 30, 2017.

Kevin Clark, Paul Vicol, Kevin Swersky, and David J. Fleet. Directly fine-tuning diffusion models on differentiable rewards. In *The Twelfth International Conference on Learning Representations*, 2024.

Rongyao Fang, Aldrich Yu, Chengqi Duan, Linjiang Huang, Shuai Bai, Yuxuan Cai, Kun Wang, Si Liu, Xihui Liu, and Hongsheng Li. Flux-reason-6m & prism-bench: A million-scale text-to-image reasoning dataset and comprehensive benchmark. *arXiv preprint arXiv:2509.09680*, 2025.

Xuelu Feng, Yunsheng Li, Ziyu Wan, Zixuan Gao, Junsong Yuan, Dongdong Chen, and Chunming Qiao. Rubricrl: Simple generalizable rewards for text-to-image generation, 2026.

Minghao Fu, Guo-Hua Wang, Tianyu Cui, Qing-Guo Chen, Zhao Xu, Weihua Luo, and Kaifu Zhang. Diffusion-sdpo: Safeguarded direct preference optimization for diffusion models, 2025.

Zigang Geng, Yibing Wang, Yeyao Ma, Chen Li, Yongming Rao, Shuyang Gu, Zhao Zhong, Qinglin Lu, Han Hu, Xiaosong Zhang, Linus, Di Wang, and Jie Jiang. X-omni: Reinforcement learning makes discrete autoregressive image generative models great again. *CoRR*, abs/2507.22058, 2025.

Dhruba Ghosh, Hanna Hajishirzi, and Ludwig Schmidt. Geneval: An object-focused framework for evaluating text-to-image alignment. *arXiv preprint arXiv:2310.11513*, 2023.

Yuxiang Guo, Jiang Liu, Ze Wang, Hao Chen, Ximeng Sun, Yang Zhao, Jialian Wu, Xiaodong Yu, Zicheng Liu, and Emad Barsoum. Imagedoctor: Diagnosing text-to-image generation via grounded image reasoning, 2025.

Xiwei Hu, Rui Wang, Yixiao Fang, Bin Fu, Pei Cheng, and Gang Yu. Ella: Equip diffusion models with llm for enhanced semantic alignment. *arXiv preprint arXiv:2403.05135*, 2024.

Yushi Hu, Benlin Liu, Jungo Kasai, Yizhong Wang, Mari Ostendorf, Ranjay Krishna, and Noah A. Smith. Tifa: Accurate and interpretable text-to-image faithfulness evaluation with question answering. *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 20349–20360, 2023.

Zijing Hu, Fengda Zhang, and Kun Kuang. D-fusion: Direct preference optimization for aligning diffusion models with visually consistent samples. In *Forty-second International Conference on Machine Learning*, 2025.

Zitong Huang, Kaidong Zhang, Yukang Ding, Chao Gao, Rui Ding, Ying Chen, and Wangmeng Zuo. Mind the generative details: Direct localized detail preference optimization for video diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 35998–36008, June 2026.

Elena Izzo, Luca Parolari, Davide Vezzaro, and Lamberto Ballan. 7bench: a comprehensive benchmark for layout-guided text-to-image models. *arXiv preprint arXiv:2508.12919*, 2025.

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. In *Advances in Neural Information Processing Systems*, volume 36, pp. 36652–36663, 2023.

Black Forest Labs. Flux. `https://github.com/black-forest-labs/flux`, 2024.

Black Forest Labs. FLUX.2: Frontier Visual Intelligence. `https://bfl.ai/blog/flux-2`, 2025.

Black Forest Labs, Stephen Batifol, Andreas Blattmann, Frederic Boesel, Saksham Consul, Cyril Diagne, Tim Dockhorn, Jack English, Zion English, Patrick Esser, Sumith Kulal, Kyle Lacey, Yam Levi, Cheng Li, Dominik Lorenz, Jonas Müller, Dustin Podell, Robin Rombach, Harry Saini, Axel Sauer, and Luke Smith. Flux.1 kontext: Flow matching for in-context image generation and editing in latent space, 2025.

Zejian Li, Yize Li, Chen Meng, Zhongni Liu, Ling Yang, Shengyuan Zhang, Guang Yang, Changyuan Yang, Zhiyuan Yang, and Lingyun Sun. Inversion-dpo: Precise and efficient post-training for diffusion models. *Proceedings of the 33rd ACM International Conference on Multimedia*, 2025.

Youwei Liang, Junfeng He, Gang Li, Peizhao Li, Arseniy Klimovskiy, Nicholas Carolan, Jiao Sun, Jordi Pont-Tuset, Sarah Young, Feng Yang, Junjie Ke, Krishnamurthy Dj Dvijotham, Katie Collins, Yiwen Luo, Yang Li, Kai J Kohlhoff, Deepak Ramachandran, and Vidhya Navalpakkam. Rich human feedback for text-to-image generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024.

Zhiqiu Lin, Deepak Pathak, Baiqi Li, Jiayao Li, Xide Xia, Graham Neubig, Pengchuan Zhang, and Deva Ramanan. Evaluating text-to-visual generation with image-to-text generation. In *European Conference on Computer Vision*, 2024.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matthew Le. Flow matching for generative modeling. In *The Eleventh International Conference on Learning Representations*, 2023.

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 *Advances in Neural Information Processing Systems*, volume 38, Main Conference, pp. 40783–40818, 2025.

Xingchao Liu, Chengyue Gong, and qiang liu. Flow straight and fast: Learning to generate and transfer data with rectified flow. In *The Eleventh International Conference on Learning Representations*, 2023.

Zhuohan Liu, Wujian Peng, Yitong Chen, and Zuxuan Wu. Compositional text-to-image generation via region-aware bimodal direct preference optimization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 36604–36614, June 2026.

Yihong Luo, Tianyang Hu, and Jing Tang. Reinforcing diffusion models by direct group preference optimization. In *The Fourteenth International Conference on Learning Representations*, 2026.

Yoonjin Oh, Yongjin Kim, Hyomin Kim, Donghwan Chi, and Sungwoong Kim. Ospo: Object-centric self-improving preference optimization for text-to-image generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7620–7631, June 2026.

OpenAI. Gpt-5.6 sol. `https://developers.openai.com/api/docs/models/gpt-5.6-sol`, 2026. Accessed: 2026-09-26.

Maxime Oquab, Timothée Darcet, Théo Moutakanni, Huy Vo, Marc Szafraniec, Vasil Khalidov, Pierre Fernandez, Daniel Haziza, Francisco Massa, Alaaeldin El-Nouby, Mahmoud Assran, Nicolas Ballas, Wojciech Galuba, Russell Howes, Po-Yao Huang, Shang-Wen Li, Ishan Misra, Michael Rabbat, Vasu Sharma, Gabriel Synnaeve, Hu Xu, Hervé Jegou, Julien Mairal, Patrick Labatut, Armand Joulin, and Piotr Bojanowski. Dinov2: Learning robust visual features without supervision. *arXiv preprint arxiv:2304.07193*, 2024.

Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul F Christiano, Jan Leike, and Ryan Lowe. Training language models to follow instructions with human feedback. In *Advances in Neural Information Processing Systems*, volume 35, pp. 27730–27744, 2022.

Kaihang Pan, Wendong Bu, Yuruo Wu, Yang Wu, Kai Shen, Yunfei Li, Hang Zhao, Juncheng Li, Siliang Tang, and Yueting Zhuang. Focusdiff: Advancing fine-grained text-image alignment for autoregressive visual generation through rl, 2025.

Mihir Prabhudesai, Anirudh Goyal, Deepak Pathak, and Katerina Fragkiadaki. Aligning text-to-image diffusion models with reward backpropagation, 2024.

Qwen Team. Qwen3.8-Max: A new bar for coding and cowork, August 2026. URL `https://qwen.ai/blog?id=qwen3.8`. Accessed: 2026-09-26.

Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D Manning, Stefano Ermon, and Chelsea Finn. Direct preference optimization: Your language model is secretly a reward model. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023.

Xincheng Shuai, Ziye Li, Henghui Ding, and Dacheng Tao. Glyphprinter: Region-grouped direct preference optimization for glyph-accurate visual text rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7674–7683, June 2026.

Jiayang Sun, Pin Wang, Hongbo Wang, Xinyue Liu, Huaibo Huang, and Ran He. Towards fine-grained attribution: Instance-aware preference optimization for aligning diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 43155–43164, June 2026.

Z-Image Team. Z-image: An efficient image generation foundation model with single-stream diffusion transformer. *arXiv preprint arXiv:2511.22699*, 2025.

Bram Wallace, Meihua Dang, Rafael Rafailov, Linqi Zhou, Aaron Lou, Senthil Purushwalkam, Stefano Ermon, Caiming Xiong, Shafiq Joty, and Nikhil Naik. Diffusion model alignment using direct preference optimization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8228–8238, June 2024.

Zengbin Wang, Xuecai Hu, Yong Wang, Feng Xiong, Man Zhang, and Xiangxiang Chu. Everything in its place: Benchmarking spatial intelligence of text-to-image models. *arXiv preprint arXiv:2601.20354*, 2026.

Chenfei Wu, Jiahao Li, Jingren Zhou, Junyang Lin, Kaiyuan Gao, Kun Yan, Sheng ming Yin, Shuai Bai, Xiao Xu, Yilei Chen, Yuxiang Chen, Zecheng Tang, Zekai Zhang, Zhengyi Wang, An Yang, Bowen Yu, Chen Cheng, Dayiheng Liu, Deqing Li, Hang Zhang, Hao Meng, Hu Wei, Jingyuan Ni, Kai Chen, Kuan Cao, Liang Peng, Lin Qu, Minggang Wu, Peng Wang, Shuting Yu, Tingkun Wen, Wensen Feng, Xiaoxiao Xu, Yi Wang, Yichang Zhang, Yongqiang Zhu, Yujia Wu, Yuxuan Cai, and Zenan Liu. Qwen-image technical report. *arXiv preprint arXiv:2508.02324*, 2025a.

Chenyuan Wu, Pengfei Zheng, Ruiran Yan, Shitao Xiao, Xin Luo, Yueze Wang, Wanli Li, Xiyan Jiang, Yexin Liu, Junjie Zhou, Ze Liu, Ziyi Xia, Chaofan Li, Haoge Deng, Jiahao Wang, Kun Luo, Bo Zhang, Defu Lian, Xinlong Wang, Zhongyuan Wang, Tiejun Huang, and Zheng Liu. Omnigen2: Exploration to advanced multimodal generation. *arXiv preprint arXiv:2506.18871*, 2025b.

Xiaoshi Wu, Yiming Hao, Keqiang Sun, Yixiong Chen, Feng Zhu, Rui Zhao, and Hongsheng Li. Human preference score v2: A solid benchmark for evaluating human preferences of text-to-image synthesis. *arXiv preprint arXiv:2306.09341*, 2023.

Ziyi Wu, Anil Kag, Ivan Skorokhodov, Willi Menapace, Ashkan Mirzaei, Igor Gilitschenski, Sergey Tulyakov, and Aliaksandr Siarohin. DenseDPO: Fine-grained temporal preference optimization for video diffusion models. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025c.

Xiaoying Xing, Avinab Saha, Junfeng He, Susan Hao, Paul Vicol, Moonkyung Ryu, Gang Li, Sahil Singla, Sarah Young, Yinxiao Li, Feng Yang, and Deepak Ramachandran. Focus-n-fix: Region-aware fine-tuning for text-to-image generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 18486–18496, June 2025.

Jiazheng Xu, Xiao Liu, Yuchen Wu, Yuxuan Tong, Qinkai Li, Ming Ding, Jie Tang, and Yuxiao Dong. Imagereward: Learning and evaluating human preferences for text-to-image generation. In *Advances in Neural Information Processing Systems*, volume 36, pp. 15903–15935, 2023.

Jiazheng Xu, Yu Huang, Jiale Cheng, Yuanming Yang, Jiajun Xu, Yuan Wang, Wenbo Duan, Shen Yang, Qunlin Jin, Shurun Li, Jiayan Teng, Zhuoyi Yang, Wendi Zheng, Xiao Liu, Dan Zhang, Ming Ding, Xiaohan Zhang, Xiaotao Gu, Shiyu Huang, Minlie Huang, Jie Tang, and Yuxiao Dong. Visionreward: Fine-grained multi-dimensional human preference learning for image and video generation, 2026.

Yankai Yang, Yancheng Long, Wei Chen, Xingyu Lu, Hongyang Wei, Bin Wen, Fan Yang, Tingting Gao, Han Li, and Shuo Yang. Spatialflow-grpo: Where spatial credit drives image editing, 2026.

Huaisong Zhang, Hao Yu, Yuxuan Zhang, Jiahe Wang, Xinrui Chen, Haoxiang Cao, Feng Lu, Wendong Zhang, Changqian Yu, and Chun Yuan. Where, what, why, and importance: Structured defect grounding for text-to-image feedback, 2026.

Xinchen Zhang, Ling Yang, Guohao Li, YaQi Cai, xie jiake, Yong Tang, Yujiu Yang, Mengdi Wang, and Bin CUI. Itercomp: Iterative composition-aware feedback learning from model gallery for text-to-image generation. In *The Thirteenth International Conference on Learning Representations*, 2025.

Appendix Contents

Additional Experimental Results **15**

A.1 Human Validation of Regional Preferences 15

A.2 Robustness to Different VLMs 15

A.3 Visual Quality Evaluation with HPS v2.1 17

Analysis of Regional Preference Decomposition **17**

Implementation Details **20**

C.1 Training Configuration 20

C.2 Regional Preference Mining Details 20

C.3 Ablation Details 22

Regional Preference Statistics **22**

D.1 Dataset Overview 22

D.2 Preference Structure Distribution 22

Additional Qualitative Examples **23**

Limitations and Future Work **23**

A Additional Experimental Results

A.1 Human Validation of Regional Preferences

To assess the reliability of the automatically mined regional preferences, we conduct a human evaluation on both RePair-FLUX and RePair-Ideogram. To ensure that the evaluation covers preference pairs with different numbers of regional relations, we stratify the samples according to the number of verified regions K. For each $K \in \{1, 2, 3, 4, 5\}$, we randomly sample 20 image pairs, resulting in 100 evaluated pairs and 300 regional relations for each model. For every regional relation, the annotator are presented with the original text prompt, the paired generations, and the corresponding regional bounding boxes, while the automatically generated failure descriptions, severity scores, and preference labels are hidden. Each regional relation is independently evaluated by a human annotator.

We evaluate the mined supervision from two aspects. **Region Validity** measures whether the localized region corresponds to a meaningful semantic component exhibiting a generation difference relevant to the prompt, while **Preference Agreement** measures whether the automatically mined regional preference agrees with the human judgment for that region. As shown in Table 6, RePair-FLUX achieves a region validity of 0.977 and a preference agreement of 0.973, while RePair-Ideogram achieves 0.967 and 0.967, respectively. The consistently high agreement across both models indicates that the proposed mining pipeline can reliably identify meaningful regional differences and assign preference relations that are well aligned with human judgments.

A.2 Robustness to Different VLMs

To examine whether RePair depends on a particular vision-language model for regional preference mining, we conduct a controlled robustness study using GPT-5.6 Sol OpenAI (2026), Claude Sonnet 5 Anthropic (2026), and Qwen3.8-27B Qwen Team (2026) as three different VLM annotators. All

Table 6: **Human validation of automatically mined regional preferences.** We sample 20 preference pairs for each number of regional relations $K \in \{1, 2, 3, 4, 5\}$, resulting in 100 image pairs and 300 evaluated regional relations for each model.

Dataset	# Pairs	# Regions	Region Validity	Preference Agreement
RePair-FLUX	100	300	0.977	0.973
RePair-Ideogram	100	300	0.967	0.967

Table 7: **Reliability of regional preferences mined by different VLMs.** For each VLM, we sample 20 preference pairs for each number of regional relations $K \in \{1, 2, 3, 4, 5\}$, resulting in 100 image pairs and 300 evaluated regional relations.

Mining VLM	# Pairs	# Regions	Region Validity	Preference Agreement
GPT-5.6 Sol OpenAI (2026)	100	300	0.977	0.973
Claude Sonnet 5 Anthropic (2026)	100	300	0.967	0.960
Qwen3.8-27B Qwen Team (2026)	100	300	0.950	0.950

three VLMs operate on the same FLUX.2 candidate generations and independently perform regional failure discovery and preference verification using the same prompts, matching rules, and consistency filters. To control for differences in mining yield, we uniformly sample 1K preference pairs from the data mined by each VLM while approximately matching the distribution of the number of regional relations per pair. We then train three RePair variants using identical model configurations, optimization settings, and training budgets, such that the mining VLM is the only factor that varies.

We first examine the reliability of the resulting regional supervision. Following the same human-validation protocol in Appendix A.1, we stratify the mined preference pairs according to the number of verified regional relations K. For each VLM and each $K \in \{1, 2, 3, 4, 5\}$, we randomly sample 20 image pairs, resulting in 100 evaluated pairs and 300 regional relations per VLM. We then manually assess Region Validity and Preference Agreement for all sampled relations. Table 7 reports the corresponding results.

We further evaluate whether the choice of mining VLM affects downstream preference optimization. Following the stratified evaluation protocol used in our controlled studies, we construct a 100-prompt evaluation set from OneIG-Bench EN by randomly sampling 20 prompts from each of its five evaluation subsets, corresponding to Alignment, Text, Reasoning, Style, and Diversity. The same 100 prompts, generation settings, and evaluation protocol are used for all three variants. We report the five OneIG-Bench evaluation dimensions, including Alignment, Text, Reasoning, Style, and Diversity, together with the overall score in Table 8.

Table 8: **Downstream robustness to different VLMs for regional preference mining.** Each variant is trained on 1K preference pairs mined from the same candidate pool. Evaluation uses a stratified 100-prompt subset of OneIG-Bench EN, obtained by sampling 20 prompts from each of its five evaluation subsets.

Mining VLM	Alignment	Text	Reasoning	Style	Diversity	Overall
GPT-5.6 Sol OpenAI (2026)	0.868	0.875	0.311	0.437	0.206	**0.539**
Claude Sonnet 5 Anthropic (2026)	0.863	0.879	0.309	0.440	0.201	0.538
Qwen3.8-27B Qwen Team (2026)	0.855	0.863	0.313	0.439	0.207	0.535

As shown in Tables 7 and 8, all three VLMs produce regional preferences with high human validity and preference agreement, with both metrics remaining at or above 0.950. The corresponding RePair variants also achieve highly comparable downstream performance, with overall OneIG-Bench scores ranging only from 0.535 to 0.539. These results suggest that current strong VLMs are generally capable of reliably identifying and comparing localized generation differences, making them suitable for regional preference mining in RePair.

A.3 Visual Quality Evaluation with HPS v2.1

To examine whether regional preference optimization compromises the overall visual quality of generated images, we further evaluate the models on the official HPS v2.1 benchmark Wu et al. (2023). Since Ideogram 4 requires structured JSON inputs, we use an external LLM to expand each HPS prompt and convert it into the corresponding JSON format before generation. The same converted prompts are used for both the Ideogram 4 base model and its RePair variant, such that this evaluation serves as a controlled comparison of visual quality rather than a cross-model benchmark. As shown in Table 9, RePair consistently improves HPS v2.1 scores across all four evaluation categories for both FLUX.2 and Ideogram 4. The average score increases from 28.8867 to 29.0613 on FLUX.2 and from 22.3267 to 22.6055 on Ideogram 4. These results indicate that the compositional improvements brought by RePair do not come at the expense of general visual quality.

Table 9: **Evaluation on HPS v2.1.** Higher scores indicate better human preference alignment.

Model	Concept-Art	Photo	Anime	Paintings	Average ↑
FLUX.2	29.0004	27.1512	30.2323	29.1630	28.8867
FLUX.2 + RePair	**29.1474**	**27.3307**	**30.4336**	**29.3334**	**29.0613**
Ideogram 4	21.9925	21.8568	23.9854	21.4722	22.3267
Ideogram 4 + RePair	**22.3424**	**22.0654**	**24.1804**	**21.8337**	**22.6055**

B Analysis of Regional Preference Decomposition

We analyze how image-level and regional preference objectives differ in their optimization behavior when different regions within the same image pair have heterogeneous preference directions. Specifically, when the image-level preference direction disagrees with the preference direction of a local region, image-level supervision introduces a self-gradient component determined by the global label rather than the local preference label.

Let θ denote the model parameters and let $\Delta_k(\theta)$ be the regional preference margin induced by θ. For each region k, define the gradient vector

$$g_k = \nabla_\theta \Delta_k(\theta) \in \mathbb{R}^d. \tag{8}$$

We further define the gradient interaction matrix:

$$G_{ik} = \langle g_i, g_k \rangle = \nabla_\theta \Delta_i(\theta)^\top \nabla_\theta \Delta_k(\theta). \tag{9}$$

We assume positive aggregation weights $\alpha_k > 0$ and regional weights $w_k > 0$, with $\eta > 0$ and $\beta > 0$. For the analyzed region i, we consider the non-degenerate case $g_i \neq 0$, such that $G_{ii} = \|g_i\|^2 > 0$. The diagonal term $G_{ii} = \|g_i\|^2 > 0$ measures the sensitivity of region i to its own update direction, while the off-diagonal terms characterize interactions between regions induced by shared model parameters. Unlike analyses that assume independent regions, we do not require $G_{ik} = 0$ for $i \neq k$.

We compare two preference objectives. The image-level objective is defined as

$$\mathcal{L}_{\text{img}}(\theta) = \ell\left(y_{\text{img}} \sum_{k=1}^{K} \alpha_k \Delta_k(\theta)\right), \tag{10}$$

where $y_{\text{img}} \in \{+1, -1\}$ denotes the global preference direction and $\ell(t) = -\log \sigma(\beta t)$ is the DPO-style loss satisfying $\ell'(t) < 0$.

The regional preference objective is

$$\mathcal{L}_{\text{reg}}(\theta) = \sum_{k=1}^{K} w_k \ell\left(y_k \Delta_k(\theta)\right), \tag{11}$$

where each region is optimized according to its own preference direction.

Both objectives follow gradient flow:

$$\dot{\theta} = -\eta \nabla_\theta \mathcal{L}. \tag{12}$$

The instantaneous change of the regional preference margin is

$$\dot{\Delta}_i = \langle g_i, \dot{\theta} \rangle = -\eta \langle g_i, \nabla_\theta \mathcal{L} \rangle. \tag{13}$$

Proposition 1 (Preference-direction of self-gradient components)**.** *Let $y_{\text{img}} \in \{+1, -1\}$ be the image-level preference direction. For a region i whose regional preference direction disagrees with the global one,*

$$y_i = -y_{\text{img}}, \tag{14}$$

the self-gradient components induced by image-level and regional objectives follow different preference directions.

For image-level supervision, let

$$m_{\text{img}} = \sum_k \alpha_k \Delta_k. \tag{15}$$

The margin update can be decomposed as

$$\dot{\Delta}_i^{\text{img}} = S_i^{\text{img}} + C_i^{\text{img}}, \tag{16}$$

where

$$S_i^{\text{img}} = -\eta \ell'(y_{\text{img}} m_{\text{img}}) y_{\text{img}} \alpha_i G_{ii}, \tag{17}$$

and

$$C_i^{\text{img}} = -\eta \ell'(y_{\text{img}} m_{\text{img}}) y_{\text{img}} \sum_{k \neq i} \alpha_k G_{ik}. \tag{18}$$

Because

$$-\eta \ell'(y_{\text{img}} m_{\text{img}}) \alpha_i G_{ii} > 0, \tag{19}$$

the self term satisfies

$$\text{sign}(S_i^{\text{img}}) = y_{\text{img}}. \tag{20}$$

Therefore, when

$$|S_i^{\text{img}}| > |C_i^{\text{img}}|, \tag{21}$$

the overall update follows the global preference direction and becomes inconsistent with the local regional preference.

For regional supervision, the corresponding decomposition is

$$\dot{\Delta}_i^{\text{reg}} = S_i^{\text{reg}} + C_i^{\text{reg}}, \tag{22}$$

where

$$S_i^{\text{reg}} = -\eta w_i \ell'(y_i \Delta_i) y_i G_{ii}, \tag{23}$$

and

$$C_i^{\text{reg}} = -\eta \sum_{k \neq i} w_k \ell'(y_k \Delta_k) y_k G_{ik}. \tag{24}$$

Similarly,

$$-\eta w_i \ell'(y_i \Delta_i) G_{ii} > 0, \tag{25}$$

which gives

$$\text{sign}(S_i^{\text{reg}}) = y_i. \tag{26}$$

Hence, when

$$|S_i^{\text{reg}}| > |C_i^{\text{reg}}|, \tag{27}$$

the regional update remains consistent with its own preference direction.

Proof. For the image-level objective, applying the chain rule gives

$$\nabla_\theta \mathcal{L}_{\text{img}} = \ell'(y_{\text{img}} m_{\text{img}}) y_{\text{img}} \sum_{k=1}^{K} \alpha_k g_k. \tag{28}$$

Taking the inner product with g_i:

$$\langle g_i, \nabla_\theta \mathcal{L}_{\text{img}} \rangle = \ell'(y_{\text{img}} m_{\text{img}}) y_{\text{img}} \sum_{k} \alpha_k G_{ik} \tag{29}$$

$$= \ell'(y_{\text{img}} m_{\text{img}}) y_{\text{img}} \alpha_i G_{ii} \tag{30}$$

$$+ \ell'(y_{\text{img}} m_{\text{img}}) y_{\text{img}} \sum_{k \neq i} \alpha_k G_{ik}. \tag{31}$$

Since

$$\dot{\Delta}_i^{\text{img}} = -\eta \langle g_i, \nabla_\theta \mathcal{L}_{\text{img}} \rangle,$$

the two terms correspond to S_i^{img} and C_i^{img}.

Because $\ell'(t) < 0$ for all t and $G_{ii} > 0$, we have

$$-\eta \ell'(y_{\text{img}} m_{\text{img}}) \alpha_i G_{ii} > 0.$$

Therefore,

$$\operatorname{sign}(S_i^{\text{img}}) = y_{\text{img}}.$$

When $y_i = -y_{\text{img}}$, the image-level self-gradient component is inconsistent with the local preference direction.

This conclusion does not require $m_{\text{img}} \neq 0$, because

$$\ell'(0) = -\beta/2 < 0.$$

For the regional objective,

$$\nabla_\theta \mathcal{L}_{\text{reg}} = \sum_{k=1}^{K} w_k \ell'(y_k \Delta_k) y_k g_k. \tag{32}$$

Taking the inner product with g_i gives

$$\langle g_i, \nabla_\theta \mathcal{L}_{\text{reg}} \rangle = w_i \ell'(y_i \Delta_i) y_i G_{ii} \tag{33}$$

$$+ \sum_{k \neq i} w_k \ell'(y_k \Delta_k) y_k G_{ik}. \tag{34}$$

Therefore,

$$\dot{\Delta}_i^{\text{reg}} = -\eta w_i \ell'(y_i \Delta_i) y_i G_{ii} - \eta \sum_{k \neq i} w_k \ell'(y_k \Delta_k) y_k G_{ik}. \tag{35}$$

The first term corresponds to S_i^{reg}. Because

$$-\eta w_i \ell'(y_i \Delta_i) G_{ii} > 0,$$

its sign is determined by y_i. Therefore, the regional self-gradient component is aligned with the local preference direction. □

Discussion of gradient interactions. The proposition does not require independent regional parameters or orthogonal gradients. Interactions between regions are explicitly captured by the off-diagonal terms G_{ik}. The dominance conditions for image-level and regional objectives involve different cross-region terms. The image-level cross term is uniformly modulated by the global preference direction y_{img}, whereas the regional cross term is determined by the individual regional directions y_k. The key distinction between the two objectives lies in the self-gradient component: image-level supervision injects a self-gradient component whose sign is determined by y_{img}, regardless of the local label y_i. In contrast, regional supervision assigns each region a self-gradient component aligned with its own preference direction y_i. Although cross-region interactions caused by shared parameters remain unavoidable, regional preference optimization preserves the correct local directional signal before these interactions are introduced.

Implication. Image-level preference can be viewed as a restricted case in which all regions share a consistent preference direction. When regional preferences are heterogeneous, image-level scalarization compresses distinct semantic-spatial preference directions into a single global comparison. The optimization analysis further shows that this compression also affects gradient signals: image-level supervision injects a self-gradient component whose sign is determined by y_{img}, regardless of the local preference label y_i. In contrast, regional supervision preserves region-specific self-gradient directions through y_i. By maintaining each verified relation as (R_k, y_k), RePair avoids compressing heterogeneous local preferences into a single global label and enables each semantic-spatial requirement to receive its own preference signal.

C Implementation Details

C.1 Training Configuration

We use FLUX.2-klein-base-9B and Ideogram 4 as the base models. We fine-tune the attention and feed-forward projection layers of the denoising transformer using LoRA with rank $r = 32$ and scaling factor $\alpha = 32$, while keeping all original model parameters frozen, including the text encoder and VAE. All models are optimized with AdamW using a learning rate of 10^{-5}, $(\beta_1, \beta_2) = (0.9, 0.999)$, $\epsilon = 10^{-8}$, and a weight decay of 0.01. The learning rate is linearly warmed up during the first five scheduler updates and then maintained at the base value. We clip the gradient norm to 1.0 and set the preference optimization coefficient β to 100. Regional preference losses are uniformly averaged over all verified relations within each image pair, without auxiliary supervised fine-tuning or reference-preservation losses.

All models are trained for 10 epochs on four NVIDIA H200 GPUs using BF16 precision and gradient checkpointing. Each GPU processes one preference pair per iteration, yielding an effective batch size of four without gradient accumulation. The maximum training resolution is 1024^2 pixels, and we use a fixed random seed of 42.

We follow the default flow-matching scheduler, uniformly sampling timestep indices from 1,000 scheduler steps and applying the scheduler-defined weighting $w(t)$. The two images in each preference pair share the same timestep and noise realization. The soft regional mask is constructed by projecting each verified bounding box onto the spatial token grid, assigning unit weight inside the box and applying dilation and Gaussian smoothing around its boundary.

C.2 Regional Preference Mining Details

We construct regional preference data by extracting reliable semantic-spatial preference relations from multiple stochastic generations of the same prompt. Different from conventional preference construction that assigns a single global winner to each image pair, our mining process identifies localized differences between imperfect generations and preserves only those regions with consistent and verifiable preference directions. The process consists of candidate generation, regional failure discovery, pair construction, preference verification, and consistency filtering.

Candidate Generation. For each prompt, we generate ten candidate images at a resolution of 1024×1024 using the corresponding base model. To improve the comparability among candidates while maintaining stochastic diversity, we use correlated initialization during sampling. Specifically, candidates share a common noise component while retaining independent perturbations, with a correlation coefficient $\rho = 0.95$. The first candidate directly uses the anchor noise. This strategy reduces irrelevant global variations and encourages candidate differences to concentrate on local generation behaviors. The sampling configurations of different base models are summarized in Table 10.

Regional Failure Discovery. Each generated candidate is independently analyzed by a vision-language model given the prompt, semantic target descriptions, and the generated image. The model identifies localized generation failures and returns the corresponding bounding boxes, failure categories, severity scores, and confidence scores. Multiple errors associated with the same physical region are merged into a single regional record. To ensure reliable regional supervision, we discard

Table 10: Candidate generation configurations.

Model	Sampling Steps	Guidance Scale
FLUX.2-klein-base-9B	50	4.0
Ideogram 4	48	7.0 (first 45 steps), 3.0 (last 3 steps)

Table 11: Regional matching and filtering criteria.

Criterion	Threshold
Outside-region consistency	≥ 0.70
Region matching IoU	≥ 0.80
Matched region area ratio	≤ 3
Comparison box expansion	$\leq 1.5\times$
Failure burden difference	≥ 0.40
Region area ratio	$0.1\% \sim 45\%$
Total selected region area	$\leq 80\%$
Same-direction region overlap	≤ 0.05 IoU
Opposite-direction region overlap	0

unreliable detections and retain only regions with detection confidence above 0.75 and failure severity above 0.50. Missing detections are treated as unknown cases rather than positive evidence that the corresponding region is correct.

Regional Matching and Pair Construction. For each prompt, we enumerate candidate pairs and identify comparable failure regions between two generations. Region correspondence is established through greedy one-to-one matching based on spatial and semantic compatibility. Direct matching requires an IoU above 0.80. For regions associated with the same semantic target, we additionally allow center-distance matching when the distance is within half of the smaller bounding-box diagonal. To avoid introducing ambiguous comparisons, matched regions are further filtered according to scale consistency, failure difference, spatial constraints, and region overlap. The complete criteria are summarized in Table 11.

Regional Preference Verification. Each candidate pair is further verified by an independent vision-language model. The verifier receives the two complete images and coordinate-aligned regional crops, while the provisional preference direction, failure severity, and previous analysis results are hidden. This prevents the final preference decision from being directly determined by the initial failure mining stage. A regional relation is retained only when the verification confidence exceeds 0.85 and the absolute quality-score difference exceeds 0.25. When both candidates contain failures in the same region, a stricter quality difference threshold of 0.40 is applied. For counting-related failures, we additionally require explicit instance localization consistent with the reported counts and a count confidence above 0.90. Ambiguous regions, tied comparisons, and inconsistent preference directions are removed individually, after which the final regional preference relations are recomputed.

Outside-Region Consistency Filtering. To ensure that the mined preference relations mainly reflect regional differences rather than unrelated global changes, we further evaluate outside-region consistency using DINOv2-Base patch features. The similarity is computed as the average cosine similarity between spatially corresponding patches outside all selected preference regions, and we retain pairs with an outside-region consistency score of at least 0.7. For structured Ideogram 4 prompts, requested target regions are additionally excluded from this comparison. Bounding boxes are padded by eight units in the normalized $[0, 1000]$ coordinate system, and samples with insufficient remaining background regions are discarded. The outside-region consistency score is applied after preference verification during final dataset construction and is not used for pair ranking or VLM decision making. After the complete mining and filtering procedure, we obtain approximately 5K regional preference pairs for each base model, as summarized in Table 12.

Table 12: Statistics of the mined regional preference datasets.

Dataset	Base Model	Preference Pairs
RePair-FLUX	FLUX.2-klein-base-9B	5,025
RePair-Ideogram	Ideogram 4	4,887

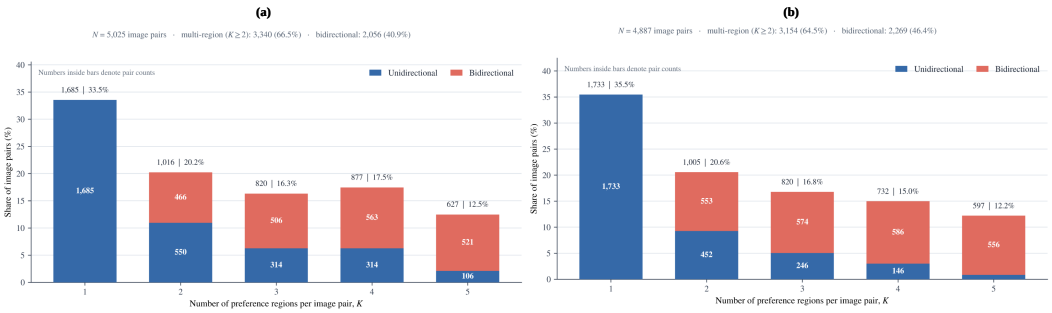

Figure 6: **Statistics of mined regional preferences.** (a) Data statistics for RePair on FLUX.2-klein-base-9B. (b) Data statistics for RePair on Ideogram4.

C.3 Ablation Details

We provide additional details of the ablation settings. For correlated candidate generation, we compare the default correlated initialization ($\rho = 0.95$) with independent sampling ($\rho = 0$), while keeping all other generation and mining configurations unchanged. For consistency verification, we remove the outside-region consistency filtering to evaluate its contribution to preference quality. To study the effect of multi-directional preference learning, we construct a single-direction variant that only retains regional relations with the dominant preference direction and discards conflicting regional preferences within the same pair. We further evaluate the influence of regional preference density by comparing the default setting, which uses all verified regional relations, with a single-region variant that randomly selects one verified region from each preference pair.

D Regional Preference Statistics

We provide additional statistics of the mined regional preference datasets. The analysis focuses on the distribution of regional preference structures, including the number of verified preference regions per image pair and the proportion of multi-directional preference relations.

D.1 Dataset Overview

Our mining pipeline produces two regional preference datasets based on FLUX.2-klein-base-9B and Ideogram 4. Each dataset contains approximately 5K verified image pairs, where each pair is associated with one or more localized preference relations. Specifically, RePair-FLUX contains 5,025 preference pairs, while RePair-Ideogram contains 4,887 preference pairs.

Each image pair is generated from the same prompt and mined from multiple stochastic candidates produced by the corresponding base model. After candidate generation, regional failure discovery, pairwise verification, and consistency filtering, only reliable semantic-spatial preference relations are retained. The final datasets contain both single-region and multi-region preference pairs, enabling the model to learn from diverse local generation successes and failures.

D.2 Preference Structure Distribution

Figure 6 illustrates the distribution of regional preference structures in the mined training sets. We categorize each image pair according to the number of verified preference regions K and whether

the regions share the same preference direction. For FLUX.2-klein-base-9B, 3,340 pairs (66.5%) contain multiple preference regions ($K \geq 2$), and 2,056 pairs (40.9%) exhibit bidirectional regional preferences, where different regions within the same image pair prefer different candidates. Similarly, the Ideogram 4 dataset contains 3,154 multi-region pairs (64.5%) and 2,269 bidirectional pairs (46.4%).

The existence of a large proportion of bidirectional preference structures demonstrates that image-level preference labels cannot fully describe the local quality differences within a generated image pair. In particular, many pairs contain regions where one candidate is preferred for one semantic requirement while the other candidate is preferred for another. These statistics motivate the regional preference formulation adopted by RePair, which preserves each verified local relation instead of collapsing them into a single global preference label.

E Additional Qualitative Examples

We provide additional qualitative comparisons to further illustrate the behavior of RePair across diverse localized generation errors. Figures 7 and 8 present examples on FLUX.2-klein-base-9B, covering counting, local structure, attribute binding, text rendering, spatial relations, physical consistency, and reasoning, while Fig. 9 shows a challenging Ideogram 4 example containing multiple error types within a single generation. Across these cases, RePair more consistently satisfies localized semantic constraints while preserving the overall composition.

F Limitations and Future Work

RePair is primarily designed for localized generation errors and is therefore less advantageous when the failure is inherently global. When the affected region spans most of the image, regional preference optimization provides limited additional structure and effectively reduces to image-level preference learning. In addition, although our human validation and multi-VLM experiments show that current strong VLMs can reliably identify and compare localized generation differences, the offline annotation, verification, and filtering stages still introduce additional computational and inference cost during data construction. Future work could explore adaptive combinations of regional and global supervision according to the spatial scope of each failure, as well as more efficient preference mining pipelines that reduce annotation overhead while preserving supervision quality.

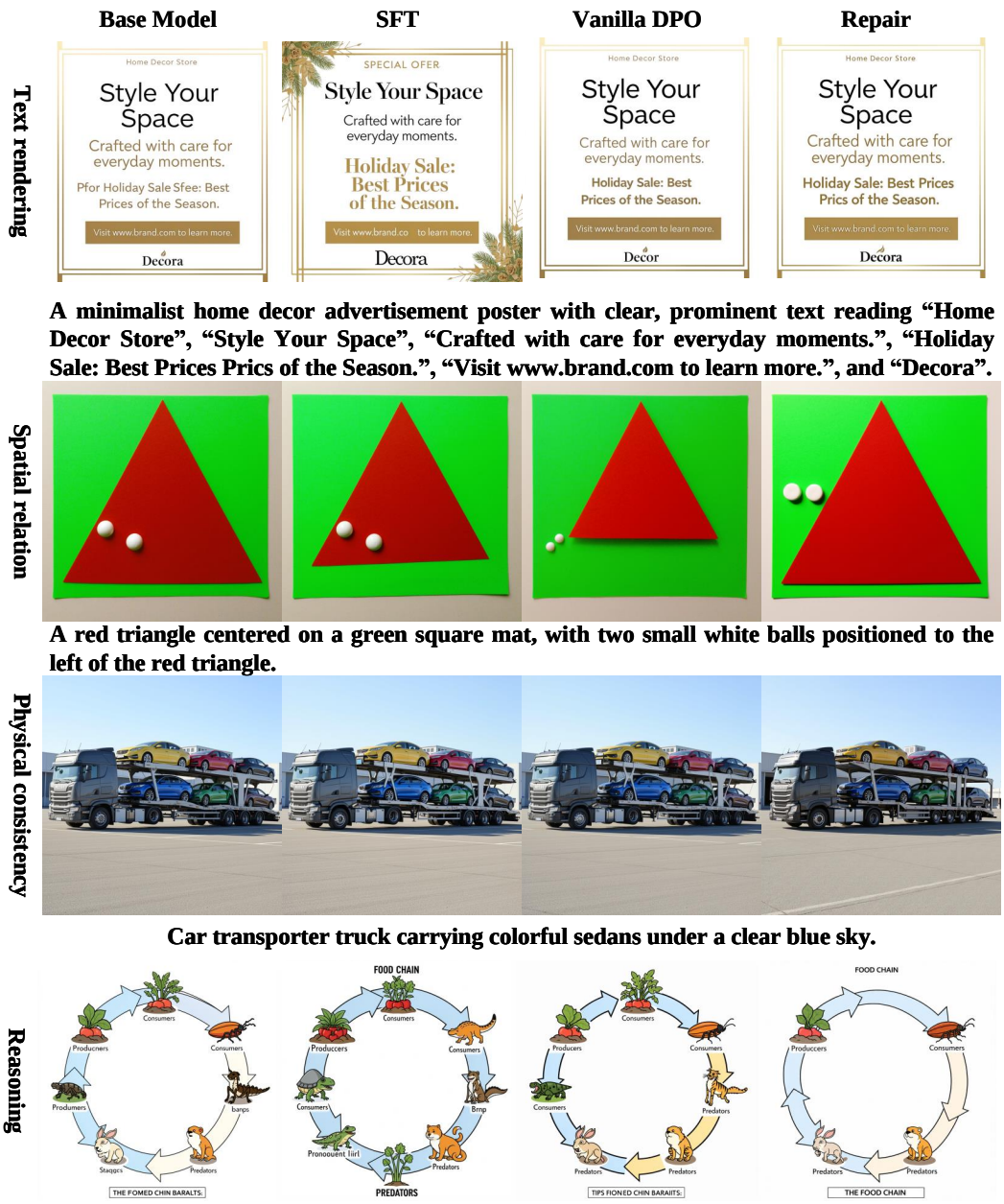

Figure 7: **Additional qualitative examples on FLUX.2-klein-base-9B.** The examples cover text rendering, spatial relation, physical consistency, and reasoning.

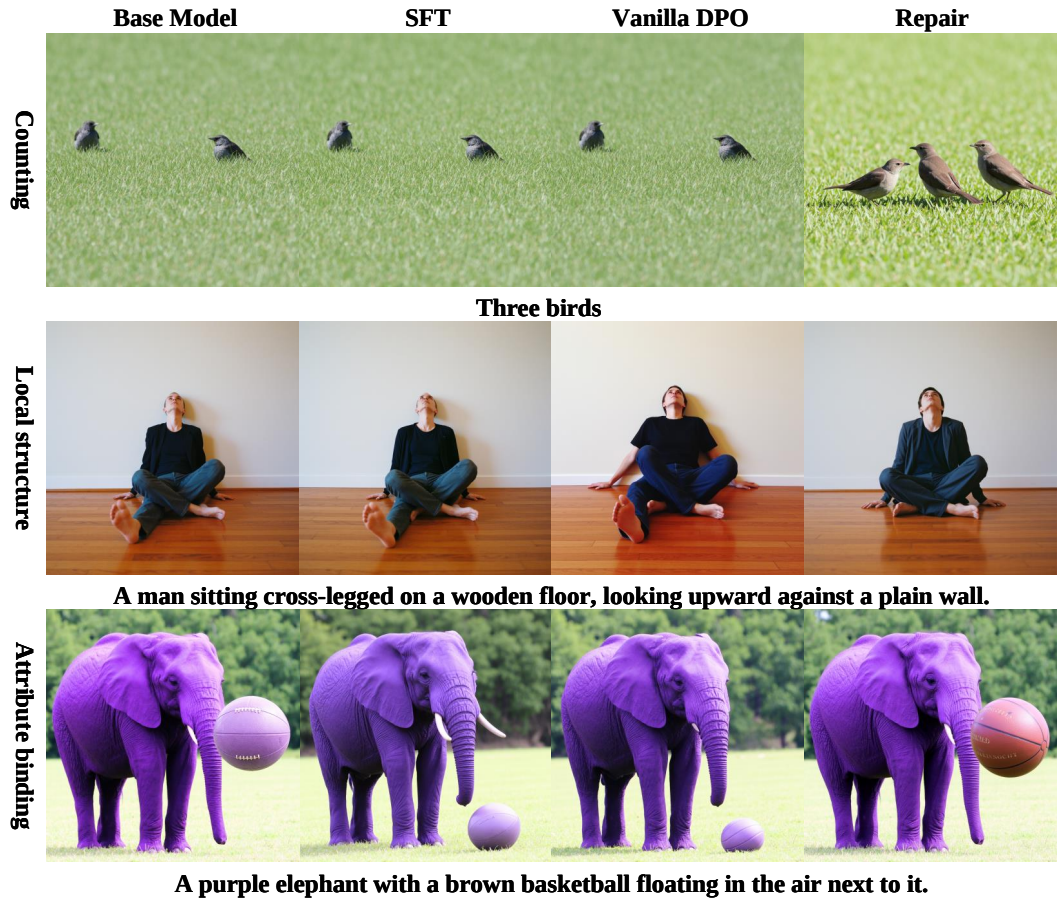

Figure 8: **Additional qualitative examples on FLUX.2-klein-base-9B.** The examples highlight improvements on counting, local structure, and attribute binding.



Figure 9: **Additional qualitative examples on Ideogram 4.** This example simultaneously covers six categories of localized errors, including text rendering, local structure, attribute binding, counting, spatial relation, and physical consistency.