BLAME ERRANT WORDS, NOT THE CANVAS: DEFECT-AWARE SPATIAL CREDIT ASSIGNMENT FOR VISUAL TEXT GENERATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Recent reinforcement learning approaches have improved visual text generation, yet generated images still suffer from incorrect wording and malformed glyphs. We identify two coupled limitations: recognition-based rewards often overlook visible glyph defects, while whole-image scalar feedback obscures the uneven quality of individual text regions. These limitations motivate our central principle: *Blame errant words, not the canvas*, to effectively preserve local quality distinctions from evaluation to policy optimization. First, we introduce TextDART, a defect-aware region-targeted reward model that localizes individual text regions and independently evaluates their semantic correctness and glyph quality, providing fine-grained, region-precise feedback. Building on TextDART, we propose SCANNER, which turns these regional judgments into spatially differentiated policy feedback. It normalizes rewards for each text target across sampled images and casts the resulting regional advantages onto the corresponding latent tokens. Extensive experiments on existing benchmarks and our newly established PolyText-Bench demonstrate that our approach achieves state-of-the-art performance in text accuracy and glyph fidelity.

1 INTRODUCTION

Text-to-image (T2I) models have made remarkable progress in visual fidelity, compositionality, and prompt adherence (Rombach et al., 2022; Podell et al., 2024; Labs, 2024; Wu et al., 2025), and are rapidly evolving from generating compelling scenes to producing usable visual content. A broad range of visual artifacts (e.g., posters, advertisements) rely on the accurate integration of language and imagery, where a single misspelled word or malformed character can compromise an otherwise compelling image. Reliable visual text generation is therefore a key capability for extending generative models from pictorial synthesis to practical content creation. Recent reinforcement learning (RL) post-training methods (Black et al., 2024; Xu et al., 2023), particularly Group Relative Policy Optimization (GRPO) (Liu et al., 2026a; Xue et al., 2025) and its variants (Li et al., 2026; 2025a; Ding & Ye, 2026; Luo et al., 2025; Chen et al., 2026), have demonstrated promising progress stability and scalability in visual alignment.

Despite this progress, visual text generation remains a challenging frontier (Chen et al., 2024; Tuo et al., 2024; Yang et al., 2023; Ma et al., 2023), where generated images still suffer from incorrect wording and malformed glyphs such as missing characters, merged strokes, and hallucinated shapes. Consequently, improvements in average image quality can conceal consequential local defects, where a single corrupted word may still render the final artifact unusable. More importantly, in RL post-training, a failure that the reward cannot distinguish is not simply left unmeasured; it may be reinforced as successful behavior (Pan et al., 2022; Skalse et al., 2022).

We attribute these persistent failures to two key properties of visual text generation that existing RL methods overlook. **(1) Errors are heterogeneous**: a failure may lie in *what* is written (a semantic error) or in *how* it looks (a glyph defect), but most existing methods supervise only the former, leaving deformations invisible to the reward. **(2) Text is regional**: it occupies localized areas, and a single image may contain multiple text instances of uneven quality, so text quality is not uniformly distributed, yet most methods score the whole image, averaging away local correctness. Accordingly,

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.

these observations motivate our central principle: *Blame errant words, not the canvas*. Realizing this principle requires preserving local quality distinctions from evaluation to policy optimization, rather than reducing them to a single image-level advantage as in standard image-level GRPO. This raises two questions: **what should fine-grained rewards measure**, and **where should reliable credit go**?

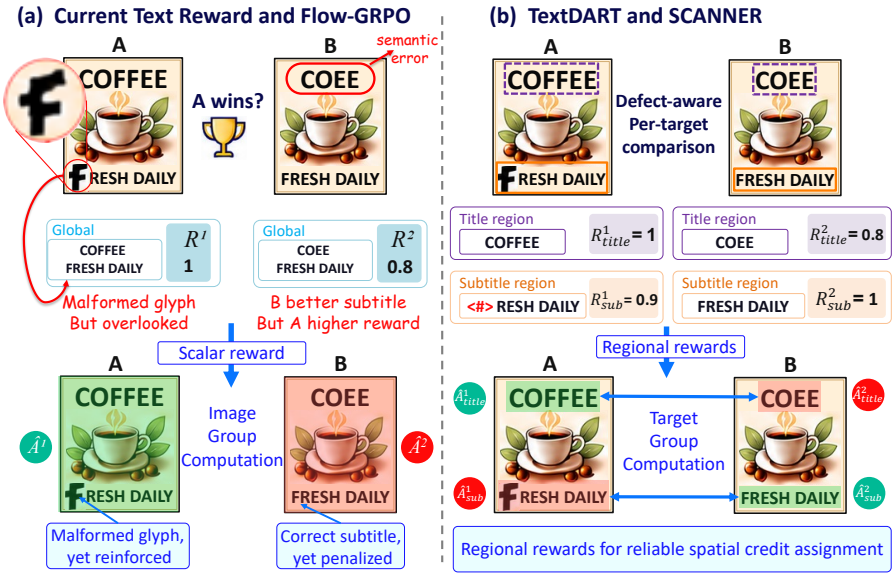


Figure 1: **Comparison of Our Method Against Current Reward Model and Flow-GRPO.** (a) A outranks B despite its malformed glyph, and image-group advantages reinforce the defect while penalizing B's correct subtitle. (b) TextDART provides defect-aware scores for individual regions, and SCANNER computes target-group advantages for spatial credit assignment, penalizing A's malformed subtitle while reinforcing B's correct one (green/red: positive/negative).

The first question concerns how rewards are computed, as shown in Fig. 1(a). Existing methods largely rely on OCR (Cui et al., 2026a) or vision-language models to recognize the rendered string and compute an edit-distance-based reward (Li et al., 2024; Cui et al., 2026b; Rahman et al., 2025). Such recognizers carry strong linguistic priors: malformed glyphs are frequently read as valid words, and obvious stroke merging or omission is hallucinated as normal text. Moreover, the normalized edit distance only reflects semantic correctness and says nothing about glyph quality. While TextPecker (Zhu et al., 2026) introduces perceptual signals for defective text, it operates on the whole image, flattening all text regions into a single score. To obtain fine-grained rewards, we propose a **D**efect-**A**ware **R**egion-**T**argeted reward model, **TextDART**. In Fig. 1(b), rather than scoring the whole image, TextDART first localizes every text region, then evaluates glyph quality and semantic correctness independently, yielding rewards that are both defect-aware and region-precise.

The second question persists even with *localized, defect-sensitive* feedback in hand. Standard GRPO assigns the same credit to every latent token at every denoising step (Liu et al., 2026a; Xue et al., 2025). As a result, uniform credit propagation lets flawed regions drag down well-formed ones, and the temporal dynamics of text emergence are ignored. Prior attempts either allocate regional signals (Yang et al., 2026; Zhang et al., 2026; Ouyang et al., 2026) while ignoring that text layout is undetermined in early steps, or compute rewards at intermediate steps (Deng et al., 2026; Tong et al., 2026; He et al., 2026), which is costly and still not region-aware. To resolve this, we propose **SCANNER**, a reinforcement learning framework for **S**patial **C**redit **A**ssig**N**ment with **N**ois**E**-awa**R**eness. Utilizing the proposed TextDART, a target-group strategy normalizes the rewards per text target and casts them onto the corresponding latent tokens. Along the denoising trajectory, the global advantage is dynamically fused with these regional signals via a noise-dependent schedule, shifting credit from the global view to local focus as text regions gradually resolve, thereby providing reliable guidance for policy optimization.

In summary, our main contributions are as follows:

- We pinpoint the absence of localized, defect-sensitive feedback as a critical barrier to visual text generation. To this end, we introduce *TextDART*, a defect-aware region-targeted reward model, to disentangle semantic correctness and glyph quality across individual text regions.

- We propose *SCANNER*, a spatial credit assignment method that normalizes and projects rewards onto target latent tokens, dynamically shifting focus from global advantages to regional signals via a noise-dependent schedule for reliable policy optimization.
- Extensive experiments on four existing benchmarks and our newly established *PolyText-Bench* demonstrate our method achieves state-of-the-art performance in text accuracy and glyph fidelity.

2 Related Work

2.1 Visual Text Generation and Evaluation

Visual text generation has been approached by improving rendering through layout or glyph conditioning and glyph-aware training data (Chen et al., 2023; Tuo et al., 2024), and more recently by post-training pretrained generators with rewards. The dominant reward is OCR-based string matching (Liu et al., 2026a), which is blind to visual defects. Zhu et al. (2026) penalize glyph deformations with perceptual reward but still collapses all text into a single whole-image score, so defects cannot be localized. Building on these observations, we focus on region-level evaluation of both semantic correctness and glyph quality, enabling fine-grained feedback.

2.2 RL and Spatial Credit Assignment for Diffusion and Flow Models

RL post-training aligns diffusion and flow models through preference optimization or GRPO (Wallace et al., 2024; Liu et al., 2026a; Shao et al., 2024), but a whole-image scalar advantage cannot distinguish local contributions. One remedy reallocates this same advantage according to attention maps (Shen et al., 2026) or perceptual allocation maps (Ni et al., 2026; Xing et al., 2025): spatial weights are adjusted, yet no region can be rewarded or penalized on its own. Pixel-level feedback takes a different route, building local advantages from defect heatmaps (Liang et al., 2024; Guo et al., 2026; Kordzanganeh et al., 2024; Zhang et al., 2026), but the same spatial position need not correspond to the same semantic target across samples. Region-level feedback instead constructs advantages per semantic target and updates the corresponding regions (Yang et al., 2026; Xu et al., 2026), yet the region-target correspondence is unreliable in early high-noise steps. Building on these observations, we focus on reliable credit assignment per text target that adapts to the denoising process, leading to more effective policy optimization.

3 Preliminaries

Diffusion and flow models transport noise to data along Gaussian probability paths $\mathbf{x}_t = \alpha_t \mathbf{x}_0 + \sigma_t \boldsymbol{\epsilon}$, $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$, and admit an ODE sampler $\mathrm{d}\mathbf{x}_t/\mathrm{d}t = \mathbf{v}_\theta(\mathbf{x}_t, t, c)$ with conditioning prompt c (Ho et al., 2020; Song et al., 2020; Lipman et al., 2023; Liu et al., 2022). Flow-GRPO (Liu et al., 2026a) enables online RL on such models by replacing the deterministic sampler with a stochastic SDE counterpart and applying group-relative policy optimization (Shao et al., 2024). For the linear path $\alpha_t = 1 - t$, $\sigma_t = t$, its SDE counterpart preserves the ODE marginals under an exact velocity field. On a grid $0 = t_0 < \cdots < t_T = 1$ with $h_k = t_k - t_{k-1} > 0$, the discretized transition is

$$\mathbf{x}_{t_{k-1}} \sim \mathcal{N}(\boldsymbol{\mu}_{\theta,k}, n_k^2 \mathbf{I}), \quad \boldsymbol{\mu}_{\theta,k} = \mathbf{x}_{t_k} - h_k \left[\mathbf{v}_{\theta,k} + \frac{g_{t_k}^2}{2t_k} \left(\mathbf{x}_{t_k} + (1 - t_k) \mathbf{v}_{\theta,k} \right) \right], \tag{1}$$

where $\mathbf{v}_{\theta,k} = \mathbf{v}_\theta(\mathbf{x}_{t_k}, t_k, c)$, $g_{t_k} > 0$ is the policy-independent noise strength, and $n_k = g_{t_k} \sqrt{h_k}$ the transition noise.

Given a prompt c, the old policy samples G trajectories with terminal rewards R_i, normalized into group-relative advantages

$$\widehat{A}_i = (R_i - \bar{R})/(s_R + \delta), \tag{2}$$

with $\bar{R}$ and s_R the group mean and standard deviation and $\delta > 0$ for numerical stability. Letting $r_{i,k}(\theta) = p_\theta(\mathbf{x}^i_{t_{k-1}} \mid \mathbf{x}^i_{t_k}, c) / p_{\theta_{\text{old}}}(\mathbf{x}^i_{t_{k-1}} \mid \mathbf{x}^i_{t_k}, c)$ denote the probability ratio, the policy maximizes

$$\mathcal{J}(\theta) = \mathbb{E}\Big[\tfrac{1}{GT} \textstyle\sum_{i=1}^{G} \sum_{k=1}^{T} \min\big(r_{i,k} \widehat{A}_i,\ \mathrm{clip}(r_{i,k}, 1-\varepsilon, 1+\varepsilon) \widehat{A}_i \big) - \beta D_{i,k} \Big], \tag{3}$$

where the expectation is over prompts and old-policy trajectories, ε the clipping threshold, and $D_{i,k}$ the conditional KL to the reference policy with weight β, which reduces to a squared difference

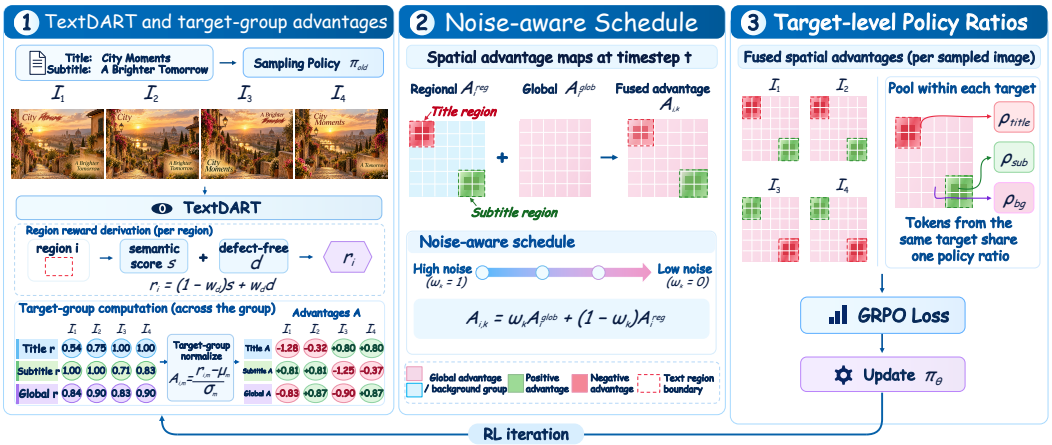


Figure 2: **Overview of TextDART and SCANNER.** (a) TextDART localizes text and evaluates semantic correctness and defects. Rewards for the same target are normalized across samples to construct spatial advantages. (b) A noise-aware schedule shifts credit from global to regional advantages as denoising proceeds. (c) Policy ratios are pooled within each text target and the background, enabling spatially aligned GRPO updates.

of transition means under identical covariances. The same image-level advantage $\widehat{A}_i$ guides all denoising steps without distinguishing regional contributions.

4 METHOD

Given a prompt that specifies the target text, we aim to generate images with accurate content and defect-free text. However, recognition-based rewards may overlook text defects due to hallucination, while image-level rewards erase regional quality differences. Reliable optimization therefore requires attributing credit to individual errant words rather than the whole image. To this end, as shown in Fig. 2, we introduce TextDART for defect-aware text recognition and localization, and SCANNER for spatial credit assignment from its regional feedback.

4.1 TEXTDART: DEFECT-AWARE REGION-TARGETED REWARDS

Localized Defect Perception. To perceive text defects together with their locations, we fine-tune Qwen3-VL-8B on the training set of TextDART-123K (detailed in A.7) to jointly predict text quadrilaterals and transcriptions with structural-anomaly markers:

$$\{(q_{i,o}, t_{i,o})\}_{o=1}^{O_i} = \mathcal{F}_{\text{text}}(I_i), \tag{4}$$

where $\mathcal{F}_{\text{text}}$ denotes TextDART, I_i the i-th sampled image, O_i the number of regions originally detected by $\mathcal{F}_{\text{text}}$, and $q_{i,o}$ and $t_{i,o}$ the quadrilateral and defect-aware transcription of region o. The model takes only the image as input, keeping visual assessment independent of the target text, which is then used for matching and scoring.

Regional Rewards. A single image-level score fails to distinguish text regions of varying quality. We therefore match regional transcriptions to target text through many-to-many matching and compute rewards separately (both detailed in A.1). Regions matched to the same target are merged, while targets covered by one region are also merged consistently throughout the rollout. This yields a shared set of M merged target texts $\bar{y}_m$, each with region $\bar{q}_{i,m}$ and transcription $\bar{t}_{i,m}$ in image i when present. Motivated by Zhu et al. (2026), the regional reward $r_{i,m}$ combines the semantic correctness score $s_{i,m}$ and the defect-free score $d_{i,m}$, with higher scores indicating better quality.

$$s_{i,m} = 1 - \text{GNED}(\bar{t}_{i,m}, \bar{y}_m), \quad d_{i,m} = \left[1 - \kappa \frac{a(\bar{t}_{i,m})}{L(\bar{t}_{i,m})}\right]_0^1, \quad r_{i,m} = (1 - w_d)s_{i,m} + w_d d_{i,m}, \tag{5}$$

where GNED denotes the generalized normalized edit distance, a the weighted defect count, L the transcription length, κ the defect-penalty scale, w_d the mixing weight, and $[\cdot]_0^1$ the clipping operator on $[0, 1]$. A target absent from image i receives $r_{i,m} = 0$, and regions matched to no target are

merged into a single extra entry $r_{i,\text{ext}} = d_{i,\text{ext}}$ computed on their concatenated transcription, which is counted as a reward entry only when such extra text is present.

4.2 SCANNER: SPATIAL CREDIT ASSIGNMENT WITH NOISE-AWARENESS

SCANNER converts the regional rewards of TextDART into token-level credit through three components: target-group advantages that compare the same text target across samples, a noise-aware schedule that shifts credit from global to regional signals along denoising, and target-level policy ratios that keep updates consistent within each target.

Target-group Advantages. Standard GRPO gives all image tokens one advantage, so flawed and well-formed regions receive identical credit. Regional rewards require a comparison unit: averaging within an image reduces to image-level comparison, while fixed latent positions are sensitive to layout variation. We therefore group regions aligned with the same merged target across images generated from one prompt and normalize rewards within each target group. Since these regions render identical content, the advantage isolates generation quality:

$$A_{i,m} = \frac{r_{i,m} - \bar{r}_m}{\max(\sigma^r_m, \delta)}, \quad \bar{r}_m = \text{mean}_{i'}\, r_{i',m}, \quad \sigma^r_m = \text{std}_{i'}\, r_{i',m}, \tag{6}$$

where $A_{i,m}$ denotes the advantage for target m in image i, $\bar{r}_m$ and σ^r_m the reward mean and sample standard deviation over the G images i' sampled for the same prompt, and $\delta > 0$ the standard-deviation floor. The extra entry $r_{i,\text{ext}}$ is normalized in the same way with a clean reference of 1 for its absence, and its advantage is assigned only to existing extra-text regions.

To cast the per-region target advantage into the token-level objective, we project each merged region $\bar{q}_{i,m}$ onto the latent grid and normalize its contribution by support size, so that every target receives equal total credit regardless of how many tokens it occupies:

$$A^{\text{reg}}_i(p) = \frac{\lambda N}{|\mathcal{M}_i|} \sum_{m \in \mathcal{M}_i} \frac{[p \in \Omega_{i,m}]}{|\Omega_{i,m}|} A_{i,m}, \tag{7}$$

where $A^{\text{reg}}_i(p)$ is the regional advantage at latent position p, N is the number of latent tokens, $\Omega_{i,m}$ is the token support of $\bar{q}_{i,m}$, $\mathcal{M}_i \subseteq \{1, \ldots, M\} \cup \{\text{ext}\}$ contains the targets present in image i, and λ is the regional scale, with $|\cdot|$ and $[\cdot]$ denoting cardinality and the indicator function. Tokens in overlapping regions are assigned to the smaller region, while background tokens—and all tokens when $\mathcal{M}_i$ is empty—receive zero advantage.

Noise-aware Schedule. Regional advantages rely on text locations, which are undetermined in early denoising steps, so credit should shift from global feedback toward regional signals as denoising proceeds. For the global branch, rather than averaging regional rewards, which weights each target equally regardless of its length, we score the concatenation of all detected text as a single string, weighting every character equally:

$$R_i = (1 - w_d)s(\text{concat}_o\, t_{i,o}, \text{concat}_j\, y_j) + w_d d(\text{concat}_o\, t_{i,o}), \tag{8}$$

where R_i denotes the global reward, y_j the j-th original target, and s and d the scoring functions in Eq. 5, with concat denoting space-separated concatenation in detected-region order for transcriptions and in target order for targets. We then mix the two advantages at each step, with the global weight decaying as the transition noise decreases:

$$\begin{aligned} \omega_k &= \frac{\widetilde{n}_k - \min_{l \in \mathcal{S}} \widetilde{n}_l}{\max_{l \in \mathcal{S}} \widetilde{n}_l - \min_{l \in \mathcal{S}} \widetilde{n}_l}, \\ A_{i,k}(p) &= \omega_k A^{\text{glob}}_i + (1 - \omega_k) A^{\text{reg}}_i(p), \end{aligned} \tag{9}$$

where ω_k denotes the global mixing weight, $\widetilde{n}_k$ the running minimum in generation order of the transition noise standard deviation $n_k = g_{t_k}\sqrt{h_k}$ from Eq. 1, A^{glob}_i the global advantage from Eq. 2, $A_{i,k}(p)$ the fused advantage, and $\mathcal{S}$ the optimized SDE steps indexed by k and l. If the noise range is numerically zero, ω_k is set to 0.5.

Target-level Policy Ratios. With token-level advantages in place, the ratio granularity still matters: token-level ratios allow separate clipping decisions within the same text region, whereas an image-level ratio couples unrelated regions. To keep updates consistent within each merged target while retaining regional differences, we pool the Guard-adjusted ratio exponent (Wang et al., 2025) over its token support:

$$z_{i,k,p} = n_k(\ell^{\theta}_{i,k,p} - \ell^{\text{old}}_{i,k,p}) + \frac{\Delta_{i,k,p}}{2n_k},$$
$$\rho_{i,k,m} = \exp\left(\frac{1}{|\Omega_{i,m}|}\sum_{p\in\Omega_{i,m}} z_{i,k,p}\right), \tag{10}$$

where $z_{i,k,p}$ denotes the adjusted token exponent, $\rho_{i,k,m}$ the ratio for merged target m, including the extra-text entry ext, $\ell^{\theta}_{i,k,p}$ and $\ell^{\text{old}}_{i,k,p}$ the channel-averaged Gaussian log densities of the same sampled transition under the current and old policies, and $\Delta_{i,k,p}$ the channel-averaged squared difference between their transition means. Background tokens form a separate group, and gradients propagate through the pooled exponent.

Text rewards do not directly constrain the surrounding scene. We therefore regularize toward a frozen reference with greater background weight while maximizing the clipped policy objective:

$$\mathcal{J}_{\text{SCANNER}}(\theta) = \mathbb{E}\left[\frac{1}{G|\mathcal{S}|}\sum_{i=1}^{G}\sum_{k\in\mathcal{S}}\left\{\frac{1}{Nh_k}\sum_{p=1}^{N}\mathcal{C}\big(\rho_{i,k}(p), A_{i,k}(p)\big) - \beta D^{\text{KL}}_{i,k}\right\}\right], \tag{11}$$

where $\mathcal{J}_{\text{SCANNER}}$ denotes the training objective, and $\rho_{i,k}(p)$ the pooled ratio assigned to token p. The clipped surrogate is $\mathcal{C}(\rho, A) = \min\{\rho A, \text{clip}(\rho, 1-\varepsilon, 1+\varepsilon)A\}$ for ratio ρ and advantage A, with clipping threshold ε and regularization weight β from Eq. 3. The penalty $D^{\text{KL}}_{i,k}$ denotes the text/background-weighted KL divergence between the current and reference transition distributions, which reduces to the weighted squared velocity difference under equal transition noise. The expectation is over prompts and old-policy trajectories, and $1/h_k$ denotes the Guard step-size factor.

5 Experiments

5.1 Experimental Setup.

Implementation Details. We perform RL post-training with LoRA on FLUX.1-dev (Labs, 2024) for English and on Qwen-Image-2512 (Wu et al., 2025) for both English and Chinese, using multi-region text prompts with specified targets. We compare against the pretrained backbones and Flow-GRPO (Liu et al., 2026a) equipped with PP-OCRv5 (Cui et al., 2026a) or TextPecker (Zhu et al., 2026) rewards, on LongText-Bench (Geng et al., 2025), OneIG-Bench (Chang et al., 2026), CVTG-2k (Du et al., 2025) (English only), GenTextEval-Bench (Zhu et al., 2026), and our PolyText-Bench. Further experimental details are provided in Appendix A.2.

PolyText-Bench. Existing benchmarks prescribe target strings and thus cover only text-critical generation. Yet models often render text spontaneously, and defective text, once present, degrades the entire image, making text quality in text-optional scenes a direct reflection of image quality. PolyText-Bench therefore covers both text-critical and text-optional settings with 600 prompts, evenly split between English and Chinese and between the two settings. All metrics are computed by TextDART from the full image: critical semantic (CS) and glyph (CG) scores average over requested targets, with missing targets scored zero, while optional glyph quality (OG) averages regional scores over text-containing images, excluding text-free outputs so that avoiding text yields no advantage. More details are provided in A.6.

5.2 Evaluation of TextDART.

Fig. 3 compares TextDART with PP-OCRv5 and TextPecker under challenging visual conditions. TextDART preserves valid content in small and low-contrast text that TextPecker mismarks as defective, identifies the malformed glyph in flawed WordArt that both baselines decode as a normal



Figure 3: **Localization and Defect-aware Recognition under Challenging Visual Conditions.** TextDART accurately localizes and reads text while marking <#> as malformed glyphs, whereas PP-OCRv5 hallucinates incorrect content and TextPecker misses or mismarks defects.

Table 1: **Quantitative Comparison of Visual Text Generation Methods.** (↑/↓: higher/lower is better.) Best in **bold**; "–" denotes no reward model or unsupported language.

Reward	RL Algorithm	LongText-Bench	OneIG-Bench				CVTG-2k		GenTextEval-Bench		PolyText-Bench		
		TextScore ↑	TextScore ↑	ED ↓	CR ↑	WAC ↑	WordAcc ↑	NED ↑	Qua ↑	Sem ↑	CS ↑	CG ↑	OG ↑
English Text Generation													
						FLUX.1-dev							
–	–	0.620	0.543	85.31	0.138	0.379	0.503	0.694	0.661	0.352	0.426	0.464	0.281
OCRv5	Flow-GRPO	0.769	0.751	64.55	0.220	0.506	0.686	0.818	0.838	0.589	0.703	0.726	0.365
TextPecker	Flow-GRPO	0.801	0.770	63.19	0.221	0.533	0.772	0.879	0.836	0.582	0.789	0.809	0.405
TextDART	**SCANNER**	**0.827**	**0.858**	**51.06**	**0.309**	**0.598**	**0.834**	**0.914**	**0.839**	**0.613**	**0.882**	**0.905**	**0.653**
						Qwen-Image-2512							
–	–	0.952	0.986	24.54	0.461	0.895	0.844	0.922	0.944	0.787	0.887	0.899	0.252
OCRv5	Flow-GRPO	0.947	0.989	19.58	0.509	0.890	0.832	0.914	0.946	0.795	0.876	0.876	0.273
TextPecker	Flow-GRPO	0.960	0.988	22.14	0.484	0.891	0.870	0.936	0.956	0.801	0.908	0.918	0.269
TextDART	**SCANNER**	**0.972**	**0.995**	**13.44**	**0.585**	**0.919**	**0.878**	**0.939**	**0.970**	**0.844**	**0.953**	**0.970**	**0.496**
Chinese Text Generation													
						Qwen-Image-2512							
–	–	0.966	0.983	13.06	0.554	0.851	–	–	0.955	0.853	0.931	0.945	0.198
OCRv5	Flow-GRPO	0.965	0.986	10.88	0.584	0.850	–	–	0.959	0.855	0.931	0.935	0.224
TextPecker	Flow-GRPO	0.974	0.985	11.83	0.574	0.853	–	–	0.962	0.866	0.937	0.945	0.190
TextDART	**SCANNER**	**0.980**	**0.992**	**8.56**	**0.691**	**0.856**	–	–	**0.975**	**0.892**	**0.956**	**0.970**	**0.403**

character, and reads the creative "NOS" composition as one coherent region that PP-OCRv5 fragments. Detailed quantitative comparison in A.3.1 further shows that TextDART achieves the best results in recognition, defect perception, and localization.

5.3 Quantitative Evaluation.

As shown in Tab. 1, our method achieves the best results on all metrics across both backbones and languages. The largest gains appear on FLUX.1-dev (CR from 0.138 to 0.309, WordAcc from 0.503 to 0.834), and clear margins hold on the stronger Qwen-Image-2512 (CR from 0.461 to 0.585 in English, ED from 13.06 to 8.56 in Chinese), where the OCRv5 reward instead degrades several metrics below the base model and TextPecker's image-level reward brings only modest gains. A user study further shows that human evaluators prefer our method over the baselines in A.3.2.

5.4 Qualitative Evaluation.

Fig. 4 compares TextDART with SCANNER against the pretrained backbone and Flow-GRPO with OCRv5 or TextPecker rewards. On text-critical prompts, baseline errors concentrate on secondary regions (the side panels in the first row and the ribbon banner in the third). The gap is largest on long text: only ours renders the full descriptive block in the second row legibly, while preserving

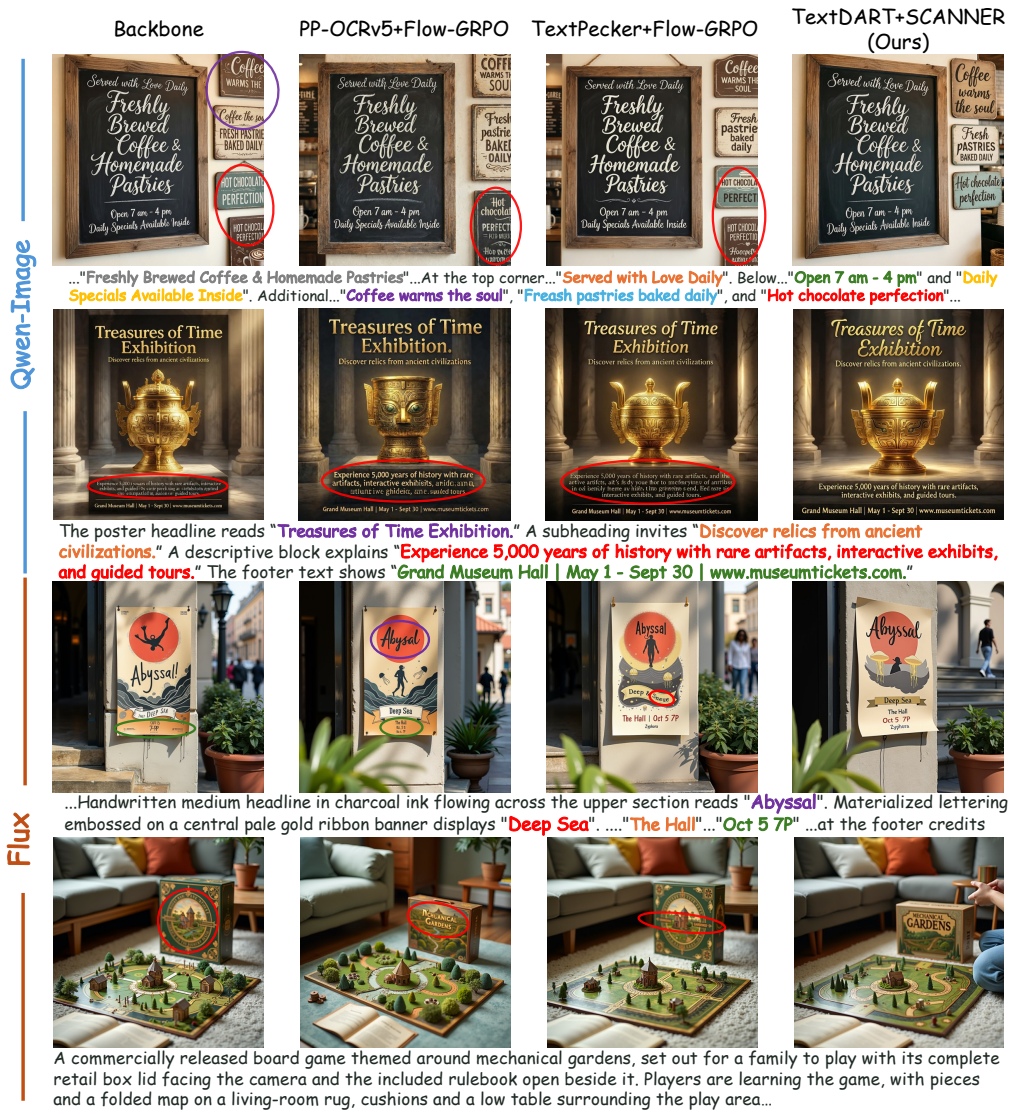


Figure 4: **Qualitative Comparison of TextDART with SCANNER against Baselines.** Target strings are color-coded in the prompts, and circles highlight text errors of the baselines. Our method renders every target correctly while preserving the background and aesthetics of the backbone, and produces legible, well-styled spontaneous text even when no text is prescribed.

the layout and lighting of the poster. In the text-optional case (last row), ours spontaneously renders a clean, well-styled title on the box lid, while the baselines garble or deform it.

5.5 Ablation Study

Contributions of Reward Components. Tab. 2 augments the reward and the optimizer step by step. Semantic rewards alone bring large gains over the base model, and defect scoring improves further. Collapsing regional rewards into an image-level scalar instead degrades Flow-GRPO (ED from 18.82 to 19.41), as flawed regions drag down well-formed ones under a broadcast value. SCANNER achieves the best results with the same rewards (ED 13.44, WordAcc 0.878), confirming that fine-grained rewards require spatially aligned credit assignment.

Ablation of Credit Assignment We ablate SCANNER along its credit-assignment pipeline. For advantage grouping in Fig. 5(a), token-level normalization oscillates, region grouping plateaus due to identical regional credit, and target grouping converges best by isolating per-target quality. For the global reward in Fig. 5(b), equal-weight regional averaging lags behind merged-text reward because per-region normalization overweights short-text errors while diluting long-text ones. Tab. 3

Table 2: **Ablation of Reward Components on Qwen-Image.** Sem and Def denote semantic and defect-aware scoring, respectively, and Reg denotes equal-weight aggregation of region-wise rewards into an image-level scalar. The best results in each language are marked in **bold**, and "–" marks unsupported languages.

Sem	Def	Reg	RL Algorithm	LongText-Bench	OneIG-Bench				CVTG-2k		GenTextEval-Bench		PolyText-Bench		
				TextScore ↑	TextScore ↑	ED ↓	CR ↑	WAC ↑	WordAcc ↑	NED ↑	Qua ↑	Sem ↑	CS ↑	CG ↑	OG ↑
English Text Generation															
–	–	–	Base	0.952	0.986	24.54	0.461	0.895	0.844	0.922	0.944	0.787	0.887	0.899	0.252
✓	–	–	Flow-GRPO	0.958	0.988	22.50	0.500	0.892	0.848	0.923	0.950	0.798	0.903	0.916	0.264
✓	✓	–	Flow-GRPO	0.960	0.991	18.82	0.510	0.901	0.846	0.922	0.945	0.803	0.904	0.913	0.287
✓	✓	✓	Flow-GRPO	0.961	0.993	19.407	0.541	0.917	0.867	0.934	0.952	0.801	0.914	0.924	0.289
✓	✓	✓	**SCANNER**	**0.972**	**0.995**	**13.44**	**0.585**	**0.919**	**0.878**	**0.939**	**0.970**	**0.844**	**0.953**	**0.970**	**0.496**
Chinese Text Generation															
–	–	–	Base	0.966	0.983	13.06	0.554	0.851	–	–	0.955	0.853	0.931	0.945	0.198
✓	–	–	Flow-GRPO	0.971	0.985	11.84	0.569	0.849	–	–	0.962	0.866	0.935	0.946	0.213
✓	✓	–	Flow-GRPO	0.975	0.988	10.48	0.621	0.854	–	–	0.963	0.867	0.944	0.951	0.222
✓	✓	✓	Flow-GRPO	0.972	0.987	11.434	0.610	0.852	–	–	0.962	0.867	0.941	0.949	0.218
✓	✓	✓	**SCANNER**	**0.980**	**0.992**	**8.56**	**0.691**	**0.856**	–	–	**0.975**	**0.892**	**0.956**	**0.970**	**0.403**

Table 3: **Ablation of Credit Assignment on Qwen-Image.** Global-only and regional-only advantages against their noise-aware combination. The best results in each language are marked in **bold**, and "–" marks unsupported languages.

Advantage Assignment	LongText-Bench	OneIG-Bench				CVTG-2k		GenTextEval-Bench		PolyText-Bench		
	TextScore ↑	TextScore ↑	ED ↓	CR ↑	WAC ↑	WordAcc ↑	NED ↑	Qua ↑	Sem ↑	CS ↑	CG ↑	OG ↑
English Text Generation												
Global-only	0.960	0.991	18.821	0.510	0.901	0.846	0.922	0.945	0.803	0.904	0.913	0.287
Regional-only	0.966	0.992	17.260	0.551	0.902	0.857	0.928	**0.970**	0.840	0.951	0.961	0.368
Noise-aware schedule	**0.972**	**0.995**	**13.436**	**0.585**	**0.919**	**0.878**	**0.939**	**0.970**	**0.844**	**0.953**	**0.970**	**0.496**
Chinese Text Generation												
Global-only	0.975	0.988	10.484	0.621	0.854	–	–	0.963	0.867	0.944	0.951	0.222
Regional-only	0.978	0.991	9.426	0.652	0.855	–	–	0.973	0.891	**0.956**	0.967	0.293
Noise-aware schedule	**0.980**	**0.992**	**8.561**	**0.691**	**0.856**	–	–	**0.975**	**0.892**	**0.956**	**0.970**	**0.403**

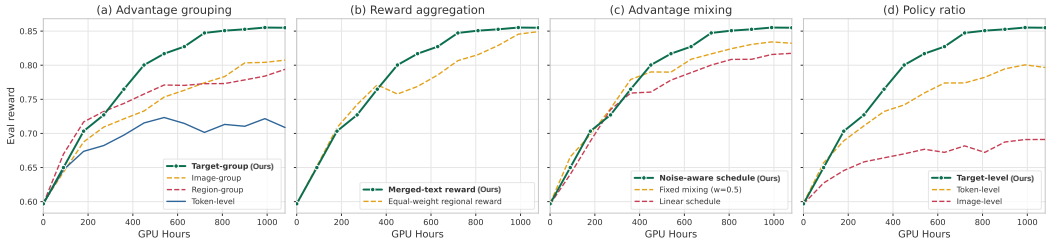


Figure 5: **Component Ablations of SCANNER.** Evaluation reward throughout training when varying (a) advantage grouping, (b) global reward aggregation, (c) global–regional advantage mixing, and (d) policy ratio granularity. Each curve varies a single component with all other settings fixed, and our design (solid green) attains the highest final reward in every comparison.

shows that regional credit dominates global credit, while their noise-aware fusion performs best by leveraging global feedback before locations stabilize. Thus, the noise-aware schedule in Fig. 5(c) outperforms fixed and linear alternatives. Finally, target-level policy ratios yield the highest final reward, whereas image-level clipping allows defective regions to suppress global updates and token-level ratios recover only partially (Fig. 5(d)). Additional ablation results are provided in A.3.3.

6 Conclusion

In this work, we address RL post-training for visual text generation, where recognition-based rewards overlook glyph defects and image-level credit obscures uneven text-region quality. Following the principle of blaming errant words rather than the whole canvas, we introduce TextDART, a defect-aware region-targeted reward model, and SCANNER, which assigns spatially aligned credit to each text target. Experiments show our method achieves better text accuracy and glyph fidelity than state-of-the-art baselines.

AI USE STATEMENT

In this work, we used generative AI tools for the required-disclosure task of generating synthetic data sets. Specifically, a language model expanded structured scene specifications into prompts, and open-source and closed-source generative models rendered synthetic images for TextDART-123K. Model-based pre-annotation was also used during dataset construction and subsequently verified through human annotation and two rounds of expert inspection. We have not used generative AI tools to develop theoretical models or conceptual frameworks, formulate mathematical claims, provide critical ingredients for proving mathematical claims, assist in the writing of proofs, propose or refine hypotheses, design or provide feedback on research methodology or experiments, implement methods, assist with translation, clean or reformat data sets, support qualitative or thematic data analysis, or interpret results; no other required-disclosure tasks are applicable to this work. Additionally, we used generative AI tools to edit and polish parts of the manuscript to improve readability. We have reviewed all AI-assisted work. Generated prompts and images were checked through the dataset annotation and expert-inspection pipeline described in Appendix A.7, and all AI-assisted text was reviewed by the authors to ensure that it preserved the intended technical meaning without altering claims, equations, citations, or reported results. We take responsibility for the final content of this work, including text, claims, and artifacts produced with the aid of generative AI.

ETHICS STATEMENT

This work involves a blind randomized user study with 10 participants, who compared anonymized outputs from different models according to text accuracy and aesthetics. The real-world portion of TextDART-123K was legally collected in-house, and all dataset entries underwent human annotation and expert inspection. A potential ethical concern is that improved visual-text generation could be misused to create misleading or deceptive text-bearing images. We encourage responsible use of the method in compliance with applicable laws and platform policies. We are not aware of other ethical concerns arising from this work.

REPRODUCIBILITY STATEMENT

We ensure reproducibility by summarizing the complete training procedure in Algorithm 1 and providing the matching, reward-computation, and credit-assignment details in Appendix A.1. Training and evaluation settings, baseline configurations, and hyperparameters are reported in Appendix A.2, particularly Table 4. The construction and metric computation of PolyText-Bench are described in Appendix A.6, while the construction, annotation protocol, and statistics of TextDART-123K are documented in Appendix A.7. The user-study protocol is described in Appendix A.3.2.

REFERENCES

Aesthetic predictor v2.5. https://github.com/discus0434/aesthetic-predictor-v2-5, 2025. Accessed: 2025-06-10.

Shuai Bai, Yuxuan Cai, Ruizhe Chen, Keqin Chen, Xionghui Chen, Zesen Cheng, Lianghao Deng, Wei Ding, Chang Gao, Chunjiang Ge, Wenbin Ge, Zhifang Guo, Qidong Huang, Jie Huang, Fei Huang, Binyuan Hui, Shutong Jiang, Zhaohai Li, Mingsheng Li, Mei Li, Kaixin Li, Zicheng Lin, Junyang Lin, Xuejing Liu, Jiawei Liu, Chenglong Liu, Yang Liu, Dayiheng Liu, Shixuan Liu, Dunjie Lu, Ruilin Luo, Chenxu Lv, Rui Men, Lingchen Meng, Xuancheng Ren, Xingzhang Ren, Sibo Song, Yuchong Sun, Jun Tang, Jianhong Tu, Jianqiang Wan, Peng Wang, Pengfei Wang, Qiuyue Wang, Yuxuan Wang, Tianbao Xie, Yiheng Xu, Haiyang Xu, Jin Xu, Zhibo Yang, Mingkun Yang, Jianxin Yang, An Yang, Bowen Yu, Fei Zhang, Hang Zhang, Xi Zhang, Bo Zheng, Humen Zhong, Jingren Zhou, Fan Zhou, Jing Zhou, Yuanzhi Zhu, and Ke Zhu. Qwen3-vl technical report. arXiv preprint arXiv:2511.21631, 2025.

Kevin Black, Michael Janner, Yilun Du, Ilya Kostrikov, and Sergey Levine. Training diffusion models with reinforcement learning. In International Conference on Learning Representations, volume 2024, pp. 4965–4987, 2024.

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. Advances in Neural Information Processing Systems, 38, 2026.

Chubin Chen, Sujie Hu, Jiashu Zhu, Meiqi Wu, Jintao Chen, Yanxun Li, Nisha Huang, Chengyu Fang, Jiahong Wu, Xiangxiang Chu, and Xiu Li. Taming preference mode collapse via directional decoupling alignment in diffusion reinforcement learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 12775–12786, June 2026.

Jingye Chen, Yupan Huang, Tengchao Lv, Lei Cui, Qifeng Chen, and Furu Wei. Textdiffuser: Diffusion models as text painters. Advances in Neural Information Processing Systems, 36:9353–9387, 2023.

Jingye Chen, Yupan Huang, Tengchao Lv, Lei Cui, Qifeng Chen, and Furu Wei. Textdiffuser-2: Unleashing the power of language models for text rendering. In European Conference on Computer Vision, pp. 386–402. Springer, 2024.

Cheng Cui, Yubo Zhang, Ting Sun, Xueqing Wang, Hongen Liu, Manhui Lin, Yue Zhang, Tingquan Gao, Changda Zhou, Jiaxuan Liu, Zelun Zhang, Jing Zhang, Jun Zhang, and Yi Liu. Pp-ocrv5: A specialized 5m-parameter model rivaling billion-parameter vision-language models on ocr tasks, 2026a. URL https://arxiv.org/abs/2603.24373.

Mingxuan Cui, Jingpu Yang, Fengxian Ji, Qian Jiang, Zhecheng Shi, Jiaming Wang, Zirui Song, Fajri Koto, and Xiuying Chen. Textalign: Preference alignment for text rendering with hierarchical rewards. arXiv preprint arXiv:2605.19320, 2026b.

Haoyou Deng, Keyu Yan, Chaojie Mao, Xiang Wang, Yu Liu, Changxin Gao, and Nong Sang. Densegrpo: From sparse to dense reward for flow matching model alignment. arXiv preprint arXiv:2601.20218, 2026.

Zheng Ding and Weirui Ye. Treegrpo: tree-advantage grpo for online rl post-training of diffusion models. In International Conference on Learning Representations, volume 2026, pp. 10877–10889, 2026.

Nikai Du, Zhennan Chen, Zhizhou Chen, Shan Gao, Xi Chen, Zhengkai Jiang, Jian Yang, and Ying Tai. Textcrafter: Accurately rendering multiple texts in complex visual scenes. arXiv e-prints, pp. arXiv–2503, 2025.

Zigang Geng, Yibing Wang, Yeyao Ma, Chen Li, Yongming Rao, Shuyang Gu, Zhao Zhong, Qinglin Lu, Han Hu, Xiaosong Zhang, et al. X-omni: Reinforcement learning makes discrete autoregressive image generative models great again. arXiv preprint arXiv:2507.22058, 2025.

Google. Gemini 3.1 Pro preview, 2026. URL https://ai.google.dev/gemini-api/docs/models/gemini-3.1-pro-preview. Accessed: 2026-09-20.

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. In International Conference on Learning Representations, volume 2026, pp. 137704–137724, 2026.

Xiaoxuan He, Siming Fu, Yuke Zhao, Wanli Li, Jian Yang, Dacheng Yin, Fengyun Rao, and Bo Zhang. Tempflow-grpo: When timing matters for grpo in flow models. In International Conference on Learning Representations, volume 2026, pp. 132869–132895, 2026.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. Advances in neural information processing systems, 33:6840–6851, 2020.

Yuval Kirstain, Adam Polyak, Uriel Singer, Shahbuland Matiana, Joe Penna, and Omer Levy. Pick-a-pic: An open dataset of user preferences for text-to-image generation. Advances in neural information processing systems, 36:36652–36663, 2023.

Mo Kordzanganeh, Danial Keshvary, and Nariman Arian. Pixel-wise rl on diffusion models: Reinforcement learning from rich feedback. arXiv preprint arXiv:2404.04356, 2024.

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

Junzhe Li, Yutao Cui, Tao Huang, Weijie Kong, Yiming Cheng, Chuxuan Zeng, Yinping Ma, Chun Fan, Miles Yang, Zhao Zhong, et al. Mixgrpo: Unlocking flow-based grpo efficiency with mixed ode-sde. arXiv preprint arXiv:2507.21802, 2025a.

Yanyu Li, Xian Liu, Anil Kag, Ju Hu, Yerlan Idelbayev, Dhritiman Sagar, Yanzhi Wang, Sergey Tulyakov, and Jian Ren. Textcraftor: Your text encoder can be image quality controller. In 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 7985–7995. IEEE, 2024.

Yuming Li, Yikai Wang, Zhongyu Zhao, Ming Lu, Qi She, Shanghang Zhang, et al. Branch-grpo: Stable and efficient grpo with structured branching in diffusion models. In International Conference on Learning Representations, volume 2026, pp. 20839–20866, 2026.

Zhang Li, Yuliang Liu, Qiang Liu, Zhiyin Ma, Ziyang Zhang, Shuo Zhang, Biao Yang, Zidun Guo, Jiarui Zhang, Xinyu Wang, et al. Monkeyocr: Document parsing with a structure-recognition-relation triplet paradigm. arXiv preprint arXiv:2506.05218, 2025b.

Youwei Liang, Junfeng He, Gang Li, Peizhao Li, Arseniy Klimovskiy, Nicholas Carolan, Jiao Sun, Jordi Pont-Tuset, Sarah Young, Feng Yang, et al. Rich human feedback for text-to-image generation. In 2024 ieee/cvf conference on computer vision and pattern recognition (cvpr), pp. 19401–19411. IEEE, 2024.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling, 2023. URL https://arxiv.org/abs/2210.02747.

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. Advances in neural information processing systems, 38:40783–40818, 2026a.

Shih-Yang Liu, Xin Dong, Ximing Lu, Shizhe Diao, Peter Belcak, Mingjie Liu, Min-Hung Chen, Hongxu Yin, Yu-Chiang Frank Wang, Kwang-Ting Cheng, et al. Gdpo: Group reward-decoupled normalization policy optimization for multi-reward rl optimization. arXiv preprint arXiv:2601.05242, 2026b.

Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow. arXiv preprint arXiv:2209.03003, 2022.

Yifu Luo, Haoyuan Sun, Xinhao Hu, Penghui Du, Keyu Fan, Bo Li, Sinan Du, Xu Wan, Zhiyu Chen, Bo Xia, et al. Principled rl for flow matching emerges from the chunk-level policy optimization. arXiv preprint arXiv:2510.21583, 2025.

Jian Ma, Mingjun Zhao, Chen Chen, Ruichen Wang, Di Niu, Haonan Lu, and Xiaodong Lin. Glyphdraw: Learning to draw chinese characters in image synthesis models coherently. arXiv preprint arXiv:2303.17870, 2(3):4, 2023.

Ziqi Ni, Yuanzhi Liang, Rui Li, Yi Zhou, Haibin Huang, Chi Zhang, and Xuelong Li. Seeing what matters: Visual preference policy optimization for visual generation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 27260–27269, 2026.

OpenAI. GPT-5.5 Instant: Smarter, clearer, and more personalized, 2026. URL https://openai.com/index/gpt-5-5-instant/. Accessed: 2026-09-20.

Zhuohan Ouyang, Zhe Qian, Wenhuo Cui, and Chaoqun Wang. Region-constrained group relative policy optimization for flow-based image editing. arXiv preprint arXiv:2604.09386, 2026.

Alexander Pan, Kush Bhatia, and Jacob Steinhardt. The effects of reward misspecification: Mapping and mitigating misaligned models. arXiv preprint arXiv:2201.03544, 2022.

Bowen Ping, Chengyou Jia, Minnan Luo, Hangwei Qian, and Ivor Tsang. Flow-factory: A unified framework for reinforcement learning in flow-matching models. arXiv preprint arXiv:2602.12529, 2026. URL https://arxiv.org/abs/2602.12529.

Dustin Podell, Zion English, Kyle Lacey, Andreas Blattmann, Tim Dockhorn, Jonas Müller, Joe Penna, and Robin Rombach. Sdxl: Improving latent diffusion models for high-resolution image synthesis. In International Conference on Learning Representations, volume 2024, pp. 1862–1874, 2024.

Kazi Mahathir Rahman, Showrin Rahman, and Sharmin Sultana Srishty. Textdiffuser-rl: Efficient and robust text layout optimization for high-fidelity text-to-image synthesis. arXiv preprint arXiv:2505.19291, 2025.

Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In 2022 IEEE/CVF conference on computer vision and pattern recognition (CVPR), pp. 10674–10685. ieee, 2022.

Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Xiao Bi, Haowei Zhang, Mingchuan Zhang, YK Li, Yang Wu, et al. Deepseekmath: Pushing the limits of mathematical reasoning in open language models. arXiv preprint arXiv:2402.03300, 2024.

Jinjie Shen, Wei Deng, Xian Hu, Daiguo Zhou, and Jian Luan. Star: Spatiotemporal adaptive reward allocation for text-to-image rl post-training. arXiv preprint arXiv:2606.17979, 2026.

Joar Skalse, Nikolaus Howe, Dmitrii Krasheninnikov, and David Krueger. Defining and characterizing reward gaming. Advances in neural information processing systems, 35:9460–9471, 2022.

Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. arXiv preprint arXiv:2011.13456, 2020.

Yunze Tong, Mushui Liu, Canyu Zhao, Wanggui He, Shiyi Zhang, Hongwei Zhang, Peng Zhang, Jinlong Liu, Ju Huang, Jiamang Wang, et al. Alleviating sparse rewards by modeling step-wise and long-term sampling effects in flow-based grpo. arXiv preprint arXiv:2602.06422, 2026.

Yuxiang Tuo, Wangmeng Xiang, Jun-Yan He, Yifeng Geng, and Xuansong Xie. Anytext: Multilingual visual text generation and editing. In International Conference on Learning Representations, volume 2024, pp. 56783–56799, 2024.

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 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 8228–8238. IEEE, 2024.

Jing Wang, Jiajun Liang, Jie Liu, Henglin Liu, Gongye Liu, Jun Zheng, Wanyuan Pang, Ao Ma, Zhenyu Xie, Xintao Wang, Meng Wang, Pengfei Wan, and Xiaodan Liang. Grpo-guard: Mitigating implicit over-optimization in flow matching via regulated clipping, 2025. URL https://arxiv.org/abs/2510.22319.

Haoran Wei, Chenglong Liu, Jinyue Chen, Jia Wang, Lingyu Kong, Yanming Xu, Zheng Ge, Liang Zhao, Jianjian Sun, Yuang Peng, et al. General ocr theory: Towards ocr-2.0 via a unified end-to-end model. arXiv preprint arXiv:2409.01704, 2024.

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, 2025. URL https://arxiv.org/abs/2508.02324.

Xiaoying Xing, Avinab Saha, Junfeng He, Susan Hao, Paul Vicol, Moonkyung Ryu, Gang Li, Sahil Singla, Sarah Young, Yinxiao Li, et al. Focus-n-fix: Region-aware fine-tuning for text-to-image generation. In 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 18486–18496. IEEE, 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. Advances in Neural Information Processing Systems, 36:15903–15935, 2023.

Shaodong Xu, Zexian Li, Zhendong Wang, Litong Gong, Tiezheng Ge, Wengang Zhou, Bo Zheng, and Houqiang Li. Edit-grpo: A locality-preserving policy optimization framework for image editing. arXiv preprint arXiv:2605.16951, 2026.

Zeyue Xue, Jie Wu, Yu Gao, Fangyuan Kong, Lingting Zhu, Mengzhao Chen, Zhiheng Liu, Wei Liu, Qiushan Guo, Weilin Huang, et al. Dancegrpo: Unleashing grpo on visual generation. arXiv preprint arXiv:2505.07818, 2025.

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. arXiv preprint arXiv:2606.26872, 2026.

Yukang Yang, Dongnan Gui, Yuhui Yuan, Weicong Liang, Haisong Ding, Han Hu, and Kai Chen. Glyphcontrol: Glyph conditional control for visual text generation. Advances in Neural Information Processing Systems, 36:44050–44066, 2023.

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. arXiv preprint arXiv:2606.06113, 2026.

Hanshen Zhu, Yuliang Liu, Xuecheng Wu, An-Lan Wang, Hao Feng, Dingkang Yang, Chao Feng, Can Huang, Jingqun Tang, and Xiang Bai. Textpecker: Rewarding structural anomaly quantification for enhancing visual text rendering. arXiv preprint arXiv:2602.20903, 2026.

A Appendix

A.1 Algorithmic Details

Many-to-many Matching. Detected regions and prompt targets rarely stand in one-to-one correspondence: a target may span several regions, one region may contain several targets, and a region may render text that no target requests. The natural one-to-one baseline assigns each region to its single most similar target,

$$j^*_{i,o} = \arg \max_{1 \le j \le K} \left[1 - \mathrm{GNED}(t_{i,o}, y_j)\right], \tag{12}$$

where K denotes the number of targets in the prompt and $j^*_{i,o}$ the target selected for region o of image i. This single-target-per-region rule fails in the last two cases. First, an extra-text region is forced onto its closest target, corrupting that target's transcription and reward with content the prompt never requested. Second, a region covering several targets can be credited to only one of them. If a region reads *SUNRISE CAFE* for the targets *SUNRISE* and *CAFE*, greedy matching scores the region against one target with semantic score 0.5 (reward 0.75 at $w_d = 0.5$), while the other target receives zero reward despite being correctly rendered. We therefore match regions and targets in a many-to-many manner, allowing one region to correspond to multiple targets, multiple regions to contribute to one target, and regions without any counterpart to remain unmatched as extra text.

We match readable text as sequences of Chinese characters and lowercased alphanumeric runs after removing anomaly markers, and region o is initially associated with

$$\begin{aligned} \mathcal{J}^{(0)}_{i,o} = \mathcal{C}_{i,o} \cup \mathcal{F}_{i,o} \cup \{j : \mathcal{U}(t_{i,o}) \cap \mathcal{U}(y_j) \neq \varnothing,\ \mathrm{GNED}(t_{i,o}, y_j) \le 0.5\} \\ \cup \{j : \mathrm{GNED}(t_{i,o}, y_j) \le 0.4\}, \end{aligned} \tag{13}$$

where $\mathcal{C}_{i,o}$ collects the targets occurring contiguously within the region (coverage), $\mathcal{F}_{i,o}$ the shortest target containing the region as a contiguous fragment, and $\mathcal{U}(\cdot)$ the readable units. The last two terms tolerate spelling and recognition errors that defeat exact containment. Regions that remain unmatched are attached to their nearest matched neighbor when their transcription occurs in, or flanks, that neighbor's target, recovering line-broken fragments, while regions without readable units stay unmatched.

Merging targets independently per image would make the reward of the same target index refer to different content across samples. We therefore pool coverage relations over all G images of the prompt and merge through connected components,

$$\mathcal{H} = \bigcup_{i=1}^{G} \bigcup_{o=1}^{O_i} \big\{\{j, j'\} : j, j' \in \mathcal{C}_{i,o},\ j < j'\big\}, \qquad \{\mathcal{T}_m\}_{m=1}^{M} = \mathrm{CC}\big(\{1, \dots, K\}, \mathcal{H}\big), \tag{14}$$

where $\mathcal{T}_m$ contains the original-target indices of merged target m and CC denotes connected components over the target-pair edges $\mathcal{H}$. Only coverage creates merging edges, while fragment and similarity matches associate regions without forcing a merge. If one image renders two targets in a single region and another renders them separately, both are thus scored against the same concatenated target, which the target-group advantages of SCANNER require. The ablation in Fig. 7(a) confirms that this matching improves training over the greedy variant.

Detailed Reward Computation Given the shared partition, each merged target collects its detections in image i,

$$\begin{aligned} \mathcal{O}_{i,m} &= \{o : \mathcal{J}_{i,o} \cap \mathcal{T}_m \neq \varnothing\}, & \bar{y}_m &= \mathrm{join}_{j \in \mathcal{T}_m}\, y_j, \\ \mathcal{Q}_{i,m} &= \bigcup_{o \in \mathcal{O}_{i,m}} q_{i,o}, & \bar{t}_{i,m} &= \mathrm{join}_{o \in \mathcal{O}_{i,m}}\, t_{i,o}, \end{aligned} \tag{15}$$

where $\mathcal{J}_{i,o}$ denotes the final matched-target indices of region o, $\mathcal{O}_{i,m}$ the detections contributing to merged target m, and $\mathcal{Q}_{i,m}$ their spatial union. The operator join concatenates in target order for $\bar{y}_m$ and in spatial reading order for $\bar{t}_{i,m}$, skipping transcriptions that duplicate a nested detection.

Following Zhu et al. (2026), the semantic score compares the collected transcription with the merged target through GNED, which tokenizes text into Chinese characters and lowercased non-Chinese

runs and solves a unit-level assignment,

$$\mathrm{GNED}(x,y) = \frac{\min\limits_{\pi\in\Pi_{n_x,n_y}} \sum\limits_{(v,w)\in\pi} \dfrac{\mathrm{ED}(x_v,y_w)}{\max(|x_v|,|y_w|)} + |n_x - n_y|}{\max(n_x,n_y)}, \tag{16}$$

where x and y denote the tokenized transcription and target with n_x and n_y units, Π_{n_x,n_y} the one-to-one unit matchings of cardinality $\min(n_x,n_y)$, and ED the character-level edit distance, with unmatched units incurring unit cost. The minimum is computed with the Hungarian algorithm. Since GNED is insensitive to the order of whole units, it measures textual agreement rather than reading order.

Recognizable text can still contain malformed glyphs, so semantic agreement alone is insufficient. The defect-free score counts the structural-anomaly markers with severity-dependent weights,

$$a(x) = n_{\mathrm{char}}(x) + 3\,n_{\mathrm{word}}(x) + 5\,n_{\mathrm{block}}(x) + n_{\mathrm{bare}}(x), \qquad L(x) = L_{\mathrm{read}}(x) + a(x), \tag{17}$$

where n_{char}, n_{word}, and n_{block} count the character-, word-, and block-level markers <#>, <###>, and <######>, n_{bare} the remaining bare # symbols, and L_{read} the non-whitespace characters after marker removal. Substituting into Eq. 5 gives $d_{i,m} = d(\bar{t}_{i,m})$, with $d(x) = 0$ when $L(x) = 0$, so a transcription consisting entirely of defect markers receives zero defect-free score at the training setting $\kappa = 1$.

Extra text has no prescribed content against which to measure semantic correctness, while a missing target has a reference but no detected location. We retain both cases distinctly. Unmatched detections form the extra entry,

$$\mathcal{O}_{i,\mathrm{ext}} = \{1,\dots,O_i\} \setminus \bigcup_{m=1}^{M} \mathcal{O}_{i,m}, \qquad (\bar{q}_{i,\mathrm{ext}}, \bar{t}_{i,\mathrm{ext}}) = \Big(\bigcup_{o\in\mathcal{O}_{i,\mathrm{ext}}} q_{i,o},\ \mathrm{join}_{o\in\mathcal{O}_{i,\mathrm{ext}}}\, t_{i,o} \Big), \tag{18}$$

and an unobserved target borrows its location from peer images of the same prompt,

$$\bar{q}_{i,m} = \begin{cases} \mathcal{Q}_{i,m}, & \mathcal{O}_{i,m} \neq \varnothing, \\ \mathrm{bbox}\Big(\bigcup_{i'\neq i} \mathcal{Q}_{i',m}\Big), & \mathcal{O}_{i,m} = \varnothing, \end{cases} \tag{19}$$

where bbox denotes the smallest enclosing rectangle, and the inferred region excludes tokens already occupied by observed regions when projected onto the latent grid. The reward entries are then

$$r_{i,m} = \begin{cases} (1-w_d)s_{i,m} + w_d d_{i,m}, & \mathcal{O}_{i,m} \neq \varnothing, \\ 0, & \mathcal{O}_{i,m} = \varnothing, \end{cases} \qquad r_{i,\mathrm{ext}} = d(\bar{t}_{i,\mathrm{ext}}) \ \text{ only if } \mathcal{O}_{i,\mathrm{ext}} \neq \varnothing, \tag{20}$$

with the spatially available entries

$$\mathcal{M}_i = \{m : \bar{q}_{i,m} \neq \varnothing\} \cup \{\mathrm{ext} : \mathcal{O}_{i,\mathrm{ext}} \neq \varnothing\}. \tag{21}$$

A target absent from every sampled image has no inferred region and is excluded from spatial assignment. These distinctions let omitted targets and malformed extra text receive localized feedback without penalizing the entire canvas.

SCANNER Algorithm We summarize the complete training procedure in Algorithm 1. For each prompt with specified target texts, we first sample a group of G images with the SDE sampler and record the transition noise n_k of each optimized step. TextDART then localizes text regions with defect-aware transcriptions, which are matched to the target texts and scored into regional rewards $r_{i,m}$ and a global reward R_i. Subsequently, the regional rewards are normalized within each target group and projected onto the latent grid, while the global reward is normalized within the prompt, and the two advantages are fused at each step with the noise-aware weight ω_k. Finally, we pool the Guard-adjusted ratio exponent over the token support of each target and update the policy with the clipped objective and the background-weighted KL regularization.

Algorithm 1 SCANNER Training Process

Require: Reference policy v^{ref}, training policy v_θ initialized from v^{ref};
Require: Reward model TextDART, prompt set $\mathcal{C}$ with specified target texts;
Require: Group size G, optimized SDE steps $\mathcal{S}$, regional scale λ, regularization weight β.
Initialize: Data collection policy $v^{\mathrm{old}} \leftarrow v^{\mathrm{ref}}$.

for each training iteration **do**
  Sample a batch of prompts with target texts from $\mathcal{C}$
  // Rollout and defect-aware regional reward
  for each prompt in the batch **do**
    Sample G images $\{I_i\}_{i=1}^{G}$ with the SDE sampler (Eq. 1), recording n_k for $k \in \mathcal{S}$
    Predict regions and defect-aware transcriptions $\{(q_{i,o}, t_{i,o})\}_{o=1}^{O_i}$ by TextDART (Eq. 4)
    Match regions to targets and compute regional rewards $r_{i,m}$ (Eq. 5) and global reward R_i (Eq. 8)
  end for
  // Target-group advantages.
  Normalize rewards within each target group into $A_{i,m}$ (Eq. 6) and project onto the latent grid as $A_i^{\mathrm{reg}}(p)$ (Eq. 7)
  // Noise-aware Schedule.
  Normalize R_i within the prompt into A_i^{glob} and fuse $A_{i,k}(p) = \omega_k A_i^{\mathrm{glob}} + (1-\omega_k) A_i^{\mathrm{reg}}(p)$ (Eq. 9)
  // Target-level policy ratios.
  for each image i and step $k \in \mathcal{S}$ **do**
    Compute Guard-adjusted exponents $z_{i,k,p}$ and pool them into target ratios $\rho_{i,k,m}$ (Eq. 10)
  end for
  Update θ by maximizing $\mathcal{J}_{\mathrm{SCANNER}}(\theta)$ (Eq. 11)
  $v^{\mathrm{old}} \leftarrow v_\theta$
end for
Output: Optimized policy v_θ

A.2 Experimental Setting Details

A.2.1 Hyperparameters

We fine-tune both backbones with LoRA on 32 H20 GPUs across four nodes, using DeepSpeed ZeRO-2 and gradient checkpointing. Table 4 lists all settings. Samples for the same prompt share initial noise, with independent perturbations during the selected Flow-SDE transitions. Qwen uses the official negative prompt during training and evaluation, while FLUX uses none. Our setup otherwise follows Zhu et al. (2026).

A.2.2 Details of Comparison with Baselines

We reproduce the RL baselines with the open-source Flow-Factory framework (Ping et al., 2026), starting from the same pretrained checkpoint for each backbone. All trained methods share the same training prompts, LoRA configuration, optimizer settings, resolution, sampling batch size, group size, and SDE window. We train Qwen-Image-2512 for 320 steps and FLUX.1-dev for 360 steps, requiring approximately 1,920 and 1,080 GPU hours, respectively.

Baseline Implementations. The base model is evaluated without RL fine-tuning. For OCRv5 with Flow-GRPO, we compute an image-level reward from the normalized edit distance between PP-OCRv5 transcriptions and the target text. For TextPecker with Flow-GRPO, we combine its semantic and defect scores with equal weights and a defect-penalty scale of 1. Both RL baselines use Fast Flow-GRPO with GRPO-Guard, and reference regularization uses the backbone-specific coefficients in Table 4 with uniform spatial weighting.

Evaluation Protocol. All methods, including the pretrained model, are evaluated on the same benchmark prompts with the per-backbone evaluation settings in Table 4. FLUX is evaluated in English and Qwen in English and Chinese, except on the English-only CVTG-2k. Each benchmark

Table 4: **Training and Evaluation Hyperparameters.** Settings for TextDART with SCANNER.

Parameter	Qwen-Image-2512	FLUX.1-dev
Model and optimization		
Training language	English / Chinese	English
Resolution	1024×1024	1024×1024
LoRA rank / alpha	64 / 128	64 / 128
Mixed precision	BF16	FP16
Optimizer	AdamW	AdamW
Learning rate	10^{-4}	10^{-4}
Weight decay	10^{-4}	10^{-4}
Adam betas	$(0.9, 0.999)$	$(0.9, 0.999)$
Adam epsilon	10^{-8}	10^{-8}
Maximum gradient norm	1.0	1.0
Per-device training batch size	4	3
EMA decay	Disabled	0.9
Random seed	42	42
Sampling and evaluation		
Fast Flow-GRPO	Enabled	Enabled
Rollout steps	20	28
Evaluation steps	50	28
Guidance scale (training / evaluation)	4.0 / 4.0	3.5 / 3.5
Samples per prompt	16	24
Flow-SDE noise level	1.2	0.9
SDE window size	5	9
Reward and spatial credit assignment		
Defect mixing weight w_d	0.5	0.5
Defect-penalty scale κ	1.0	1.0
Regional scale λ	0.5	0.5
Target reward standard-deviation floor	10^{-4}	10^{-4}
Advantage clipping bounds	$[-5, 5]$	$[-5, 5]$
GRPO-Guard	Enabled	Enabled
Policy clipping threshold	10^{-4}	10^{-4}
Reference-penalty coefficient	10^{-3}	10^{-4}
Reference weight (text / background)	1 / 2	1 / 2

uses a common evaluator and metric configuration across methods, irrespective of the training reward. On PolyText-Bench, we generate four images per prompt and set the evaluation defect-penalty scale to 5.

A.3 Extended Experiments

A.3.1 Evaluation of TextDART

We evaluate TextDART on the test set of TextDART-123K (details in A.7), where every method receives only the complete image. The first two tasks follow the evaluation protocol of Zhu et al. (2026): Canonical Text Recognition (CTR) measures the recovery of legible content through unit-level recall and normalized edit distance over all images, including defective ones, while Text Structural Anomaly Perception (TSAP) measures whether malformed text is explicitly identified, using micro precision, recall, and F_1 over counted anomaly markers, with all methods rescored under this common protocol. We additionally introduce Text Region Localization (TRL), which compares the union of predicted and ground-truth polygons at IoU thresholds of 0.5 and 0.75, accommodating different box granularities. We compare against specialist OCR systems (Cui et al., 2026a; Wei et al., 2024; Li et al., 2025b), the structure-aware recognizer (Zhu et al., 2026) and multimodal models (Bai et al., 2025; OpenAI, 2026; Google, 2026). GOT-OCR-2.0 and TextPecker produce only transcriptions, so their localization entries are unavailable. TextDART achieves the best results on every metric in Tab. 5. The CTR gap over the strongest proprietary models is moderate, as general models already read clean text well. The decisive gap lies in defect perception: OCR

Table 5: **Evaluation of Text Recognizers on Canonical Text Recognition (CTR), Text Structural Anomaly Perception (TSAP), and Text Region Localization (TRL).** The best results are marked in **bold**. Results of non-supporting models are marked as –.

Method	CTR		TSAP			TRL					
	R↑	NED↓	P↑	R↑	F1↑	P@.5↑	R@.5↑	F1@.5↑	P@.75↑	R@.75↑	F1@.75↑
English											
PP-OCRv5	0.7967	0.1191	0.0000	0.0000	0.0000	0.9070	0.9025	0.9048	0.6281	0.6250	0.6266
MonkeyOCR	0.8637	0.0788	0.0000	0.0000	0.0000	0.7262	0.5900	0.6510	0.4338	0.3525	0.3890
GOT-OCR-2.0	0.6642	0.1903	0.0000	0.0000	0.0000	–	–	–	–	–	–
Gemini-3.1-Pro-Preview	0.8948	0.0523	0.5909	0.3588	0.4465	0.8433	0.8475	0.8454	0.4428	0.4450	0.4439
GPT-5.5	0.8645	0.0669	0.5803	0.1680	0.2606	0.7487	0.7450	0.7469	0.4271	0.4250	0.4261
TextPecker	0.8943	0.0600	0.7509	0.4862	0.5902	–	–	–	–	–	–
Qwen3-VL-8B	0.8941	0.0576	0.0769	0.0005	0.0010	0.8650	0.8650	0.8650	0.5875	0.5875	0.5875
TextDART (Ours)	**0.9346**	**0.0362**	**0.7799**	**0.7874**	**0.7836**	**0.9750**	**0.9750**	**0.9750**	**0.9225**	**0.9225**	**0.9225**
Chinese											
PP-OCRv5	0.9530	0.0407	0.0000	0.0000	0.0000	0.9279	0.9217	0.9247	0.5638	0.5600	0.5619
MonkeyOCR	0.9363	0.0589	0.0000	0.0000	0.0000	0.7500	0.5650	0.6445	0.4624	0.3483	0.3973
GOT-OCR-2.0	0.8939	0.0929	0.0000	0.0000	0.0000	–	–	–	–	–	–
Gemini-3.1-Pro-Preview	0.9621	0.0343	0.5841	0.3219	0.4150	0.8781	0.8883	0.8832	0.5568	0.5633	0.5601
GPT-5.5	0.8115	0.1837	0.3626	0.1480	0.2102	0.7880	0.7683	0.7781	0.4342	0.4233	0.4287
TextPecker	0.8814	0.1141	0.7974	0.7123	0.7525	–	–	–	–	–	–
Qwen3-VL-8B	0.9502	0.0458	0.6000	0.0021	0.0041	0.9228	0.9167	0.9197	0.6946	0.6900	0.6923
TextDART (Ours)	**0.9639**	**0.0342**	**0.9033**	**0.7713**	**0.8321**	**0.9933**	**0.9917**	**0.9925**	**0.9399**	**0.9383**	**0.9391**

systems score zero on TSAP, as they cannot report malformed glyphs at all. Even the best baseline TextPecker trails by 0.19 and 0.08 F_1 in English and Chinese, and the unfine-tuned backbone scores nearly zero, showing that defect awareness comes from TextDART-123K fine-tuning rather than the backbone itself. The localization margin also widens at the stricter IoU 0.75, indicating precise boundaries rather than coarse coverage. These results validate TextDART as a reliable reward source for region-level credit assignment.

A.3.2 User Study

We conduct a blind randomized user study with 10 participants to complement the automatic evaluation with human judgments. For each sample, outputs from TextDART + SCANNER, TextPecker + Flow-GRPO, OCRv5 + Flow-GRPO, and the Qwen-Image-2512 base model are anonymized and presented in a randomly shuffled order. Participants are unaware of the model identities and independently select the best result according to two criteria: **text accuracy**, which evaluates whether the rendered text is correct and legible, and **text aesthetics**, which evaluates the visual quality and naturalness of the rendered text. We report the average win rate aggregated over all samples and annotators.

As shown in Table 6, TextDART + SCANNER receives the highest human preference on both dimensions, achieving a win rate of **46.8%** for text accuracy and **39.4%** for text aesthetics. Notably, OCRv5 + Flow-GRPO receives substantially lower human preference despite explicitly optimizing recognition-based rewards, suggesting that improved recognizer scores do not necessarily translate into better perceived text rendering quality.

Table 6: **Blind Randomized User Study on Visual Text Generation.** We report the average win rate across all samples and annotators. The best result in each row is highlighted in bold.

Metric	Qwen-Image-2512	OCRv5 + Flow-GRPO	TextPecker + Flow-GRPO	TextDART + SCANNER
Text Accuracy ↑	21.30%	9.40%	22.50%	**46.80%**
Text Aesthetics ↑	24.80%	13.10%	22.70%	**39.40%**

A.3.3 Additional Ablation Study



Figure 6: **Ablation on Spatially Weighted KL Regularization.** Global KL leaves residual misspellings, text-free penalties cause hallucinated text, and equal weights induce reward hacking (covering text with plain blocks, row 3). Our background-weighted variant (R:G = 1:2) renders all target text correctly while preserving the scene.

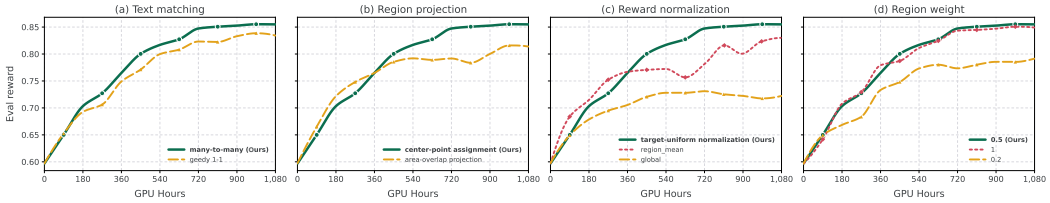


Figure 7: **Supplementary Ablations.** Training curves of evaluation reward for (a) region–target matching, (b) image-to-latent projection, (c) normalization of regional contributions, and (d) the regional scale λ. Our choices (green) converge the highest in all four comparisons.

Spatially Weighted KL Regularization. Fig. 6 compares global KL regularization with spatially weighted variants of different text-region (R) and background (G) weights. Global KL applies the penalty uniformly, which counteracts the text reward on text tokens and leaves residual misspellings (rows 1 and 4). Removing the penalty on text regions (R:G = 0:1 and 0:2) frees text optimization but weakens the anchor on content, leading to hallucinated text such as the extra letter and the garbled sign in the second row. With equal weights (R:G = 1:1), reward hacking emerges, where the model covers text regions with plain blocks to escape defect penalties (third row). Our choice (R:G = 1:2) keeps text regions free to improve while anchoring the background more strongly, yielding correct text and faithful scenes without hacking behavior.

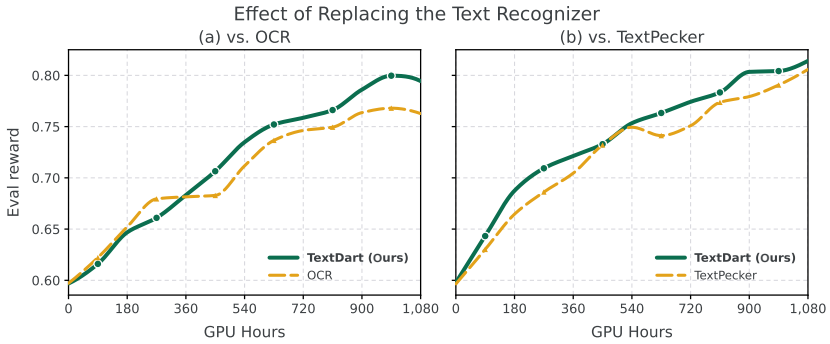


Figure 8: **Effect of Replacing the Text Recognizer.** Training curves with the downstream pipeline fixed. (a) The OCR reward saturates and declines late in training, while TextDART keeps improving. (b) TextDART also surpasses TextPecker with a steadier curve, showing the benefit of region-precise feedback.

Region–Target Matching. Fig. 7(a) compares region–target matching strategies. Greedy one-to-one matching plateaus well below ours, as it mismatches regions and targets whenever one region covers multiple targets or one target spans several regions, corrupting the regional rewards. Many-to-many matching with merging preserves the correct correspondence and reaches a consistently higher reward.

Projection of Regional Advantages onto Latent Tokens. Fig. 7(b) compares image-to-latent projection rules. Area-overlap projection rises slightly faster at first but plateaus with oscillation, as partially covered tokens at region boundaries receive advantages from neighboring content. Assigning each token by its center point keeps region boundaries clean and converges markedly higher.

Area Normalization of Regional Contributions. Fig. 7(c) compares normalization schemes for regional contributions. Without normalization (global), larger regions dominate the updates and small text remains under-corrected, yielding the lowest curve. Normalizing per detected region (region-mean) improves over it but stays unstable, since the weighting depends on how text is partitioned into regions rather than on semantic targets. Equalizing contributions per merged target (ours) is both stable and the highest.

Sensitivity to Regional Loss Weight. Fig. 7(d) varies the regional scale λ. Reducing it to 0.2 clearly slows learning, as regional signals become too weak against the global branch, while 0.5 and 1.0 behave almost identically with a slight edge for 0.5. Our method is thus robust to λ as long as regional credit is not under-weighted.

A.3.4 Effect of Replacing the Text Recognizer

Fig. 8 isolates the contribution of the text recognizer by replacing OCRv5 or TextPecker with TextDART, while keeping downstream reward computation and Flow-GRPO optimization unchanged. Against OCR (Fig. 8a), both variants learn similarly at first, but the OCR-based variant saturates and even declines late in training, as it misreads malformed glyphs and stops providing useful signal once most text is nearly correct, whereas TextDART keeps improving through defect-aware feedback. Against TextPecker (Fig. 8b), the gap is smaller since glyph defects are already penalized, yet TextDART remains consistently higher with a steadier curve, indicating that region-precise feedback optimizes more effectively than whole-image defect scoring.

A.3.5 SCANNER Generalization across Reward Models

Tab. 7 applies SCANNER to existing reward models. Both adaptations first localize text regions with PP-OCRv5 and then follow the same region-wise reward computation and SCANNER training procedure as TextDART. OCRv5 directly transcribes each detected region. TextPecker does not localize text from a full image; instead, we provide the full image together with each detected box and query it separately to obtain defect-aware transcriptions. Since each image incurs one request

Table 7: **Generalization of SCANNER across Reward Models.** Both adaptations follow the same region-wise reward computation paradigm as TextDART and the same subsequent SCANNER training procedure. The best results are marked in **bold**. Results for unsupported languages are marked as "–".

Reward Model	RL Algorithm	LongText- Bench	OneIG-Bench				CVTG-2k		GenTextEval-Bench		PolyText-Bench		
		TextScore ↑	TextScore ↑	ED ↓	CR ↑	WAC ↑	WordAcc ↑	NED ↑	Qua ↑	Sem ↑	CS ↑	CG ↑	OG ↑
OCRv5	Flow-GRPO	0.769	0.751	64.55	0.220	0.506	0.686	0.818	0.838	0.589	0.703	0.726	0.365
	SCANNER	0.809	0.771	66.285	0.230	**0.552**	**0.809**	**0.905**	**0.862**	0.591	0.797	0.840	0.368
TextPecker	Flow-GRPO	0.801	0.770	63.19	0.221	0.533	0.772	0.879	0.836	0.582	0.789	0.809	0.405
	SCANNER	**0.813**	**0.804**	**58.10**	**0.253**	0.548	0.807	0.901	0.840	**0.593**	**0.864**	**0.897**	**0.583**

Table 8: **Multi-reward Optimization.** Each text reward (OCRv5, TextPecker, or TextDART) is combined with the PickScore (Kirstain et al., 2023) and aesthetic (aes, 2025) rewards (PA) under its respective RL algorithm, and Base denotes the model without RL fine-tuning. Combining TextDART with PA under SCANNER achieves the best results on all text metrics in both languages while maintaining good visual appeal (Fig. 9), showing that superior text quality is attained together with, rather than at the cost of, general image quality. The best results within each language group are marked in **bold**, and – marks unsupported languages.

Variant	LongText-Bench	OneIG-Bench				CVTG-2k		GenTextEval-Bench		PolyText-Bench		
	TextScore ↑	TextScore ↑	ED ↓	CR ↑	WAC ↑	WordAcc ↑	NED ↑	Qua ↑	Sem ↑	CS ↑	CG ↑	OG ↑
English Text Generation												
Base	0.952	0.986	24.539	0.461	0.895	0.844	0.922	0.944	0.787	0.887	0.899	0.252
OCRv5 + PA / Flow-GRPO	0.951	0.992	18.804	0.507	0.906	0.848	0.924	0.950	0.796	0.880	0.889	0.262
TextPecker + PA / Flow-GRPO	0.957	0.992	17.210	0.505	0.903	0.849	0.925	0.945	0.796	0.880	0.883	0.212
TextDART + PA / SCANNER	**0.969**	**0.994**	**14.646**	**0.571**	**0.907**	**0.880**	**0.940**	**0.982**	**0.855**	**0.957**	**0.976**	**0.434**
Chinese Text Generation												
Base	0.966	0.983	13.059	0.554	0.851	–	–	0.955	0.853	0.931	0.945	0.198
OCRv5 + PA / Flow-GRPO	0.967	0.986	11.273	0.584	0.851	–	–	0.956	0.855	0.935	0.936	0.209
TextPecker + PA / Flow-GRPO	0.969	0.987	10.845	0.571	0.856	–	–	0.960	0.870	0.935	0.943	0.177
TextDART + PA / SCANNER	**0.983**	**0.993**	**8.209**	**0.706**	**0.857**	–	–	**0.975**	**0.897**	**0.958**	**0.973**	**0.414**

per detected box, this adaptation introduces additional inference overhead and is used only for this generalization study. SCANNER improves Flow-GRPO on most metrics under both rewards, with large gains on text-specific metrics such as WordAcc from 0.686 to 0.809 with OCRv5 and ED from 63.19 to 58.10 with TextPecker, showing that spatially aligned credit assignment is largely orthogonal to the choice of reward model. Two further observations stand out. First, the gains are generally larger with the defect-aware TextPecker than with OCRv5, whose ED slightly degrades, as regional credit can only amplify what the reward perceives and a recognition-only reward remains blind to glyph defects. Second, both adapted variants still fall short of the full combination of TextDART and SCANNER in Tab. 1, confirming that reward quality and credit assignment are complementary rather than interchangeable.

A.3.6 Multi-Reward Optimization

We further investigate whether TextDART can be combined with general-purpose rewards under SCANNER to jointly optimize textual fidelity and overall image quality. Following Liu et al. (2026b), we normalize each reward within its prompt group and sum the resulting advantages. This integrates naturally with SCANNER: the PickScore and aesthetic rewards (PA) are image-level, so their advantages are uniform over all latent tokens, while the text reward contributes its spatially aligned advantage map, and their sum enters the same target-level policy objective unchanged. We combine each text reward (OCRv5, TextPecker, or TextDART) with PA and perform RL fine-tuning.

As shown in Tab. 8, TextDART + PA with SCANNER achieves the strongest results on all metrics in both English and Chinese, outperforming OCRv5 + PA and TextPecker + PA under Flow-GRPO. Qualitative results in Fig. 9 further show that this combination renders text correctly while preserving visual quality and composition, indicating that TextDART with SCANNER serves as a flexible component in multi-reward optimization.



Figure 9: **Qualitative Results of Multi-reward Optimization on Qwen-Image-2512.** Each text reward is combined with the PickScore(Kirstain et al., 2023) and aesthetic(aes, 2025) rewards (PA). Red dashed boxes mark text errors. TextDART with SCANNER renders all target text correctly while producing the most appealing layout, color, and composition, whereas the baselines garble text or degrade the scene.

A.4 Additional Qualitative Results for Localization and Defect-Aware Recognition

Fig. 10 presents more cases of TextDART's region-level recognition, where each detected region is transcribed with defect markers and scored for glyph quality and semantic correctness.

A.5 Additional Qualitative Results for Visual Text Generation

Results on English Prompts. Fig. 11 presents additional English results on Qwen-Image (rows 1–3) and Flux (rows 4–6), where the trends observed in the main text hold throughout. Baseline errors concentrate on small secondary text, such as the wish plaque in row 1, the warning sign in row 2, and the street signs in row 3, where unrequested regions are filled with garbled strokes. The baselines also miss or duplicate content, rendering only a single letter as the poster title in row 4, producing incomplete signage in row 5, and repeating the venue line in row 6. Our method renders every requested target correctly across all rows while keeping spontaneous text free of garbled writing.

Table 9: **Comparison of Visual Text Generation Benchmarks.** The comparison concerns the released prompt sets and their original text-evaluation protocols. Multi-block prompts specify multiple spatially separated text blocks. Glyph-defect scoring explicitly assesses malformed glyphs beyond transcription accuracy. Text-optional evaluation assesses text that may emerge without prescribing its presence or content. For OneIG-Bench, we consider its text-rendering subset. ✓: included; ×: not included.

Benchmark	Multi-block prompts	English + Chinese	Glyph-defect scoring	Text-optional evaluation
LongText-Bench (Geng et al., 2025)	✓	✓	×	×
OneIG-Bench (Chang et al., 2026)	✓	✓	×	×
CVTG-2K (Du et al., 2025)	✓	×	×	×
GenTextEval (Zhu et al., 2026)	✓	✓	✓	×
PolyText-Bench (ours)	✓	✓	✓	✓

Table 10: **Construction Branches (left) and Image Sources (right) of the TextDART-123K Training Split.** Percentages use all 122,200 training entries. The graphic design collections category covers existing text-rich design images, including posters, logos, infographics, and artistic lettering.

Branch	Entries	Share
Controlled synthesis	18,856	15.4%
Caption regeneration	30,423	24.9%
Graphic design collections	26,915	22.0%
Long-tail	13,885	11.4%
Small-text	10,273	8.4%
Text-free	21,848	17.9%
Total	122,200	100.0%

Source	Entries	Share
Open-source models	47,828	39.1%
Closed-source models	49,284	40.4%
Real-world imagery	25,088	20.5%
Total	122,200	100.0%

Results on Chinese Prompts. Fig. 12 presents additional Chinese results on Qwen-Image. The backbone and the OCRv5 reward repeat the middle descriptive line several times in row 1, and all baselines garble the small price tags in row 2 and the long top banner in row 5. Rows 3 and 4 exhibit the same failure on spontaneous and secondary text, where the baselines append garbled small print or corrupt the call-to-action slogans, and row 6 contains wrong or repeated characters in the poster header. Our method renders every requested target correctly, from multi-line posters to dense magazine spreads, and produces no garbled extra text.

A.6 Additional Details on PolyText-Bench

Comparison with Existing Benchmarks. Table 9 compares PolyText-Bench with existing visual-text benchmarks. While multi-block prompts and, in most cases, both English and Chinese are already covered, glyph defects beyond transcription accuracy are scored only by GenTextEval, and none of them evaluates text that may emerge without being prescribed. PolyText-Bench is the only benchmark that covers all four aspects.

Prompt Construction. We build prompts from structured scene specifications that combine a physical setting, a text-bearing object, a theme, a composition, and a color palette, and expand them with a language model into full descriptions covering material, artwork, props, background, and lighting, so that text appears on functional objects in concrete scenes rather than as isolated lettering. Critical prompts additionally specify one to five target strings with roles such as titles, menu entries, product labels, and headings, together with their described locations and visual treatments. Optional prompts instead describe the object's ordinary real-world function, such as a published book offered for sale or a menu consulted by a diner, without prescribing any wording, so that text is a natural possibility rather than an explicit instruction.

Dataset Statistics. PolyText-Bench contains 600 prompts organized as 300 critical–optional pairs balanced between English and Chinese, where each pair shares a scene category and theme and differs in whether text is prescribed. The pairs span ten scene categories, including book and magazine

Table 11: **Construction Dimensions of the Controlled Prompts.** Distributions are recorded on the specifications of the controlled subset and listed in the order of the corresponding categories. *Single-line vs. multi-line regions. †Recorded over the first three categories.

Dimension	Categories	Distribution (%)
Scene	Street views, posters, infographics, etc. (17 in total)	5.8–5.9 each
Script	Chinese, English, mixed, digits and symbols	–
Text content	Frequency-balanced characters and lexical units, spanning short labels to long phrases	–
Composition	Single, multiple, and dense regions, 1–10 lines each	49.3 / 50.7*
Geometry	Horizontal, vertical, rotated, perspective	42.4 / 21.7 / 21.6 / 14.3
Typography	Standard, handwriting or calligraphy, bold, artistic lettering	39.9 / 40.0 / 20.1†
Scale	Large display, medium, small	30.1 / 39.9 / 30.0
Text presence	Text-containing, text-free	–

covers, posters, menus, product packaging, storefronts, tickets and programs, information boards, screen displays, and album and game covers, with category sizes ranging from 10 to 44 pairs and both languages represented in every category. The 300 critical prompts contain 783 target strings (404 English and 379 Chinese, 2.6 per prompt on average), with lengths of 3–76 characters for English and 1–19 for Chinese, counting literal characters including spaces and punctuation. Each prompt is assigned four fixed sampling seeds shared within its pair, yielding 2,400 images for a bilingual model and 1,200 for English-only evaluation.

Metric Computation. All metrics are computed from the outputs of TextDART, which takes only the image as input, keeping recognition independent of the target text. For critical images, detected regions are matched to the requested targets with the many-to-many matching of Sec. 4.1. CS averages the semantic score s and CG the defect-free score d of Eq. 5 over all requested targets, with unmatched targets scored zero so that omitting text cannot reduce the denominator. For optional images, OG averages d over the detected regions and then over images containing text, while text-free outputs are excluded rather than assigned a perfect score, so that avoiding text yields no advantage. All scores lie in $[0, 1]$, with higher values indicating better quality. Since TextDART also serves as the training reward of SCANNER, we additionally conduct a user study as independent human validation, detailed in A.3.2..

A.7 Details of TextDART-123K

Existing text-centric corpora cover mostly clean, high-frequency text and annotate only readable content, providing no supervision for defect-aware transcription. TextDART-123K is therefore built from complementary branches (Table 10), including controlled prompt synthesis, caption-based regeneration, graphic design collections, long-tail and small-text supplementation, and text-free negatives. The controlled prompts follow the coverage dimensions in Table 11. Images are rendered by multiple open-source and closed-source generators and are complemented by real-world imagery legally collected in-house, exposing rendering outcomes from clean text to severely distorted glyphs. Each image goes through OCR localization, region-wise TextPecker pre-annotation (Zhu et al., 2026), human annotation, and two rounds of expert inspection, totaling approximately 12,500 person-hours. Annotations record the text actually rendered rather than the prompt intent. Every visible text piece is enclosed by a quadrilateral, transcriptions are literal with three structural-anomaly markers, and a readable but incorrect word is retained as rendered, since semantic comparison is performed separately by the regional reward. Text-free images receive empty region lists to supervise the empty-output case. We reserve 1,000 text-containing entries as the test split and use the remaining 122,200 entries for training, with disjoint sample IDs. Table 12 lists the instruction template of TextDART, which mirrors the annotation protocol above. The template never exposes the requested target strings, consistent with the image-only assessment in Sec. 4.1. Table 13 summarizes the training split. Notably, 44.9% of training entries mix marked and unmarked regions, providing direct supervision for regional quality differences that image-level rewards obscure. The test split contains 1,000 text-containing entries with 4,383 annotated regions, balanced by design across language (600 Chinese and 400 English) and anomaly presence (half anomalous per language).

Table 12: **Instruction Template of TextDART** for defect-aware text localization and transcription.

Component	Content
Role	You are an expert in perceiving structural anomalies of rendered text in AI-generated images.
Task	Given a generated image, locate and transcribe every piece of visible text, including distorted, broken, or entirely unreadable text, and report each piece as a quadrilateral region together with its transcription.
Marking	<#> for a single structurally flawed and unrecognizable Chinese character or English letter. <###> for a single structurally flawed and unrecognizable English word (never a phrase, line, or sentence). <#####> for an entirely unreadable secondary text block whose characters cannot be counted individually. Transcribe literally what is rendered: write out every readable character, never guess plausible words from context, never silently repair broken text, and keep Traditional or Simplified Chinese as rendered. Separate multiple lines within one quadrilateral by a single space.
Grounding	Every visible text piece is covered by at least one quadrilateral $[x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4]$, whose vertices start at the top-left corner in the semantic reading direction and proceed clockwise, including vertical, rotated, or curved text. Text that wraps over lines, curves, or is broken up by graphics is emitted as several quadrilaterals instead of one loose box.
Output	A single JSON object: {"regions": [{"quad": [x1, y1, ..., x4, y4], "recognized_text": "text with markers"}]}.

Table 13: **Composition of the TextDART-123K Training Split.** The upper rows report counts and the lower rows report ratios. Image-level ratios use the 100,352 text-containing entries, and the region-level ratio uses all 495,778 annotated regions.

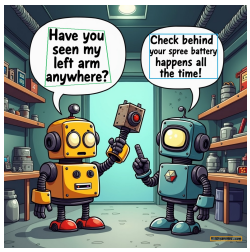
Statistic	Train (122K)
Sample entries	122,200
Text-containing	100,352
Text-free	21,848
Annotated regions	495,778
Entries with anomaly markers	55,587 (55.4%)
Multi-region entries	81.1%
Entries with $\geq$6 regions	25.0%
Entries mixing marked and unmarked regions	44.9%
Regions with anomaly markers	35.1%
Artistic-text entries	25.7%
Script: CJK-only / Latin-only / mixed	54.7 / 13.2 / 30.9%

A.8 Discussion

Suppression of hallucinated text under prompt enhancement. TextDART penalizes glyph defects strictly, while the small text that generative models hallucinate spontaneously is almost always malformed. With prompt enhancement enabled, prompts become richer in detail and models are more prone to adding small text on their own. Such text is left unconstrained by recognition-based rewards and typically appears with malformed glyphs. Models trained with SCANNER instead tend to render the text requested by the prompt accurately, and even spontaneously added small text remains legible, so their advantage grows further under evaluation with prompt enhancement.

Limitations and future work. (1) Our reward currently evaluates only the content correctness and glyph quality of text, and does not assess text layout, that is, whether the position, size, and relative arrangement of text regions match the prompt. Since TextDART natively outputs text quadrilaterals, incorporating layout consistency into the reward is a direct extension, for example by constructing layout rewards from the relative positions between predicted boxes and described targets, or by combining with box-conditioned generation for controllable typesetting. (2) TextDART-123K currently covers English and Chinese, and defect annotation and reward training for more languages remain to be explored. (3) The principle of blaming errant words rather than the whole canvas is not limited to text. Any localizable defect can be incorporated into the same framework of regional rewards

and spatial credit assignment, and extending this framework to temporal credit assignment in video generation is a natural next step.



prompt: A playful cartoon-style illustration showing two robots humorously conversing inside a futuristic, tech-themed workshop filled with mechanical parts and gadgets. At the upper-left side of the panel, Robot A, holding up a rusty, detached metal arm, looks concerned and asks in an oval-shaped speech bubble with bold, playful lettering, "Have you seen my left arm anywhere?". On the lower-right side, Robot B responds casually, pointing toward a cluttered shelf behind him, with a rectangular speech bubble and italicized friendly font replying, "Check behind your spare battery, happens all the time!". Robot B's speech bubble partially overlaps slightly beneath Robot A's bubble, guiding the reading flow naturally diagonally downward. Each bubble contains fewer than 10 words, clearly delivering a humorous exchange between the robot characters.
text_ recognized: Have you seen my left arm anywhere?
d_score: 1.000 s_score: 1.000 reward: 1.000
text_ recognized: Check behind your spree battery happens all the time!
d_score: 1.000 s_score: 0.942 reward: 0.971
text_ recognized: ®<#>ani<#>e<#>.com
d_score: 0.750 reward: 0.750



prompt: In a bookstore, a display reads 'New Arrivals', a checkout counter with 'Scan Items' written on it.
text _ recognized: New Arrivals.
d_score: 1.000 s_score: 0.944 reward: 0.972
text_ recognized: D.Scan items
d_score: 1.000 s_score: 0.833 reward: 0.917
text_ recognized: <######>
d_score: 0.000 reward: 0.000



prompt: A bustling street market with a sign saying 'Open Every Day from 9 AM' in large green letters, a stall banner reading 'Fresh Fruits and Vegetables' in bold red, and a chalkboard outside displaying 'Special Discount Today: 20 percent Off' in italic blue.
text_ recognized: OPEN EVERY DAY FROM 9 AM
d_score: 1.000 s_score: 1.000 reward: 1.000
text_ recognized: FOO<#><#> MARMKE R
d_score: 0.833 reward: 0.833
text_ recognized: <###> <###>
d_score: 0.000 reward: 0.000
text_ recognized: FRESH FRUITS AND VEGETABLES BY HE <#>BS VENTS
d_score: 0.972 s_score: 0.500 reward: 0.736
text_ recognized: Special Discount Today 20% Off
d_score: 1.000 s_score: 0.750 reward:0.875



prompt: 一张复古风格的旅游宣传册安然地放置在乡村风格的木桌上。宣传册展开并有些折痕，中心上方醒目地印着怀旧标题“探索威尼斯的历史奇迹！”。标题下方，较小的文字以优雅的草书字体书写着“在运河、文化与咖啡中开启难忘的旅程”。更下方，整齐的衬线字体详细描述了行程内容，清晰可见地写着：“清晨在迷人咖啡馆品尝正宗的意式浓缩咖啡，登上浪漫的贡多拉船穿行于如画的水道，欣赏隐藏教堂中的文艺复兴杰作”，激发行动欲望并鼓励立即预订。威尼斯著名运河和地标的插图地图装饰着宣传册背景，为内容增添了背景信息和视觉吸引力。
text_ recognized: 探索威尼斯的历史奇迹!
d_score: 1.000 s_score: 1.000 reward: 1.000
text_ recognized: 在运河、文化与咖啡中开启难忘的旅程
d_score: 1.000 s_score: 1.000 reward: 1.000
text_ recognized: 清晨在迷人咖啡馆品尝正宗的意式浓缩咖啡，登上浪漫的贡提拉船穿行<#>如画的水道，欣赏隐<#>教堂中的文艺复<#>作
d_score: 0.939 s_score: 0.865 reward: 0.902
text_ recognized: <###>
d_score: 0.000 reward: 0.000



prompt: 一个专业且安全的视觉呈现的在线银行应用程序，背景为醒目的蓝色和灰色。引人注目的口号“智能银行，生活更好”立即吸引注意，与为现代专业人士提供的 “FinEase”标志相得益彰。
text_ recognized: FinEase
d_score: 1.000 s_score: 1.000 reward: 1.000
text_ recognized: 智能银行，生活更好
d_score: 1.000 s_score: 0.889 reward:0.945
text_ recognized: Modern p<#><#>fesisanns <#><#>of<#>silakis <#>e<#><#><#>e
d_score: 0.750 reward: 0.750



prompt: 为独立乐队设计一个极简专辑封面，使用柔和渐变和简单图标。放置乐队名称“回声波”、专辑名称“迷失在声音中”，以及发行年份“2024”。
text_ recognized: 独<#>
d_score: 0.500 reward: 0.500
text_ recognized: 回声波
d_score: 1.000 s_score: 1.000 reward: 1.000
text_ recognized: 迷失在声音中
d_score: 1.000 s_score: 1.000 reward: 1.000
text_ recognized: <#><#>nased.is <#>et V<#><#>ere
d_score: 0.737 reward: 0.737
text_ recognized: 2024
d_score: 1.000 s_score: 1.000 reward: 1.000

Figure 10: **Localization and Recognition Results of TextDART.** Each detected region is annotated with its defect-aware transcription and glyph-quality score (d_score), and regions matched to a target additionally receive a semantic-correctness score (s_score).

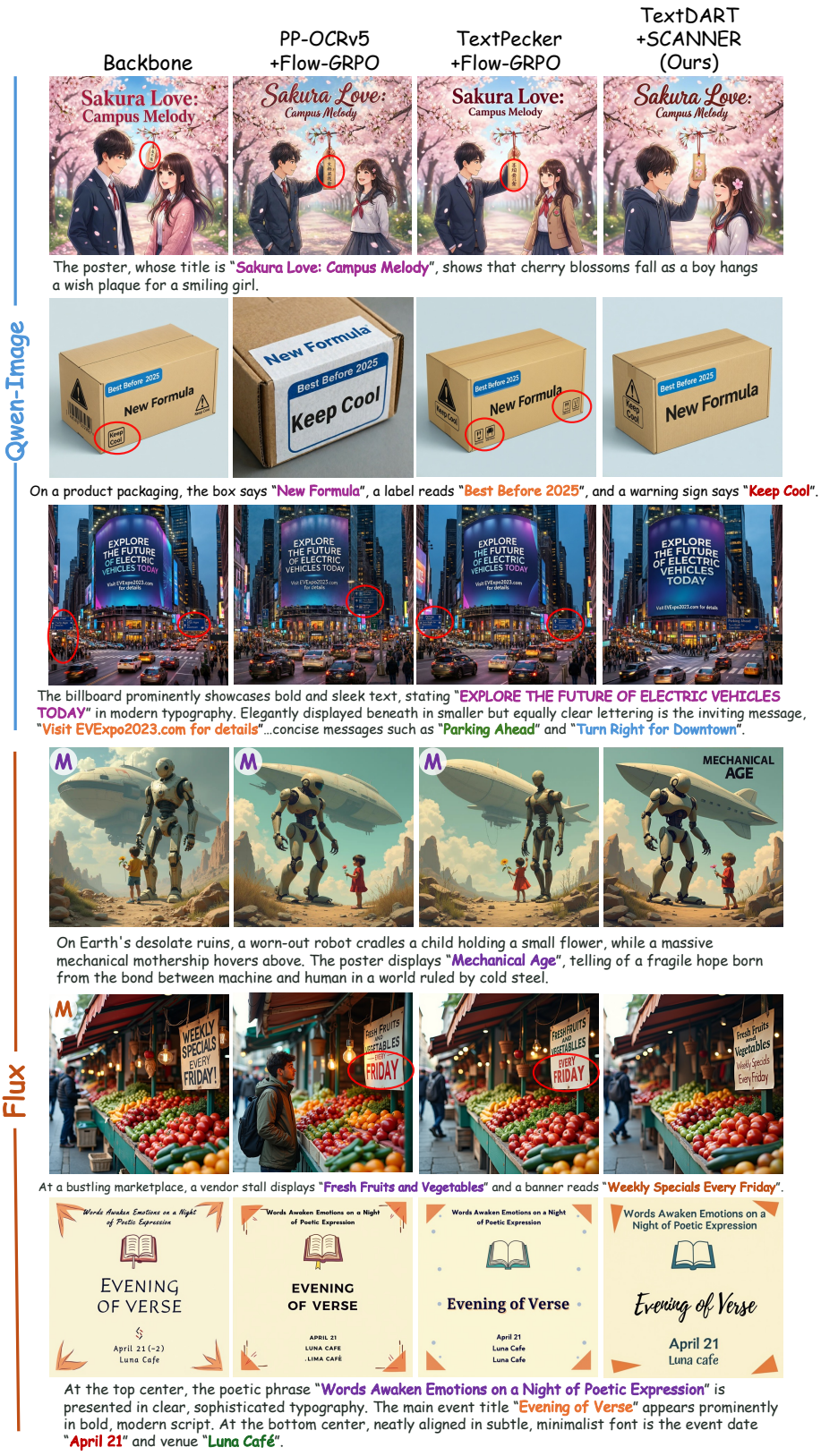
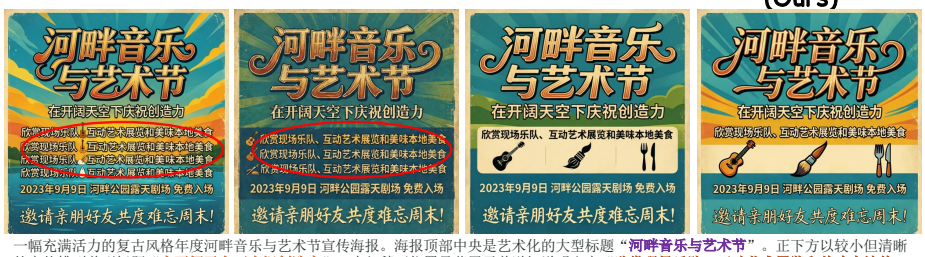


Figure 11: **Qualitative Comparison on English Prompts.** Rows 1–3 are generated on Qwen-Image and rows 4–6 on Flux, and circles highlight text errors of the baselines.

Backbone	PP-OCRv5 +Flow-GRPO	TextPecker +Flow-GRPO	TextDART +SCANNER (Ours)



一幅充满活力的复古风格年度河畔音乐与艺术节宣传海报。海报顶部中央是艺术化的大型标题“**河畔音乐与艺术节**”。正下方以较小但清晰的字体排列着副标题“**在开阔天空下庆祝创造力**”。中部稍下位置显著展示着详细说明文字“**欣赏现场乐队、互动艺术展览和美味本地美食**”。海报下三分之一区域，清晰可见的活动信息均匀分布在底部：“**2023年9月9日 河畔公园露天剧场 免费入场**”。左下角以手写体斜体字呈现欢快提示文字“**邀请亲朋好友共度难忘周末！**”



热闹而繁忙的户外农贸市场里，热情的顾客们悠闲地浏览着色彩斑斓的摊位，摊位上摆满了新鲜的农产品和迷人的手工艺品。每个商品都配有清晰书写的便签标签，经过精心摆放以便观看。一个格外显眼且吸引人的标签清晰地写着“**有机野生蜂蜜 原始未过滤**”，瞬间抓住了观者的目光，旁边整齐地附着着一行更小友好的说明文字“**本地用心采集**”。围绕这个招牌的其他简洁易读的标签也经过精心布置，如“**手工精选传家番茄**”，“**手工全谷面包**”。游客们热情地与摊主互动，快乐地探索商品，为这个充满吸引力的农贸市场增添了温馨的社区氛围。



在地球荒凉的废墟上，一台破旧的机器人怀抱着一个手拿小花的孩子，而一艘巨大的机械母舰悬停在上空。海报上写着“**机械时代**”，讲述在一个由冰冷钢铁统治的世界中，机器与人类之间的纽带孕育出的脆弱希望。



为了一个充满活力的健身活动，请设计一张大胆的单页宣传单，采用充满活力的视觉效果和醒目的红色装饰。将“**PowerGym**“显著地置于顶部，以“**立即加入**”激励读者，并以“**首月免费**”的特别优惠吸引新会员。



深色木纹咖啡桌上，一本关于山地旅行的杂志平铺展开，顶部横幅区域以常规印刷字体呈现中等标题大小的“**極境山脈探險誌：藍紫青綠之旅全記錄**”，与山峰插图顶部留白相融合 中央特色板块采用常规印刷字体，中等尺寸的“**山行**”叠加于登山者身旁的云雾背景之上 左侧栏位使用常规印刷字体，中等大小的“**雲頂步道指南**”垂直排列于岩壁纹理旁，与主图形成视觉平衡 底部条带区域以实体化字母形式展示中等尺寸的“**二零二六秋**”，嵌入青绿色山脚阴影中，与纸张边缘平行。



文化中心正在举办的木偶剧活动正式宣传海报，完整活动发布版张贴在入口展示架上，来访者停在门口查看后准备入场。主题摄影或插画融入常规活动宣传版式，旁边可见盆栽、金属门框与走向接待处的人群。环境中近景，主载体占画面约一半，正面基本平行于镜头，道具退到两侧。整体配色以砖红、炭灰与淡金色为主，光线符合现场环境，反光避开主体正面，保留真实材质和自然使用痕迹。