

Test-Time Alignment of Text-to-Image Diffusion Models via Null-Text Embedding Optimisation

Supplementary Material

Table 6. Influence of Search Width on Null-TTA rewards.

| Search Width | PickScore | Aesthetic | HPSv2 | ImageReward |
|--------------|--------------|--------------|--------------|--------------|
| 1 | 0.226 | 5.340 | 0.339 | 0.585 |
| 3 | 0.228 | 5.449 | 0.346 | 0.741 |
| 10 | 0.228 | 5.453 | 0.347 | 0.684 |
| 25 | 0.229 | 5.335 | 0.352 | 0.700 |
| 50 | 0.228 | 5.370 | 0.351 | 0.671 |

Table 7. Comparison of optimised scores for different values of annealing coefficient (γ).

| γ | PickScore | AestheticScore | HPSv2 | ImageReward |
|----------|--------------|----------------|--------------|--------------|
| 0.004 | 0.225 | 5.404 | 0.318 | 0.664 |
| 0.008 | 0.228 | 5.449 | 0.346 | 0.741 |
| 0.016 | 0.226 | 5.266 | 0.399 | 0.700 |

A. Implementation Details

For Null-TTA, we set $(\lambda_1, \lambda_2, \sigma_\phi^2)$ as $(100, 0.002, 0.01)$ for HPSv2 [4], PickScore [2] and $(2, 0.002, 0.01)$ for Aesthetic score [3]. As stated in Sec 3.3, we gradually decrease λ_2 as denoising proceeds with annealing coefficient $\max(2 - (1 + \gamma)^{T-t}, 0)$. Similarly, as reward approximation by Tweedie’s formula becomes accurate as denoising proceeds, we increase the number of the optimisation steps gradually. optimisation step at each iteration is computed as $n_{\text{opt}} = n_{\text{min}} + \lambda_{\text{opt}}^t * (n_{\text{max}} - n_{\text{min}})$ where $\lambda_{\text{opt}}^t = \min((1 + \gamma)^{T-t} - 1, 1)$. We set n_{min} to 5 and conduct experiments with varying n_{max} . γ is set to 0.008 and exploration number for greedy search K is set to 3. We used Adam optimiser [1] with learning rate 0.01 for optimisation. Note that all the results are averaged results from 3 different random seeds.

B. Ablations

We ablate the key components of Null-TTA to understand how optimisation stability and reward-generalisation trade-offs are influenced by the search width, annealing schedule, and the balance between the reward and regularisation terms. Across all experiments (Tables 6, 7, 8), Null-TTA consistently exhibits stable behaviour: target rewards improve while cross-reward generalisation is largely preserved. Each hyperparameter controls a different aspect of the optimisation dynamics, and together they illustrate why Null-TTA avoids the reward over- and under-optimisation

commonly observed in prior TTA methods. All the experiments were done with $(n_{\text{min}}, n_{\text{max}}) = (5, 25)$ and HPSv2 as target reward.

B.1. Effect of Search Width.

Table 6 reports the effect of search width used in the latent-space greedy search step. The search width does not influence the optimisation of the null-text embedding itself; instead, it controls the extent of latent-space search after optimisation.

A small search width (e.g., 3) provides a mild but useful search, allowing the sampler to more reliably follow the optimisation-improved update direction. This leads to modest improvements in both the target reward (HPSv2) and the cross rewards. Excessively large search width broadens the search too aggressively. This can pull samples away from the semantic manifold and reintroduce undesired latent variations, eroding cross-reward metrics even as the target reward continues to improve – an effect reminiscent of mild over-optimisation. Notably, even without search, Null-TTA already produces strong performance, showing that Null-TTA does not depend on greedy search to achieve alignment. We set search width to 3 as the default, as it balances search with stability.

B.2. Effect of the annealing coefficient γ .

Table 7 evaluates the role of the annealing coefficient γ . Importantly, γ influences both (i) the rate at which the optimisation strength increases, and (ii) the decay speed of the regularisation terms in Eq. 26.

When γ is too large, the optimisation steps grow rapidly while the regularisation decays quickly, causing the model to emphasise the (potentially inaccurate) early reward approximation. This can lead to early-stage over-optimisation and reduced stability. In contrast, very small γ values slow both optimisation growth and regularisation decay; it limits the ability to sufficiently optimise the target reward even when the reward approximation becomes accurate at later timesteps. Null-TTA maintains stable behaviour across all tested values of γ , and moderate settings (e.g., $\gamma = 0.008$) achieve the best balance between strong HPSv2 improvement and cross-reward retention.

Algorithm 1: Null-Text Test-Time Alignment with Greedy Search (Null-TTA)

Input : Pre-trained diffusion model ϵ_θ , Reward Function $R(\cdot)$, Prompt \mathcal{P} , Total steps T , Step bounds n_{min}, n_{max} , Growth rate γ , Weights $\lambda_1, \lambda_2, \sigma_\phi^2$, Search Width K , Learning rate η .

Output: Aligned image x_0

Initialize latent $x_T \sim \mathcal{N}(0, \mathbf{I})$;

Get initial embeddings: $\phi \leftarrow \text{TextEncoder}(\emptyset)$, $c \leftarrow \text{TextEncoder}(\mathcal{P})$;

Initialize optimizable embedding: $\phi' \leftarrow \phi$;

for $t = T, \dots, 1$ **do**

// 1. Dynamic Hyperparameter Scheduling

Calculate adaptive regularization weight: $\tilde{\lambda}_{2,t} \leftarrow \lambda_2 \cdot \max(2 - (1 + \gamma)^{T-t}, 0)$;

Calculate adaptive optimization steps: $N_t \leftarrow n_{min} + \lceil \min((1 + \gamma)^{T-t} - 1, 1) \cdot (n_{max} - n_{min}) \rceil$;

// 2. Null-Text Optimization (Eq. 26)

for $n = 1, \dots, N_t$ **do**

 Compute unconditional noise: $\epsilon_{uncond} \leftarrow \epsilon_\theta(x_t, t, \phi')$;

 Compute original unconditional noise (fixed): $\epsilon_{orig} \leftarrow \epsilon_\theta(x_t, t, \phi)$;

 Compute original CFG noise: $\tilde{\epsilon}_{orig} \leftarrow \epsilon_{orig} + s(\epsilon_\theta(x_t, t, c) - \epsilon_{orig})$;

 Compute CFG noise: $\tilde{\epsilon} \leftarrow \epsilon_{uncond} + s(\epsilon_\theta(x_t, t, c) - \epsilon_{uncond})$;

 Estimate clean image (Tweedie’s Formula): $\hat{x}_0(x_t, \phi') \leftarrow \frac{x_t - \sqrt{1 - \tilde{\alpha}_t} \tilde{\epsilon}}{\sqrt{\tilde{\alpha}_t}}$;

 Calculate Objective $J(\phi')$:

$$J(\phi') = \lambda_1 R(\hat{x}_0(x_t, \phi')) - \frac{\tilde{\lambda}_{2,t}(1 - \alpha_t)}{2\alpha_t(1 - \tilde{\alpha}_t)} \|\tilde{\epsilon} - \tilde{\epsilon}_{orig}\|^2 - \frac{\tilde{\lambda}_{2,t}}{2\sigma_\phi^2} \|\phi' - \phi\|^2$$

 Compute Gradient: $g \leftarrow \nabla_{\phi'} J(\phi')$ (via Autograd or Zeroth-Order Estimation);

 Update Embedding: $\phi' \leftarrow \phi' + \eta \cdot g$;

// Gradient Ascent

// 3. Greedy Search

 Generate K candidates $\{x_{t-1}^{(k)}\}_{k=1}^K$ from $p_\theta(x_{t-1}|x_t, \phi')$;

 Evaluate candidates via Tweedie’s estimate: $r^{(k)} \leftarrow R(\hat{x}_0(x_{t-1}^{(k)}))$ for $k = 1 \dots K$;

 Select best trajectory: $x_{t-1} \leftarrow x_{t-1}^{(k^*)}$ where $k^* = \arg \max_k r^{(k)}$;

return $x_{final} = \text{Decoder}(x_0)$;

B.3. Effect of the reward{regularisation balance.

Table 8 examines how the balance between the reward term and the two regularisation components in Eq. 26 affects optimisation behaviour. Increasing the regularisation strength (larger λ_2 or smaller σ_ϕ^2) leads to more conservative updates: target–reward improvement becomes limited, but cross-reward generalisation improves, indicating an under-optimised yet stable solution. Conversely, weakening the regularisation results in more aggressive updates that favour the target reward but can destabilise cross-reward metrics, resembling the early symptoms of reward over-optimisation.

Our default configuration achieves a balanced trade-off, providing strong target reward gains while maintaining high cross-reward performance. This demonstrates that Null-TTA jointly improves alignment and generalisation when the reward–regularisation ratio is within a reasonable range, and that Null-TTA remains robust even as this ratio varies.

Table 8. Effect of the reward/regularisation ratio on the target reward (HPSv2) and the cross reward (ImageReward). All rows except SD-v1.5 share the same target weight $\lambda_\alpha = 100$.

| $(\lambda_1, \lambda_2, \sigma_\phi^2)$ | HPSv2 | ImageReward | Note |
|---|--------------|--------------|---|
| (100, 0.0004, 0.01) | 0.343 | 0.642 | Weak reg ($\lambda_{reg} \downarrow$) |
| (100, 0.0100, 0.01) | 0.339 | 0.787 | Strong reg ($\lambda_{reg} \uparrow$) |
| (100, 0.0020, 0.050) | 0.340 | 0.303 | Weak reg ($\sigma_\phi^2 \uparrow$) |
| (100, 0.0020, 0.002) | 0.337 | 0.781 | Strong reg ($\sigma_\phi^2 \downarrow$) |
| (100, 0.0020, 0.010) | 0.346 | 0.741 | Ours (default) |
| SD-v1.5 | 0.279 | 0.332 | Baseline |

C. Detailed Procedure for Null-TTA

We provide the detailed algorithmic procedure of our proposed Null-TTA in Algorithm 1.

References

- [1] Diederik P Kingma. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 1
- [2] Yuval Kirstain, Adam Polyak, Uriel Singer, Shahbuland Matiana, Joe Penna, and Omer Levy. Pick-a-pic: An open dataset of user preferences for text-to-image generation. *Advances in neural information processing systems*, 36:36652–36663, 2023. 1
- [3] Christoph Schuhmann. Laion-ai aesthetic predictor (v1): A linear estimator on top of clip to predict the aesthetic quality of pictures. <https://github.com/LAION-AI/aesthetic-predictor>, 2022. MIT License. Commit 6d122ad (2022-08-15). Accessed: 2025-11-13. 1
- [4] Xiaoshi Wu, Keqiang Sun, Feng Zhu, Rui Zhao, and Hongsheng Li. Human preference score: Better aligning text-to-image models with human preference. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 2096–2105, 2023. 1