

Guia de Instruções para Máxima Fidelidade à Referência

Objetivo: Fidelidade Absoluta sem Alterações no Aplicativo

Este documento fornece instruções de *prompting* e configuração para garantir que o seu aplicativo de geração de imagens (baseado em modelos de difusão como Stable Diffusion ou Gemini) siga estritamente as referências fornecidas, sem a necessidade de alterar o código-fonte do aplicativo. A chave é usar as ferramentas de controle que o modelo já oferece.

1. O Princípio da Força de Denoising (Denoising Strength)

Em fluxos de trabalho de **Image-to-Image (Img2Img)**, a **Força de Denoising** (ou *Denoising Strength*) é o parâmetro mais crucial para a fidelidade.

- **Definição:** É o quanto o modelo de IA tem permissão para alterar a imagem de referência.
- **Fidelidade Máxima:** Para replicar a imagem de referência com o mínimo de alteração, o valor deve ser **muito baixo** (e.g., 0.1 a 0.3).
 - 0.0: Nenhuma alteração.
 - 0.1 - 0.3: Pequenas alterações, mantendo a estrutura, cor e composição da referência. Ideal para refinar ou adicionar pequenos detalhes.
 - 0.5 - 0.7: Alterações moderadas. O modelo muda o estilo, mas mantém a composição geral.
 - 0.8 - 1.0: Alterações drásticas. O modelo usa a imagem apenas como uma vaga inspiração.

Instrução para o Gem: Ao receber uma referência, o Gem deve instruir o usuário a usar a **menor Força de Denoising possível** que ainda permita a alteração desejada (se houver).

2. O Uso Cirúrgico do ControlNet

O **ControlNet** é a ferramenta definitiva para forçar a fidelidade estrutural. Ele transforma a imagem de referência em um “mapa” que o modelo deve seguir.

Tipo de Referência	Módulo ControlNet Recomendado	Configuração de Fidelidade
Pose e Estrutura Corporal	OpenPose	Control Weight (Peso de Controle) Alto (e.g., 1.0 a 1.5). Isso força o modelo a seguir o esqueleto da referência com precisão.
Composição e Profundidade	Depth (Profundidade)	Control Weight Alto. Garante que a escala e a posição dos objetos na cena sejam idênticas às da referência.
Estilo e Cor	Reference Only ou IP Adapter	Control Weight Alto. Força o modelo a adotar o estilo, paleta de cores e textura da imagem de referência.

Instrução para o Gem: O Gem deve sempre recomendar o uso do ControlNet apropriado com um **Peso de Controle (Control Weight) elevado** para garantir que o “blueprint” da referência seja a lei.

3. Prompting para Reforçar a Fidelidade

O prompt de texto deve ser usado para reforçar a intenção de fidelidade.

- **Prompt de Reforço:** Adicionar termos como `exact replication`, `perfect fidelity`, `maintain original composition`, `photorealistic copy` no início do prompt.
- **Prompt Negativo:** Usar prompts negativos para evitar que o modelo adicione elementos não solicitados ou mude o estilo: `--no artistic license`, `no style change`, `no added elements`, `no creative interpretation`.

4. Otimização de Parâmetros (Seed e Sampler)

- **Seed (Semente)**: Se o objetivo é recriar uma imagem anterior, o **Seed** deve ser **exatamente o mesmo**. Se o objetivo é manter a fidelidade à referência, mas com uma variação, o Seed pode ser alterado.
- **Sampler (Amostrador)**: Samplers determinísticos (como DPM++ 2M Karras ou Euler A com um Seed fixo) tendem a produzir resultados mais consistentes.

Resumo da Estratégia

Para garantir que o aplicativo siga as referências sem alterações no código:

1. **Img2Img**: Usar **Força de Denoising Baixa** (0.1 - 0.3).
2. **ControlNet**: Usar **Peso de Controle Alto** (1.0+) com o módulo ControlNet apropriado para a referência (OpenPose para pose, Depth para composição).
3. **Prompting**: Reforçar a fidelidade com termos como `exact replication` e prompts negativos para evitar interpretação criativa.

4. Otimização de Parâmetros Adicionais

Pesos de Prompt (Prompt Weights)

Os pesos de prompt permitem ajustar a influência de diferentes partes do prompt de texto.

- **Sintaxe**: Geralmente, usa-se parênteses e dois pontos para aumentar ou diminuir o peso de uma palavra ou frase. Ex: `(photorealistic:1.2)` aumenta o peso, `(cartoon:0.8)` diminui.
- **Fidelidade à Referência**: Se o prompt de texto descreve um elemento que **deve** estar na imagem, o Gem deve aplicar um **peso alto** a essa descrição para forçar o modelo a aderir a ela.

CFG Scale (Classifier-Free Guidance Scale)

O CFG Scale (ou Guidance Scale) controla o quanto o processo de geração segue o prompt de texto.

- **CFG Baixo (e.g., 3-7):** O modelo tem mais liberdade criativa e o resultado é mais diverso, mas menos fiel ao prompt.
- **CFG Alto (e.g., 7-12):** O modelo segue o prompt de texto de forma mais estrita, resultando em maior fidelidade ao texto.

Instrução para o Gem: Para **máxima fidelidade** ao prompt de texto (que deve descrever a referência), o Gem deve recomendar um **CFG Scale moderado a alto** (e.g., 7-10).

Resumo da Estratégia para Fidelidade Absoluta

A fidelidade absoluta à referência é alcançada pela combinação de três fatores de controle:

1. **Controle de Conteúdo (Img2Img): Força de Denoising Baixa** (0.1 - 0.3) para preservar a imagem de referência.
2. **Controle Estrutural (ControlNet): Peso de Controle Alto** (1.0+) para forçar a estrutura da referência.
3. **Controle de Texto (CFG/Prompt Weights): CFG Scale Moderado/Alto** (7-10) e **Pesos de Prompt Altos** para forçar a adesão à descrição de texto da referência.

O Gem deve atuar como um **Engenheiro de Fidelidade**, garantindo que esses três vetores de controle estejam sempre otimizados para a replicação exata.