Paper Digest Meeting

CFG++: Manifold-Constrained Classifier Free Guidance for Diffusion Models

Hyungjin Chung* Jeongsol Kim* Geonyeong Park* Hyelin Nam* Jongchul Ye

2024.07.17

Jaihoon Kim

KAIST Visual AI Group

Problem Definition

Problem: DDIM with CFG lacks invertibility and results in issues like mode collapse.

Goal: High-quality, edit-friendly (invertible) reverse process utilizing a low CFG weight.

Text Alignment

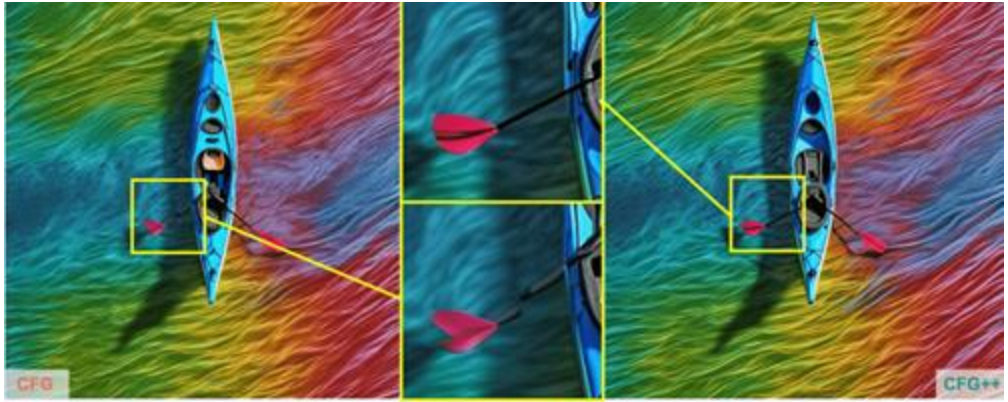


Image Editing

Key Ideas

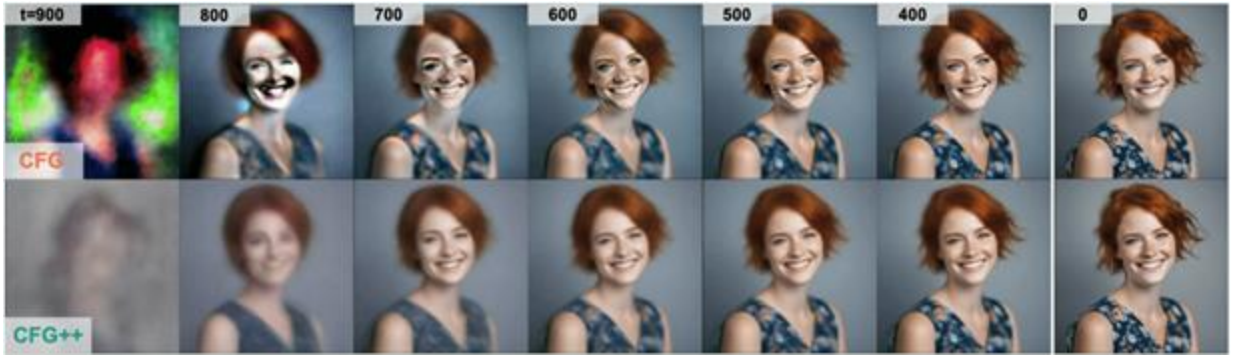
Insight: CFG extrapolates beyond the unconditional and the conditional posterior means, which results in unexpected sudden shift in the image and intense color saturation.

CFG++ - Interpolation

$$\hat{\boldsymbol{x}}_{\boldsymbol{c}}^{\lambda}(\boldsymbol{x}_t) = (1-\lambda)\hat{\boldsymbol{x}}_{\varnothing}(\boldsymbol{x}_t) + \lambda\hat{\boldsymbol{x}}_{\boldsymbol{c}}(\boldsymbol{x}_t),$$

CFG - Extrapolation

$$\hat{\boldsymbol{x}}_{\boldsymbol{c}}^{\omega}(\boldsymbol{x}_t) = (1-\omega)\hat{\boldsymbol{x}}_{\varnothing}(\boldsymbol{x}_t) + \omega\hat{\boldsymbol{x}}_{\boldsymbol{c}}(\boldsymbol{x}_t)$$



Method

Reparameterization of DPS using SDS loss.

$$\boldsymbol{x}_{t-1} \simeq \sqrt{\bar{\alpha}_{t-1}}\left(\hat{\boldsymbol{x}}_{\varnothing} - \gamma_t \nabla_{\hat{\boldsymbol{x}}_{\varnothing}} \ell(\hat{\boldsymbol{x}}_{\varnothing})\right) + \sqrt{1-\bar{\alpha}_{t-1}}\hat{\boldsymbol{\epsilon}}_{\varnothing}.$$

$$\boldsymbol{x}_{t-1} \quad = \quad \sqrt{\bar{\alpha}_{t-1}}\left(\hat{\boldsymbol{x}}_{\varnothing} - \gamma_t \nabla_{\hat{\boldsymbol{x}}_{\varnothing}} \ell_{sds}(\hat{\boldsymbol{x}}_{\varnothing})\right) + \sqrt{1-\bar{\alpha}_{t-1}}\hat{\boldsymbol{\epsilon}}_{\varnothing}$$

Algorithm 1 Reverse Diffusion with CFG

Require: $\boldsymbol{x}_T \sim \mathcal{N}(0, \mathbf{I}_d), 0 \leq \omega \in \mathbb{R}$

for $i = T$ **to** 1 **do**
  $\hat{\boldsymbol{\epsilon}}_c^{\omega}(\boldsymbol{x}_t) = \hat{\boldsymbol{\epsilon}}_{\varnothing}(\boldsymbol{x}_t) + \omega[\hat{\boldsymbol{\epsilon}}_c(\boldsymbol{x}_t) - \hat{\boldsymbol{\epsilon}}_{\varnothing}(\boldsymbol{x}_t)]$
  $\hat{\boldsymbol{x}}_c^{\omega}(\boldsymbol{x}_t) \leftarrow (\boldsymbol{x}_t - \sqrt{1-\bar{\alpha}_t}\hat{\boldsymbol{\epsilon}}_c^{\omega}(\boldsymbol{x}_t))/\sqrt{\bar{\alpha}_t}$
  $\boldsymbol{x}_{t-1} = \sqrt{\bar{\alpha}_{t-1}}\hat{\boldsymbol{x}}_c^{\omega}(\boldsymbol{x}_t) + \sqrt{1-\bar{\alpha}_{t-1}}\hat{\boldsymbol{\epsilon}}_c^{\omega}(\boldsymbol{x}_t)$
end for
return $\boldsymbol{x}_0$

Algorithm 2 Reverse Diffusion with CFG++

Require: $\boldsymbol{x}_T \sim \mathcal{N}(0, \mathbf{I}_d), \lambda \in [0, 1]$

for $i = T$ **to** 1 **do**
  $\hat{\boldsymbol{\epsilon}}_c^{\lambda}(\boldsymbol{x}_t) = \hat{\boldsymbol{\epsilon}}_{\varnothing}(\boldsymbol{x}_t) + \lambda[\hat{\boldsymbol{\epsilon}}_c(\boldsymbol{x}_t) - \hat{\boldsymbol{\epsilon}}_{\varnothing}(\boldsymbol{x}_t)]$
  $\hat{\boldsymbol{x}}_c^{\lambda}(\boldsymbol{x}_t) \leftarrow (\boldsymbol{x}_t - \sqrt{1-\bar{\alpha}_t}\hat{\boldsymbol{\epsilon}}_c^{\lambda}(\boldsymbol{x}_t))/\sqrt{\bar{\alpha}_t}$
  $\boldsymbol{x}_{t-1} = \sqrt{\bar{\alpha}_{t-1}}\hat{\boldsymbol{x}}_c^{\lambda}(\boldsymbol{x}_t) + \sqrt{1-\bar{\alpha}_{t-1}}\hat{\boldsymbol{\epsilon}}_{\varnothing}(\boldsymbol{x}_t)$
end for
return $\boldsymbol{x}_0$

Experiments

Text Alignment: Enhanced T2I results

Image Editing - Better Invertibility, editability