

# 解决潜在扩散模型逆问题的重采样投影优化

旷 荣

成都理工大学数学科学学院, 四川 成都

收稿日期: 2026年3月5日; 录用日期: 2026年4月11日; 发布日期: 2026年4月23日

## 摘 要

随着深度学习在科学研究领域日益普及, 它被用于解决越来越多的问题, 也成为求解反问题的热门研究方向。近年来, 潜扩散模型已被证明能够生成高质量图像, 而扩散模型在隐空间中的应用也比在像素空间中更为高效。因此, 我们对先前提出的重采样算法进行了进一步改进: 将浅层空间中的迭代优化改为“零空间投影”, 并将优化步长调整为每5步执行一次。实验表明, 我们的方法在原有基础上取得了进一步的良好效果, 且在一些细节上实现了更显著的提升。

## 关键词

扩散模型, 逆问题

# The Latent Diffusion Model Solves the Resampling Projection Optimization of the Inverse Problem

Rong Kuang

College of Mathematical Sciences, Chengdu University of Technology, Chengdu Sichuan

Received: March 5, 2026; accepted: April 11, 2026; published: April 23, 2026

## Abstract

As deep learning is becoming more and more popular in the field of scientific research, it is used to solve more and more problems, and it has also become a popular research field for solving anti-problems. Recently, the latent diffusion model has been shown to generate high-quality images, and the application of the diffusion model in subconscious space is more efficient than that in pixel space. Therefore, we have made further modifications to the proposed resampling algorithm, we have changed the iterative optimization in shallow space to zero space projection, and we have changed

the number of optimized steps to every 5 steps, and our experiment has achieved further good results on the basis of the original. Moreover, our experimental results have seen greater improvements in some details.

## Keywords

Diffusion Model, Inverse Problem

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

反问题在许多领域有广泛的应用,包括但不限于图像去模糊和去噪、计算成像、医学成像以及遥感。在解决这些逆问题时,其主要目的是从给定的由前向算子  $A$  产生的观测测量  $y \in R^m$  中恢复未知的干净无噪声数据  $x \in R^n$

$$y = A(x_*) + \eta, \eta \sim (0, \delta^2) \quad (1)$$

其中  $A: R_n \rightarrow R^m$  是前向算子(可以是线性或非线性),  $\eta$  是相应的加性随机高斯噪声。一般来说,逆问题可以分为适定问题和不适定问题,而我们更感兴趣的是研究当逆问题( $m < n$ )不适定的情况,因为它与许多现实世界的场景有关。为了得到有意义的解,需要某种正则化器(或先验)。在经典数理物理方程的确定解问题中,适定问题是指需要满足以下三个要求来指定解的问题:

- 1) 解存在;
- 2) 解唯一;
- 3) 解连续依赖于固定的解条件,即解是稳定的。

只要这三个要求中的任意一个不满足,该问题就是不适当的。

随着扩散模型的近期出现(Sohl-Dickstein *et al.*, 2015, Song & Ermon, 2019, Ho *et al.*, 2020, Dhariwal & Nichol, 2021, Rombach *et al.*, 2022) [1]-[5]; 以及基于流的模型(Lipman *et al.*, 2022, Liu *et al.*, 2022, Albergo *et al.*, 2023, Ma *et al.*, 2024) [6]-[9]将其应用于解决线性逆问题并取得了显著成果。Song *et al.* (2021) [10]为使用扩散模型解决非线性逆问题奠定了理论基础:将统一的扩散建模纳入 SDE 框架,并提出了一种逆时间采样算法,该算法首次被证明可用于解决线性逆问题。Kawar *et al.* (2022) [11]提出了一个 DDRM 框架:将贝叶斯推断与扩散模型相结合,以解决非线性观测算子(如部分傅里叶变换)的逆问题,该方法可应用于超分辨率和磁共振重建。Chung *et al.* (2023) [12]提出了一种 DPS 算法,通过近似后验分布处理非线性前向模型,支持病理方程的求解。对于 DPS 算法而言其在迭代过程不进行严格投影,而是在每个扩散去噪步骤中,向模型的得分函数添加一个基于测量误差的梯度项作为引导。这个梯度项近似于对数似然的梯度,使生成过程朝向与观测数据  $y$  更一致的方向进行。而我们的算法会根据当前的迭代,在像素空间以投影的方式进行尝试在一定的步数之后将生成的量“拉回”符合当前轨迹周围,防止过于偏离而造成严重的误差并且每 5 步进行一次优化;尽管方法上存在一些差异,但所有方法的终极目标是一致的:在利用扩散模型强大生成先验的同时,使重建结果与观测数据  $y$  保持一致。它们都在探索如何平衡“真实性”和“保真度”这两个核心指标。在文献中,使用手工设计先验(例如稀疏性)的传统方法正在被丰富的、学习得到的先验(如深度生成模型)逐步取代。最近,由于在图

像生成中使用扩散模型(Dhariwal & Nichol, 2021, Karras *et al.*, 2022, Song *et al.*, 2023, Lou & Ermon, 2023) [13]-[16], 将扩散模型作为结构先验的做法受到了广泛关注。扩散模型通常比生成对抗网络(GAN)更容易、更稳定地训练, 使得它们成为更易获得的生成先验(Dhariwal & Nichol, 2021) [4]。使用扩散模型作为先验的最常见方法是求助于后验采样。然而, 尽管这些技术取得了显著成功, 它们也存在一些局限性。主要的挑战是, 大多数现有工作直接在像素空间训练这些模型, 这需要大量的计算资源和大量的训练数据(Rombach *et al.*, 2022) [5]。此外, 在像素空间进行训练可能导致模型对高频细节过于敏感, 从而影响其在低分辨率或部分观测条件下的鲁棒性。为了解决这些问题, 一些研究开始探索在潜在空间训练扩散模型的可能性(Saharia *et al.*, 2022) [17]。这种方法不仅降低了计算和数据的需求, 还在保持生成质量的同时增强了模型的泛化能力。不过, 潜在空间的表达能力与压缩程度之间的权衡仍然是一个关键问题, 需要进一步研究以优化模型性能。

## 2. 背景

### 2.1. 去噪扩散概率模型(DDPM)

扩散模型是一类生成模型, 其核心思想是通过逐渐加入高斯噪声来扰动训练数据, 然后学习逆转这一过程, 从纯噪声中恢复数据。去噪扩散概率模型(DDPM) (Ho *et al.*, 2020) [3]通常分为两个阶段: 前向扩散和反向生成。在正向扩散阶段, 随机高斯噪声逐步添加到图像  $x_0 \sim p_{\text{data}}(x)$ , 直到数据变为纯噪声  $x_T \sim N(0, I)$ 。其正向过程的公式可以定义如下:

$$x_t = \sqrt{1 - \beta_t} x_{t-1} + \sqrt{\beta_t} z_t, z_t \sim N(0, I) \quad (2)$$

其中,  $x_t$  是时刻  $t$  的噪声图像,  $\beta_t$  是预定义的方差序列,  $z_t$  是在时刻  $t$  添加的随机高斯噪声。使用参数化技术, 该公式可以进一步修改为:

$$x_t = \sqrt{\alpha_t} x_0 + \sqrt{1 - \alpha_t} \varepsilon, \varepsilon \sim N(0, I) \quad (3)$$

其中  $\alpha_t = \prod_{i=1}^t (1 - \beta_i)$ ,  $\varepsilon$  是重新参数化的噪声。

在逆向生成阶段, 模型尝试从噪声中重建原始数据。这个过程依赖于马尔可夫链的构建, 其中每一步的状态转移由一个简单的高斯分布决定。扩散模型中的逆过程表示为:

$$x_{t-1} = \frac{1}{\sqrt{1 - \beta_t}} \left( x_t - \frac{\beta_t}{\sqrt{1 - \alpha_t}} \varepsilon_\theta(x_t, t) \right) + \frac{1 - \alpha_{t-1}}{1 - \alpha_t} \beta_t \quad (4)$$

其中  $\varepsilon_\theta(x_t, t)$  是一个经过训练的神经网络, 用于预测时间步  $t$  上噪声图像  $x_t$  的噪声  $\varepsilon$ 。噪声近似模型  $\varepsilon_\theta(x_t, t)$  可以通过最小化以下目标来训练, 以近似前向过程的噪声:

$$\min E_{x_0 \sim q(x_0), \varepsilon \sim N(0, I)} \|\varepsilon - \varepsilon_\theta(x_t, t)\|_2^2 \quad (5)$$

### 2.2. 基于分数的生成建模与 SDEs

引入随机微分方程(SDE)方法, 以更好地刻画扩散过程中的连续状态变化, 从而构建由连续时间变量  $t$  索引的扩散过程  $\{x_t\}$ 。通过 SDEs, 前向过程和逆向过程可以在单一框架下统一分析, 同时能够使用数值求解器高效模拟噪声图像演化路径。前向过程可以形式化地表示为:

$$dx = f(x, t)dt + g(t)dw \quad (6)$$

其中:

- $f(x, t)$  是漂移系数函数

- $g(t)$  是扩散系数函数
- $w$  是标准维纳过程

当  $t = 0$  时, 获得数据分布; 当  $t = T$  时, 获得高斯分布, 即  $x_T \sim N(0, I)$ 。该扩散过程的目的是基于 SDE 方法逐步去噪噪声图像, 并实现对原始数据分布的近似, 即学习的反向过程。该过程可以形式化为: 其中:

- $dx$  是状态变量微分
- $f(x, t)$  是漂移系数函数
- $g(t)$  是扩散系数函数
- $\nabla_{x_t} \log p(x_t)$  是得分函数
- $d\bar{w}$  是反向时间维纳过程微分

在实践中, 我们使用  $s_\theta$  来近似得分函数(由  $\theta$  参数化的神经网络), 它可以通过去噪评分匹配(Vincent, 2011) [18]来训练:

$$\hat{\theta} = \arg \min_{\theta} E \left[ \left\| s_{\theta}(x_t, t) - \nabla_{x_t} p(x_t | x_0) \right\|_2^2 \right] \quad (8)$$

$t$  从  $[0, T]$  均匀采样,  $x_t \sim p(x_t | x_0)$ ,  $x_0 \sim p_{\text{data}}(x)$ , 只要我们有得分函数  $s_\theta$  的近似值, 就可以近似 SDE 的逆过程。逆过程描述了一个无条件过程, 而解逆问题被描述为一个条件生成, 其中数据从后验  $p(x_t | y)$  中采样, 并且根据贝叶斯规则, 我们可以将条件得分写为:

$$\nabla_{x_t} \log p(x_t | y) = \nabla_{x_t} \log p(x_t) + \nabla_{x_t} \log p(y | x_t) \quad (9)$$

这给出了在求解逆问题时 SDE 的形式:

$$dx = \left[ f(x, t) - g(t)^2 \left( \nabla_{x_t} \log p(x_t) + \nabla_{x_t} \log p(y | x_t) \right) \right] dt + \sqrt{\beta_t} d\bar{w} \quad (10)$$

通过数值求解该方程, 可以在生成过程中精确控制噪声轨迹, 从而提高图像生成的质量并加快采样速度。通过对正向和逆向过程的联合分布建模, 理论上可以提供对数据分布的更细粒度估计。这一特性允许在图像生成任务中更准确地控制生成结果的细节, 并提高模型对复杂数据结构的适应能力。这种方法不仅拓展了扩散模型的理论基础, 也为实际应用提供了更高的灵活性和稳定性。

### 2.3. 潜空间优化

由于图像  $x_0$  在像素空间中的扩散模型计算强度很大, 一个有效的替代方法是使用自编码器将高维像素空间中的图像压缩到低维潜在空间, 从而提高工作的效率。该方法表述如下:

$$\xi: R^n \mapsto R^d, D: R^d \mapsto R^n, x \sim D(\xi(x))$$

其中  $\xi$  和  $D$  分别表示编码器和解码器, 并且  $d \ll n$ 。通过将图像编码到潜在空间  $z = \xi(x)$  中, 可以在简化的表示上训练扩散模型, 显著降低计算成本, 并使高分辨率图像建模成为可能。潜在空间。潜在扩散模型主要分为三个模块:

**图像感知压缩:** 在感知图像压缩中, 我们使用了训练好的 VQ-GAN。它包括一个编码器和一个解码器。

**LDM:** LDM 主要是在获得压缩图像潜变量后, 添加噪声并去噪扩散过程中的噪声。因为 LDM 在潜在空间中作用, 特征的大小远小于图像空间, 所以 LDM 的推理速度要快得多。

**条件机制:** 类似于其他条件模型, 我们可以在 LDM 中添加不同条件, 让模型生成我们预设的内容。具体来说, 我们通过在 U-Net 中引入交叉注意力来引入条件。

### 3. 方法

对方程(10)的后验抽样的问题在于似然项  $\nabla_{x_t} \log p(y|x_t)$  无法直接写出其解析表达式。为了解决这个问题，现有的方法主要包括：

使用交替投影到测量子空间，从而避免直接使用似然(Kawar et al., 2022; Chung et al., 2022; Wang et al., 2022) [11] [19] [20]；

对于无法在假设下提出似然估计的情况，提出了相关假设(Chung et al., 2023; Song et al., 2023) [12] [21]。例如，Chung et al. (2023) [12]提出了后验采样(DPS)，使用了似然的拉普拉斯近似。

$$x_{t-1} = \sqrt{\bar{\alpha}_{t-1}} \hat{x}_0(x_t) + \sqrt{1 - \bar{\alpha}_{t-1} - \eta \sigma_t^2} s_\theta(x_t, t) + \eta \sigma_t \varepsilon \quad (11)$$

$$x_{t-1} = x'_{t-1} - \varsigma \nabla_{x_t} \|y - A(\hat{x}_0(x_t))\|_2^2 \quad (12)$$

其中  $\varsigma$  是可调节步长， $\eta$  是随机扰动强度， $\sigma_t$  是随机扰动。引入随机扰动是为了避免反向过程的确定性行为，其中  $\eta$  控制随机扰动的比例， $0 < \eta < 1$ 。由于该公式在像素空间中的作用有限，它对应于将其应用于潜在空间。

$$z'_{t-1} = \sqrt{\bar{\alpha}_{t-1}} \hat{z}_0(z_t) + \sqrt{1 - \bar{\alpha}_{t-1} - \eta \sigma_t^2} s_\theta(z_t, t) + \eta \sigma_t \varepsilon \quad (13)$$

$$z_{t-1} = z'_{t-1} - \varsigma \nabla_{z_t} \|y - A(D(\hat{z}_0(z_t)))\|_2^2 \quad (14)$$

这个公式可以被视为在像素空间中称为潜在 DPS 的 DPS 算法的推导。

在重采样算法中[22]，数据一致性被集成到 LDM 的反采样过程中。重采样是一种新颖的算法，旨在利用预训练的潜在扩散模型作为先验，来解决包括线性和非线性在内的通用逆问题(如图像超分辨率、去模糊、修复、CT 重建等)。对于之前基于梯度更新的方法在潜在空间中，由于解码器 D 的高度非凸性，容易陷入局部最优，导致重建结果模糊或与测量数据不一致。因此重采样摒弃了“软”的梯度约束，改为在反向采样过程的某些关键步骤  $t$  上，直接求解一个严格的优化问题(公式 16)从而找到一个与观测测量  $y$  严格一致的潜在变量  $\hat{z}_0(y)$ ；并且与 Latent-DPS 展示的梯度更新不同的是重采样算法提出在 10 步内解决一个优化问题；但在我们的工作中我们修改了优化步数使其从 10 步优化为 5 步。基于像素空间的扩散模型的有限适用性可以通过交替使用更高效的 LDM 作为生成先验来解决，其公式如下：

$$\hat{x}_0(y) \in \arg \min_x \frac{1}{2} \|y - A(x)\|_2^2 \quad (15)$$

$$\hat{z}_0(y) \in \arg \min_z \frac{1}{2} \|y - A(D(z))\|_2^2 \quad (16)$$

$\hat{x}_0(y)$  和  $\hat{z}_0(y)$  是与像素空间和浅层空间中数据  $y$  项一致的值。对于这个优化问题，可以通过在像素空间和潜在空间中使用迭代求解器来解决。基于上述思想，我们改进了模型的性能思想，并在保持潜在空间数据一致性收敛的硬特征的前提下，将像素空间中的优化公式替换为零空间投影：

$$\hat{x}_0(y) = \hat{x}_{0t} = (I - H^+ H) x_{0t} + H^+ y \quad (17)$$

其中  $H^+$  是  $H$  的伪逆； $H$  为线性测量矩阵(或算子)，描述了原始信号  $x$  如何被观测为  $y$ ， $\hat{x}_{0t}$  是当前时间步对精确值  $x_0$  的估计， $x_{0t}$  是使用 Tweedie 公式估计的：

$$x_{0t} = E[x_0|x_t] = \frac{1}{\sqrt{\bar{\alpha}_t}} (x_t + (1 - \bar{\alpha}_t) \nabla \log p(x_t)) \quad (18)$$

投影公式(17)正是线性最小二乘问题(15)的一个特解。它通过正交投影，将当前估计  $\hat{x}_{0_t}$  分解为两个正交分量；故在像素空间进行投影，被滤除的分量(位于  $H$  的零空间)正是需要由生成模型先验来填补和细化的部分；在上述投影公式中，去噪像素表示  $\hat{x}_0(y)$  的估计由两个部分组成：第一项对应于时间步  $t$  投影和反投影操作之间的残差修正，而第二项编码了从原始输入图像保留的关键信息。

$$H\hat{x}_0(y) \equiv H\left[\left(I - H^+H\right)x_{0_t} + H^+y\right] \equiv y \tag{19}$$

该公式表明，在投影子空间的反投影部分  $\hat{x}_0(y)$  确实与输入测量图像完全相同。为了防止在工作过程中从初始值使用该投影公式，我们选择从中间状态  $x_{t+1}$  开始。为了提高计算效率，整个计算过程我们保留原始算法中的三个阶段：第一阶段不对其使用任何优化方法，第二阶段主要优化潜在空间，最后阶段主要在像素空间进行优化；由于采样过程中存在连续性，随后步骤在优化后会在一定程度上保持一定的信息，因此我们仍然使用跳步机制进行优化，但是我们选择每 5 步使用一次优化时间步。

### 4. 实验

在本研究中，我们对来自 FFHQ 和 CelebA 数据集的自然图像进行了线性逆问题的实验。我们将提出的改进算法与直接在像素空间中运行的扩散模型进行了比较，使用了多种定量评估指标，包括学习感知图像块相似性(LPIPS)、峰值信噪比(PSNR)、结构相似性指数(SSIM)和 Fréchet Inception 距离(FID)。

**Table 1.** Quantitative evaluation on the FFHQ datasets for various typical IR tasks  
**表 1.** 在 FFHQ 数据集上对各种典型 IR 任务的定量评估

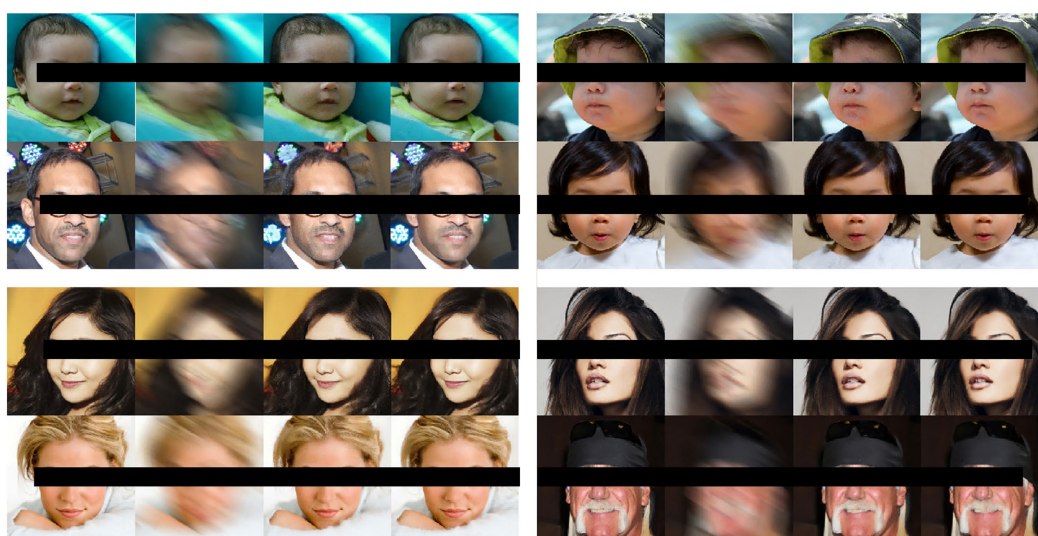
方法	PSNR↑	SSIM↑	LPIPS↓
Gaussian-deblur			
Resample	28.61	0.790	0.269
our	<b>28.80</b>	<b>0.810</b>	<b>0.244</b>
inpainting			
Resample	32.64	0.928	0.109
our	<b>32.87</b>	<b>0.932</b>	<b>0.103</b>
Motion-deblur			
Resample	28.85	0.796	0.279
our	<b>30.23</b>	<b>0.845</b>	<b>0.221</b>
SR × 4			
Resample	<b>28.93</b>	0.801	0.220
our	28.53	<b>0.818</b>	<b>0.218</b>

实验使用 FFHQ 和 CelebA-HQ 数据集进行，分辨率为  $256 \times 256 \times 3$ 。对于这两个数据集，我们各自提取了 100 张图像。对于 FFHQ 数据集和 CelebA-HQ 数据集，我们各选择了 100 张图像。在所有实验中，我们添加了标准差为  $\delta_y = 0.01$  的高斯测量噪声。我们使用了预训练的潜在扩散模型 LDMVQ4，该模型通过基于自编码器的架构将图像压缩为大小为  $64 \times 64 \times 3$  的潜在表示，并在 FFHQ 和 CelebA-HQ 上进行了训练(Rombach *et al.*, 2022) [5]，其中表 1 和表 2 分别展示了我们在两个数据集上的实验结果对比，图 1 和图 2 为我们在两个数据上对噪声或模糊图像的恢复效果的直观展示。

**Table 2.** Quantitative evaluation on the CelebA-HQ datasets for various typical IR tasks  
**表 2.** 在 CelebA-HQ 数据集上对各种典型 IR 任务的定量评估

方法	PSNR↑	SSIM↑	LPIPS↓
Gaussian-deblur			
Resample	30.09	0.843	0.193
our	<b>30.19</b>	<b>0.847</b>	<b>0.181</b>
inpainting			
Resample	33.86	0.937	0.107
our	<b>33.94</b>	<b>0.940</b>	<b>0.103</b>
Motion-deblur			
Resample	30.65	0.849	0.207
our	<b>30.73</b>	<b>0.856</b>	<b>0.188</b>
SR × 4			
Resample	<b>30.49</b>	<b>0.860</b>	<b>0.173</b>
our	30.10	0.852	0.179

所考虑的线性逆问题包括：(1) 使用大小为  $61 \times 61$ 、标准差为 3.0 的卷积核的高斯去模糊；(2) 使用随机二值掩码的图像修复；以及(3) 基于双三次下采样的超分辨率。实验结果表明，在所有任务中，在 FFHQ 上评估的模型始终优于在 CelebA 上的模型。此外，在 4 倍超分辨率(SR × 4)任务中，与高斯去模糊、运动去模糊和图像修复任务的结果相比，所有数据结果的表现均较差。



**Figure 1.** In the motion-Deblur task on the FFHQ (top) and CelebA-HQ (bottom) datasets, the results of the proposed method are compared with the baseline methods. The results are grouped by task and dataset, and the methods are arranged in the order of Reference, Measure, ReSample, and ours

**图 1.** 在 FFHQ (上)和 CelebA-HQ (下)数据集上的运动去模糊任务中，将所提出方法的结果与基线方法进行比较。结果按任务和数据集分组，方法按参考、测量、重采样和我们的顺序排列



**Figure 2.** In the inpainting task on the FFHQ (top left) and CelebA-HQ (bottom left) datasets, and the gaussian-deblur task on the FFHQ (top right) and CelebA-HQ (bottom right) datasets, the results of the proposed method are compared with the baseline methods. The results are grouped by task and dataset, and the methods are arranged in the order of Reference, Measure, ReSample, and ours

**图 2.** 在 FFHQ (左上)和 CelebA-HQ (左下)数据集上的图像修复任务, 以及在 FFHQ (右上)和 CelebA-HQ (右下)数据集上的高斯去模糊任务中, 提出的方法的结果与基线方法进行了比较。结果按任务和数据集进行分组, 方法按 Reference、Measure、ReSample 和我们的顺序排列

在对两个数据集进行的实验中, 可以观察到, 对于运动去模糊任务, 我们的方法在 FFHQ 和 CelebA-HQ 数据集上都显著优于基线方法。这在 FFHQ 上尤为明显, 其 PSNR 以相当大的幅度更高。然而, 对于  $SR \times 4$  任务, 我们的方法表现不太理想。特别是在 CelebA-HQ 数据集上, 我们的结果始终不如  $4 \times$  超分辨率任务的结果, 其存在的一定原因可能在于在我们设置的算法中超分辨率中的  $H$  是下采样,  $H^+$  是简单的上采样,  $H^+H$  是严重的低通滤波操作, 会抹除  $x_{0_t}$  中已生成的高频细节最终输出倾向于退化为  $H^+y$ , 而非基于扩散先验的高质量重建。

## 5. 总结

在本工作中, 我们在 ReSample 的基础方法上进行构建, 并引入了一种结合零空间分解的策略。为了应对潜在空间和像素空间的不同特性, 我们在潜在空间中保留原有方法, 同时将像素空间的处理替换为零空间分解。此外, 采样步骤的数量从 10 步减少到 5 步。实验结果表明,  $4 \times$  超分辨率任务不适用于所提出的方法, 而所提出的方法在大多数情况下优于修改后的算法, 取得了性能提升。需要注意的是, 目前的研究仅限于线性问题, 尚未扩展到非线性场景。此外, 由于扩散步骤数量的调整, 每张图像的推理时间比原始设置增加了约 10 到 20 分钟。

## 参考文献

- [1] Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N. and Ganguli, S. (2015) Deep Unsupervised Learning Using Nonequilibrium Thermodynamics. *International Conference on Machine Learning (ICML)*, Vol. 37, 2256-2265.
- [2] Song, Y. and Ermon, S. (2019) Generative Modeling by Estimating Gradients of the Data Distribution. *Advances in Neural Information Processing Systems*, Vancouver, 8-14 December 2019, 11895-11907.
- [3] Ho, J., Jain, A. and Abbeel, P. (2020) Denoising Diffusion Probabilistic Models. *Advances in Neural Information Processing Systems (NeurIPS)*, Volume 33, 6840-6851.
- [4] Dhariwal, P. and Nichol, A.Q. (2021) Diffusion Models Beat GANs on Image Synthesis. *Advances in Neural Information*

*Processing Systems*, Volume 34, 8780-8794.

- [5] Rombach, R., Blattmann, A., Lorenz, D., Esser, P. and Ommer, B. (2022) High-Resolution Image Synthesis with Latent Diffusion Models. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, 18-24 June 2022, 10674-10685. <https://doi.org/10.1109/cvpr52688.2022.01042>
- [6] Lipman, Y., Chen, R.T., Ben-Hamu, H., Nickel, M. and Le, M. (2022) Flow Matching for Generative Modeling.
- [7] Liu, X.C., Gong, C.Y. and Liu, Q. (2022) Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow.
- [8] Albergo, M.S., Boffi, N.M. and Vanden-Eijnden, E. (2023) Stochastic Interpolants: A Unifying Framework for Flows and Diffusions.
- [9] Ma, N., Goldstein, M., Albergo, M.S., Boffi, N.M., Vanden-Eijnden, E. and Xie, S. (2024) SiT: Exploring Flow and Diffusion-Based Generative Models with Scalable Interpolant Transformers. *Computer Vision—ECCV 2024 18th European Conference*, Milan, 29 September-4 October 2024, 23-40.
- [10] Song, Y., Shen, L.Y., Xing, L. and Ermon, S. (2021) Solving Inverse Problems in Medical Imaging with Score-Based Generative Models. *International Conference on Learning Representations (ICLR)*, Vienna, 3-7 May 2021, 18 p.
- [11] Kawar, B., Elad, M., Ermon, S. and Song, J. (2022) Denoising Diffusion Restoration Models. *Proceedings of the 36th International Conference on Neural Information Processing Systems*, New Orleans, 28 November-9 December 2022, 23593-2360.
- [12] Chung, H., Kim, J., Mccann, M.T., Klasky, M.L. and Ye, J.C. (2023) Diffusion Posterior Sampling for General Noisy Inverse Problems. *The 11th International Conference on Learning Representations*, Kigali, 1-5 May 2023, 30 p.
- [13] Dhariwal, P. and Nichol, A. (2021) Diffusion Models Beat GANs on Image Synthesis. *Proceedings of the 35th International Conference on Neural Information Processing Systems*, 6-14 December 2021, 8780-8794.
- [14] Karras, T., Aittala, M., Aila, T. and Laine, S. (2022) Elucidating the Design Space of Diffusion-Based Generative Models. *Proceedings of the 36th International Conference on Neural Information Processing Systems*, New Orleans, 28 November-9 December 2022, 26565-26577.
- [15] Song, Y., Dhariwal, P., Chen, M. and Sutskever, I. (2023) Consistency Models. *Proceedings of the 40th International Conference on Machine Learning*, Honolulu, 23-29 July 2023, 32211-32252.
- [16] Lou, A. and Ermon, S. (2023) Reflected Diffusion Models. *Proceedings of the 40th International Conference on Machine Learning*, Honolulu, 23-29 July 2023, 22675-22701.
- [17] Chan, W., Denton, E., Fleet, D., Ghasemipour, K., Lopes, R.G., Ho, J., *et al.* (2022) Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding. *Advances in Neural Information Processing Systems 35*, New Orleans, 28 November-9 December 2022, 36479-36494. <https://doi.org/10.52202/068431-2643>
- [18] Vincent, P. (2011) A Connection between Score Matching and Denoising Autoencoders. *Neural Computation*, **23**, 1661-1674. [https://doi.org/10.1162/neco\\_a\\_00142](https://doi.org/10.1162/neco_a_00142)
- [19] Chung, H., Ryu, D., Sim, B. and Ye, J.C. (2022) Improving Diffusion Models for Inverse Problems Using Manifold Constraints. *Advances in Neural Information Processing Systems 35*, New Orleans, 28 November-9 December 2022, 25683-25696. <https://doi.org/10.52202/068431-1862>
- [20] Wang, Y.H., Yu, J.W. and Zhang, J. (2022) Zero-Shot Image Restoration Using Denoising Diffusion Null-Space Model. *The 11th International Conference on Learning Representations, ICLR 2023*, Kigali, 1-5 May 2023, 31 p.
- [21] Song, J.M., Vahdat, A., Mardani, M. and Kautz, J. (2023) Pseudoinverse-Guided Diffusion Models for Inverse Problems. *International Conference on Learning Representations*, Kigali, 1-5 May 2023, 30 p.
- [22] Song, B.Q., Kwon, S.M., Zhang, Z.X., Zhen, X.T., Lei, Q. and Hu, L.J. (2024) Solving Inverse Problems with Latent Diffusion Models via Hard Data Consistency. *International Conference on Learning Representations*, Vienna, 7-11 May 2024, 31 p.