Zero-Shot Based Diffusion Model in Image Restoration

- 涂陆炜

preliminaries

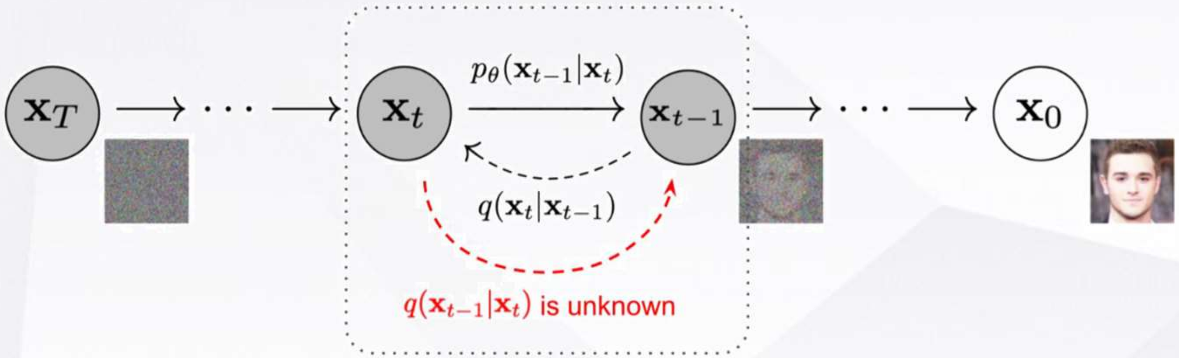
DDPM: Denoising Diffusion Probabilistic Models

True data dist. : $x_0 \sim q(x_0)$

Forward process: $q(x_{1:T}|x_0) := \prod_{t=1}^{T} q(x_t|x_{t-1})$

Reverse process: $p_\theta(x_{0:T}) := p(x_T) \prod_{t=1}^{T} p_\theta(x_{t-1}|x_t)$

Markov assumption



Algorithm 1 Training

1: **repeat**
2: $\mathbf{x}_0 \sim q(\mathbf{x}_0)$
3: $t \sim \mathrm{Uniform}(\{1, \ldots, T\})$
4: $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
5: Take gradient descent step on
$$\nabla_\theta \left\| \boldsymbol{\epsilon} - \mathbf{z}_\theta(\sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}, t) \right\|^2$$
6: **until** converged

Algorithm 2 Sampling

1: $\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
2: **for** $t = T, \ldots, 1$ **do**
3: $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ if $t > 1$, else $\mathbf{z} = \mathbf{0}$
4: $\boxed{\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}}\mathbf{z}_\theta(\mathbf{x}_t, t)\right) + \sigma_t \mathbf{z}}$
5: **end for**
6: **return** $\mathbf{x}_0$

Forward:

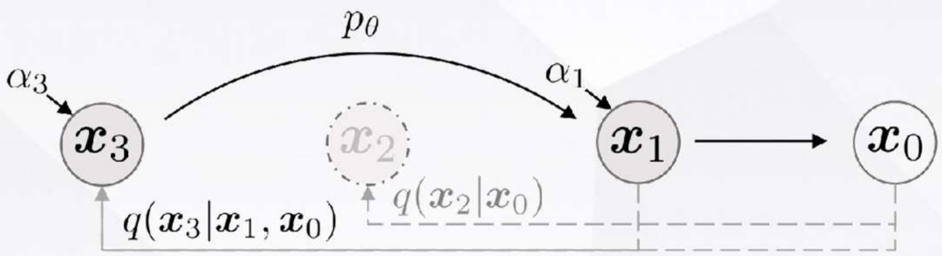
$$x_t = \sqrt{\bar{\alpha}_t}x_0 + \sqrt{1-\bar{\alpha}_t}\epsilon$$

Reverse:

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(x_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}}\epsilon_\theta\right) + \sigma_t z \quad (z \sim \mathcal{N}(0, \mathbf{I}))$$

DDIM: Denoising Diffusion Implicit Models

DDPM: Markov assumption → 慢



Reverse:

$$x_{prev} = \sqrt{\bar{\alpha}_{prev}}\left(\frac{x_t - \sqrt{1-\bar{\alpha}_t}\epsilon_t}{\sqrt{\bar{\alpha}_t}}\right) + \sqrt{1-\bar{\alpha}_{prev} - \sigma^2}\epsilon_t + \sigma^2\epsilon$$

preliminaries

中山大學 SUN YAT-SEN UNIVERSITY

stochastic differential eqaution (SDE) 从离散到连续

Forward:

$$\mathrm{d}\boldsymbol{x}_t = \underbrace{\boldsymbol{f}(\boldsymbol{x}_t, t)}_{\text{drift coeff.}}\,\mathrm{d}t + \underbrace{g(t)}_{\text{diffusion coeff.}}\,\mathrm{d}\boldsymbol{W}_t,$$

漂移系数 扩散系数 Wiener process(布朗运动噪声)

Reverse:

$$\mathrm{d}\boldsymbol{x}_t = \left(\boldsymbol{f}(\boldsymbol{x}_t, t) - g^2(t)\underbrace{\nabla_{\boldsymbol{x}_t}\log p_t(\boldsymbol{x}_t)}_{\text{score}}\right)\mathrm{d}t + g(t)\mathrm{d}\boldsymbol{W}_t,$$

方差保持型SDE（VP-SDE）：

$$\boldsymbol{s}_\theta(\boldsymbol{x}_t, t) \approx \nabla_{\boldsymbol{x}_t}\log p(\boldsymbol{x}_t|\boldsymbol{x}_0)$$ Denoising score matching

$$\boldsymbol{f}(x, t) = -\frac{1}{2}\beta(t)x, \quad g(t) = \sqrt{\beta(t)}$$

确定性ODE（常微分方程），是SDE的特例（扩散系数为零, 消除随机性）

$$g(t) = 0, \quad f(x, t) = -\frac{1}{2}\beta(t)x + \beta(t)\nabla_x \log p_t(x)$$

$$\frac{\mathrm{d}\boldsymbol{x}_t}{\mathrm{d}t} = \left(\boldsymbol{f}(\boldsymbol{x}_t, t) - \frac{g^2(t)}{2}\underbrace{\nabla_{\boldsymbol{x}_t}\log p_t(\boldsymbol{x}_t)}_{\text{score}}\right)$$

$$\boldsymbol{x}_{t-1} = \sqrt{\alpha_{t-1}}\underbrace{\left(\frac{\boldsymbol{x}_t + (1-\alpha_t)\nabla_{\boldsymbol{x}_t}\log p_t(\boldsymbol{x}_t)}{\sqrt{\alpha_t}}\right)}_{=:\hat{\boldsymbol{x}}_0 = \text{predicted } \boldsymbol{x}_0} + \sqrt{1-\alpha_{t-1}-\sigma_t^2}\underbrace{\left(-\sqrt{1-\alpha_t}\nabla_{\boldsymbol{x}_t}\log p_t(\boldsymbol{x}_t)\right)}_{\text{direction toward } \boldsymbol{x}_t}$$

Problem formulation （degradation model）：

$$\mathbf{y} = \mathcal{H}\mathbf{x} + \mathbf{n}$$

Unconditional to conditional:

$$\mathrm{d}\boldsymbol{x}_t = \left(\boldsymbol{f}(\boldsymbol{x}_t, t) - g^2(t) \underbrace{\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t)}_{\text{score}} \right) \mathrm{d}t + g(t)\mathrm{d}\boldsymbol{W}_t,$$

$$\mathrm{d}\boldsymbol{x}_t = \left(\boldsymbol{f}(\boldsymbol{x}_t, t) - g^2(t) \underbrace{\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t | \boldsymbol{y})}_{\text{conditional score}} \right) \mathrm{d}t + g(t)\mathrm{d}\boldsymbol{W}_t,$$

Reconstrution（贝叶斯）：

$$\begin{aligned} \hat{\mathbf{x}} &\sim p(\mathbf{x}|\mathbf{y}) \\ &\propto p(\mathbf{x})p(\mathbf{y}|\mathbf{x}) \end{aligned}$$

$$\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t|\boldsymbol{y}) = \underbrace{\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t)}_{\text{score}} + \underbrace{\nabla_{\boldsymbol{x}_t} \log p(\boldsymbol{y}|\boldsymbol{x}_t)}_{\text{measurements matching term}}.$$

Conditional reverse formlation:

$$\mathrm{d}\boldsymbol{x}_t = \left(\boldsymbol{f}(\boldsymbol{x}_t, t) - g^2(t) \left(\boxed{\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t)} + \boxed{\nabla \log p(\boldsymbol{y}|\boldsymbol{x}_t)} \right) \right) \mathrm{d}t + g(t)\mathrm{d}\boldsymbol{W}_t.$$

数据分布（pretrain model）
unconditional model（已知）

条件分数
time-dependent classifier

Solution:

$$p_t(\boldsymbol{y}|\boldsymbol{x}_t) = \int p(\boldsymbol{y}|\boldsymbol{x}_0)p(\boldsymbol{x}_0|\boldsymbol{x}_t)\mathrm{d}\boldsymbol{x}_0.$$

Problem formulation （degradation model）：

$$\mathbf{y} = \mathcal{H}\mathbf{x} + \mathbf{n}$$

Reconstrution（贝叶斯）：

$$\begin{aligned}\hat{\mathbf{x}} &\sim p(\mathbf{x}|\mathbf{y})\\ &\propto p(\mathbf{x})p(\mathbf{y}|\mathbf{x})\end{aligned}$$

Maximum A Posteriori (MAP) estimation:
（最大后验估计）

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}} \log p(\mathbf{y}|\mathbf{x}) + \log p(\mathbf{x})$$

Problem transformation:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \underbrace{\frac{1}{2\sigma_n^2}\|\mathbf{y} - \mathcal{H}(\mathbf{x})\|^2}_{\text{data term}} + \underbrace{\lambda\mathcal{P}(\mathbf{x})}_{\text{prior term}}$$

Solution（HQS）：

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \frac{1}{2\sigma^2}\|\mathbf{y} - \mathcal{H}(\mathbf{x})\|^2 + \lambda\mathcal{P}(\mathbf{z}) \quad s.t. \quad \mathbf{z} = \mathbf{x}$$

引入辅助变量 z

$$\begin{cases}\mathbf{z}_k = \arg\min_{\mathbf{z}} \underbrace{\frac{1}{2(\sqrt{\lambda/\mu})^2}\|\mathbf{z} - \mathbf{x}_k\|^2}_{\text{consistence}} + \underbrace{\mathcal{P}(\mathbf{z})}_{\text{prior}}\\ \mathbf{x}_{k-1} = \arg\min_{\mathbf{x}} \underbrace{\|\mathbf{y} - \mathcal{H}(\mathbf{x})\|^2}_{\text{condition}} + \underbrace{\mu\sigma_n^2\|\mathbf{x} - \mathbf{z}_k\|^2}_{\text{consistence}}\end{cases}$$

Problem formulation 2

Solution（HQS）：

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \frac{1}{2\sigma^2}\|\mathbf{y}-\mathcal{H}(\mathbf{x})\|^2 + \lambda\mathcal{P}(\mathbf{z}) \quad s.t. \quad \mathbf{z}=\mathbf{x}$$

引入辅助变量 **z**

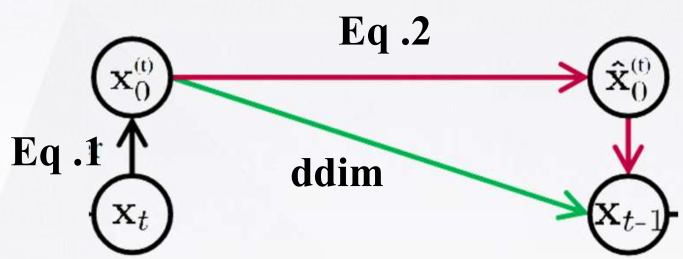
$$\begin{cases} \mathbf{z}_k = \arg\min_{\mathbf{z}} \underbrace{\frac{1}{2(\sqrt{\lambda/\mu})^2}\|\mathbf{z}-\mathbf{x}_k\|^2}_{\text{consistence}} + \underbrace{\mathcal{P}(\mathbf{z})}_{\text{prior}} & \textbf{Eq .1} \\ \mathbf{x}_{k-1} = \arg\min_{\mathbf{x}} \underbrace{\|\mathbf{y}-\mathcal{H}(\mathbf{x})\|^2}_{\text{condition}} + \underbrace{\mu\sigma_n^2\|\mathbf{x}-\mathbf{z}_k\|^2}_{\text{consistence}} & \textbf{Eq .2} \end{cases}$$

Eq .1: 相当于需要在**xk**中引入数据先验，相当于使用扩散模型预测一个符合数据分布的**x0**

$$\mathbf{x}_0^{(t)} = \frac{1}{\sqrt{\bar{\alpha}_t}}(\mathbf{x}_t + (1-\bar{\alpha}_t)\mathbf{s}_\theta(\mathbf{x}_t, t))$$

Eq .2: 相当于对齐项，也就是需要符合**y=H(x)+n**这样的退化过程（条件分布）

等同于求 $$p_t(\boldsymbol{y}|\boldsymbol{x}_t) = \int p(\boldsymbol{y}|\boldsymbol{x}_0)p(\boldsymbol{x}_0|\boldsymbol{x}_t)\mathrm{d}\boldsymbol{x}_0.$$



Zero-Shot Image Restoration Using Denoising Diffusion Null-Space Model

Range null space decomposition

$$\mathbf{A}\mathbf{A}^{\dagger}\mathbf{A} \equiv \mathbf{A}.$$

伪逆

$$\mathbf{x} \equiv \mathbf{A}^{\dagger}\mathbf{A}\mathbf{x} + (\mathbf{I} - \mathbf{A}^{\dagger}\mathbf{A})\mathbf{x}.$$

Range space **Null space**

$$\mathbf{A}\mathbf{x} \equiv \mathbf{A}\mathbf{A}^{\dagger}\mathbf{A}\mathbf{x} + \mathbf{A}(\mathbf{I} - \mathbf{A}^{\dagger}\mathbf{A})\mathbf{x} \equiv \mathbf{A}\mathbf{x} + \mathbf{0} \equiv \mathbf{y}.$$

可以取任何值，都能符合**y=Ax**

$$\mathbf{x} \equiv \mathbf{A}^{\dagger}\mathbf{A}\mathbf{x} + (\mathbf{I} - \mathbf{A}^{\dagger}\mathbf{A})\mathbf{x}. \quad \hat{\mathbf{x}} = \mathbf{A}^{\dagger}\mathbf{y} + (\mathbf{I} - \mathbf{A}^{\dagger}\mathbf{A})\bar{\mathbf{x}}.$$

Zero-Shot Image Restoration Using Denoising Diffusion Null-Space Model

From DDIM:

$$p(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) \qquad \boldsymbol{\mu}_t(\mathbf{x}_t, \mathbf{x}_0) = \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1-\bar{\alpha}_t}\mathbf{x}_0 + \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_t, \quad \sigma_t^2 = \frac{1-\bar{\alpha}_{t-1}}{1-\bar{\alpha}_t}\beta_t,$$

$$\hat{\mathbf{x}}_{0|t} = \mathbf{A}^{\dagger}\mathbf{y} + (\mathbf{I} - \mathbf{A}^{\dagger}\mathbf{A})\mathbf{x}_{0|t}.$$

$$\mathbf{x}_{0|t} = \frac{1}{\sqrt{\bar{\alpha}_t}}\left(\mathbf{x}_t - \mathcal{Z}_{\boldsymbol{\theta}}(\mathbf{x}_t, t)\sqrt{1-\bar{\alpha}_t}\right).$$

Diffusion prior

$$\mathbf{A}\hat{\mathbf{x}}_{0|t} \equiv \mathbf{A}\mathbf{A}^{\dagger}\mathbf{y} + \mathbf{A}(\mathbf{I} - \mathbf{A}^{\dagger}\mathbf{A})\mathbf{x}_{0|t} \equiv \mathbf{A}\mathbf{A}^{\dagger}\mathbf{A}\mathbf{x} + \mathbf{0} \equiv \mathbf{A}\mathbf{x} \equiv \mathbf{y}.$$

consistency

Core: x_{t-1} is the noisy version of x_{0t}

Algorithm 1 Sampling of DDNM

1: $\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
2: **for** $t = T, ..., 1$ **do**

Diffusion prior

3: $\boxed{\mathbf{x}_{0|t} = \frac{1}{\sqrt{\bar{\alpha}_t}}\left(\mathbf{x}_t - \mathcal{Z}_{\boldsymbol{\theta}}(\mathbf{x}_t, t)\sqrt{1-\bar{\alpha}_t}\right)}$
4: $\boxed{\hat{\mathbf{x}}_{0|t} = \mathbf{A}^{\dagger}\mathbf{y} + (\mathbf{I} - \mathbf{A}^{\dagger}\mathbf{A})\mathbf{x}_{0|t}}$ **consistency**
5: $\mathbf{x}_{t-1} \sim p(\mathbf{x}_{t-1}|\mathbf{x}_t, \hat{\mathbf{x}}_{0|t})$
6: **return** $\mathbf{x}_0$

Zero-Shot Image Restoration Using Denoising Diffusion Null-Space Model

With Gaussian Noise: $\mathbf{y}=\mathbf{A}\mathbf{x}+\mathbf{n}$

$$\hat{\mathbf{x}}_{0|t}=\mathbf{A}^{\dagger}\mathbf{y}+(\mathbf{I}-\mathbf{A}^{\dagger}\mathbf{A})\mathbf{x}_{0|t}.$$

$$\hat{\mathbf{x}}_{0|t}=\mathbf{A}^{\dagger}\mathbf{y}+(\mathbf{I}-\mathbf{A}^{\dagger}\mathbf{A})\mathbf{x}_{0|t}=\mathbf{x}_{0|t}-\mathbf{A}^{\dagger}(\mathbf{A}\mathbf{x}_{0|t}-\mathbf{A}\mathbf{x})+\mathbf{A}^{\dagger}\mathbf{n},$$ 不符合**consistency**

Solution: 利用扩散模型本身的denoiser来去除值域中的噪声 n: 保证含噪图 xt−1 中的噪声方差和原始定义的噪声方差一致，这样对于denoiser来说，面临的噪声就是不变的

$$\hat{\mathbf{x}}_{0|t}=\mathbf{x}_{0|t}-\Sigma_t\mathbf{A}^{\dagger}(\mathbf{A}\mathbf{x}_{0|t}-\mathbf{y}),$$ **Scale the range null space**

$$\hat{p}(\mathbf{x}_{t-1}|\mathbf{x}_t,\hat{\mathbf{x}}_{0|t})=\mathcal{N}(\mathbf{x}_{t-1};\boldsymbol{\mu}_t(\mathbf{x}_t,\hat{\mathbf{x}}_{0|t}),\mathbf{\Phi}_t\mathbf{I}).$$ **Scale the noise**

Close to I

1. **Consistency**
2. **Preserve null space**

Rescale the variance of noise

$$\mathcal{N}(\mathbf{x}_{t-1};\boldsymbol{\mu}_t(\mathbf{x}_t,\mathbf{x}_0),\sigma_t^2\mathbf{I})$$

Results without noise

ImageNet	4× SR	Deblurring	Colorization	CS 25%	Inpainting
Method	PSNR↑/SSIM↑/FID↓	PSNR↑/SSIM↑/FID↓	*Cons*↓/FID↓	PSNR↑/SSIM↑/FID↓	PSNR↑/SSIM↑/FID↓
$\mathbf{A}^{\dagger}\mathbf{y}$	24.26 / 0.684 / 134.4	18.56 / 0.6616 / 55.42	0.0 / 43.37	15.65 / 0.510 / 277.4	14.52 / 0.799 / 72.71
DGP	23.18 / 0.798 / 64.34	N/A	- / 69.54	N/A	N/A
ILVR	27.40 / **0.870** / 43.66	N/A	N/A	N/A	N/A
RePaint	N/A	N/A	N/A	N/A	31.87 / **0.968** / 12.31
DDRM	27.38 / 0.869 / 43.15	43.01 / 0.992 / 1.48	260.4 / 36.56	19.95 / 0.704 / 97.99	31.73 / 0.966 / 4.82
DDNM(ours)	**27.46 / 0.870/ 39.26**	**44.93 / 0.994 / 1.15**	**42.32 / 36.32**	**21.66 / 0.749 / 64.68**	**32.06 / 0.968 / 3.89**

CelebA	4× SR	Deblurring	Colorization	CS 25%	Inpainting
Method	PSNR↑/SSIM↑/FID↓	PSNR↑/SSIM↑/FID↓	*Cons*↓/FID↓	PSNR↑/SSIM↑/FID↓	PSNR↑/SSIM↑/FID↓
$\mathbf{A}^{\dagger}\mathbf{y}$	27.27 / 0.782 / 103.3	18.85 / 0.741 / 54.31	0.0 / 68.81	15.09 / 0.583 / 377.7	15.57 / 0.809 / 181.56
PULSE	22.74 / 0.623 / 40.33	N/A	N/A	N/A	N/A
ILVR	31.59 / **0.945** / 29.82	N/A	N/A	N/A	N/A
RePaint	N/A	N/A	N/A	N/A	35.20 / 0.981 /14.19
DDRM	**31.63 / 0.945** / 31.04	43.07 / 0.993 / 6.24	455.9 / 31.26	24.86 / 0.876 / 46.77	34.79 / 0.978 /12.53
DDNM(ours)	**31.63 / 0.945 / 22.27**	**46.72 / 0.996 / 1.41**	**26.25 / 26.44**	**27.56 / 0.909 / 28.80**	**35.64 / 0.982 /4.54**

Table 1: Quantitative results of zero-shot IR methods on **ImageNet**(*top*) and **CelebA**(*bottom*), including five typical IR tasks. We mark N/A for those not applicable and **bold** the best scores.

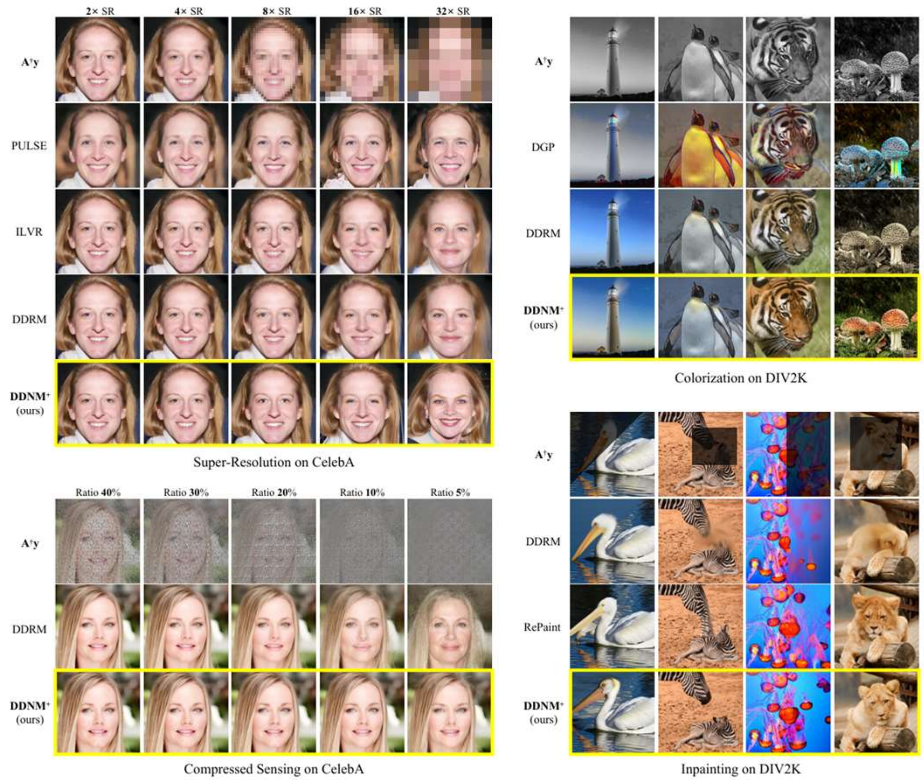
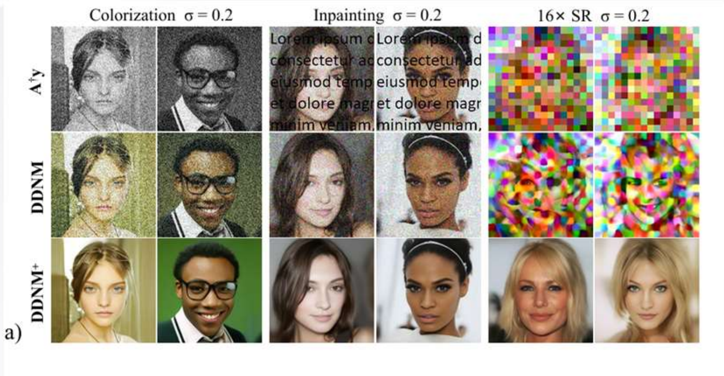


Figure 3: Qualitative results of zero-shot IR methods.

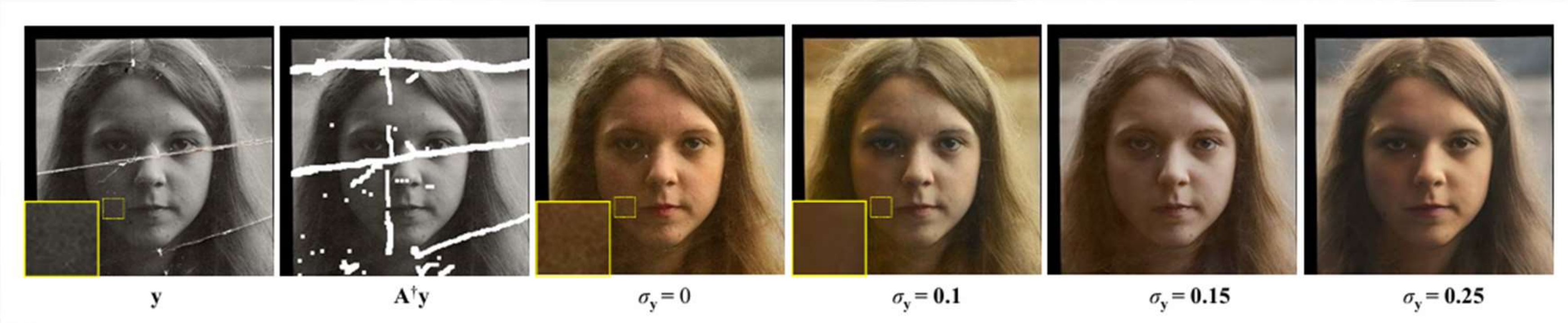
DDNM+

Results with noise



With noise

JPEG compression



Old photo restoration

Decomposed diffusion sampler for accelerating large-scale inverse problems

Algorithm 1 Conjugate Gradient (CG)

Require: $\boldsymbol{A}, \boldsymbol{y}, \boldsymbol{x}_0, M$

$\boldsymbol{r}_0 \leftarrow \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}_0$

$\boldsymbol{p}_0 \leftarrow \mathrm{b}_0$

for $i = 0 : K - 1$ **do**

$\quad \alpha_k \leftarrow \frac{\boldsymbol{r}_k^\top \boldsymbol{r}}{\boldsymbol{p}_k^\top \boldsymbol{A}\boldsymbol{p}_k}$

$\quad \boldsymbol{x}_{k+1} \leftarrow \boldsymbol{x}_k + \alpha_k \boldsymbol{p}_k$

$\quad \boldsymbol{r}_{k+1} \leftarrow \boldsymbol{r}_k - \alpha_k \boldsymbol{A}\boldsymbol{p}_k$

$\quad \beta_k \leftarrow \frac{\boldsymbol{r}_k^\top \boldsymbol{r}}{\boldsymbol{r}_k^\top \boldsymbol{r}_k}$

$\quad \boldsymbol{p}_{k+1} \leftarrow \mathrm{b}_{k+1} + \beta_k \boldsymbol{p}_k$

end for

return $\boldsymbol{x}_K$

Algorithm 2 DDS (PI MRI; VP; noiseless)

Require: $\boldsymbol{\epsilon}_{\theta^*}, N, \{\alpha_t\}_{t=1}^N, \eta, \boldsymbol{A}, M$

$\boldsymbol{x}_N \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$

for $t = N : 2$ **do**

$\quad \hat{\boldsymbol{\epsilon}}_t \leftarrow \boldsymbol{\epsilon}_{\theta^*}(\boldsymbol{x}_t)$

$\quad \triangleright$ Tweedie denoising

$\quad \hat{\boldsymbol{x}}_t \leftarrow (\boldsymbol{x}_t - \sqrt{1 - \bar{\alpha}_t}\hat{\boldsymbol{\epsilon}}_t)/\sqrt{\bar{\alpha}_t}$ **Diffusion prior**

$\quad \triangleright$ Data consistency

$\quad \hat{\boldsymbol{x}}_t' \leftarrow \mathtt{CG}(\boldsymbol{A}^*\boldsymbol{A}, \boldsymbol{A}^*\boldsymbol{y}, \hat{\boldsymbol{x}}_t, M)$ **consistency**

$\quad \boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$

$\quad \triangleright$ DDIM sampling

$\quad \boldsymbol{x}_{t-1} \leftarrow \sqrt{\bar{\alpha}_{t-1}}\hat{\boldsymbol{x}}_t' - \sqrt{1 - \bar{\alpha}_{t-1} - \eta^2\tilde{\beta}_t^2}\hat{\boldsymbol{\epsilon}}_t + \eta\tilde{\beta}_t\boldsymbol{\epsilon}$

end for

$\boldsymbol{x}_0 \leftarrow (\boldsymbol{x}_1 - \sqrt{1 - \bar{\alpha}_1}\boldsymbol{\epsilon}_{\theta^*}(\boldsymbol{x}_1))/\sqrt{\bar{\alpha}_1}$

return $\boldsymbol{x}_0$

Decomposed diffusion sampler for accelerating large-scale inverse problems

Algorithm 3 DDS (PI MRI; VP; noisy)

Require: $\epsilon_{\theta^*}, N, \{\alpha_t\}_{t=1}^{N}, \eta, \boldsymbol{A}, M, \gamma$

$\boldsymbol{x}_N \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$
for $t = N : 2$ **do**
  $\hat{\boldsymbol{\epsilon}}_t \leftarrow \boldsymbol{\epsilon}_{\theta^*}(\boldsymbol{x}_t)$
  ▷ Tweedie denoising
  $\hat{\boldsymbol{x}}_t \leftarrow (\boldsymbol{x}_t - \sqrt{1 - \bar{\alpha}_t}\hat{\boldsymbol{\epsilon}}_t)/\sqrt{\bar{\alpha}_t}$
  ▷ Data consistency
  $\boldsymbol{A}_{\mathrm{CG}} \leftarrow \boldsymbol{I} + \gamma \boldsymbol{A}^* \boldsymbol{A}$
  $\boldsymbol{y}_{\mathrm{CG}} \leftarrow \hat{\boldsymbol{x}}_t + \gamma \boldsymbol{A}^* \boldsymbol{y}$
  $\hat{\boldsymbol{x}}'_t \leftarrow \mathtt{CG}(\boldsymbol{A}_{\mathrm{CG}}, \boldsymbol{y}_{\mathrm{CG}}, \hat{\boldsymbol{x}}_t, M)$
  $\boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0}, \mathbf{I})$
  ▷ DDIM sampling
  $\boldsymbol{x}_{t-1} \leftarrow \sqrt{\bar{\alpha}_{t-1}}\hat{\boldsymbol{x}}'_t - \sqrt{1 - \bar{\alpha}_{t-1} - \eta^2 \tilde{\beta}_t^2}\hat{\boldsymbol{\epsilon}}_t + \eta \tilde{\beta}_t \boldsymbol{\epsilon}$
end for
$\boldsymbol{x}_0 \leftarrow (\boldsymbol{x}_1 - \sqrt{1 - \bar{\alpha}_1}\boldsymbol{\epsilon}_{\theta^*}(\boldsymbol{x}_1))/\sqrt{\bar{\alpha}_1}$
return $\boldsymbol{x}_0$

DPS

Decomposed diffusion sampler for accelerating large-scale inverse problems

Mask Pattern	Acc.		TV	Supervised U-Net Zbontar et al. 2018	E2E-Varnet Sriram et al. 2020	Jalal *et al.* (2100)	Score-MRI ($4000 \times 2 \times C^*$)	DPS (1000)	DDS VP (99)	DDS VP (49)	DDS VP (19)
Uniform 1D	× 4	PSNR [db]	27.32 ±0.43	31.77±0.89	32.96±0.59	32.49±2.10	33.25±1.18	30.56±0.66	**34.88**±0.74	34.61±0.32	32.73±2.04
		SSIM	0.662±0.17	0.846±0.11	0.856±0.11	0.868±0.08	0.857±0.08	0.840±0.20	0.954±0.11	**0.956**±0.08	0.927±0.08
	× 8	PSNR [db]	25.02±2.21	29.51±0.37	31.98±0.35	**32.19**±2.45	32.01±2.30	30.29±0.33	31.62±1.88	30.16±1.19	30.33±2.35
		SSIM	0.532±0.05	0.780±0.05	0.828±0.08	0.835±0.16	0.821±0.15	0.811±0.18	0.876±0.08	0.830±0.04	**0.891**±0.16
Gaussian 1D	× 4	PSNR [db]	30.55±1.77	32.66±0.26	34.15±1.40	33.98±1.25	34.25±1.33	32.47±1.09	35.12±1.37	**35.15**±0.39	34.63±1.95
		SSIM	0.789±0.06	0.866±0.12	0.878±0.19	0.881±0.12	0.885±0.08	0.838±0.20	**0.963**±0.15	0.961±0.06	0.957±0.09
	× 8	PSNR [db]	27.98±1.28	31.64±1.12	33.15±2.09	32.76±2.43	32.43±0.95	30.47±2.32	33.27±1.06	**33.43**±0.75	32.83±1.29
		SSIM	0.747±0.21	0.841±0.09	0.868±0.18	0.870±0.13	0.855±0.13	0.839±0.16	0.937±0.07	**0.947**±0.15	0.940±0.09
Gaussian 2D	× 8	PSNR [db]	29.20±2.37	24.51±0.69	20.97±1.24	30.97±1.14	31.43±1.23	29.65±1.26	33.99±1.30	**34.55**±1.69	32.55±1.54
		SSIM	0.781±0.09	0.724±0.10	0.642±0.08	0.812±0.17	0.831±0.18	0.795±0.12	0.948±0.13	**0.956**±0.15	0.916±0.14
	× 15	PSNR [db]	26.28±2.28	14.93±3.33	16.66±4.02	27.34±1.97	**29.17**±0.98	26.30±1.34	27.86±1.67	25.75±1.77	25.66±2.03
		SSIM	0.547±0.19	0.372±0.29	0.435±0.26	0.692±0.13	0.704±0.08	0.688±0.11	**0.732**±0.10	0.695±0.09	0.693±0.08
VD Poisson disk	× 8	PSNR [db]	29.52±1.26	20.89±3.09	20.70±3.08	32.60±1.88	31.98±0.51	31.05±0.46	35.31±0.79	35.36±0.41	**35.39**±0.57
		SSIM	0.562±0.11	0.576±0.10	0.592±0.18	0.833±0.05	0.816±0.07	0.811±0.08	0.897±0.07	0.875±0.09	**0.915**±0.11
	× 15	PSNR [db]	26.19±2.36	16.01±5.59	18.82±3.30	30.22±1.89	29.59±1.22	30.02±1.72	34.84±1.44	**35.18**±0.97	34.59±1.50
		SSIM	0.510±0.20	0.537±0.21	0.548±0.19	0.749±0.17	0.702±0.15	0.753±0.15	0.934±0.06	0.931±0.05	**0.940**±0.05

	Score-MRI	DDNM	DDS (ours)			
			1	3	5	10
PSNR[db]	26.48	31.36	31.51	33.78	**34.61**	32.48
SSIM	0.688	0.932	0.934	0.952	**0.956**	0.949

Without Noise

Mask Pattern	Acc.		TV	DPS (1000)	DDS VP (49)
Uniform 1D	× 4	PSNR [db]	24.19	24.40	**29.47**
		SSIM	0.687	0.656	**0.866**
	× 8	PSNR [db]	23.02	24.60	**26.77**
		SSIM	0.638	0.666	**0.827**
VD Poisson disk	× 8	PSNR [db]	23.07	23.48	**30.95**
		SSIM	0.609	0.592	**0.890**
	× 15	PSNR [db]	20.92	23.57	**29.36**
		SSIM	0.554	0.622	**0.853**

With Noise

Diffusion posterior sampling for general noisy inverse problems（DPS）
可以解决非线性逆问题

$$p_t(\boldsymbol{y}|\boldsymbol{x}_t) = \int p(\boldsymbol{y}|\boldsymbol{x}_0)p(\boldsymbol{x}_0|\boldsymbol{x}_t)\mathrm{d}\boldsymbol{x}_0.$$

DPS核心： $\nabla_{\boldsymbol{x}_t} \log p(\boldsymbol{y}|\boldsymbol{X}_t = \boldsymbol{x}_t) \approx \nabla_{\boldsymbol{x}_t} \log p(\boldsymbol{y}|\boldsymbol{X}_0 = \mathbb{E}[\boldsymbol{X}_0|\boldsymbol{X}_t = \boldsymbol{x}_t]).$

$$\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t|\boldsymbol{y}) = \nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t) + \nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{y}|\boldsymbol{x}_t)$$

$$\begin{aligned}
&p(y|x_t)\\
&= \int p(y|x_0, x_t)p(x_0|x_t)dx_0\\
&= \int p(y|x_0)p(x_0|x_t)dx_0\\
&= \mathbb{E}_{x\sim p(x_0|x_t)}[p(y|x_0)]
\end{aligned}$$

对从条件分布 p(x0|xt) 中采样的 x0，计算 p(y|x0) 的期望

$$\mathbb{E}_{x\sim p(x_0|x_t)}[p(y|x_0)] \approx p(y|\mathbb{E}_{x\sim p(x_0|x_t)}[x_0]) = p(y|\hat{x}_0)$$

将原本应该积分的随机变量 x0 用它的期望值E（x0|xt）替代

将xt下取y的概率看作与从x0（xt）下取y的概率

Diffusion posterior sampling for general noisy inverse problems（DPS）
可以解决非线性逆问题

Why？ $\mathbb{E}_{x\sim p(x_0|x_t)}[p(y|x_0)] \approx p(y|\mathbb{E}_{x\sim p(x_0|x_t)}[x_0]) = p(y|\hat{x}_0)$

詹森不等式： $J = \mathbb{E}[f(x)] - f(\mathbb{E}(x)) \leq \frac{d}{\sqrt{2\pi\sigma^2}} e^{-1/2\sigma^2} \|\nabla_x A(x)\| m_1$

$\|\nabla_x \mathcal{A}(x)\| := \max_x \|\nabla_x \mathcal{A}(x)\| \quad m_1 := \int \|x_0 - \hat{x}_0\| p(x_0|x_t)\, dx_0$

有限值　　有限值

噪声强度σ趋于无穷，该上界会收敛于0，对应于DPS能够在较大的噪声强度下表现好

对于**y=Ax+n**，**n**为高斯噪声的情况下：

$$p(y|x_0) = \frac{1}{\sqrt{(2\pi)^n \sigma^{2n}}} \exp\left[-\frac{\|y - A(x_0)\|_2^2}{2\sigma^2}\right]$$

$$\nabla_{x_t} \log p_t(y|x_t) \approx -\frac{1}{\sigma^2} \nabla_{x_t} \|y - A(\hat{x}_0(x_t))\|_2^2$$

$$\nabla_{x_t} \log p_t(x_t|y) = \underbrace{\nabla_{x_t} \log p_t(x_t)}_{\text{score}} + \underbrace{\nabla_{x_t} \log p(y|x_t)}_{\text{measurements matching term}} .$$

$$\nabla_{x_t} \log p_t(x_t|y) \simeq s_{\theta^*}(x_t, t) - \rho \nabla_{x_t} \|y - \mathcal{A}(\hat{x}_0)\|_2^2,$$

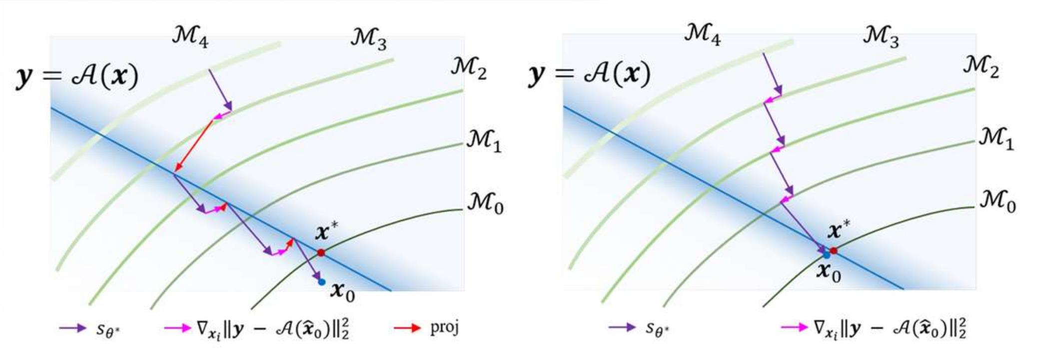
Algorithm 2 DPS - Gaussian [8]

Require: $N, y, \{\zeta_i\}_{i=1}^N, \{\tilde{\sigma}_i\}_{i=1}{}^N$

1: $x_N \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
2: **for** $i = N-1$ **to** 0 **do**
3: $\quad \hat{s} \leftarrow s_\theta(x_i, i)$
4: $\quad \hat{x}_0 \leftarrow \frac{1}{\sqrt{\bar{\alpha}_i}}(x_i + \sqrt{1-\bar{\alpha}_i}\,\hat{s})$
5: $\quad z \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
6: $\quad x'_{i-1} \leftarrow \frac{\sqrt{\alpha_i}(1-\bar{\alpha}_{i-1})}{1-\bar{\alpha}_i} x_i + \frac{\sqrt{\bar{\alpha}_{i-1}}\beta_i}{1-\bar{\alpha}_i}\hat{x}_0 + \tilde{\sigma}_i z$
7: $\quad x_{i-1} \leftarrow x'_{i-1} - \zeta_i \nabla_{x_i} \|y - \mathcal{A}(\hat{x}_0)\|_2^2$
8: **return** x_0

DPS

Advantage:



(a) Geometry of Chung et al. (2022a) (b) Geometry of DPS

Figure 3: Conceptual illustration of the geometries of two different diffusion processes. Our method prevents the sample from falling off the generative manifolds when the measurements are noisy.

Algorithm 2 DPS - Gaussian [8]

Require: $N, y, \{\zeta_i\}_{i=1}^N, \{\tilde{\sigma}_i\}_{i=1^N}$

1: $x_N \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$

2: **for** $i = N-1$ **to** 0 **do**

3: $\hat{s} \leftarrow s_\theta(x_i, i)$

Sample x_i-1 from p(x)

4: $\hat{x}_0 \leftarrow \frac{1}{\sqrt{\bar{\alpha}_i}}(x_i + \sqrt{1-\bar{\alpha}_i}\,\hat{s})$

5: $z \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$

6: $x'_{i-1} \leftarrow \frac{\sqrt{\alpha_i}(1-\bar{\alpha}_{i-1})}{1-\bar{\alpha}_i} x_i + \frac{\sqrt{\bar{\alpha}_{i-1}}\beta_i}{1-\bar{\alpha}_i}\hat{x}_0 + \tilde{\sigma}_i z$

7: $x_{i-1} \leftarrow x'_{i-1} - \zeta_i \nabla_{x_i} \|y - \mathcal{A}(\hat{x}_0)\|_2^2$

8: **return** x_0

condition

$$\begin{cases} \mathbf{z}_k = \arg\min_{\mathbf{z}} \underbrace{\frac{1}{2(\sqrt{\lambda/\mu})^2}\|\mathbf{z} - \mathbf{x}_k\|^2}_{\text{consistence}} + \underbrace{\mathcal{P}(\mathbf{z})}_{\text{prior}} & \textbf{Eq .1} \\ \mathbf{x}_{k-1} = \arg\min_{\mathbf{x}} \underbrace{\|\mathbf{y} - \mathcal{H}(\mathbf{x})\|^2}_{\text{condition}} + \underbrace{\mu\sigma_n^2\|\mathbf{x} - \mathbf{z}_k\|^2}_{\text{consistence}} & \textbf{Eq .2} \end{cases}$$

DPS

Method	SR (×4)		Inpaint (box)		Inpaint (random)		Deblur (gauss)		Deblur (motion)	
	FID ↓	LPIPS ↓	FID ↓	LPIPS ↓	FID ↓	LPIPS ↓	FID ↓	LPIPS ↓	FID ↓	LPIPS ↓
DPS (ours)	**39.35**	**0.214**	**33.12**	**0.168**	**21.19**	**0.212**	**44.05**	**0.257**	**39.92**	**0.242**
DDRM (Kawar et al., 2022)	<u>62.15</u>	<u>0.294</u>	42.93	<u>0.204</u>	69.71	0.587	<u>74.92</u>	<u>0.332</u>	-	-
MCG (Chung et al., 2022a)	87.64	0.520	<u>40.11</u>	0.309	<u>29.26</u>	<u>0.286</u>	101.2	0.340	310.5	0.702
PnP-ADMM (Chan et al., 2016)	66.52	0.353	151.9	0.406	123.6	0.692	90.42	0.441	<u>89.08</u>	<u>0.405</u>
Score-SDE (Song et al., 2021b) (ILVR (Choi et al., 2021))	96.72	0.563	60.06	0.331	76.54	0.612	109.0	0.403	292.2	0.657
ADMM-TV	110.6	0.428	68.94	0.322	181.5	0.463	186.7	0.507	152.3	0.508

Table 1: Quantitative evaluation (FID, LPIPS) of solving linear inverse problems on FFHQ 256×256-1k validation dataset. **Bold**: best, <u>underline</u>: second best.

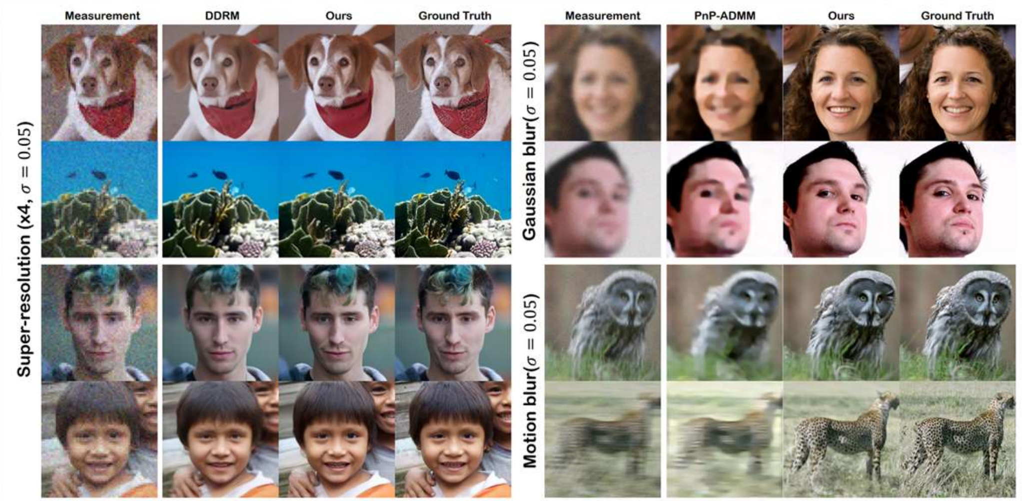


Figure 4: Results on solving linear inverse problems with Gaussian noise ($\sigma = 0.05$).

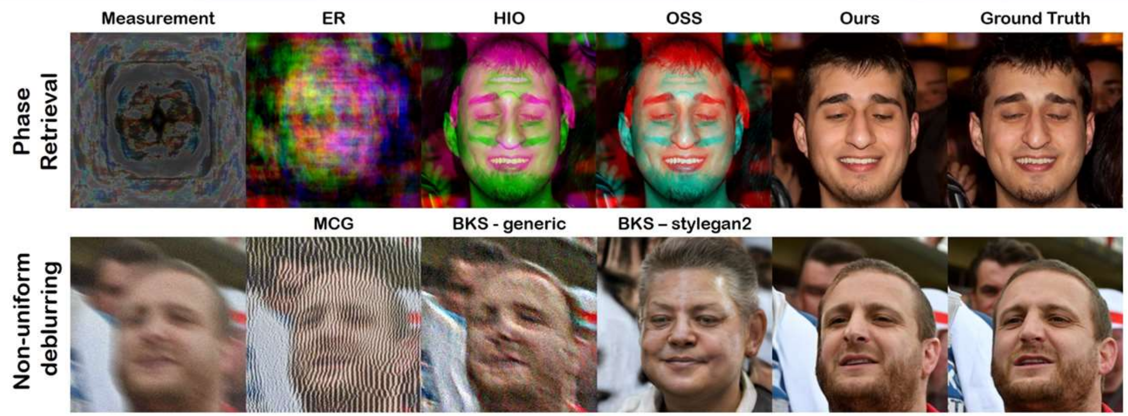


Figure 6: Results on solving nonlinear inverse problems with Gaussian noise ($\sigma = 0.05$).

THANKS