ECCV 2026 | MALMÖ | SEPT 8-12

DiffRGD: An Inference-Time Diffusion Guidance Through Riemannian Gradient Descent


Jia-Wei Liao[1]


Li-Xuan Peng


Mei-Heng Yueh


Min Sun


Cheng-Fu Chou


Jun-Cheng Chen[2]

d11922016@csie.ntu.edu.tw[1], pullpull@citi.sinica.edu.tw[2]

Inference-Time Diffusion Controlling

Image Inverse Problem

$$\underbrace{y = \mathcal{A}x + \eta}$$

Operator

Noise

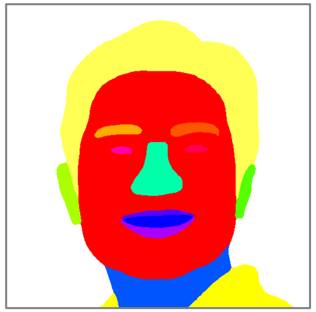
Measurement y

Reconstruction $\hat{x}$

$$\mathcal{L}_{\mathrm{IP}}(x, y) = \|\mathcal{A}x - y\|_2^2$$

Conditional Generation (Segmentation Mask)

ψ : Face Segmentation Model

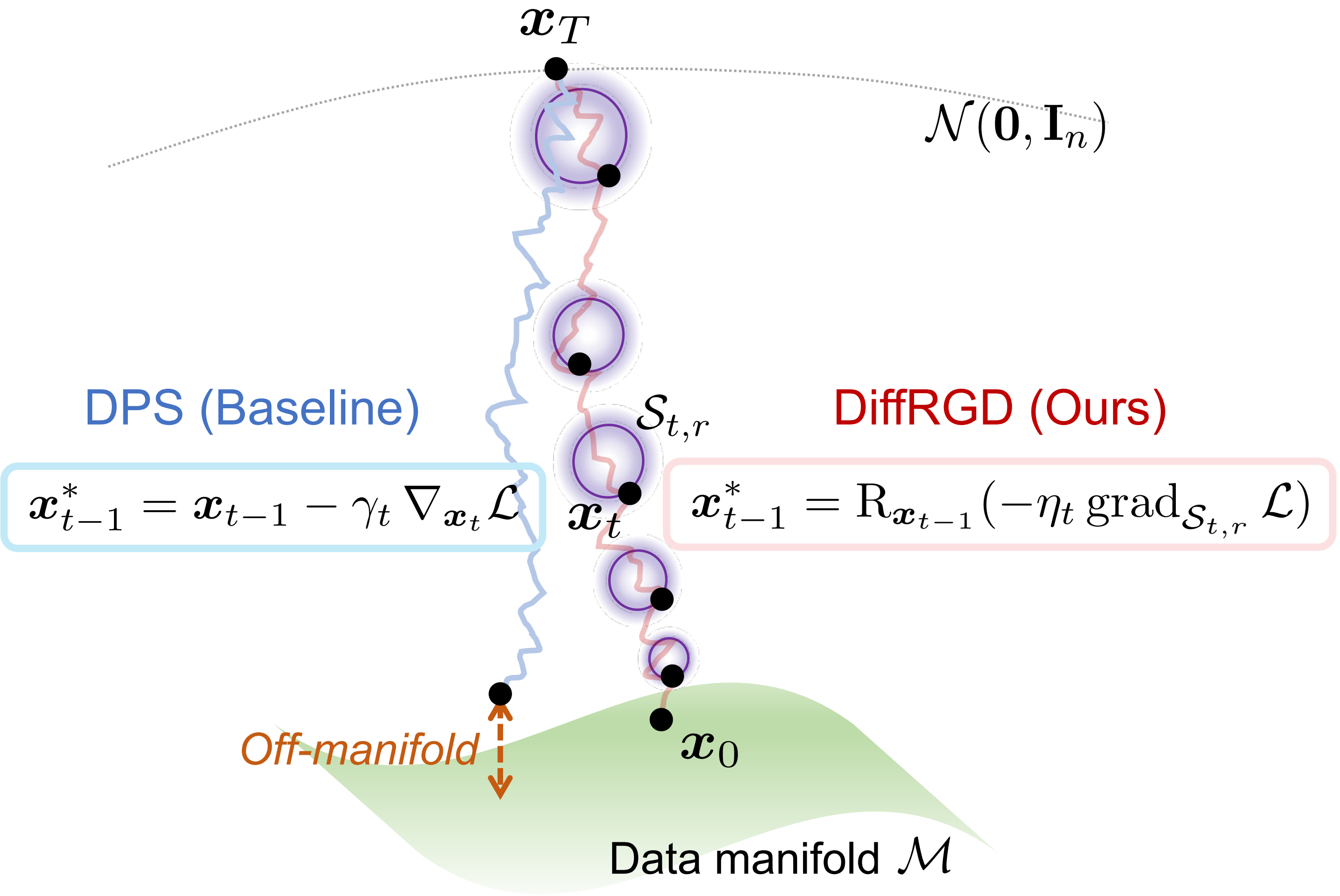


Condition y

Generation $\hat{x}$

$$\mathcal{L}_{\mathrm{Seg}}(x, y) = \|\psi(x) - y\|_2^2$$

Distribution Preserving Diffusion Guidance



Research Question: *Can we design latent-distribution-aware guidance that ensures the guided sample lies within the original latent distribution?*

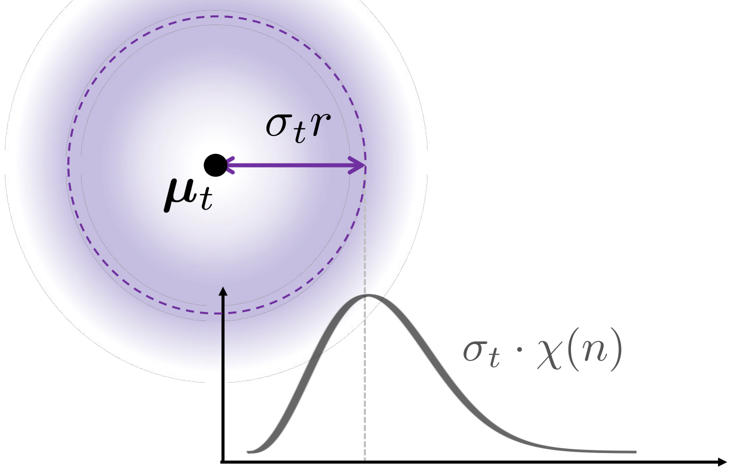
Distribution-induced Geometry

Constructing the geometric structure

Polar Decomposition of Isotropy Gaussian

$$\boldsymbol{x}_{t-1} \sim \mathcal{N}(\boldsymbol{\mu}_t, \sigma_t^2 \mathbf{I}_n) \quad \Longleftrightarrow \quad \boldsymbol{x}_{t-1} = \boldsymbol{\mu}_t + \sigma_t r \boldsymbol{u} \qquad \begin{array}{l} r \sim \chi(n) \\ \boldsymbol{u} \sim \mathrm{Unif}(\mathbb{S}^{n-1}) \end{array}$$

$$\mathcal{S}_{t,r} = \{\boldsymbol{x} \in \mathbb{R}^n \mid \|\boldsymbol{x} - \boldsymbol{\mu}_t\|_2 = \sigma_t r\}$$



Constraint Optimization

$$\boldsymbol{x}_{t-1} = \underset{\boldsymbol{x} \in \mathcal{S}_{t,r}}{\mathrm{argmin}}\, \mathcal{L}(\hat{\boldsymbol{x}}_0(\boldsymbol{x}, t-1), \boldsymbol{y})$$

Diffusion Guidance with Riemannian Gradient Descent

Tangent Space

$$\mathrm{T}_{\boldsymbol{x}}\mathcal{S}_{t,r} := \left\{\boldsymbol{v} \in \mathbb{R}^n \mid (\boldsymbol{x} - \boldsymbol{\mu}_t)^\top \boldsymbol{v} = 0\right\}$$

Projection Operator

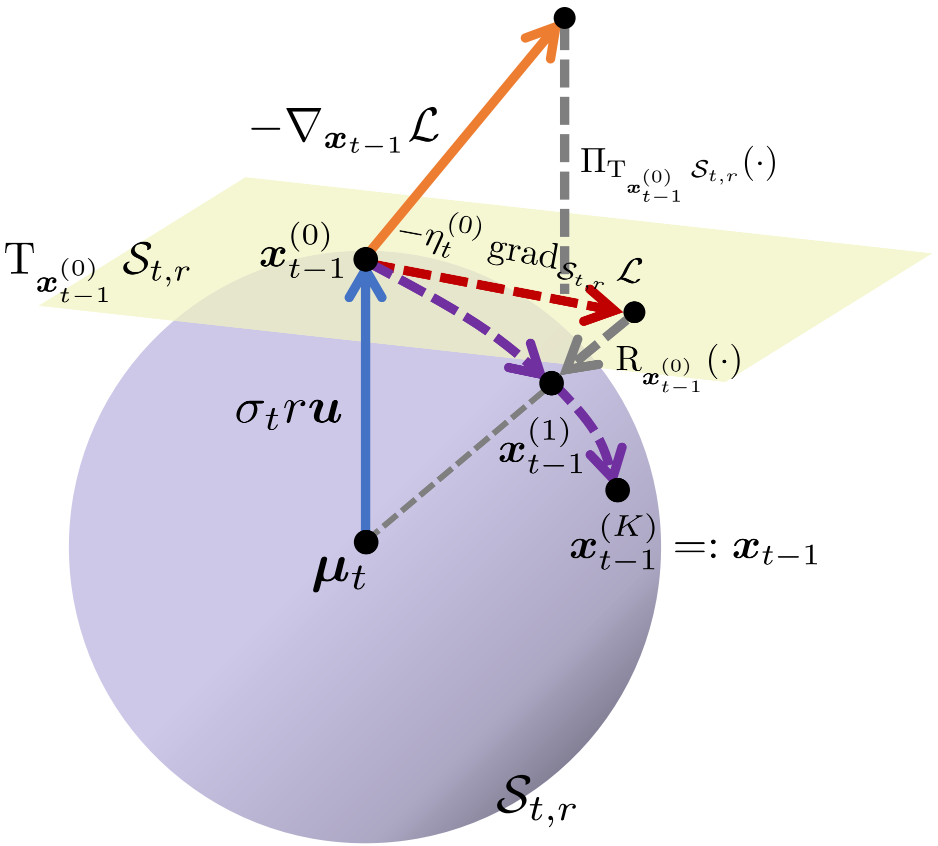
$$\Pi_{\mathrm{T}_{\boldsymbol{x}}\mathcal{S}_{t,r}}(\boldsymbol{v}) := \left(\mathbf{I}_n - \frac{(\boldsymbol{x} - \boldsymbol{\mu}_t)(\boldsymbol{x} - \boldsymbol{\mu}_t)^\top}{\|\boldsymbol{x} - \boldsymbol{\mu}_t\|_2^2}\right)\boldsymbol{v}$$

Retraction

$$\mathrm{R}_{\boldsymbol{x}}(\boldsymbol{v}) := \boldsymbol{\mu}_t + \sigma_t r \cdot \frac{\boldsymbol{x} - \boldsymbol{\mu}_t + \boldsymbol{v}}{\|\boldsymbol{x} - \boldsymbol{\mu}_t + \boldsymbol{v}\|_2}$$

Riemannian Gradient

$$\mathrm{grad}_{\mathcal{S}_{t,r}}\mathcal{L} := \Pi_{\mathrm{T}_{\boldsymbol{x}_{t-1}^{(k)}}\mathcal{S}_{t,r}}\left(\nabla_{\boldsymbol{x}_{t-1}^{(k)}}\mathcal{L}(\hat{\boldsymbol{x}}_0(\boldsymbol{x}_{t-1}^{(k)}, t-1), \boldsymbol{y})\right)$$

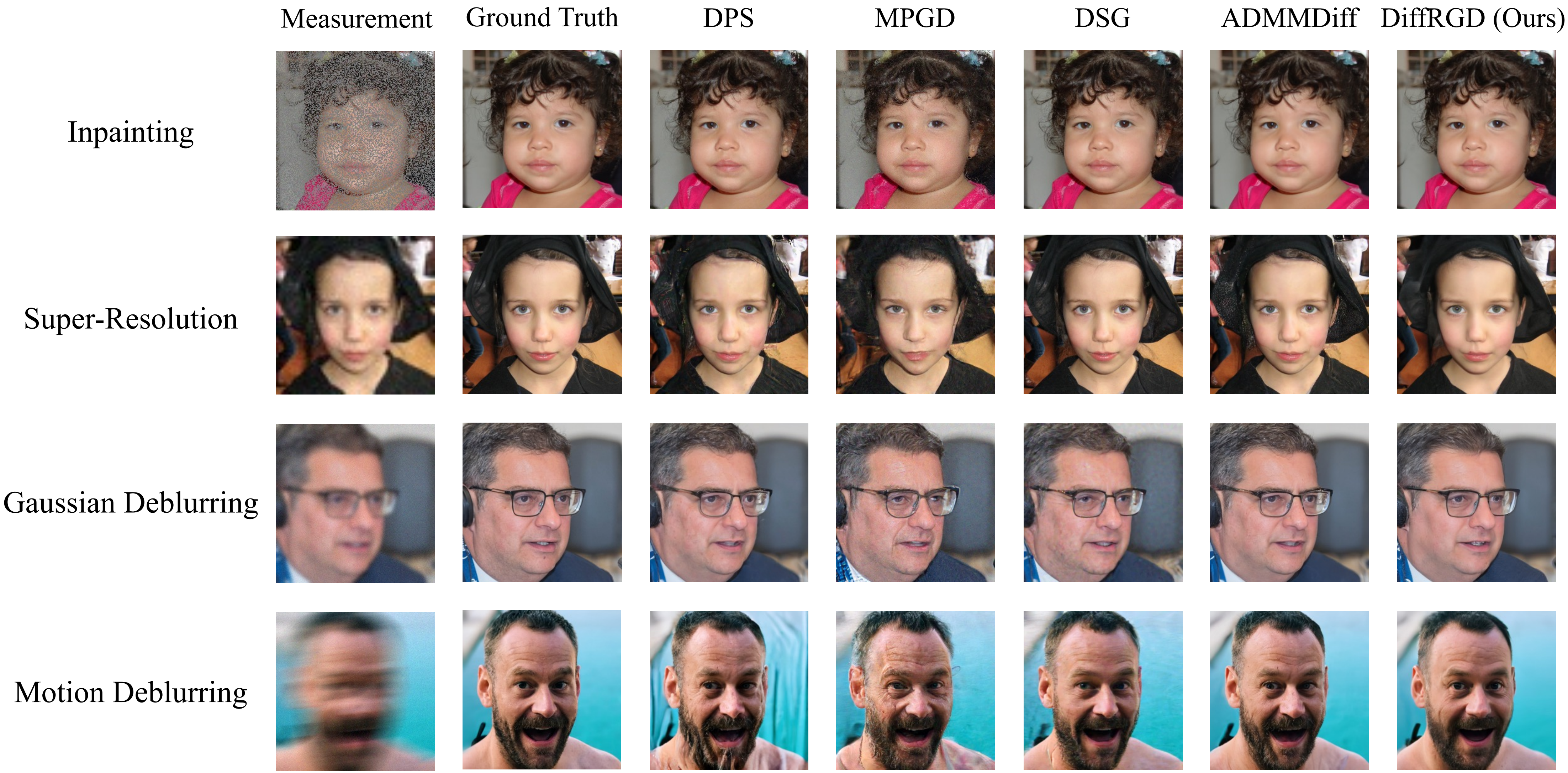


$$\boldsymbol{x}_{t-1}^{(k+1)} = \mathrm{R}_{\boldsymbol{x}_{t-1}^{(k)}}\left(-\eta_t^{(k)}\,\mathrm{grad}_{\mathcal{S}_{t,r}}\mathcal{L}\right)$$

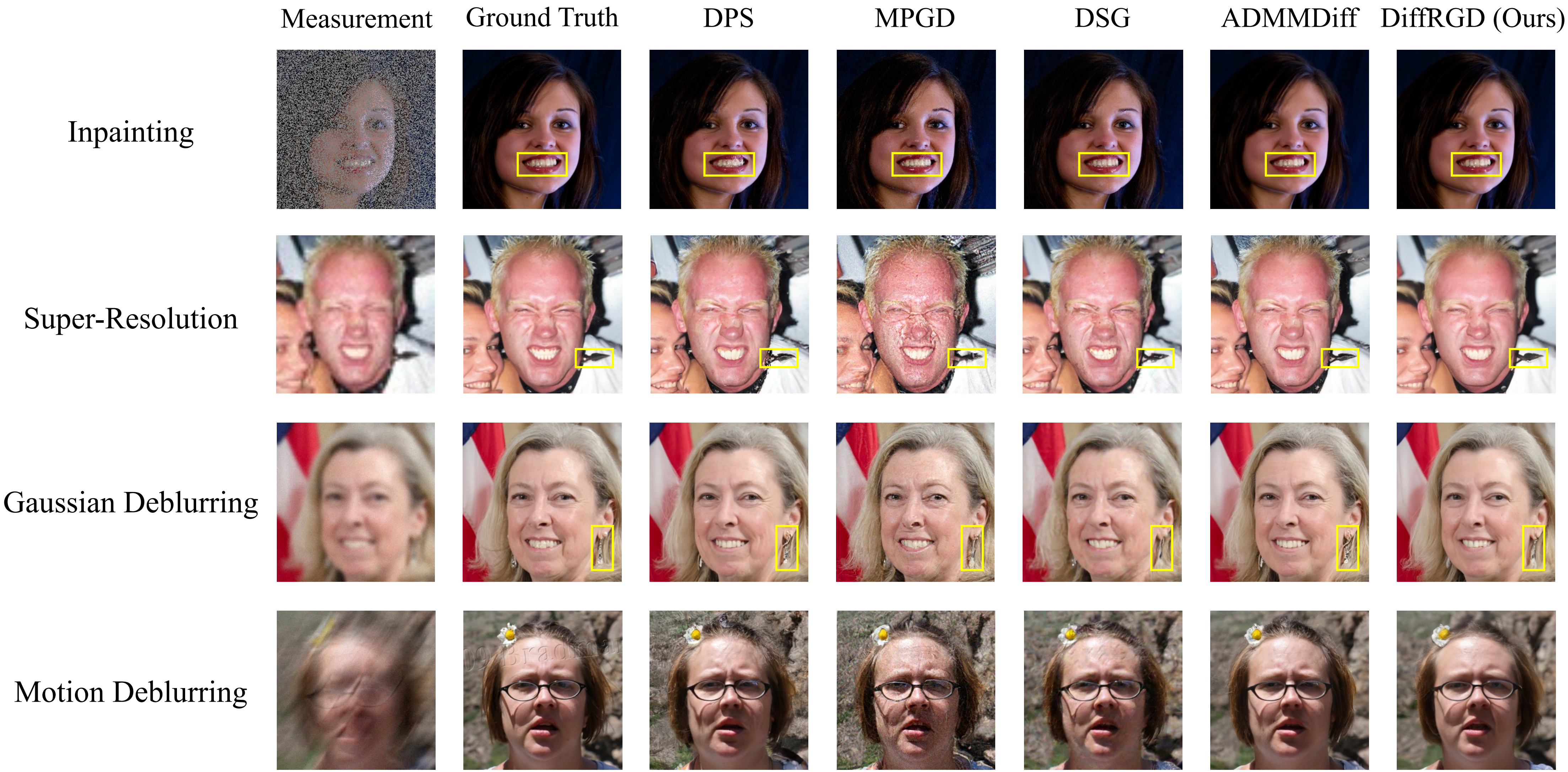
Experiments: Image Inverse Problems

Tasks	Methods	Venue	PSNR ↑	SSIM ↑	LPIPS ↓	FID ↓
Inpainting	DPS [5]	ICLR 2023	30.44	0.863	0.153	42.68
	MPGD [13]	ICLR 2024	27.51	0.724	0.256	68.24
	DSG [51]	ICML 2024	31.03	0.866	0.144	36.30
	ADMMDiff [56]	CVPR 2025	32.38	0.899	0.119	29.52
	DiffRGD (Ours)	-	**34.04***	**0.926***	**0.096***	**21.88***
Super-Resolution 4×	DPS [5]	ICLR 2023	26.03	0.727	0.260	80.05
	MPGD [13]	ICLR 2024	24.40	0.614	0.354	101.50
	DSG [51]	ICML 2024	26.71	0.737	0.256	74.67
	ADMMDiff [56]	CVPR 2025	26.48	0.712	0.297	96.69
	DiffRGD (Ours)	-	**27.77***	**0.783***	**0.220***	**63.94***
Gaussian Deblurring	DPS [5]	ICLR 2023	25.88	0.721	0.237	69.38
	MPGD [13]	ICLR 2024	24.07	0.576	0.328	95.12
	DSG [51]	ICML 2024	**27.45**	0.751	0.259	75.85
	ADMMDiff [56]	CVPR 2025	26.57	**0.757**	0.226	79.30
	DiffRGD (Ours)	-	26.80	**0.757**	**0.218***	**63.83***
Motion Deblurring	DPS [5]	ICLR 2023	24.47	0.685	0.271	80.75
	MPGD [13]	ICLR 2024	23.15	0.569	0.357	106.99
	DSG [51]	ICML 2024	26.80	0.709	0.290	87.96
	ADMMDiff [56]	CVPR 2025	**27.26**	**0.778**	**0.222**	72.92
	DiffRGD (Ours)	-	25.84	0.736	0.250	**72.26**

Experiments: Image Inverse Problems

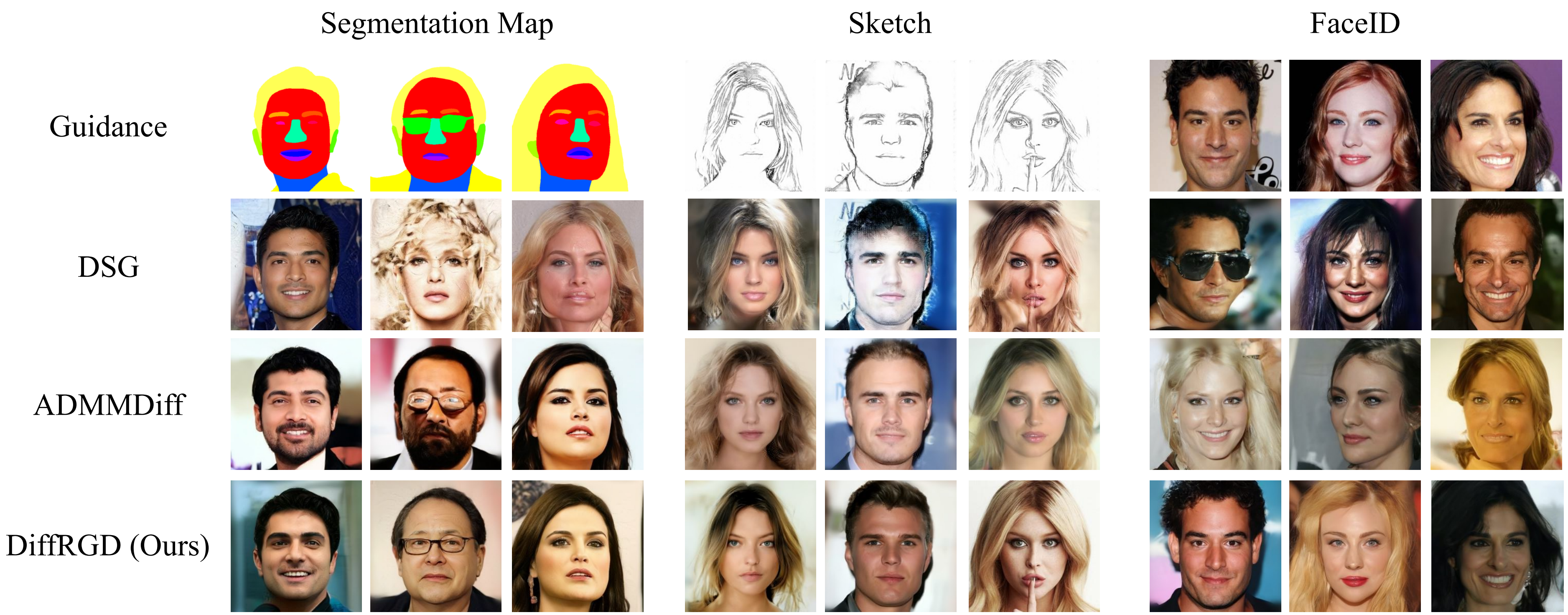


Experiments: Image Inverse Problems

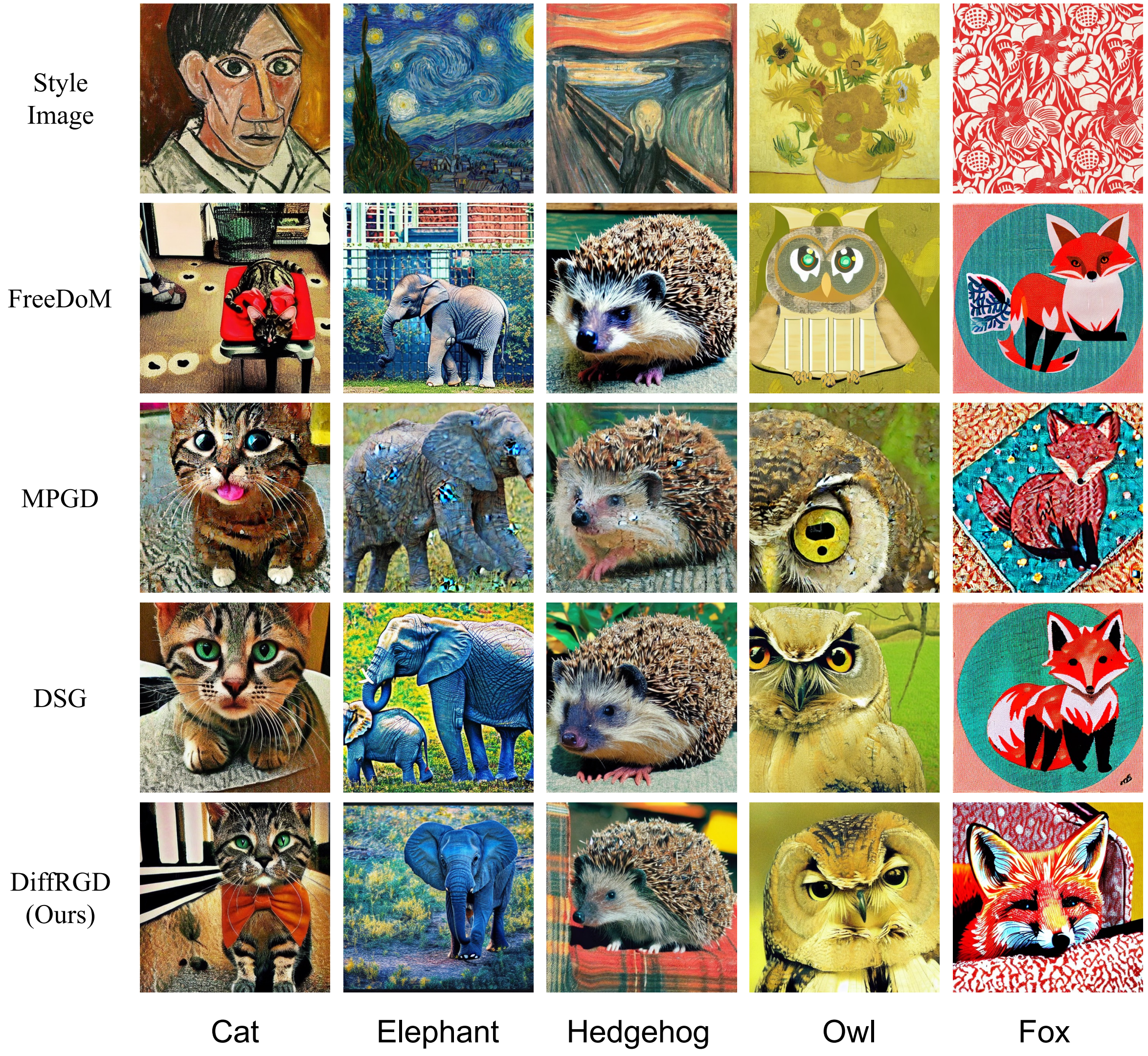


Experiments: Conditional Generation

Methods	Segmentation Map			Sketch			FaceID		
	mIoU ↑	FID ↓	KID ↓	Sketch-ℓ_2 ↓	FID ↓	KID ↓	FaceID-ℓ_2 ↓	FID ↓	KID ↓
FreeDoM [53]	0.622	156.02	0.057	30.85	101.90	0.017	0.557	127.05	0.032
DSG [51]	0.750	117.48	<u>0.028</u>	<u>21.36</u>	107.00	0.024	<u>0.340</u>	<u>95.27</u>	<u>0.014</u>
ADMMDiff [56]	<u>0.758</u>	<u>101.86</u>	0.031	30.82	<u>97.52</u>	<u>0.014</u>	0.346	100.81	<u>0.014</u>
DiffRGD (Ours)	**0.804***	**96.10***	**0.026**	**19.48***	**87.82***	**0.012**	**0.303***	**93.80**	**0.011**



Experiments: Style-Guided Generation



Takeaways

- Respecting the latent Gaussian geometry is crucial for effective inference-time diffusion guidance. Diffusion guidance should preserve the data distribution.
- DiffRGD performs guidance on a distribution-induced spherical manifold using Riemannian Gradient Descent, preserving distributional fidelity while enabling effective control.
- DiffRGD is training-free and plug-and-play, achieving strong performance across diverse image restoration and conditional generation tasks.

Thank you!