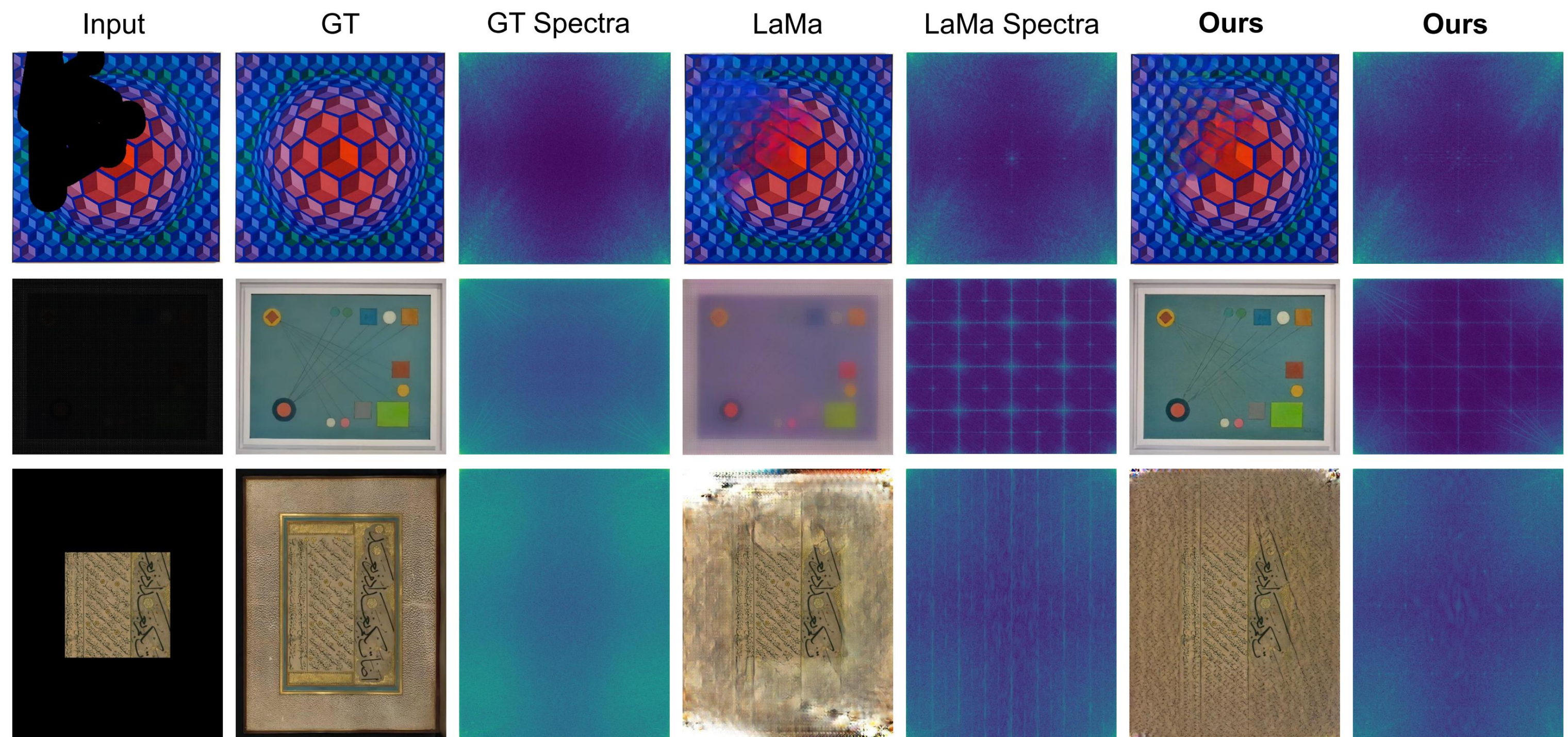


GLaMa: Joint Spatial and Frequency Loss for General Image Inpainting

Zeyu Lu, Junjun Jiang, Junqin Huang, Gang Wu, Xianming Liu

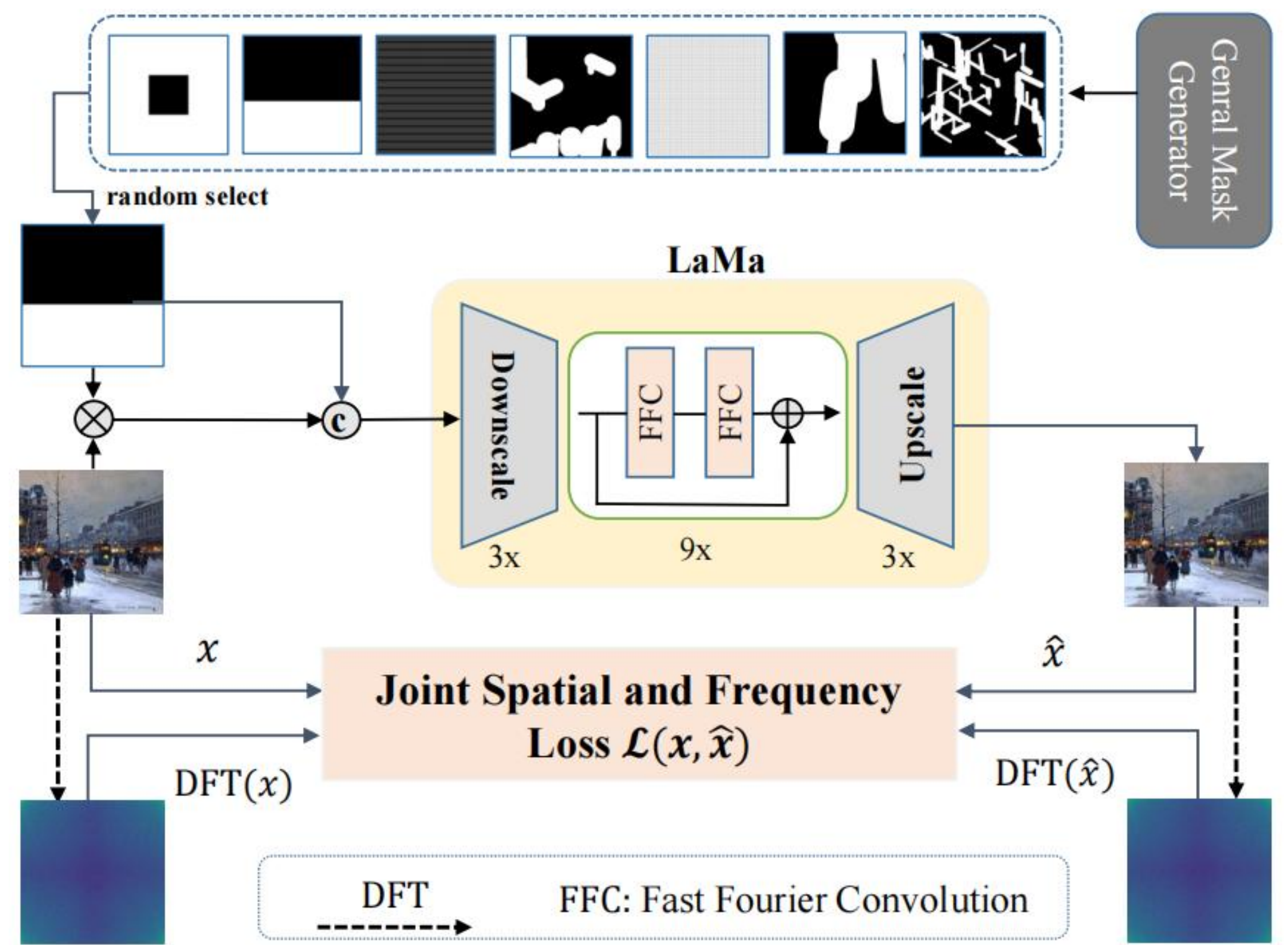


Visualization of the spectra of real and fake images



- Frequency domain gap between real and fake spectra

GLaMa Overview



- General Mask generation strategy with Joint Spatial and Frequency Loss

Joint Spatial and Frequency Loss

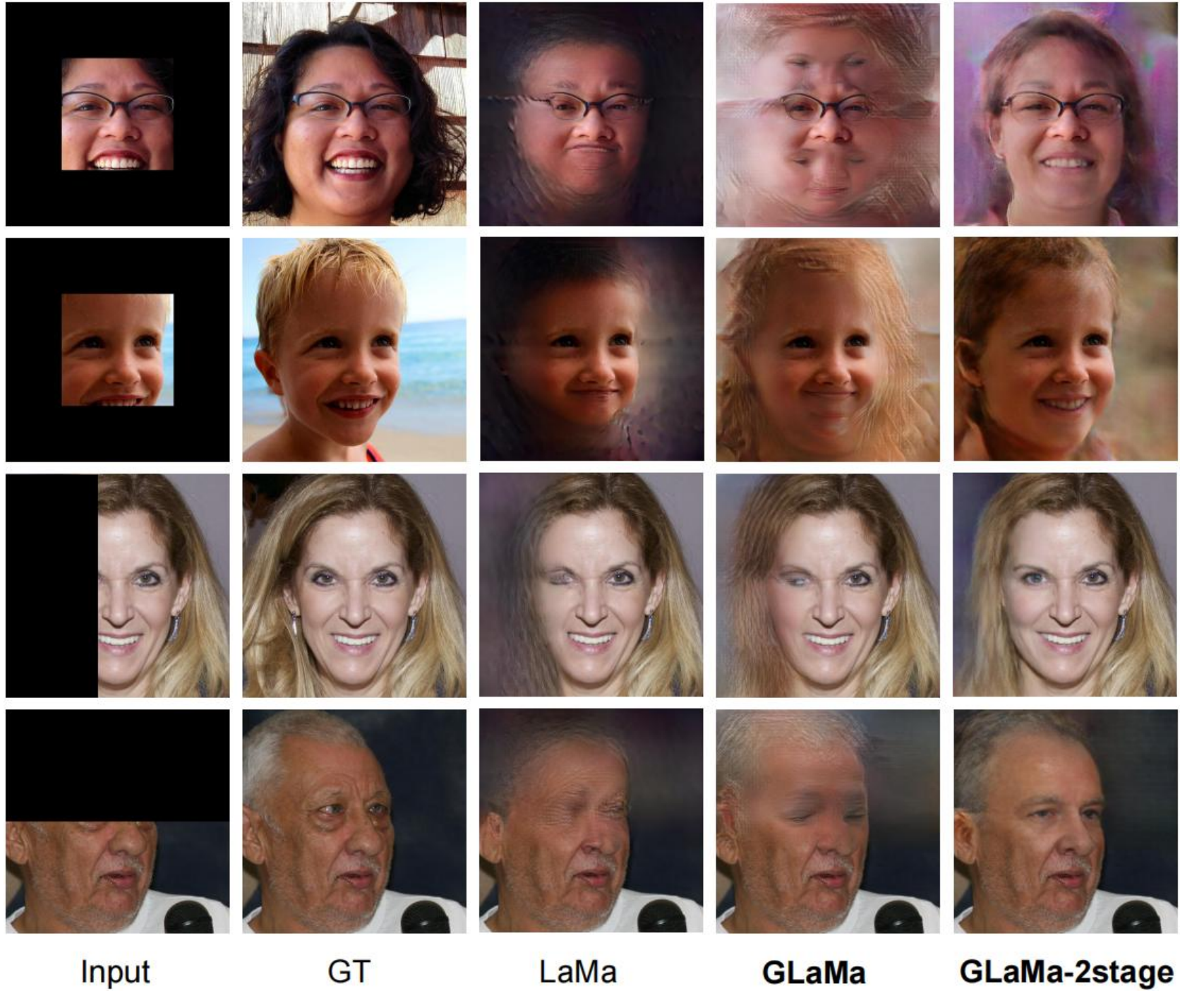
$$\mathcal{L}_{LaMa} = \lambda_1 \mathcal{L}_{L1} + \lambda_{adv} \mathcal{L}_{adv} + \lambda_{fm} \mathcal{L}_{fm} + \lambda_{PL} \mathcal{L}_{PL}$$

$$\mathcal{L}_{FFL} = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} w(u, v) |F_r(u, v) - F_f(u, v)|^2$$

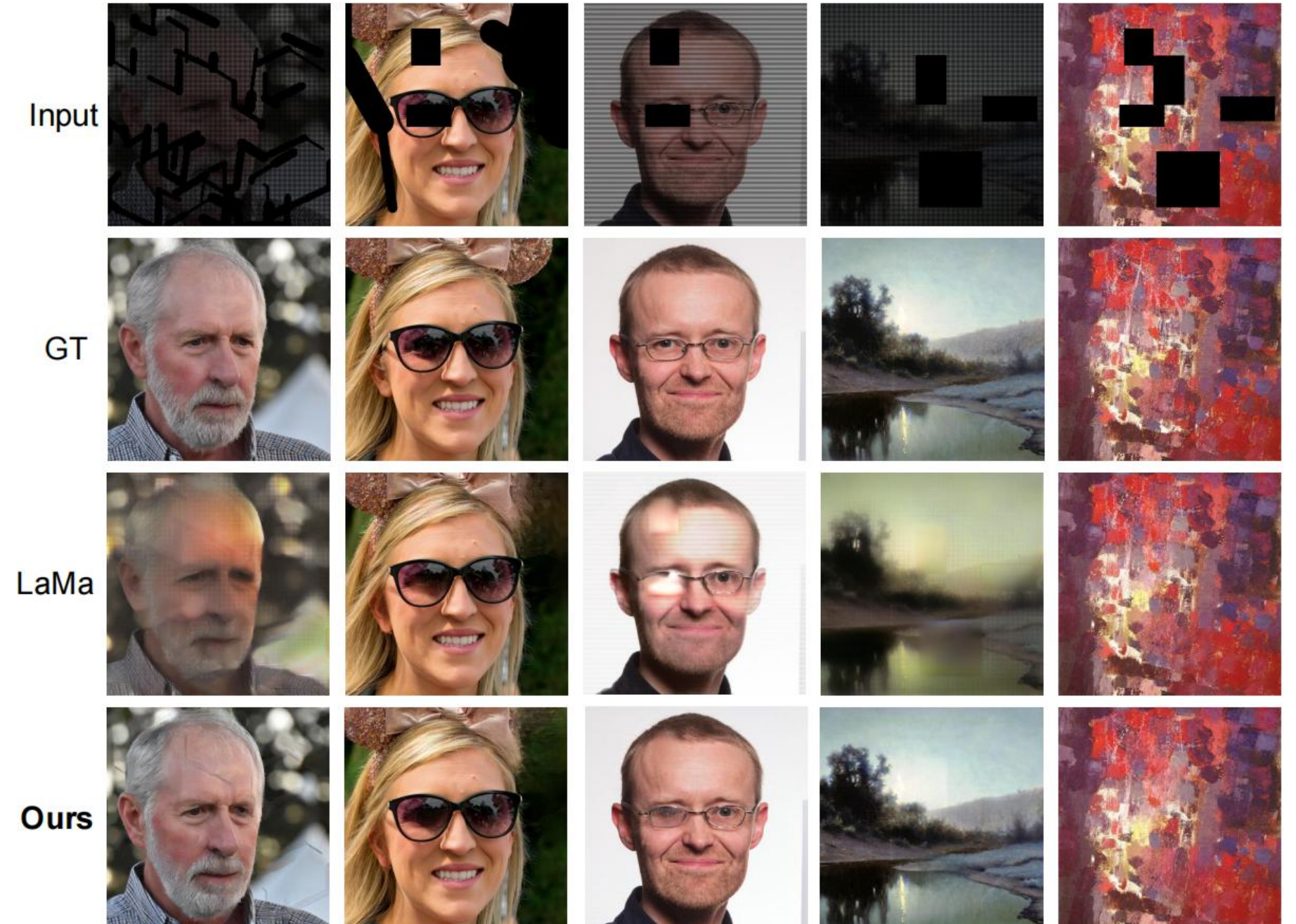
$$\mathcal{L}_{TV} = \sum_{i,j} ((x_{i,j+1} - x_{i,j})^2 + (x_{i+1,j} - x_{i,j})^2)^{\frac{\beta}{2}}$$

$$\mathcal{L} = \alpha_1 \mathcal{L}_{TV} + \alpha_2 \mathcal{L}_{FFL} + \alpha_3 \mathcal{L}_{LaMa}$$

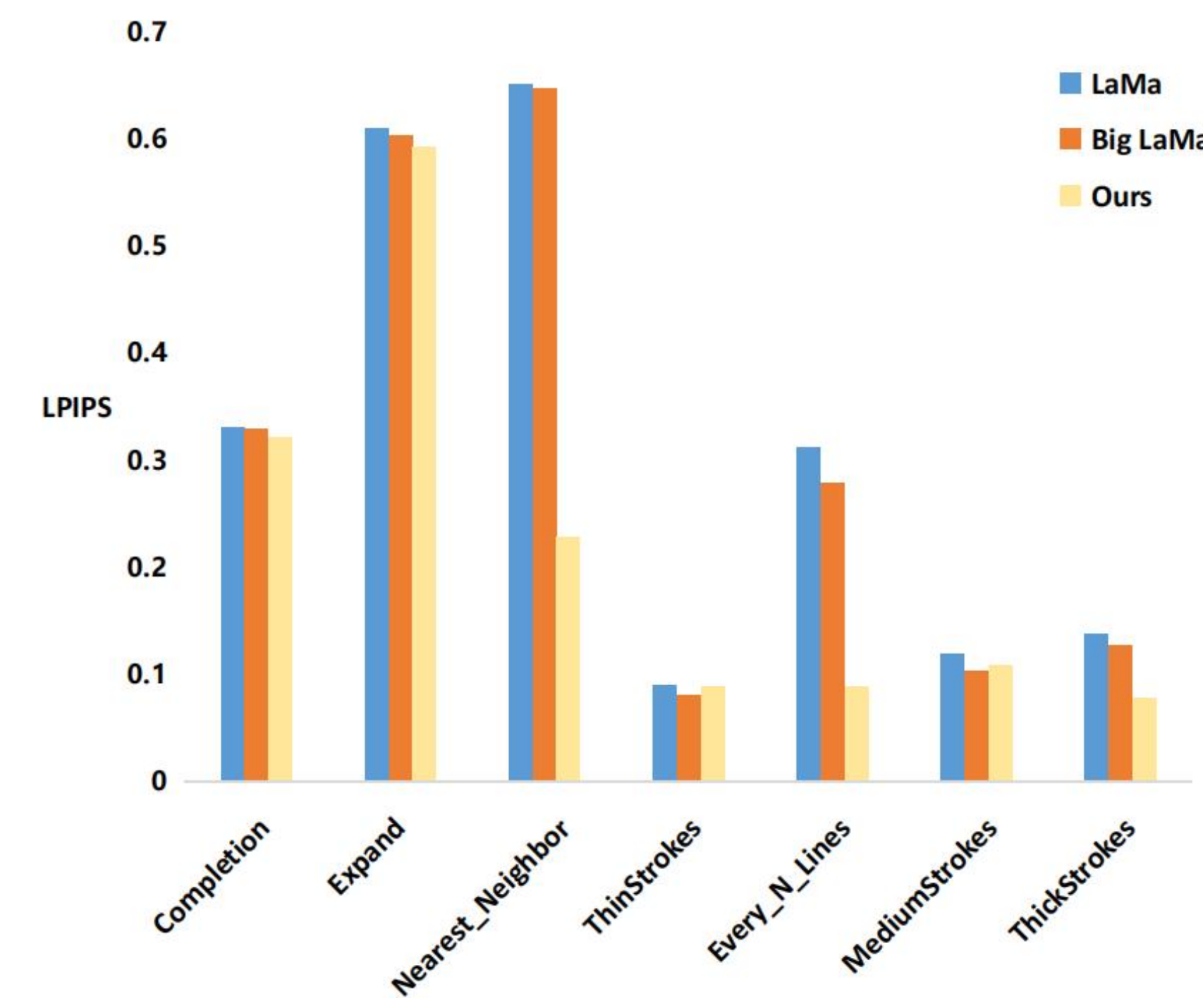
Generalization to the high Resolution (1024x1024)



Generalization to the unseen Masks



Comparison with the state-of-the-art method



Teams	FFHQ			
	FID (↓)	LPIPS (↓)	PSNR (↑)	SSIM (↑)
VIP	115.922	0.433	17.144	0.649
AIIA (Ours)	9.823	0.239	25.316	0.814
HSSLAB	13.504	0.236	25.187	0.821
KwaiInpainting	21.345	0.239	25.060	0.838
ArtificiallyInspired	4.719	0.205	25.999	0.816
SIGMA	7.203	0.248	24.860	0.795

Teams	Places2			
	FID (↓)	LPIPS (↓)	PSNR (↑)	SSIM (↑)
VIP	52.471	0.415	17.256	0.626
AIIA (Ours)	8.772	0.224	24.145	0.800
HSSLAB	9.861	0.227	24.345	0.798
KwaiInpainting	18.334	0.255	23.410	0.787
ArtificiallyInspired	7.544	0.225	23.248	0.777
SIGMA	11.496	0.270	22.562	0.748

Teams	ImageNet			
	FID (↓)	LPIPS (↓)	PSNR (↑)	SSIM (↑)
VIP	50.898	0.403	17.450	0.626
AIIA (Ours)	10.007	0.207	25.226	0.800
HSSLAB	11.770	0.227	24.303	0.798
KwaiInpainting	18.854	0.249	23.804	0.787
ArtificiallyInspired	12.059	0.217	24.278	0.777
SIGMA	19.646	0.311	22.454	0.748

Teams	WikiArt			
	FID (↓)	LPIPS (↓)	PSNR (↑)	SSIM (↑)
VIP	75.645	0.437	17.243	0.609
AIIA (Ours)	14.974	0.244	24.350	0.767
HSSLAB	14.986	0.254	24.257	0.752
KwaiInpainting	26.395	0.276	23.142	0.759
ArtificiallyInspired	8.524	0.248	23.799	0.758
SIGMA	14.125	0.286	22.717	0.720

- GLaMa is better than Big LaMa under some mask settings (the backbone of GLaMa is LaMa)

