

Introduction:

➤ Challenges:

- Most deep image inpainting methods are based on auto-encoder architecture, in which the spatial details of images will be lost in the down-sampling process, leading to the degradation of generated results.
- Texture information and structure information can not be well integrated into a serial conventional inpainting network like auto-encoder.

➤ Our Contributions:

- This is the first work to introduce parallel multi-resolution network architecture into image inpainting, which is able to maintain high-resolution inpainting in the whole process and generate promising texture patterns for the inpainted images.
- Built on parallel multi-resolution network architecture, we propose novel mask-aware representation fusion and attention-guided representation fusion, which can fuse the low- and high-resolution representations more effectively.
- Extensive experiments validate that our method can produce more reasonable and fine-detailed results than other state-of-the-art methods.

Method:

➤ Parallel Multi-Resolution Network

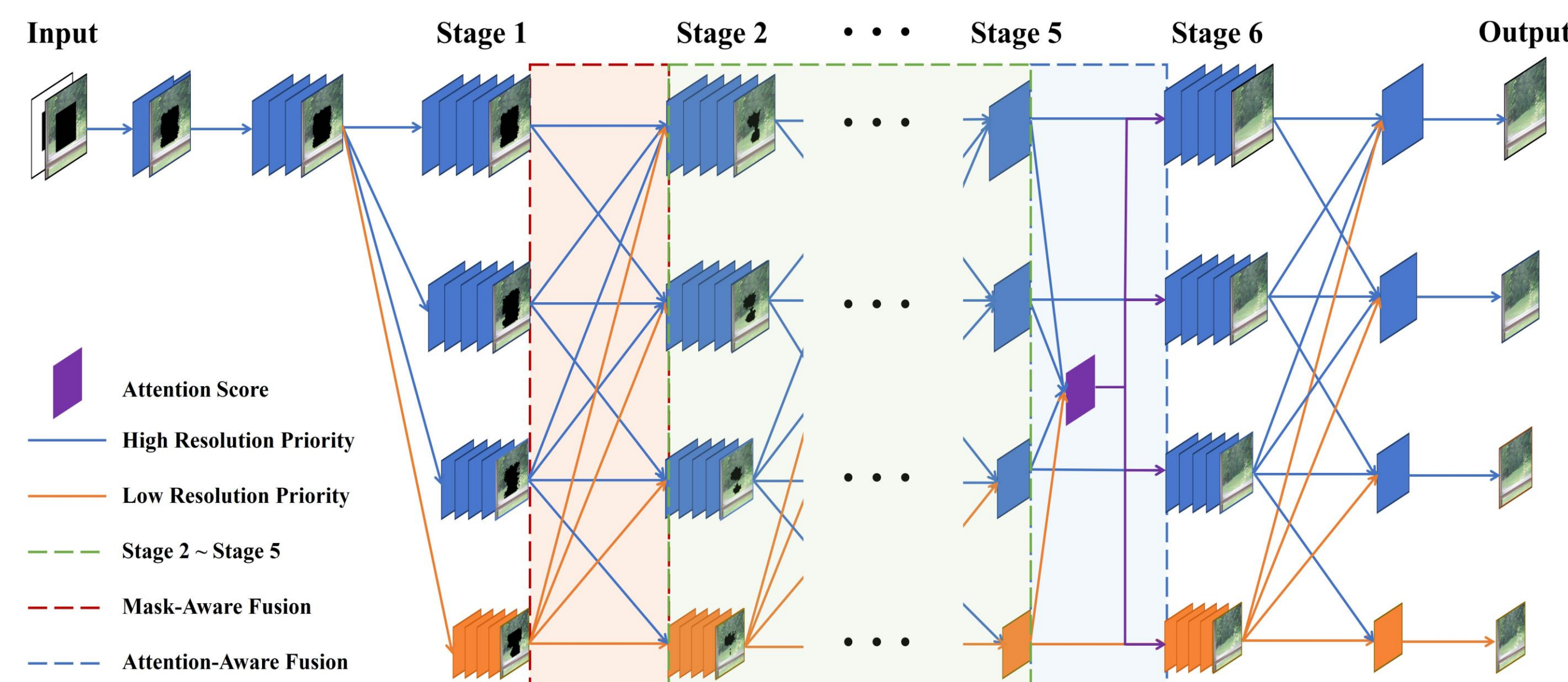


Figure 1. The architecture of Parallel Multi-Resolution Network.

➤ Inpainting Priority

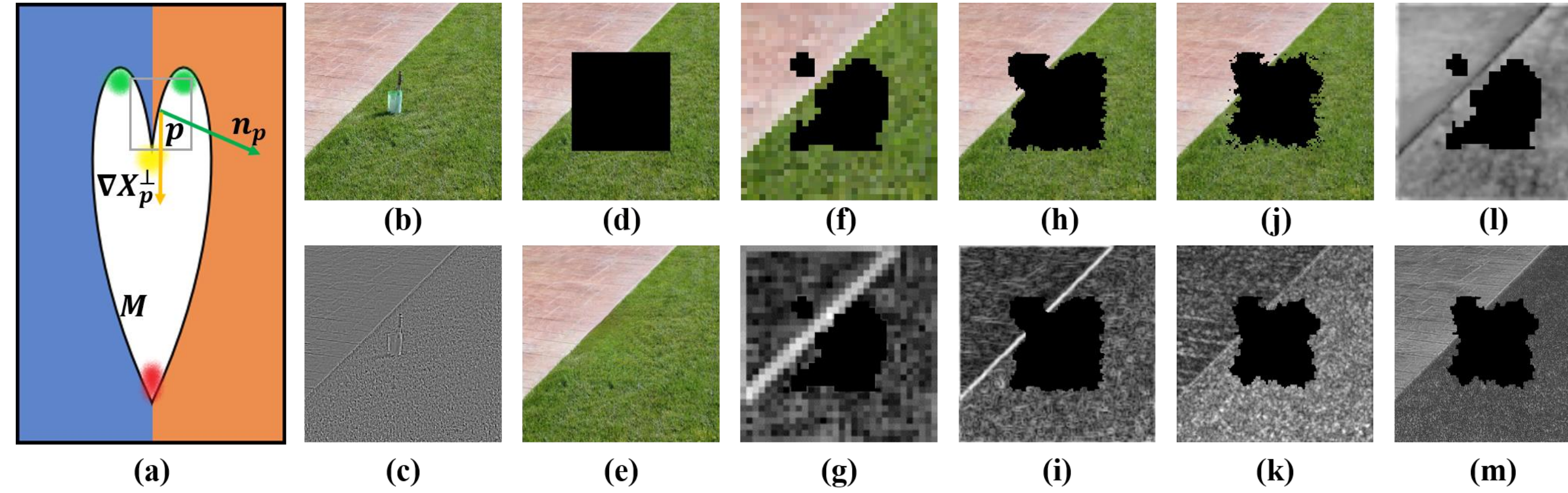


Figure 2. Illustration of common priority and resolution-specific priority.

- The mask and feature update mechanism based on Inpainting Priority:

Step 1: Mask update

$$m' = \begin{cases} 1, & \text{if } m = 1 \text{ or } q \geq \delta \cdot q^{max} \\ 0, & \text{otherwise} \end{cases}$$

The calculate of the pixel x priority q :

$$q = \text{sum}(\mathbf{M}_p) \cdot \rho^l(x)$$

Common priority term $\text{sum}(\mathbf{M}_p)$: the more confidently the pixel can be inpainted with more valid surrounding pixels.

Low-resolution priority ($l = 0$):

$$\rho^l(x) = |n_p \cdot \nabla X_p^\perp|$$

High-resolution priority ($l = 1,2,3$):

$$\rho^l(x) = |n_p \cdot \nabla(X_p - X_{p\uparrow\downarrow})|$$

Step 2: Feature update

$$x_p = \begin{cases} \frac{\Omega_p}{\text{sum}(\mathbf{M}_p)} W \cdot (X_p \odot \mathbf{M}_p) + b, & \text{if } m' = 1 \\ 0, & \text{otherwise} \end{cases}$$

Experiments:

➤ Quantitative Comparison on Places2

	Mask	GC [39]	EC [25]	SF [28]	HF [37]	MEDFE [22]	Ours
ℓ_1 (\downarrow)	0-10%	1.74	1.18	2.71	2.41	1.30	1.12
	10-20%	2.38	1.91	3.51	3.38	2.09	1.74
	20-30%	3.36	2.91	4.55	4.67	2.66	2.60
	30-40%	4.55	4.06	5.69	6.13	3.85	3.65
	40-50%	5.96	5.42	7.00	7.93	5.31	4.89
	50-60%	8.52	7.66	9.12	10.7	7.91	7.24
	Ave%	4.41	4.42	5.43	5.87	4.60	3.54
SSIM \uparrow	0-10%	0.951	0.964	0.898	0.917	0.960	0.971
	10-20%	0.913	0.921	0.855	0.859	0.925	0.934
	20-30%	0.859	0.863	0.801	0.788	0.882	0.884
	30-40%	0.799	0.802	0.745	0.714	0.819	0.827
	40-50%	0.732	0.733	0.685	0.632	0.747	0.761
	50-60%	0.640	0.646	0.610	0.536	0.649	0.669
	Ave%	0.816	0.806	0.765	0.741	0.805	0.840
PSNR \uparrow	0-10%	31.085	32.441	28.609	28.825	31.707	33.690
	10-20%	27.454	27.941	25.522	25.255	27.422	28.924
	20-30%	24.466	24.931	23.121	22.635	25.855	25.871
	30-40%	22.195	22.787	21.336	20.672	23.271	23.487
	40-50%	20.395	21.043	19.818	18.903	21.211	21.659
	50-60%	18.022	18.957	17.981	16.841	18.738	19.220
	Ave%	23.937	25.700	22.732	22.188	25.220	25.475
FID \downarrow	0-10%	1.40	1.60	3.74	1.95	1.46	1.25
	10-20%	2.60	3.18	5.00	5.20	3.27	2.30
	20-30%	4.18	5.87	6.92	11.54	7.23	4.14
	30-40%	7.20	9.90	9.47	22.36	14.34	6.57
	40-50%	11.70	15.65	13.07	39.98	25.78	10.61
	50-60%	19.88	25.55	21.70	70.91	43.90	16.99
	Ave%	7.71	13.58	9.98	25.32	22.68	6.98



Figure 3. The qualitative comparison on Places2.