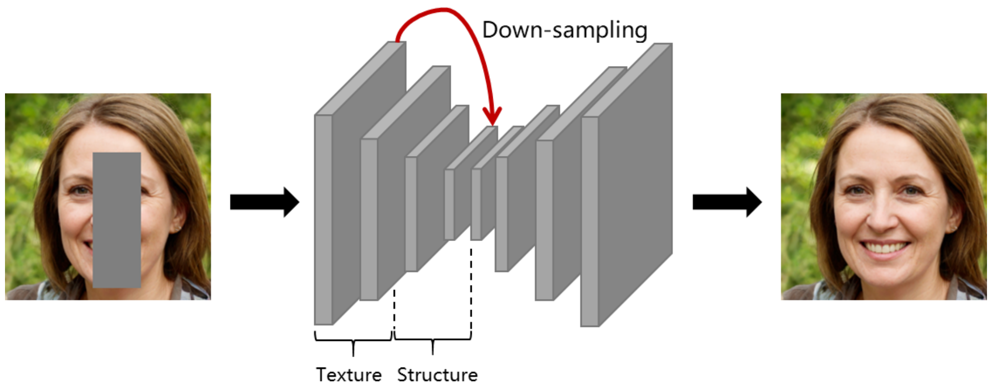
Parallel Multi-Resolution Fusion Network for Image Inpainting

Wentao Wang, Jianfu Zhang, Li Niu, Haoyu Ling, Xue Yang, Liqing Zhang

Department of Computer Science and Engineering, MoE Key Lab of Artificial Intelligence,
Shanghai Jiao Tong University
Tensor Learning Team, RIKEN AIP

Challenges

- **Issue 1.** 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.



Auto-encoder based inpainting methods

Qualitative comparison on Paris Street View

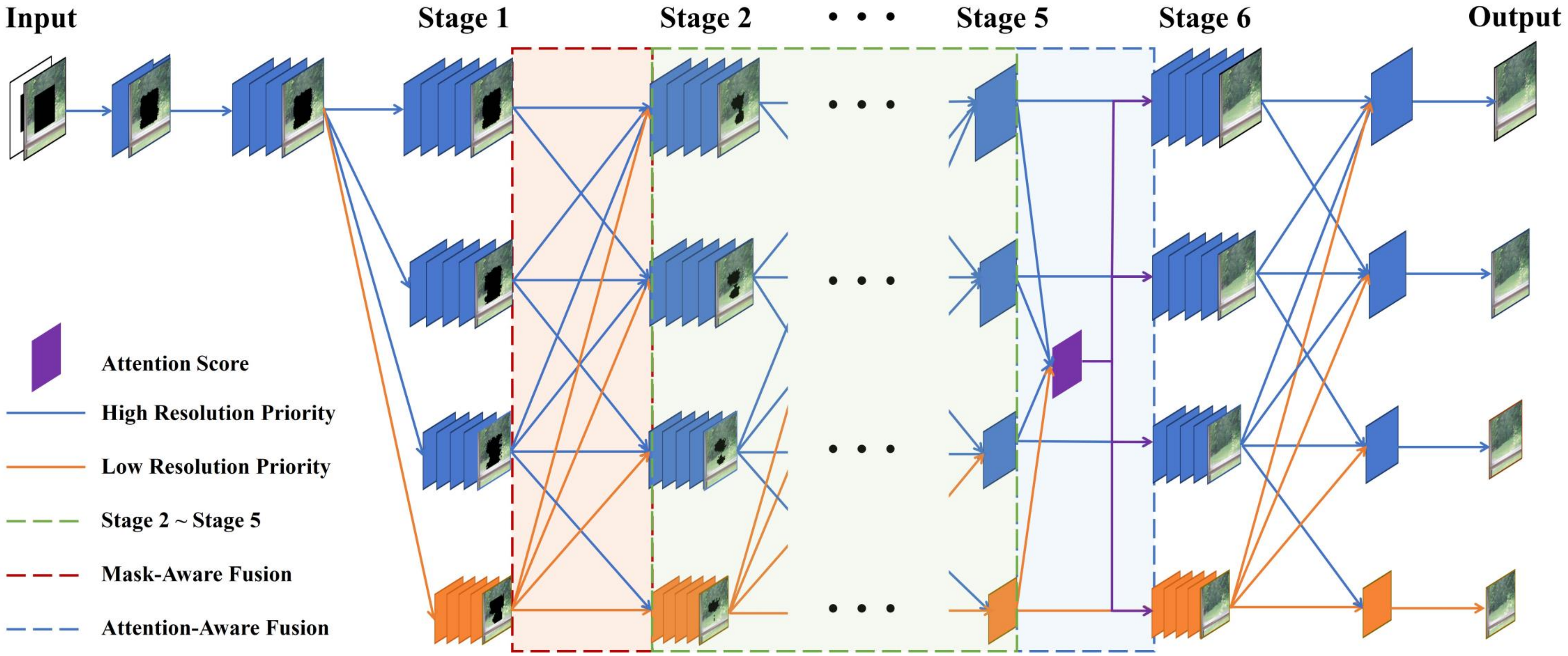
- **Issue 2.** Texture information and structure information can not be well integrated into a serial inpainting network like auto-encoder.

Our Solutions

- For **Issue 1.**
  - Parallel multi-resolution architecture always maintains high resolution representation.
- For **Issue 2.**
  - Multiple multi-resolution feature fusions make a better integration of structure and texture.
  - Two fusion methods:
    - Mask-aware representation fusion.
    - Attention-guided representation fusion.

Proposed Method

- Parallel Multi-Resolution Inpainting Network



Proposed Method

- Inpainting Priority

Step 1: Mask update

$$m' = \begin{cases} 1, & if \ \mathrm{sum}(\boldsymbol{M_p}) > 0 \\ 0, & otherwise \end{cases}$$

Step 2: Feature update

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

Partial Conv[1]

[1] Liu et al., Image inpainting for irregular holes using partial convolutions, ECCV'18

Step 1: Mask update

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

The calculate of the pixel x priority q:

$$\mathrm{q} = \mathrm{sum}(\boldsymbol{M_p}) \cdot \rho^l(\mathrm{x})$$

Common priority term : $\mathrm{sum}(\boldsymbol{M_p})$

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

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

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

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

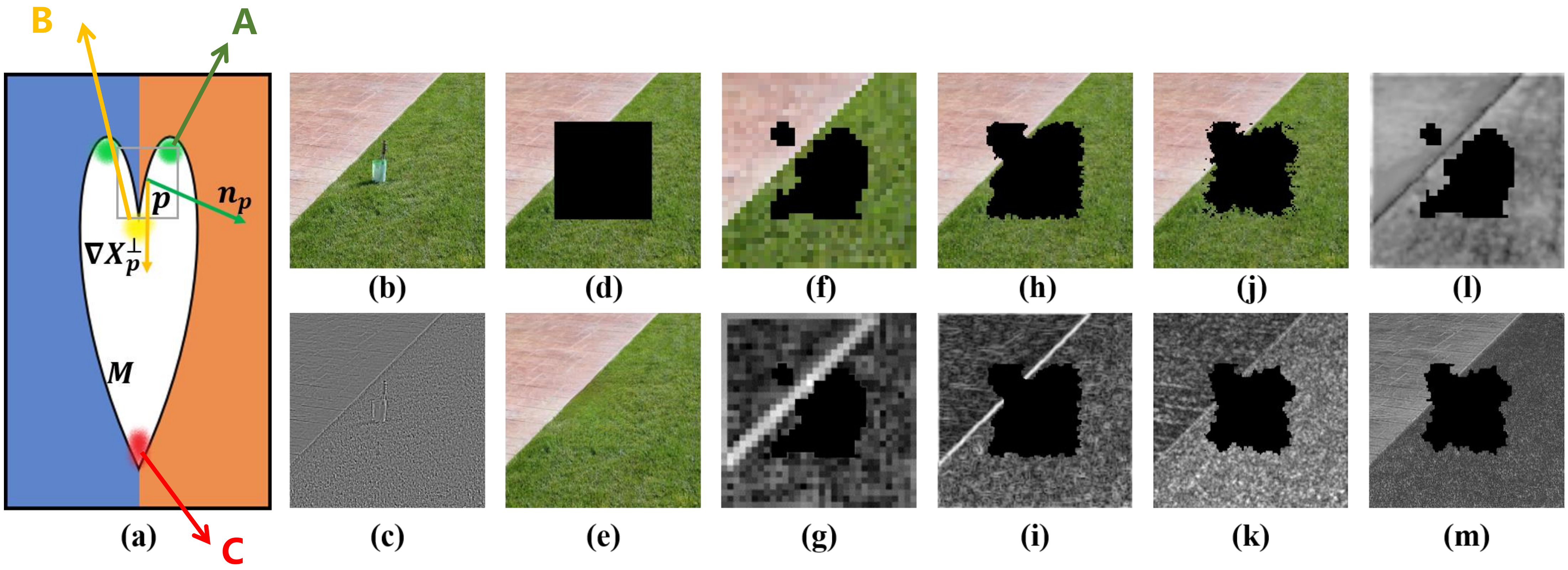
Step 2: Feature update

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

Ours

Proposed Method

- Inpainting Priority



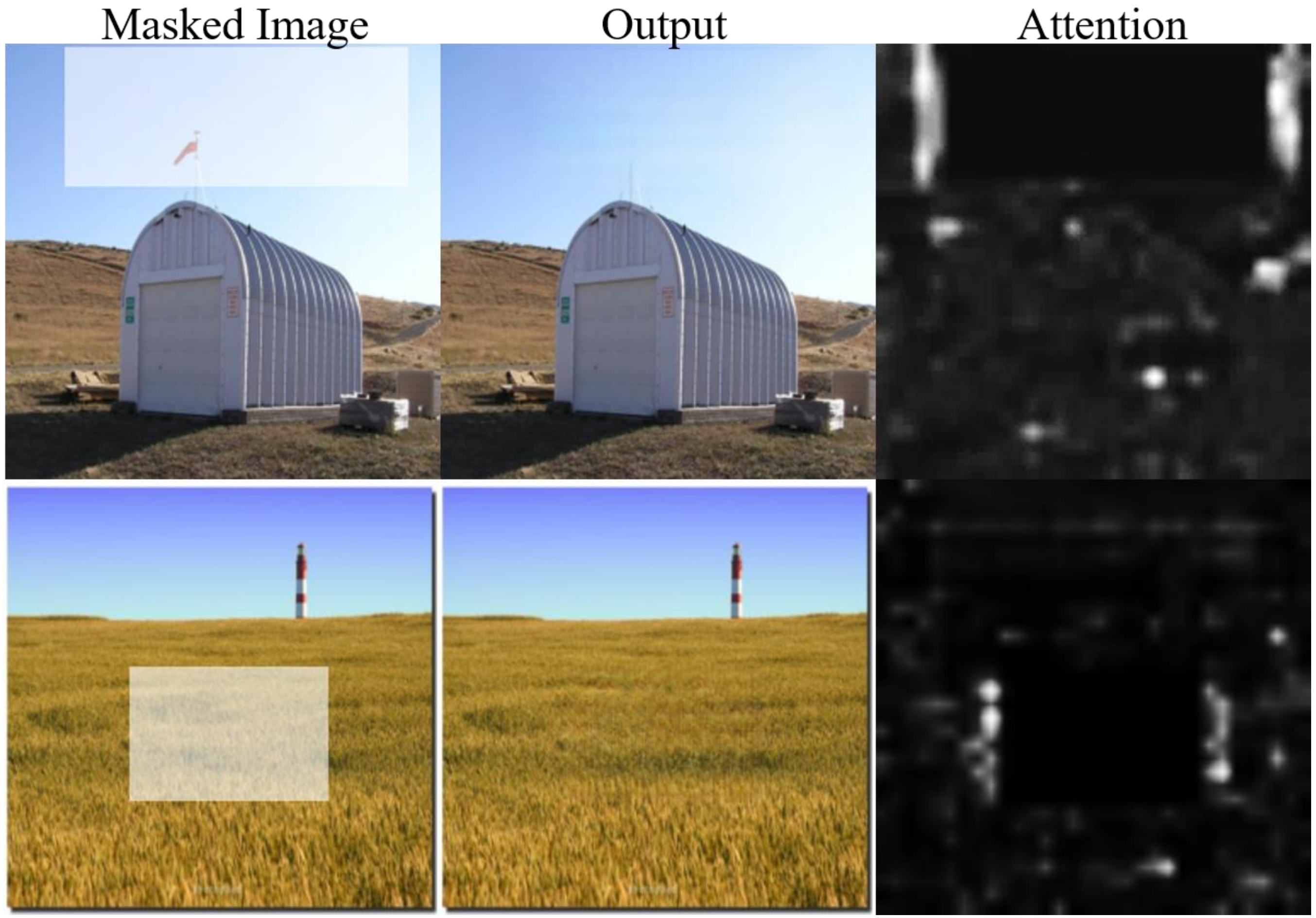
A: High common priority but low contour striking priority.

B: High contour striking priority but low common priority.

C: Both high common priority and high contour striking priority.

Proposed Method

- Attention-Guided Representation Fusion



Masked region attends relevant information from unmasked region. E.g., wheat land (resp.,sky) for wheat land (resp., sky).

Experimental Results

- Qualitative comparison on Places2

(a) Masked Image (b) GC (c) EC (d) SF (e) HF (f) MEDFE (g) Ours (h) Ground Truth

Thanks for watching!

Parallel Multi-Resolution Fusion Network for Image Inpainting

Wentao Wang, Jianfu Zhang, Li Niu, Haoyu Ling, Xue Yang, Liqing Zhang

Department of Computer Science and Engineering, MoE Key Lab of Artificial Intelligence,
Shanghai Jiao Tong University
Tensor Learning Team, RIKEN AIP