

# You Only Erase Once: Erasing Anything without Bringing Unexpected Content

- Supplementary Material -

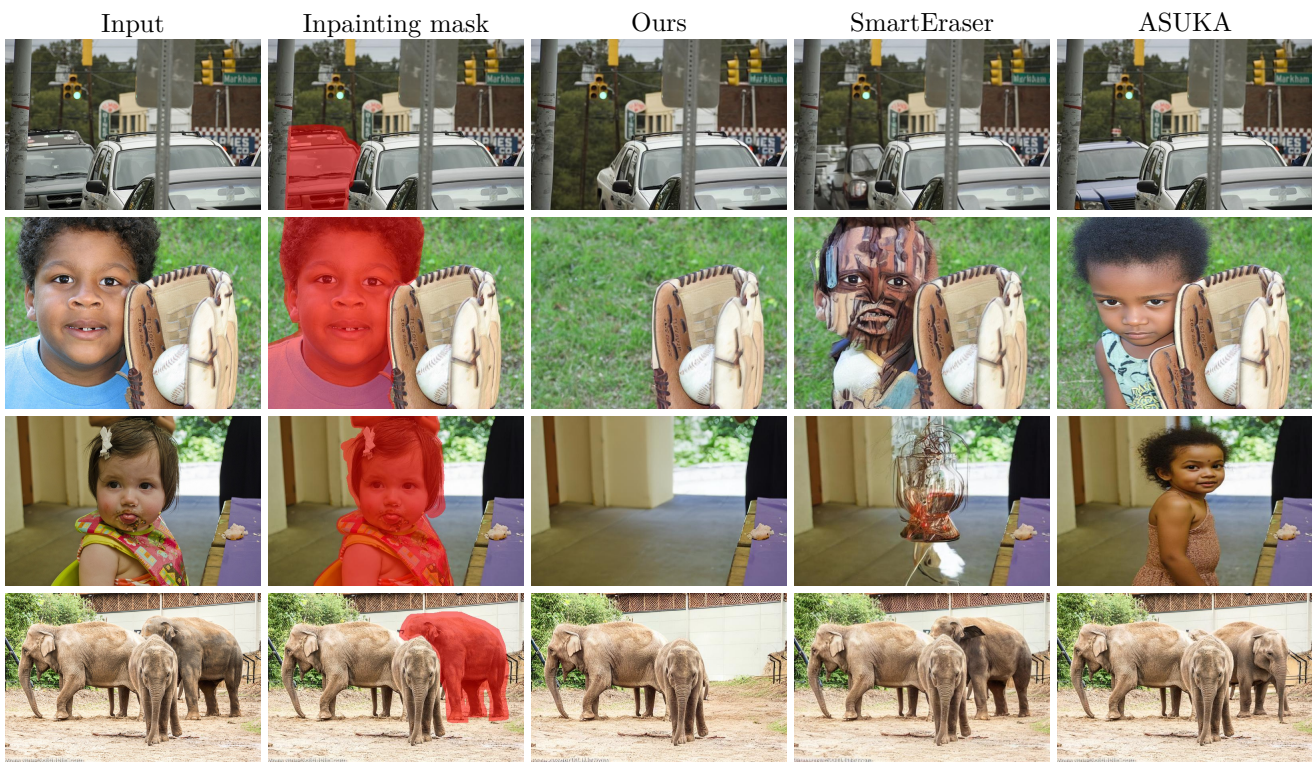


Figure 1. More comparisons with existing methods.

## 1. More Visual Comparison

Figures 1, 2, 3, 4, 5, 6 and 7 provide additional visual comparisons with SmartEraser [1] and ASUKA [2]. Our method generates context-coherent completion results after removing the selected objects, whereas other methods tend to produce unwanted content.

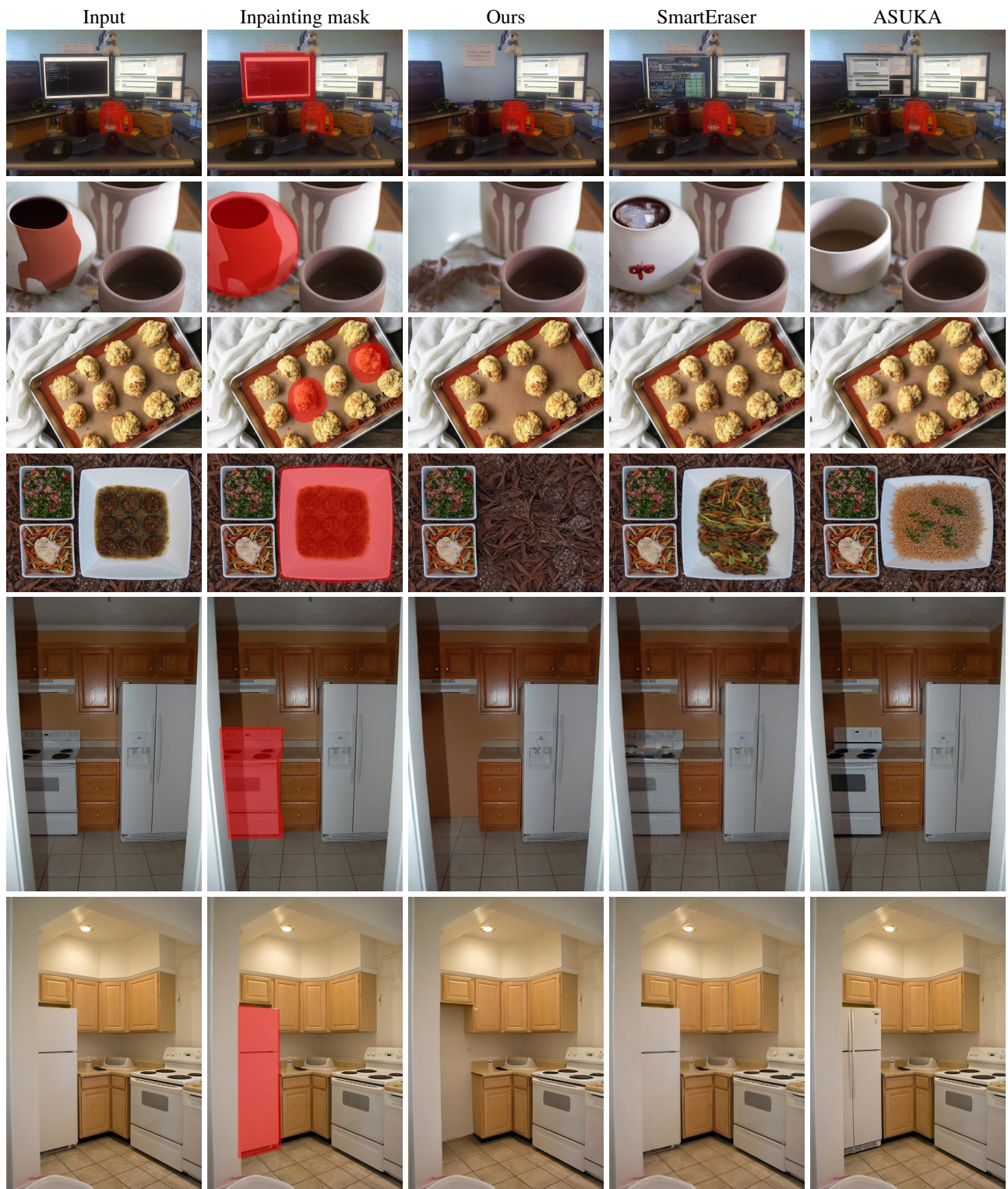


Figure 2. More comparisons with existing methods..

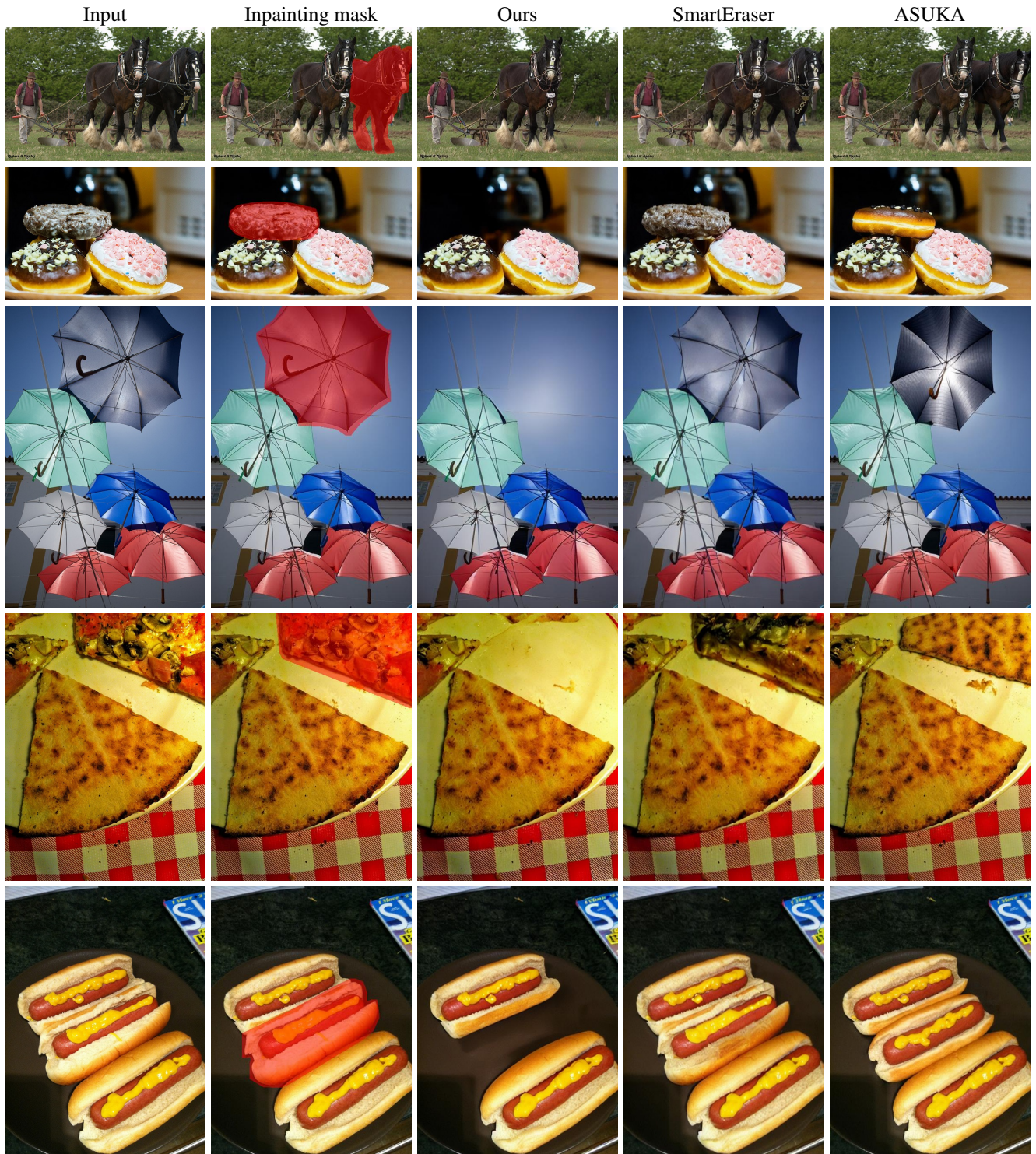


Figure 3. More comparisons with existing methods.

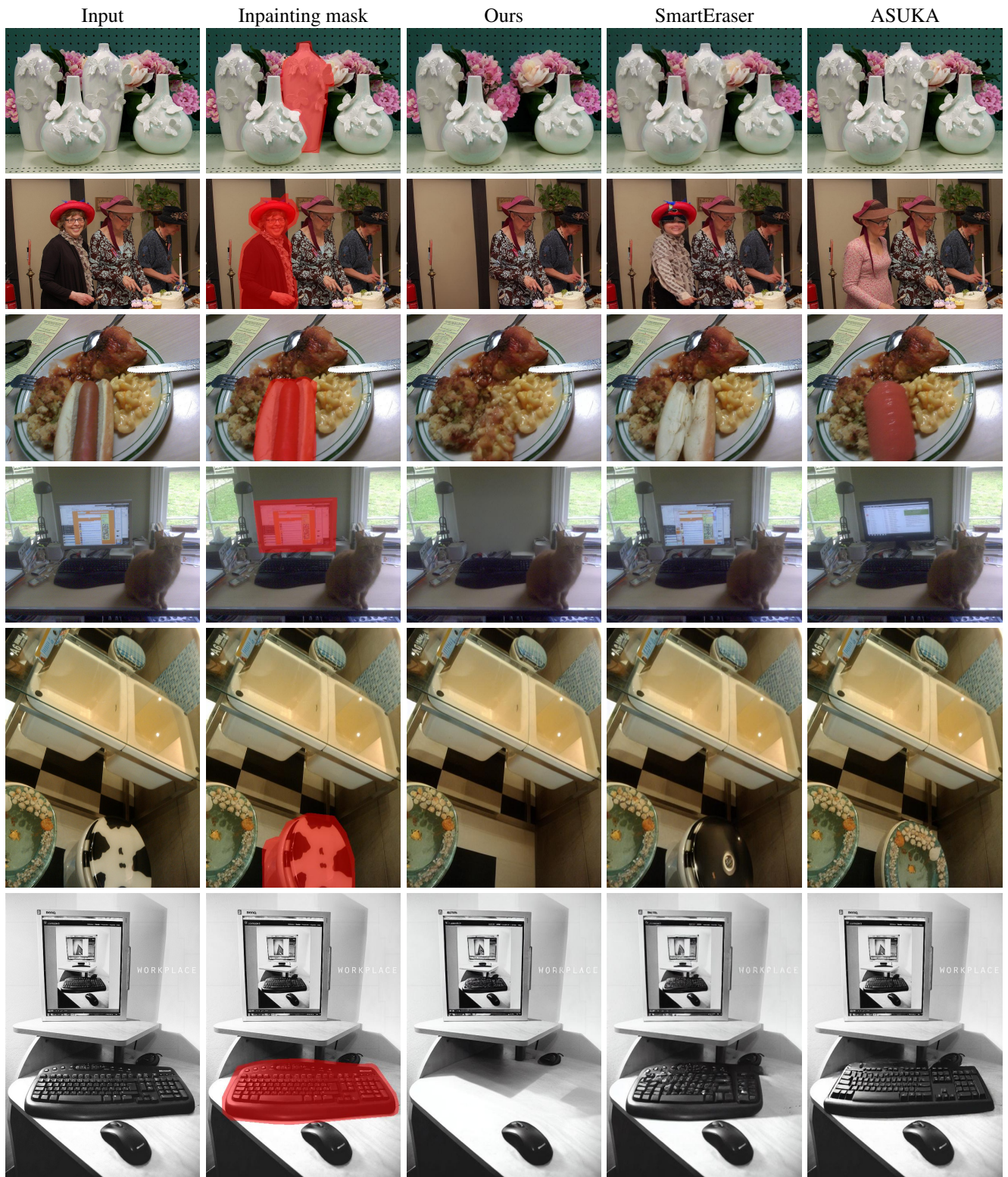


Figure 4. More comparisons with existing methods.



Figure 5. More comparisons with existing methods.

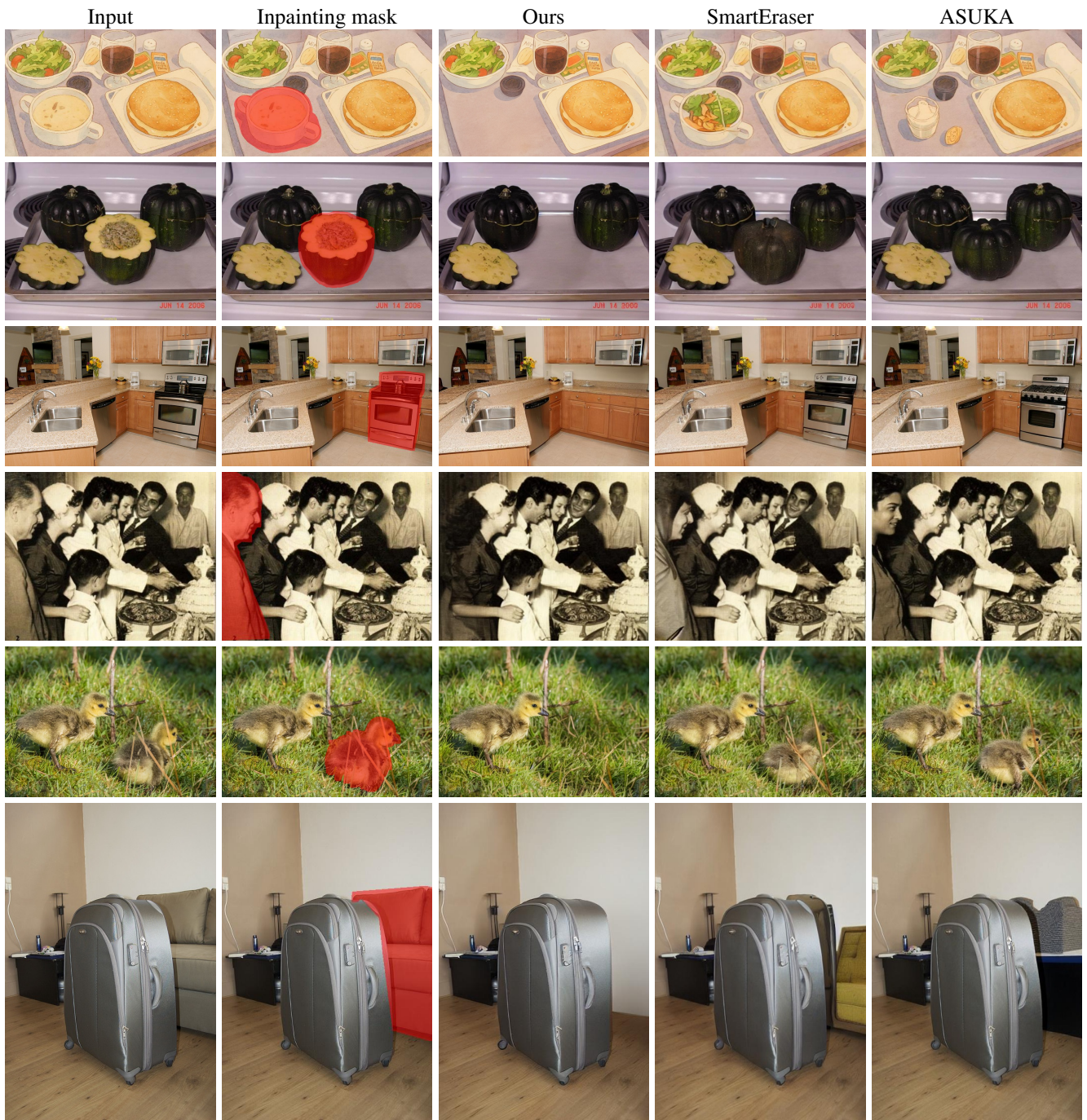


Figure 6. More comparisons with existing methods.



Figure 7. More comparisons with existing methods.



Figure 8. **Limitation.** When filling in the missing regions, the model may produce erroneous continuations that deviate from the actual content.

## 2. Ablation study on loss-related hyperparameters

In table 1, we provide sensitivity analysis for loss-related hyperparameters ( $\tau = 0.2$ ,  $\lambda = 0.9$ ,  $\lambda_{SS} = 0.5$ ,  $\lambda_{EFC} = 0.5$  as the default setting) on the EntitySeg dataset, where we only adjust one hyperparameter while keeping others as default. As shown, our method is insensitive to different hyperparameters.

Table 1. Ablation study on loss-related hyperparameters.

$\tau$	FID↓	MSN↓	MARS↓	CFD↓	$\lambda$	FID↓	MSN↓	MARS↓	CFD↓
0.1	59.3	<b>0.041</b>	0.0059	0.312	0.5	59.3	0.076	0.0059	0.321
0.2	<b>58.8</b>	0.049	<b>0.0050</b>	<b>0.311</b>	0.8	59.3	0.053	<b>0.0048</b>	0.314
0.3	59.1	0.048	0.0058	0.313	0.9	<b>58.8</b>	<b>0.049</b>	0.0050	<b>0.311</b>
0.5	59.1	0.072	0.0075	0.320	0.95	59.1	0.051	0.0063	0.319
$\lambda_{SS}$	FID↓	MSN↓	MARS↓	CFD↓	$\lambda_{EFC}$	FID↓	MSN↓	MARS↓	CFD↓
0.1	59.0	0.066	0.0063	0.315	0.1	<b>58.8</b>	0.112	0.0136	0.356
0.5	<b>58.8</b>	<b>0.049</b>	0.0050	<b>0.311</b>	0.5	<b>58.8</b>	<b>0.049</b>	0.0050	<b>0.311</b>
1.0	59.1	0.051	<b>0.0045</b>	0.312	1.0	58.9	0.054	<b>0.0046</b>	0.314

### **3. Limitation**

Figure 8 illustrates the limitation of our method. In complex and ambiguous scenes, our model may generate erroneous extensions when completing the missing regions, leading to results that deviate from the actual content. In the top example, the toys overlap heavily and their boundaries are unclear. After the tiger is removed, the completed content does not align with the true contour of the teddy bear. In the middle example, the microwave oven is embedded in a wall cabinet; if it were removed in reality, the background should be a wall. Similarly, if the refrigerator in the lower example were removed, the generated content should not be an extension of the side wall but rather the scene that would appear behind and beside the refrigerator.

## References

- [1] Longtao Jiang, Zhendong Wang, Jianmin Bao, Wengang Zhou, Dongdong Chen, Lei Shi, Dong Chen, and Houqiang Li. Smarteraser: Remove anything from images using masked-region guidance. In *CVPR*, pages 24452–24462, 2025. [1](#)
- [2] Yikai Wang, Chenjie Cao, Junqiu Yu, Ke Fan, Xiangyang Xue, and Yanwei Fu. Towards enhanced image inpainting: Mitigating unwanted object insertion and preserving color consistency. In *CVPR*, pages 23237–23248, 2025. [1](#)