Low-Resolution Editing is All You Need for High-Resolution Editing

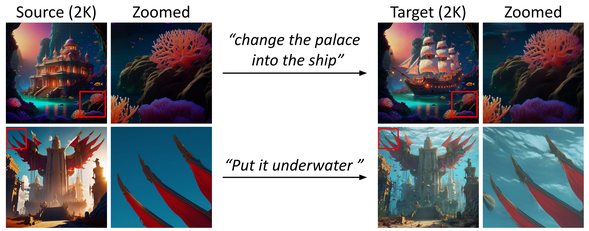
Junsung Lee[1*] **Hyunsoo Lee**[1*] **Yong Jae Lee**[2] **Bohyung Han**[1]

[1]Seoul National University [2]University of Wisconsin-Madison

CVPR JUNE 3-7, 2026 DENVER COLORADO

1 Introduction

- **High-resolution Image Editing using ScaleEdit**
  - Edit high-resolution source image into target image according to a given visual concept while **preserving fine-grained details**



- **Key Contributions of ScaleEdit**
  - Propose a **novel task**: high-resolution image editing
  - Enable zero-training editing via patch-wise test-time optimization
  - Develop a **feature transfer** function and a patch **synchronization** strategy
  - Achieve **strong results** on diverse editing scenarios

2 Background

- **Diffusion for Image-to-Image Translation**
  - Forward & Reverse Diffusion Process

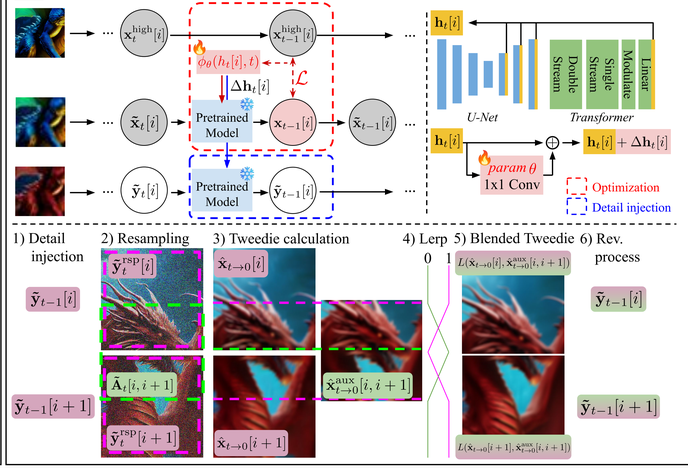
$$\mathbf{x}_{t+1}^{\text{fwd}} = f^{\text{fwd}}(\mathbf{x}_t^{\text{fwd}}, t) := \sqrt{\alpha_{t+1}}\hat{\mathbf{x}}_0(\mathbf{x}_t^{\text{fwd}}, t) + \sqrt{1-\alpha_{t+1}}\epsilon_\theta(\mathbf{x}_t^{\text{fwd}}, t)$$

$$\text{(Reverse) } \mathbf{x}_{t-1}^{\text{rev}} = f^{\text{rev}}(\mathbf{x}_t^{\text{rev}}, t) := \sqrt{\alpha_{t-1}}\hat{\mathbf{x}}_0(\mathbf{x}_t^{\text{rev}}, t) + \sqrt{1-\alpha_{t-1}}\epsilon_\theta(\mathbf{x}_t^{\text{rev}}, t)$$

- **Problem Setup**
  - Input: (a) High-res. src $I_{\text{src}}^{\text{high}}$, (b) Low-res. src $I_{\text{src}}^{\text{low}}$, (c) Low-res. tgt image $I_{\text{tgt}}^{\text{low}}$
  - Output: High-res. target image $I_{\text{tgt}}^{\text{high}}$

- **Naïve solutions: Real-ISR baselines**
  - Performs super-resolution only using $I_{\text{tgt}}^{\text{low}}$ with heavily trained model
  - Does not conditioned on $I_{\text{src}}^{\text{high}}$ & $I_{\text{src}}^{\text{low}}$, fail to transfer fine-grained details
  - ScaleEdit: **Conditioned on source images**, transfers details without training

3 Framework



- **Patch-wise Operation for Detail Transfer**
  - Divide $I_{\text{src}}^{\text{high}}$ into patches at the model-native resolution
    → Fully inherit the model's strong generative priors
  - Generate multiple detail-enhanced patches
    & Refined by a synchronization for cross-patch consistency
- **Detail Enhancement Module**
  - Transfer fine-grained details via a **learnable transfer function**
  - Manipulate the intermediate feature $\mathbf{h}_t[i]$ by optimizing $\Delta\mathbf{h}_t[i]$

$$\mathcal{L} := ||\mathbf{x}_{t-1}^{\text{high}}[i] - f^{\text{rev}}(\tilde{\mathbf{x}}_t[i], t\,; \Delta\mathbf{h}_t[i])||_2^2 \quad \text{with} \quad \Delta\mathbf{h}_t[i] := \phi_\theta(\mathbf{h}_t[i], t)$$

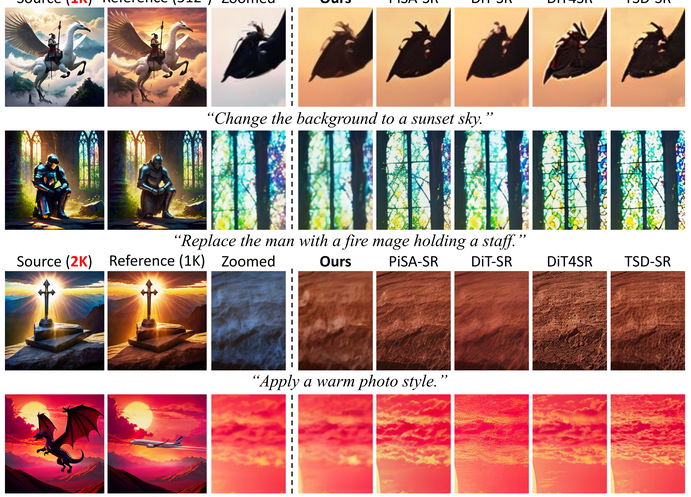
  → Practical design: $\phi_\theta(\cdot,\cdot) = \text{conv}_{1\times1}(\cdot,\cdot)$ & $\Delta\mathbf{h}_t[i] = \mathbf{0}, \forall t > \tau$
  - Inject $\Delta\mathbf{h}_t[i]$ during the sampling of $\tilde{\mathbf{y}}_t[i]$, resulting in $I_{\text{tgt}}^{\text{high}}$
- **Patch-wise Synchronization**
  - Apply **Blended-Tweedie**-based sampling for patch coherence
  - $\Delta\mathbf{h}_t[i]$ is undefined for Blended-Tweedie
    → Use **resampling** to disentangle detail injection and sync.
  - Perform (1) forward w/o detail injection & (2) backward w/ sync.

(1) $\tilde{\mathbf{y}}_t^{\text{rsp}}[i] = f^{\text{fwd}}(\tilde{\mathbf{y}}_{t-1}[i], t-1)$

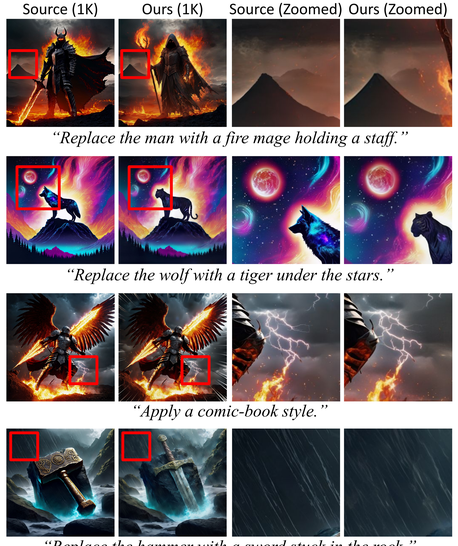
(2) $\tilde{\mathbf{y}}_{t-1}[i] = \sqrt{\alpha_{t-1}}L(\hat{\mathbf{x}}_{t\to0}[i], \hat{\mathbf{x}}_{t\to0}^{\text{aux}}[i, i+1]) + \sqrt{1-\alpha_{t-1}}\epsilon_\theta(\tilde{\mathbf{y}}_t^{\text{rsp}}[i], t)$

4 Experiments

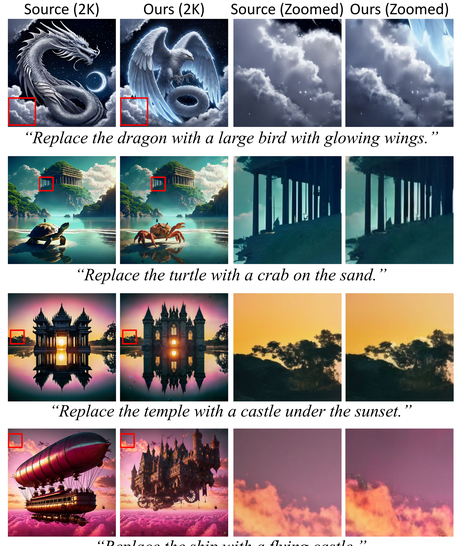
- **Qualitative Comparisons with Real-ISR baselines**



- **Results (1K-Editing)**



- **Results (2K-Editing)**



- **Results (8K-Editing)**

- **Results (1K Real Image Editing)**

- **Ablation Study**

- **Quantitative Comparisons with Real-ISR baselines**

Method	1K-editing					2K-editing				
	HaarPSI ↑	M-MSE ↓	M-SSIM ↑	M-PSNR ↑	LPIPS ↓	HaarPSI ↑	M-MSE ↓	M-SSIM ↑	M-PSNR ↑	LPIPS ↓
DiT-SR [7]	0.335	0.058	0.695	21.528	0.477	0.316	0.057	0.754	21.380	0.507
DiT4SR [12]	0.324	0.060	0.625	20.740	0.509	0.305	0.058	0.684	20.701	0.534
PiSA-SR [48]	0.328	0.058	0.668	21.273	0.465	0.312	0.056	0.755	21.320	**0.472**
TSD-SR [10]	0.329	0.061	0.649	20.766	0.489	0.312	0.059	0.715	20.796	0.514
ScaleEdit (Ours)	**0.342**	**0.054**	**0.739**	**22.132**	**0.460**	**0.331**	**0.053**	**0.806**	**21.955**	0.496