PhysRemover

A Unified Framework for Physically Realistic Object Removal

Dongxu Yue, Weixuan Jin, Yuhang Yu, Bo Li, Qi Wen, Xinrui Chen,
Jinwei Chen, Tianyi Zheng, Hao Zhang, Zhihai He, Chun Yuan

Contact: ydx25@mails.tsinghua.edu.cn

NeurIPS 2026
MAIN CONFERENCE

Tsinghua University

Paper

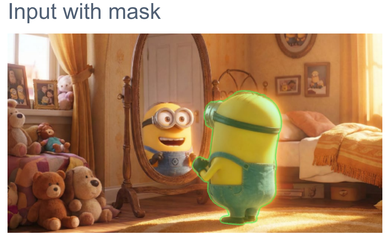
Remove the object and its physical effects. Preserve the background visible through it.

1 The challenge

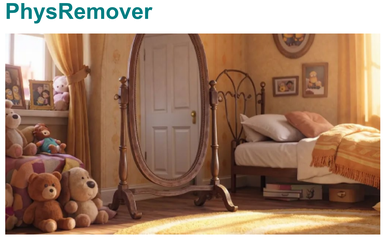
A mask identifies the object, but its influence extends beyond it. Standard inpainting can leave physical traces and erase visible background details.

Object + reflection

Input with mask



PhysRemover

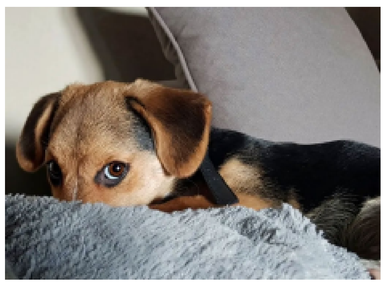


Transparent object + background

Input with mask

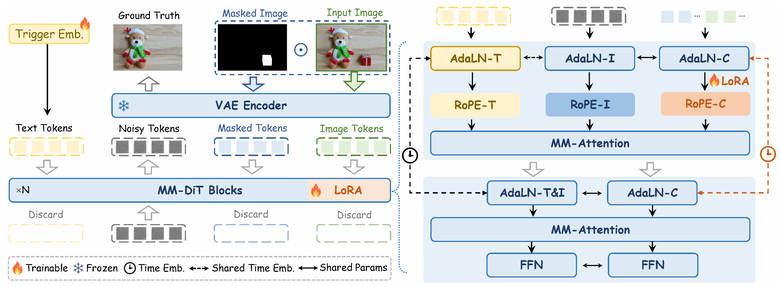


PhysRemover



Counterfactual goal: reconstruct the scene as if the target object were absent, while retaining authentic background structure.

2 How PhysRemover works



In-context contrastive guidance (ICCG)

Paired visual conditions. The original image retains object properties and background cues. The masked image specifies what to remove.

Spatial and temporal separation. Spatially aligned RoPE connects corresponding locations. Clean conditions use timestep 0, separately from the noisy target.

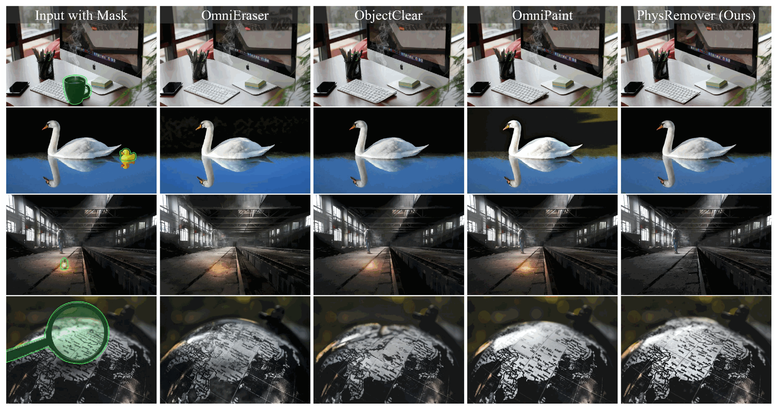
Learnable removal trigger (LRT)

Optimized task tokens replace a fixed text instruction. Object-based pivot initialization guides attention to the target and its physical effects.

Two-stage training: general inpainting pretraining, followed by counterfactual clean-background learning on PhyTrace.

3 Evidence and benchmark

Steam, reflections, illumination and transparency



Selected baselines from the paper (Table 1)

Method	FID (down)	PER (up)	PSNR (up)
ObjectClear	20.317	3.360	19.551
OmniPaint	22.001	3.334	19.331
GeoRemover	23.620	3.492	19.531
PhysRemover	**17.402**	**4.528**	**20.882**

FID / PER: PhyTrace-Test. PER evaluates physical-effects removal (1–5). PSNR: transparent regions on PhyTrace-TP, in dB.

10,000 pairs
500 test pairs · 150 transparent cases

PhyTrace: high-resolution image pairs, verified by VLM assessment and human review.

PhysRM: removal quality, physical effects and natural blending.