



Re-HOLD

Re-HOLD: Video Hand Object Interaction Reenactment via adaptive Layout-instructed Diffusion Model

CVPR25

Yingying Fan¹ Quanwei Yang² Kaisiyuan Wang³ Hang Zhou³
Yingying Li³ Haocheng Feng³ Errui Ding³ Yu Wu¹ Jingdong Wang³

¹ Wuhan University, ² University of Science and Technology of China

³ Department of Computer Vision Technology (VIS),
Baidu Inc.

Comparison Methods

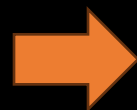
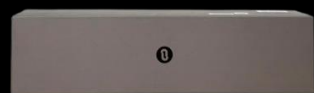
- Max Ku, Cong Wei, et al. **Anyv2v: A plug-and-play framework for any video-to-video editing tasks.** TMLR 2024
- Yuchao Gu, Yipin Zhou, et al. **Videoswap: Customized video subject swapping with interactive semantic point correspondence.** CVPR 2024
- Li Hu, Xin Gao, et al. **Animate anyone: Consistent and controllable image-to-video synthesis for character animation.** CVPR 2024
- Jingkai Zhou, Benzhi Wang, et al. **RealisDance: Equip controllable character animation with realistic hands.**
arXiv:2409.06202
- Zihui Xue, Mi Luo, et al. **HOI-Swap: Swapping Objects in Videos with Hand-Object Interaction Awareness.**
NeurIPS 2024

Self-Reenactment Results

Driving Video



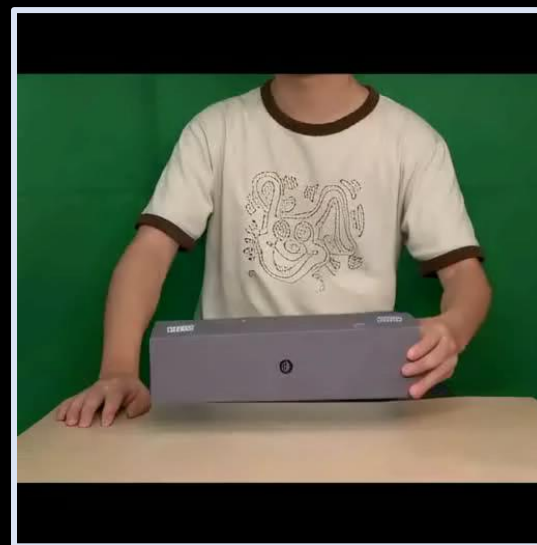
**Target
Object**



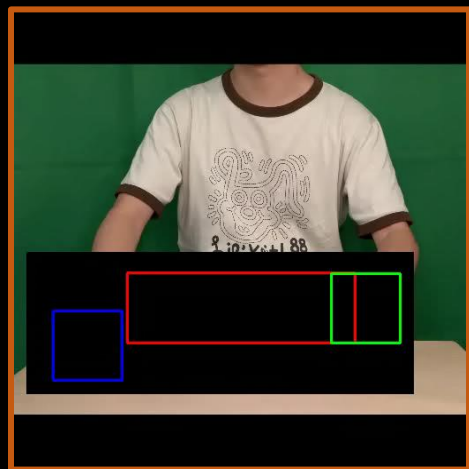
Ours



HOI-Swap



RealisDance

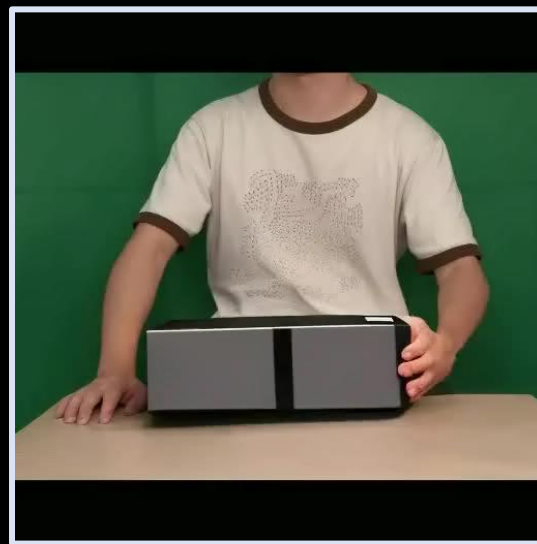


Layout Guidance

AnimateAnyone



VideoSwap



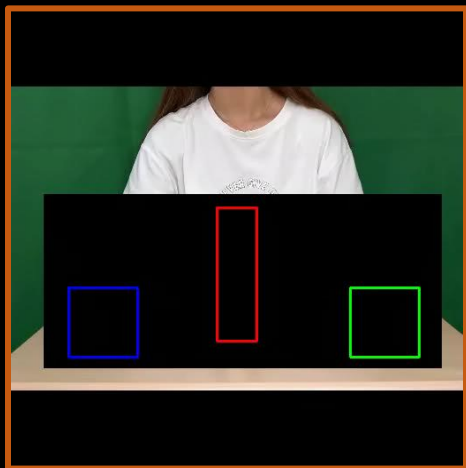
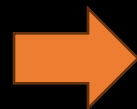
AnyV2V



Driving Video

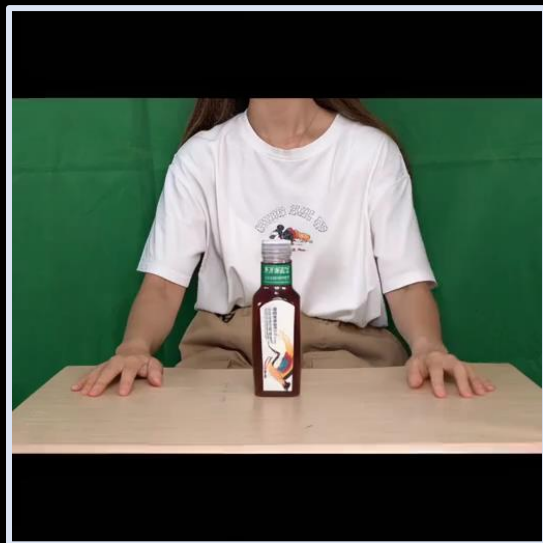


**Target
Object**



Layout Guidance

Ours



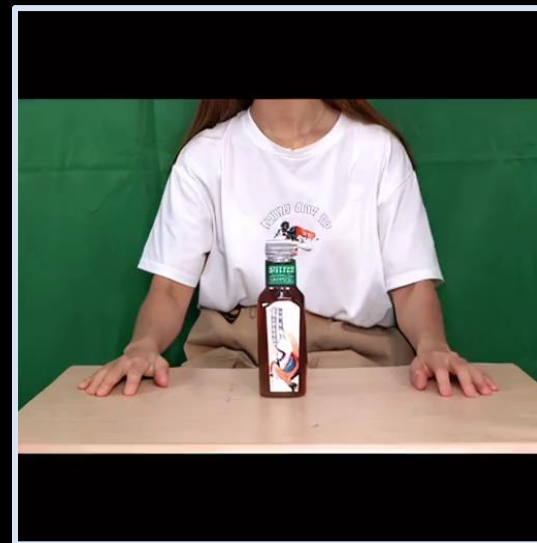
AnimateAnyone

HOI-Swap



VideoSwap

RealisDance

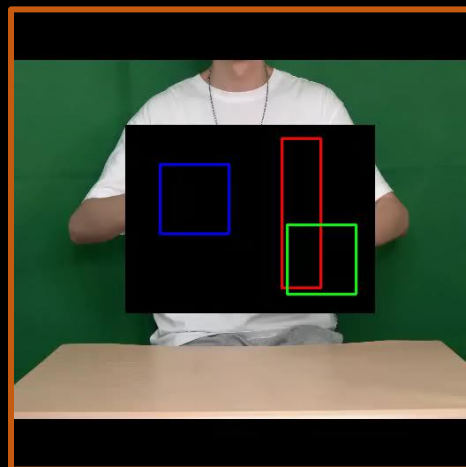
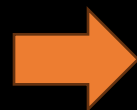


AnyV2V

Driving Video



**Target
Object**



Layout Guidance

Ours



AnimateAnyone

HOI-Swap



VideoSwap

RealisDance



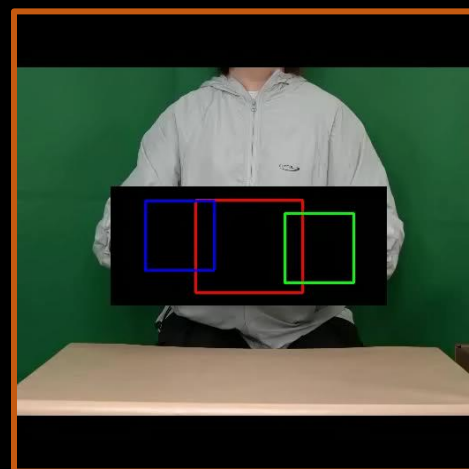
AnyV2V

Cross-Reenactment Results

Driving Video



Target Object



Layout Guidance

Ours



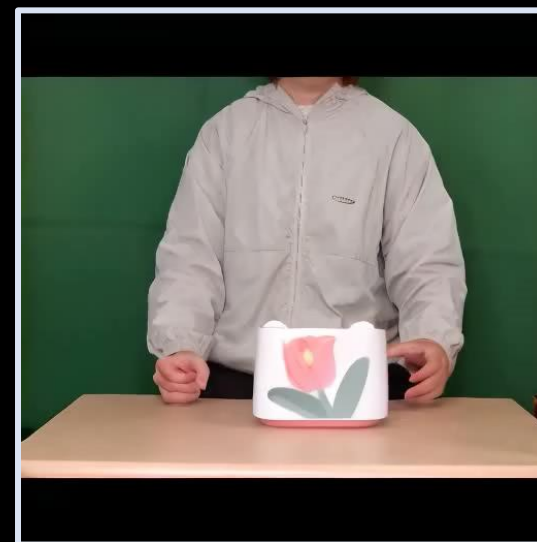
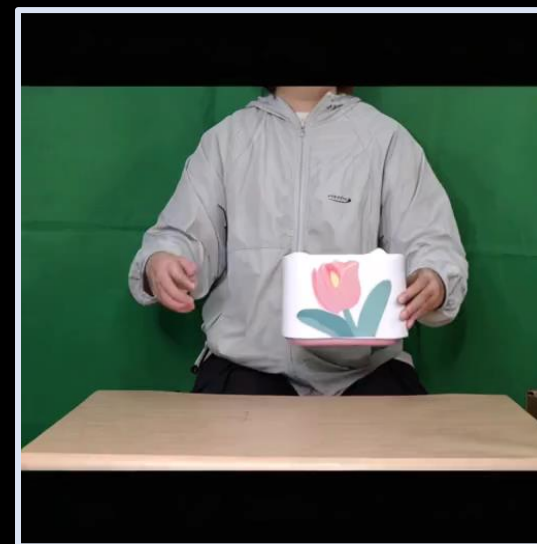
AnimateAnyone

HOI-Swap



VideoSwap

RealisDance

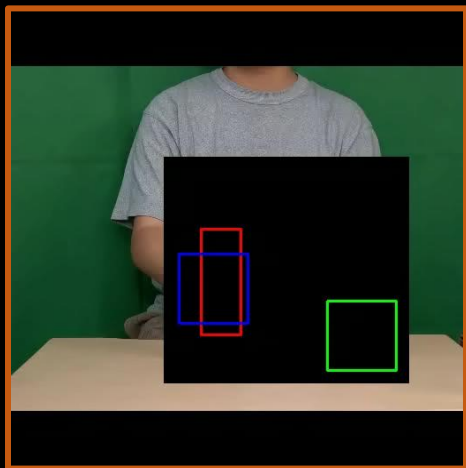


AnyV2V

Driving Video



**Target
Object**



Ours



HOI-Swap



RealisDance



AnimateAnyone



VideoSwap



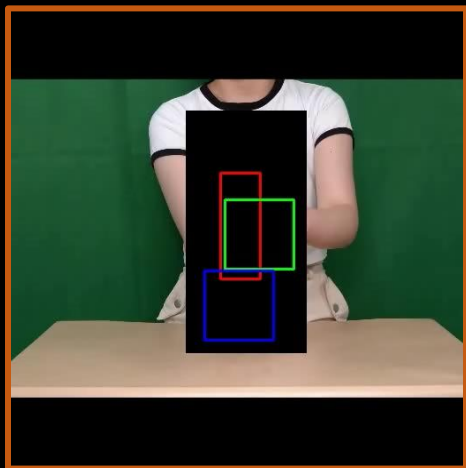
AnyV2V



Driving Video

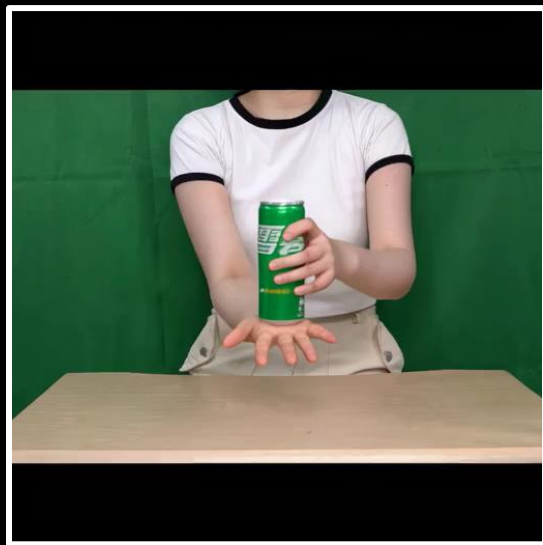


Target Object



Layout Guidance

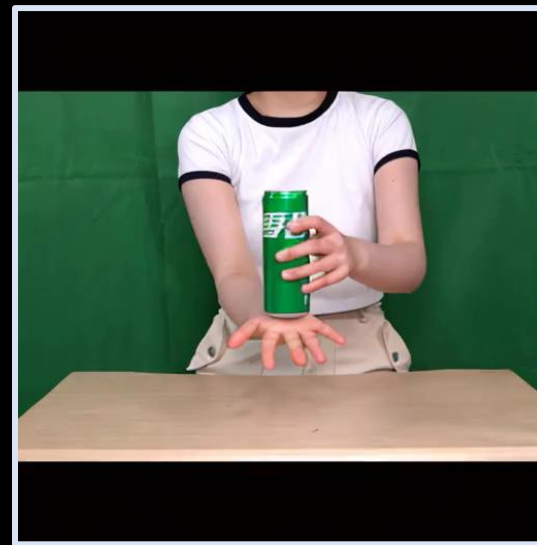
Ours



HOI-Swap



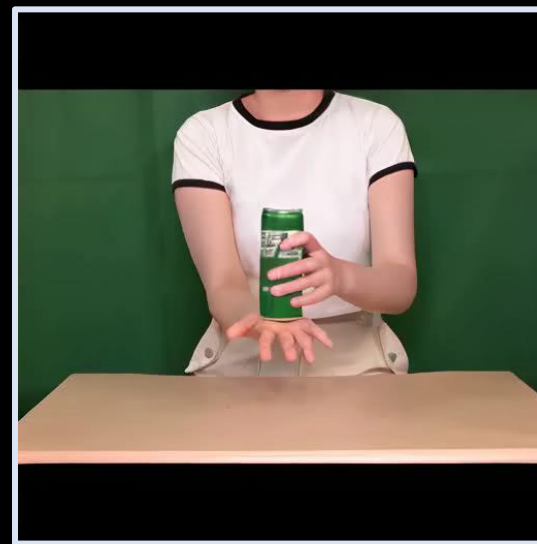
RealisDance



AnimateAnyone



VideoSwap

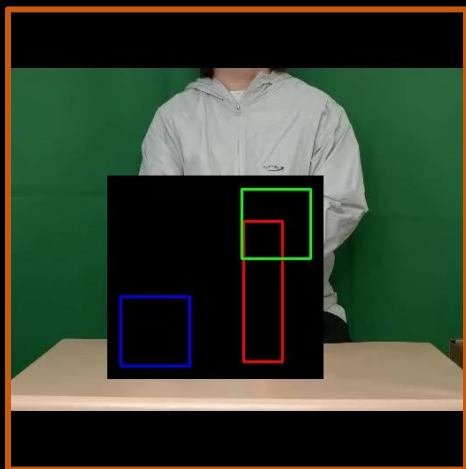


AnyV2V

Driving Video



Target Object



Layout Guidance

Ours



AnimateAnyone

HOI-Swap



VideoSwap

RealisDance



AnyV2V

Self-Reenactment Results on HOI4D



Target Object



Driving Video



Layout Guidance



Ours



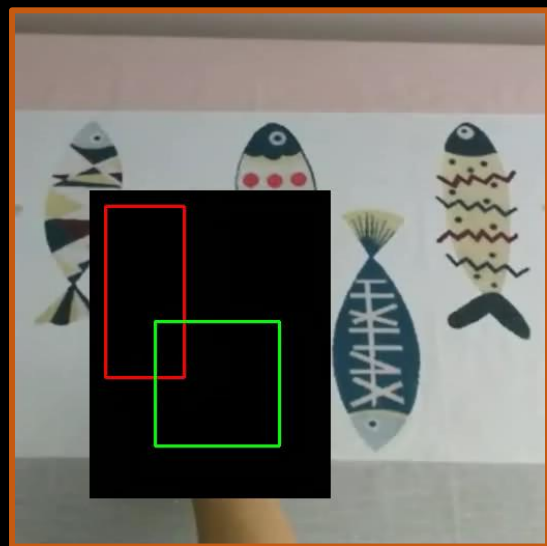
HOI-Swap



VideoSwap



AnyV2V



Cross-Reenactment Results on HOI4D



Target Object



Driving Video



Layout Guidance



Ours



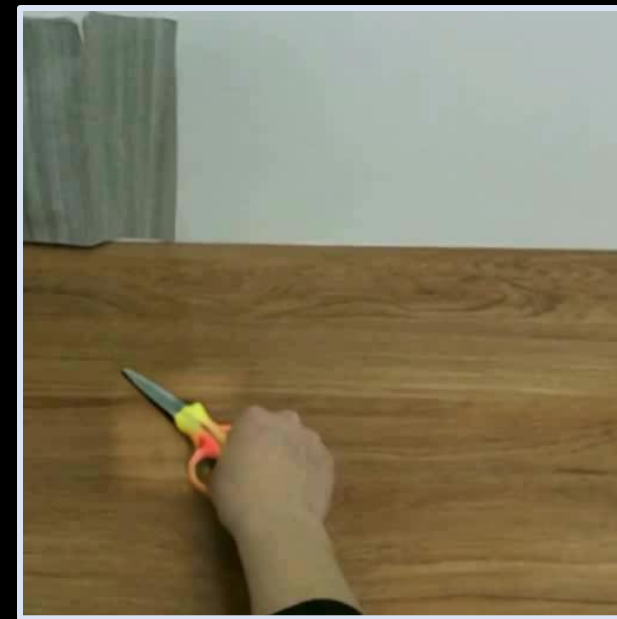
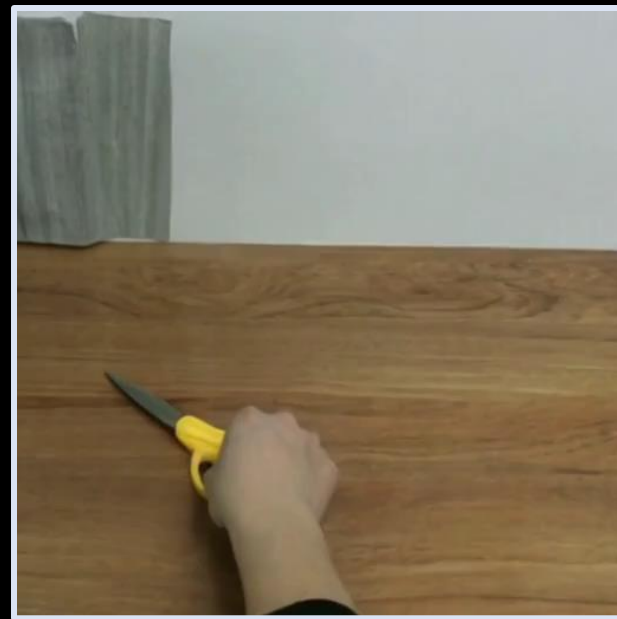
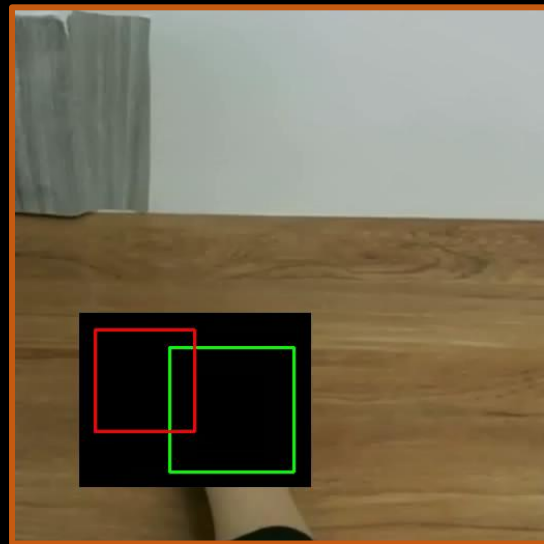
HOI-Swap



VideoSwap



AnyV2V



Ablation Study

Ablation of the layout instruction and 3D Hand Mesh

Driving Video & Frame



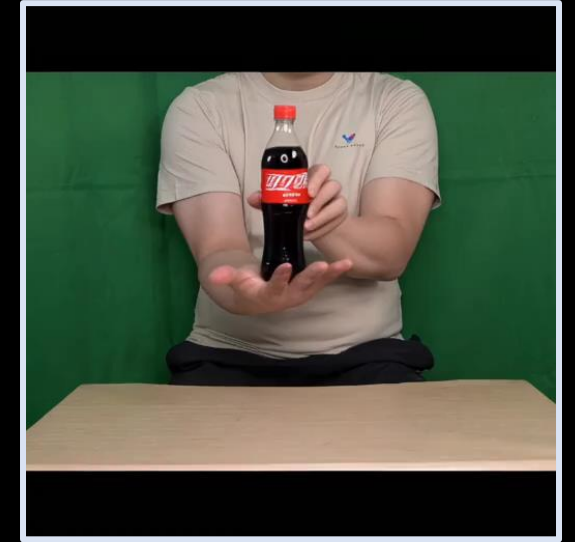
Ours



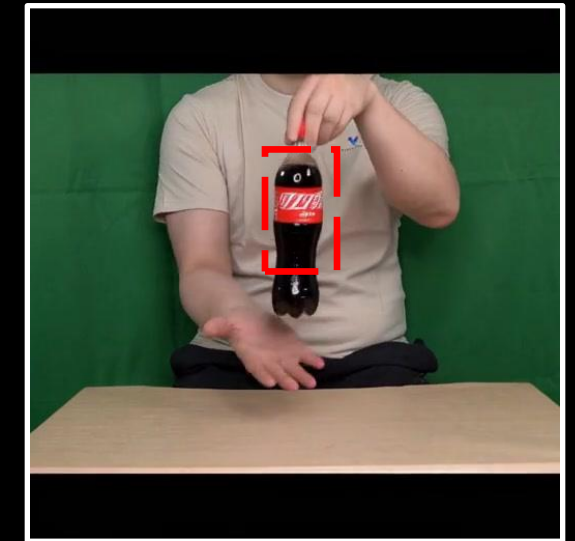
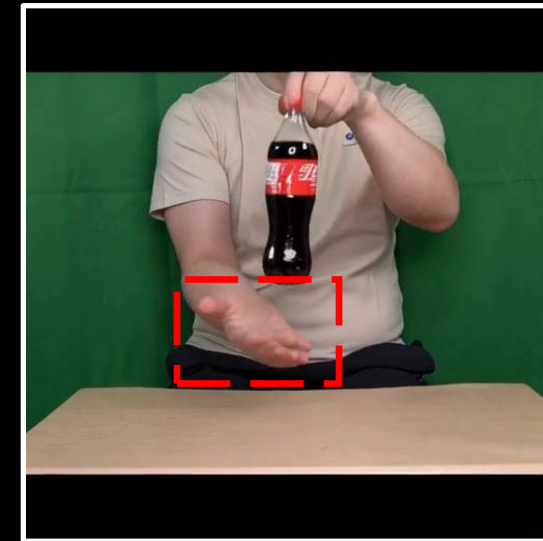
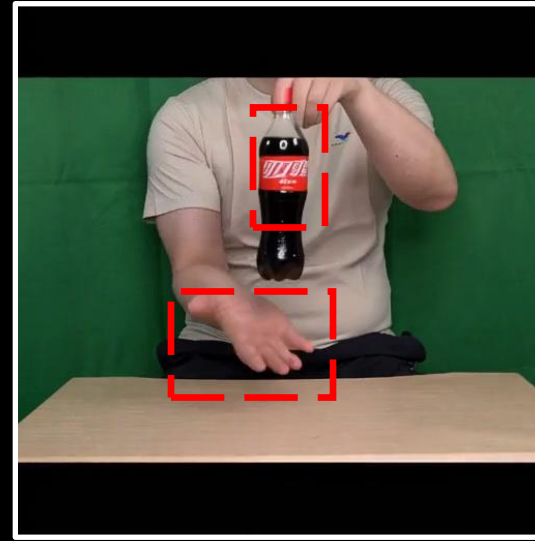
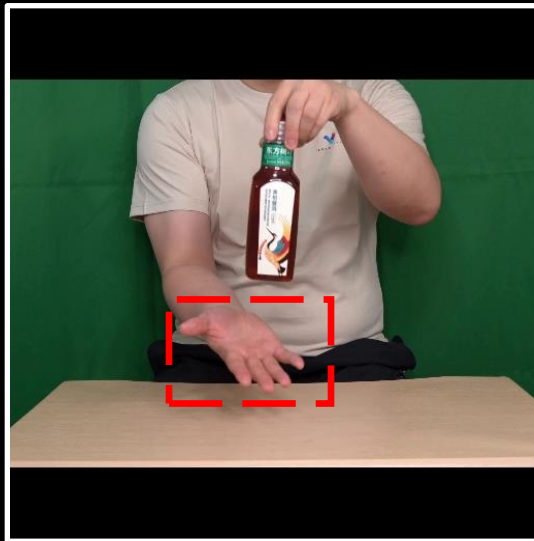
w/o Hand Mesh



w/o Layout



Target Object



Best
Result

Hand
Blurring

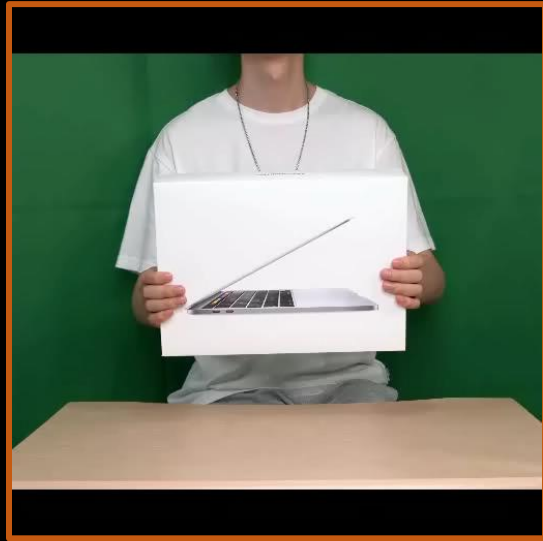
Deformed
Object

Ablation of the HOI Restoring Module

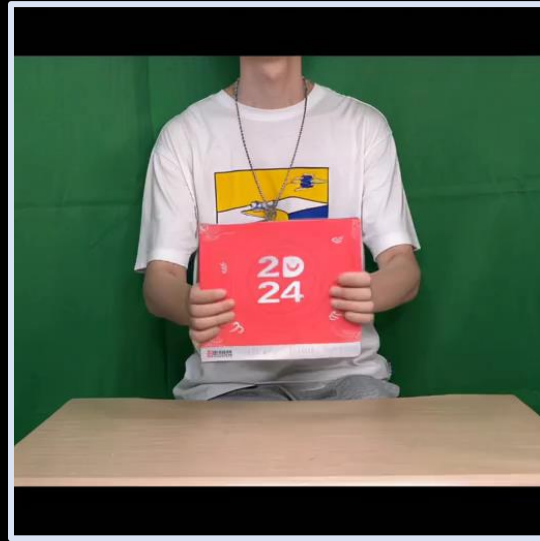
Ours

w/o Hand Attn

w/o Obj Attn



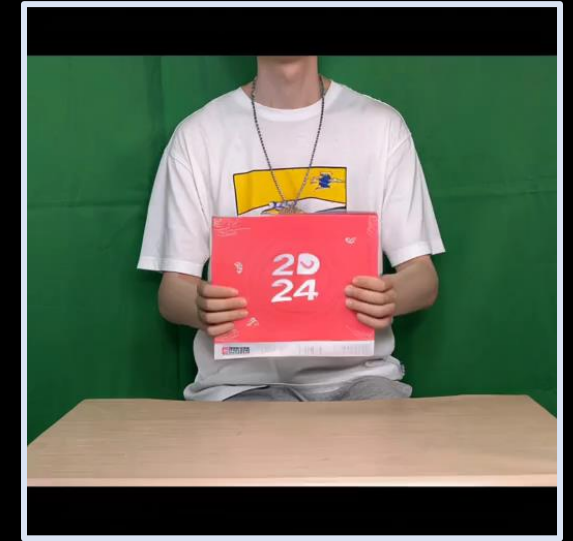
Driving Video



Best
Result



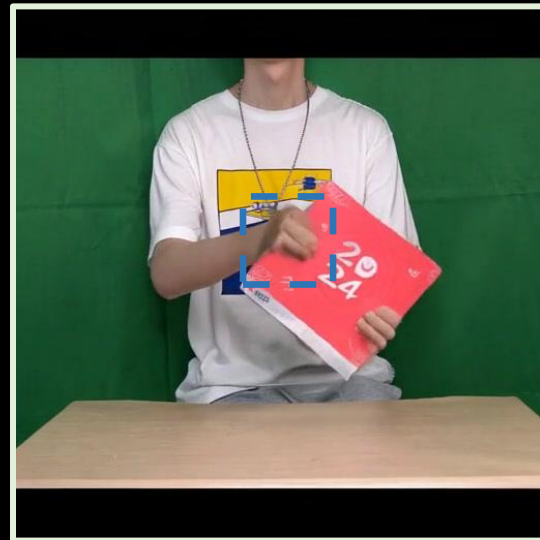
Hand
Blurring



Object
Texture Blurring



Target Object



Ablation of the adaptive layout adjustment strategy

Driving Video

Ours

w/o adaptive

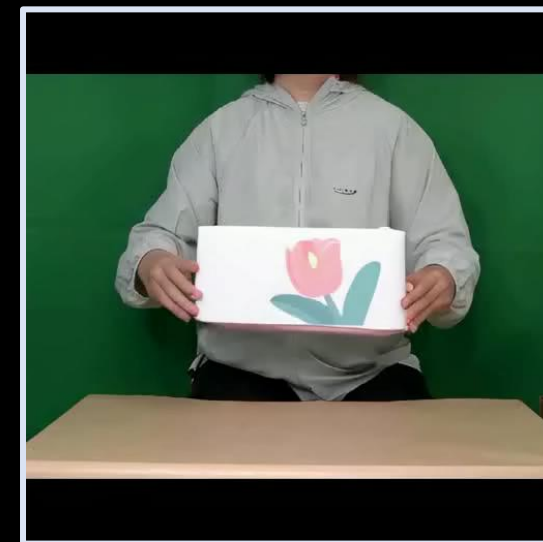
HOI-Swap



Target Object



Target Object



Best
Result

Deformed
Object

Replacement
Failure

Ablation of different Layout adjustment



Target Object

Object Size
adjustment

Driving Video

Ours

w/o adjustment

Deformed
Object



Target Object

Horizontal
position
adjustment

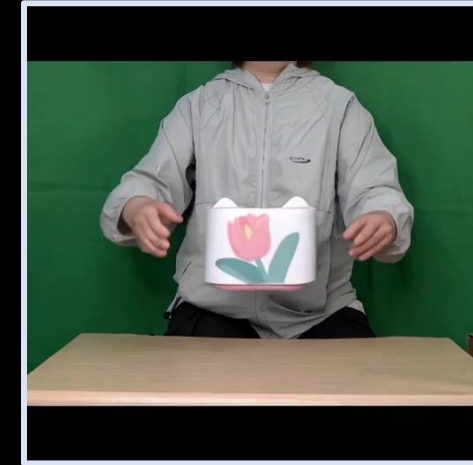
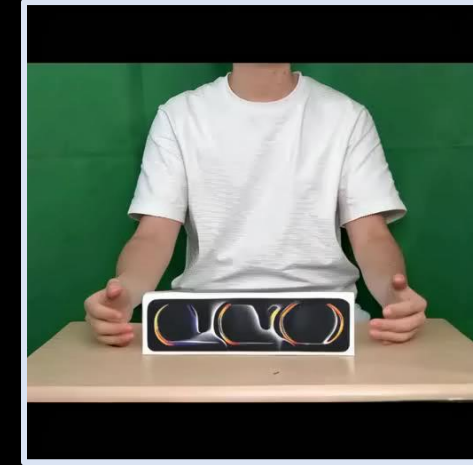
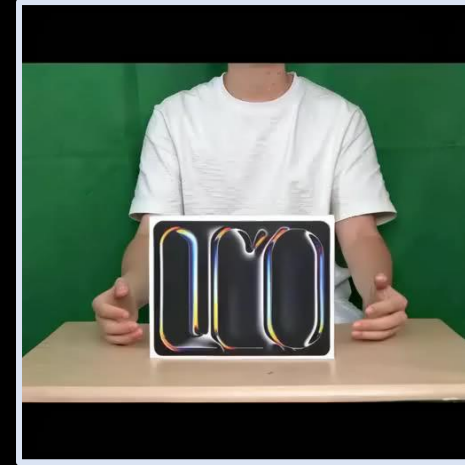
Inappropriate
Hand Position



Target Object

Vertical
position
adjustment

Inappropriate
Hand Position



Thank you for watching!