

PhysX-Anything: Simulation-Ready Physical 3D Assets from Single Image

Ziang Cao¹, Fangzhou Hong¹, Zhaoxi Chen¹
Liang Pan², Ziwei Liu¹

¹Nanyang Technological University

²Shanghai AI Lab



**NANYANG
TECHNOLOGICAL
UNIVERSITY**
SINGAPORE

S-LAB
FOR ADVANCED
INTELLIGENCE

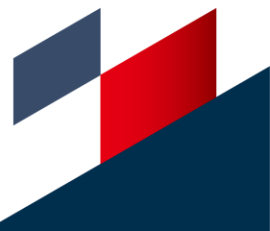
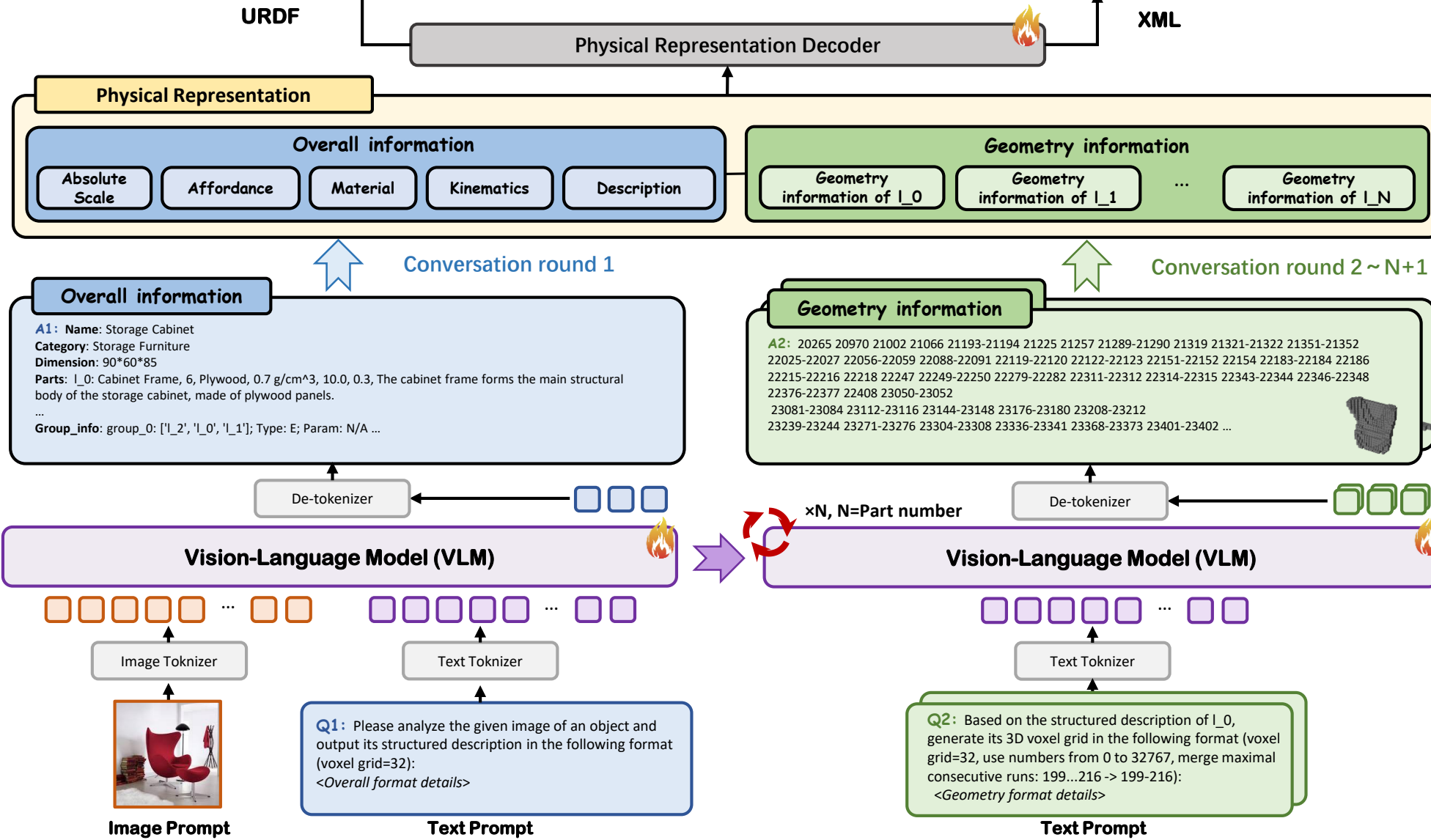
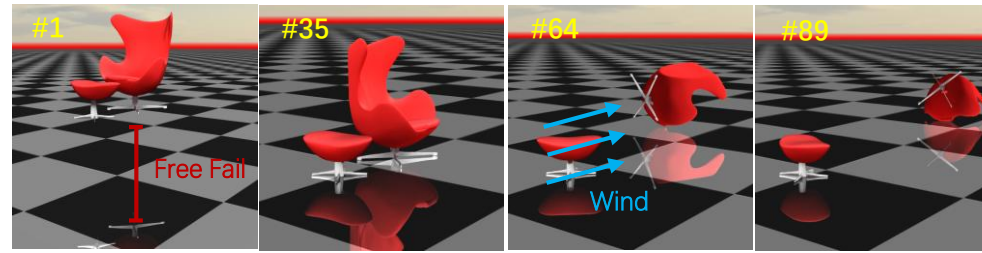
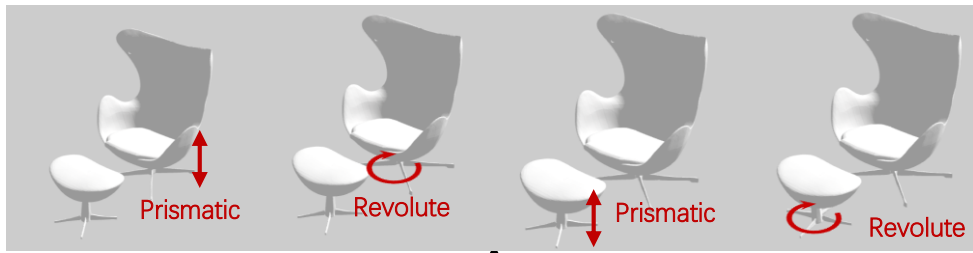


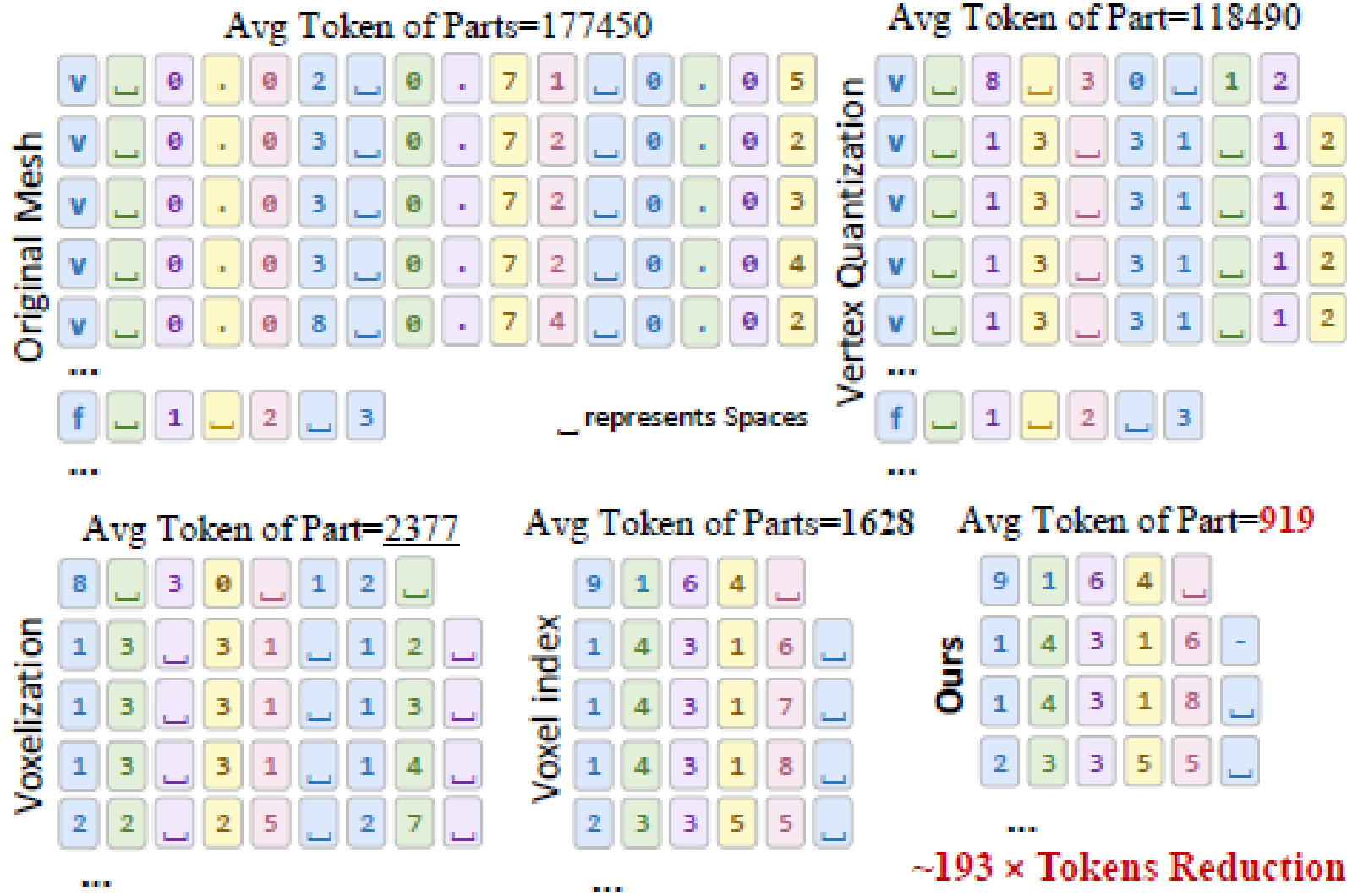
Contribution

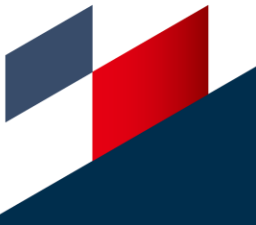
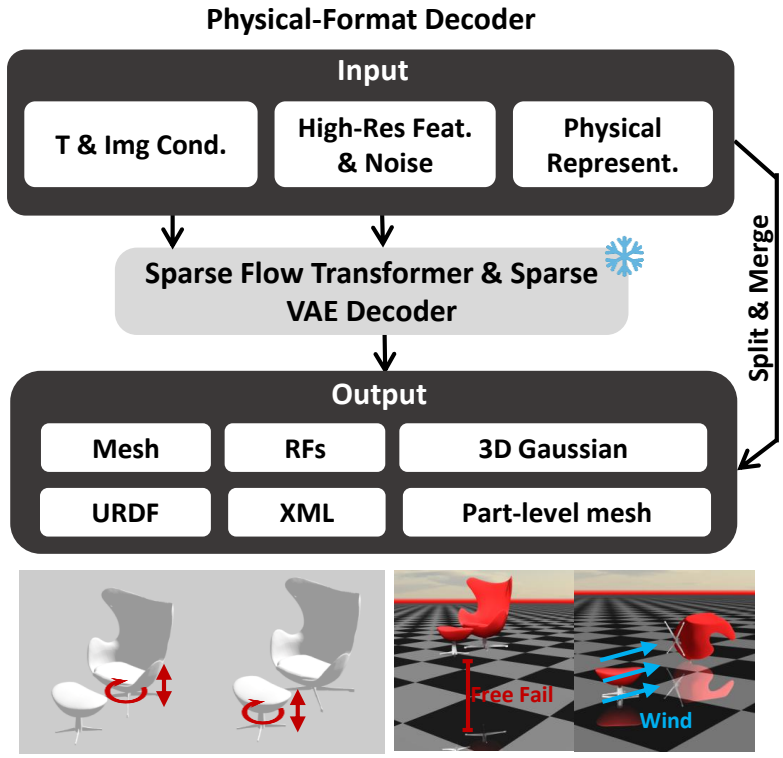
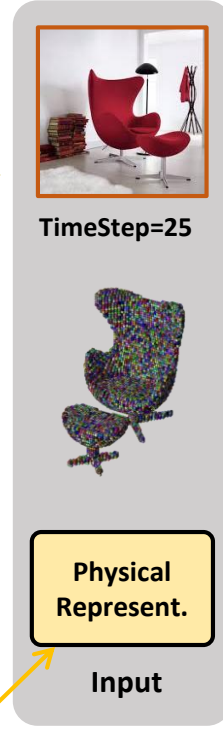
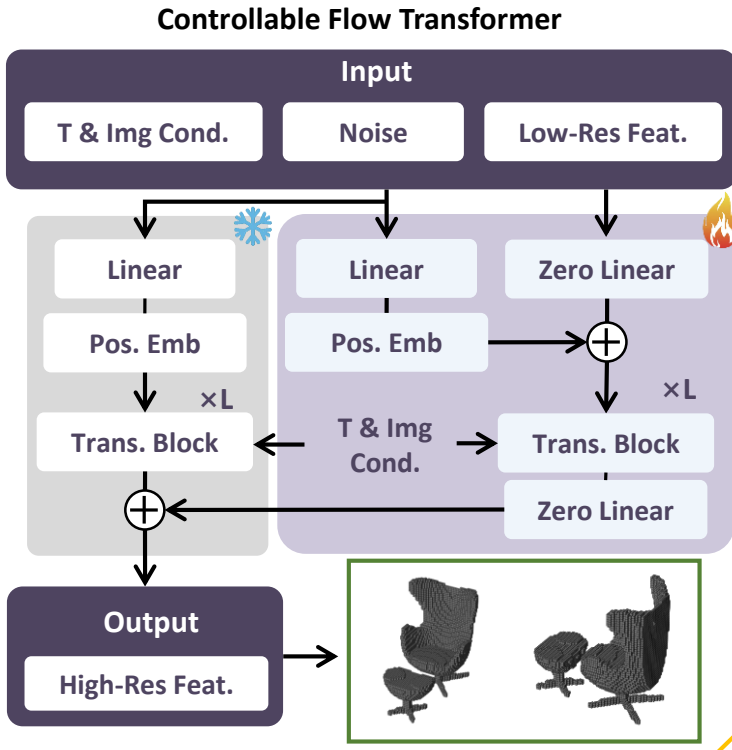
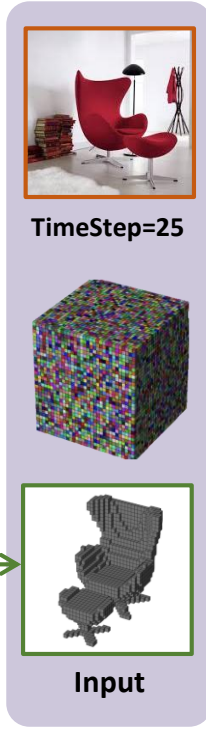
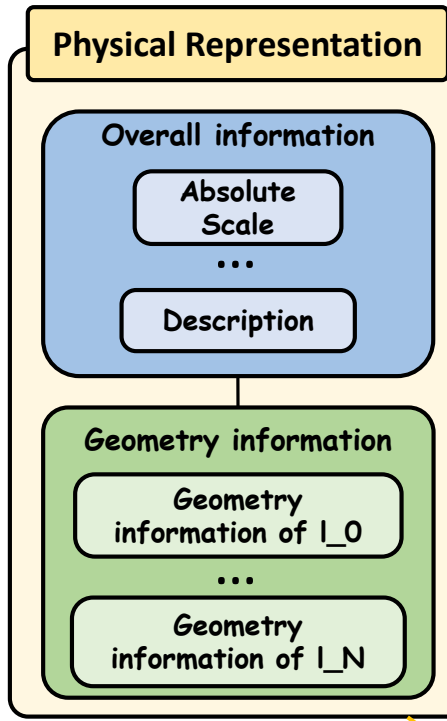
- We introduce PhysX-Anything, **the first sim-ready physical 3D generative** paradigm that, given a single in-the-wild image, produces high-quality sim-ready 3D assets, thereby pushing the frontier of physically grounded 3D content creation and unlocking new possibilities for downstream applications in simulation and embodied AI.
- We propose a unified **VLM-based generative pipeline together with a novel physical 3D representation**. Our representation compresses geometry tokens at a high rate while preserving explicit geometric structure, and avoids introducing any special tokens during fine-tuning.
- We construct a **new physically grounded 3D dataset, PhysX-Mobility**, which enriches the category diversity of existing physical 3D datasets by over 2x, including over 2K common real-world objects such as cameras, coffee machines, and staplers.

Methodology of PhysX-Anything



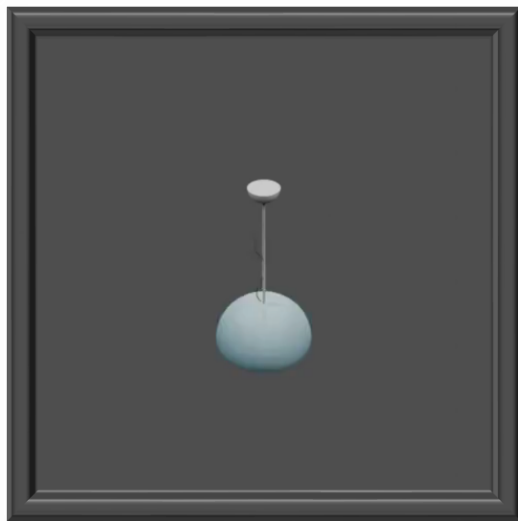


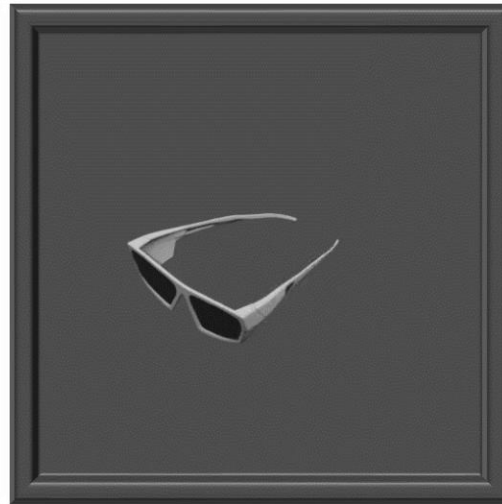


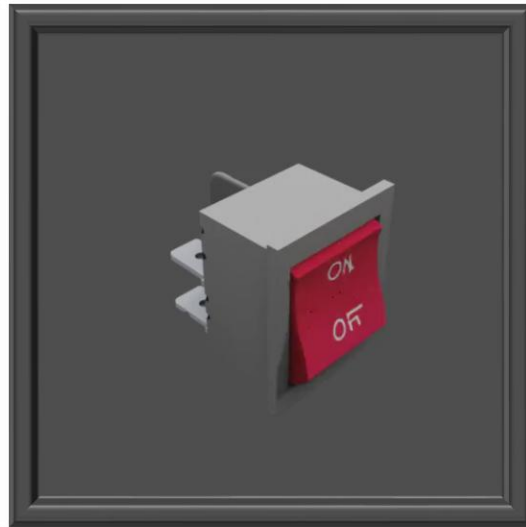


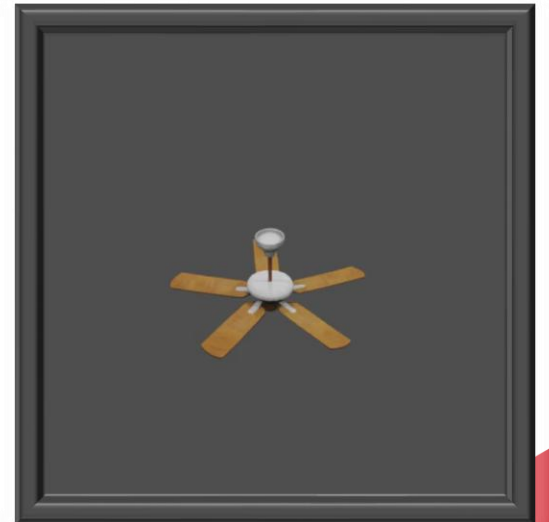
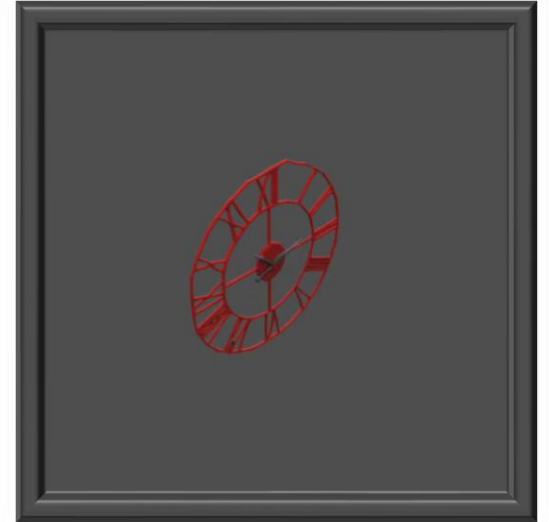
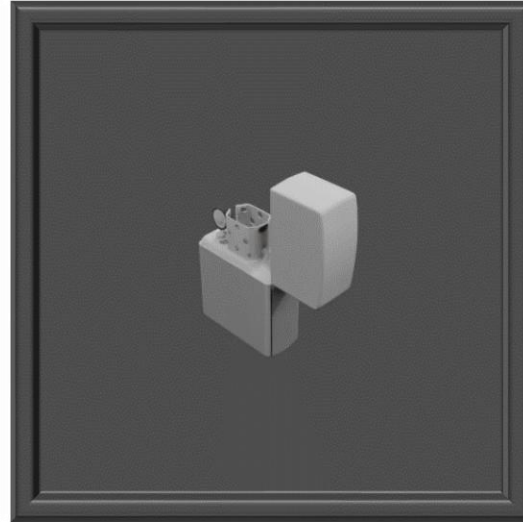
Experimental results













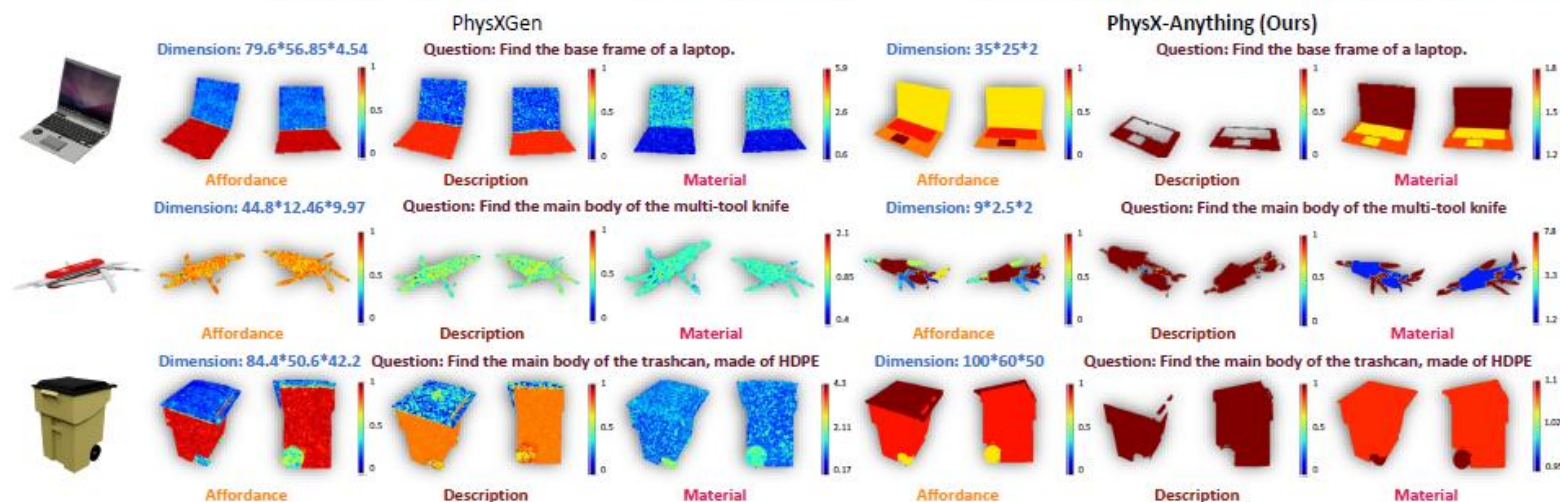
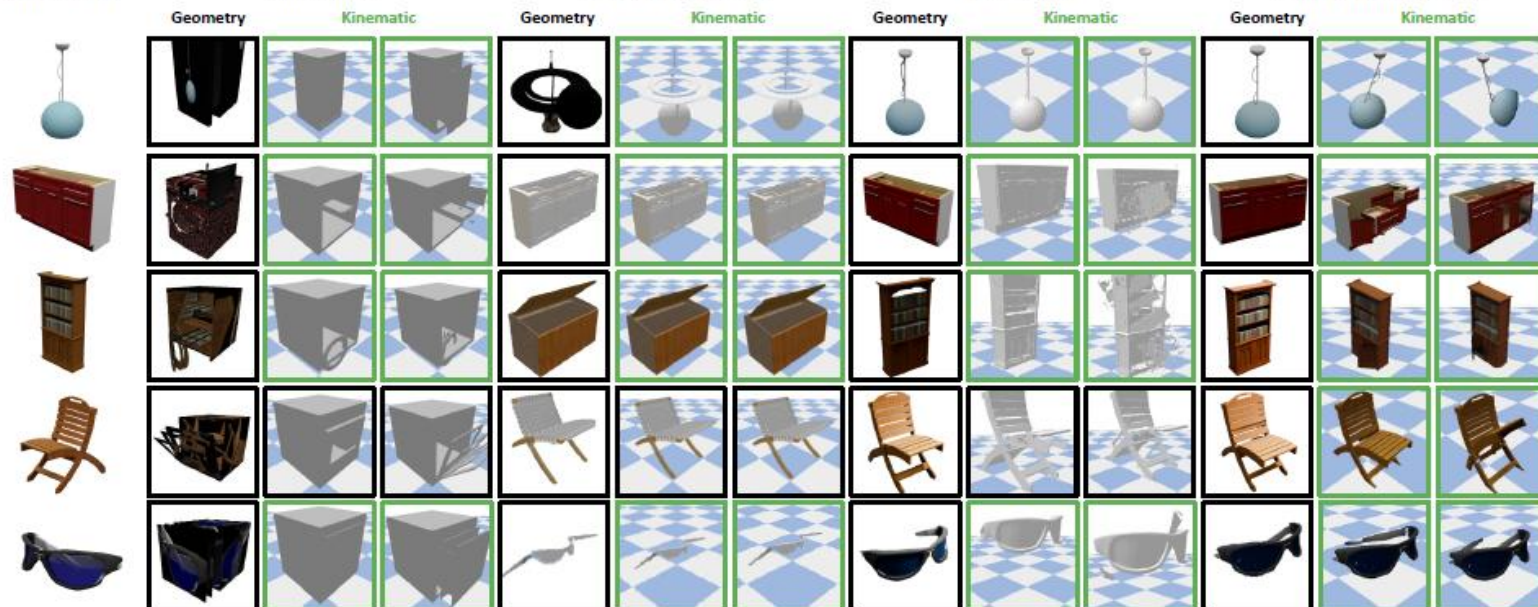
Input Image

URDFormer

Articulate-Anything

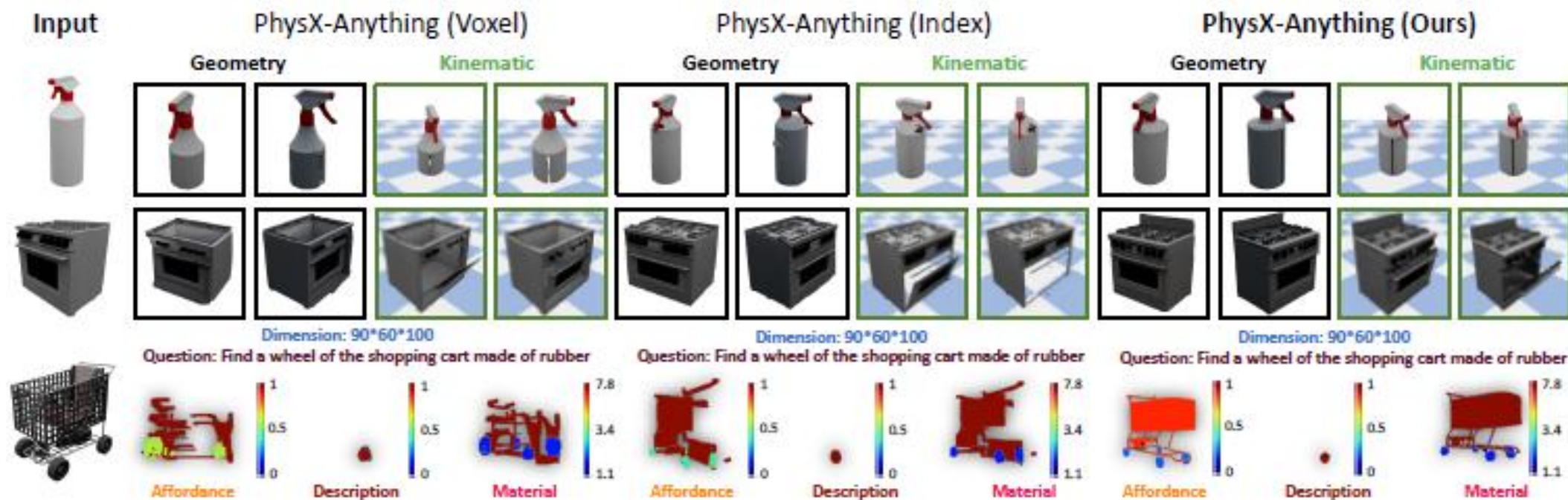
PhysXGen

PhysX-Anything (Ours)



| Methods | Geometry | | | Physical Attributes | | | | |
|------------------------------|-----------------|-----------------|--------------------|-----------------------------|---------------------|-----------------------|---------------------------------------|------------------------|
| | PSNR \uparrow | CD \downarrow | F-score \uparrow | Absolute scale \downarrow | Material \uparrow | Affordance \uparrow | Kinematic parameters (VLM) \uparrow | Description \uparrow |
| URDFormer [11] | 7.97 | 48.44 | 43.81 | – | – | – | 0.31 | – |
| Articulate-Anything [16] | 16.90 | 17.01 | 67.35 | – | – | – | 0.65 | – |
| PhysXGen [3] | 20.33 | 14.55 | 76.3 | 43.44 | 6.29 | 9.75 | 0.71 | 12.89 |
| PhysX-Anything (Ours) | 20.35 | 14.43 | 77.50 | 0.30 | 17.52 | 14.28 | 0.83 | 19.36 |

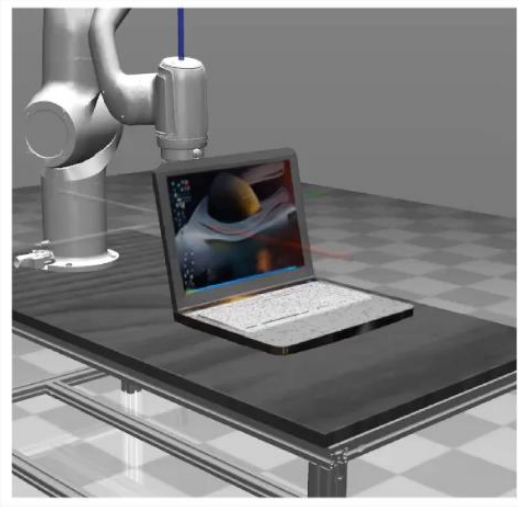




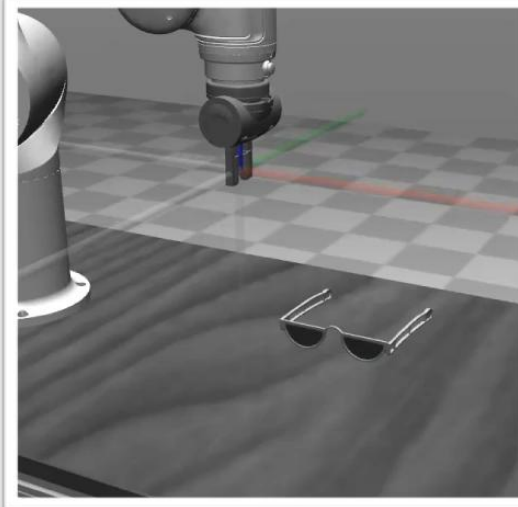
| Methods | Geometry | | | Physical Attributes | | | | |
|-----------------------|--------------|--------------|--------------|---------------------|--------------|--------------|------------------------------|---------------|
| | PSNR ↑ | CD ↓ | F-Score ↑ | Absolute scale ↓ | Material ↑ | Affordance ↑ | Kinematic parameters (VLM) ↑ | Description ↑ |
| PhysX-Anything-Voxel | 16.96 | 17.81 | 63.10 | 0.40 | 12.32 | 11.63 | 0.39 | 17.38 |
| PhysX-Anything-Index | 18.21 | 16.27 | 68.70 | 0.30 | 13.35 | 12.04 | 0.76 | 17.97 |
| PhysX-Anything (Ours) | 20.35 | 14.43 | 77.50 | 0.30 | 17.52 | 14.28 | 0.94 | 19.36 |



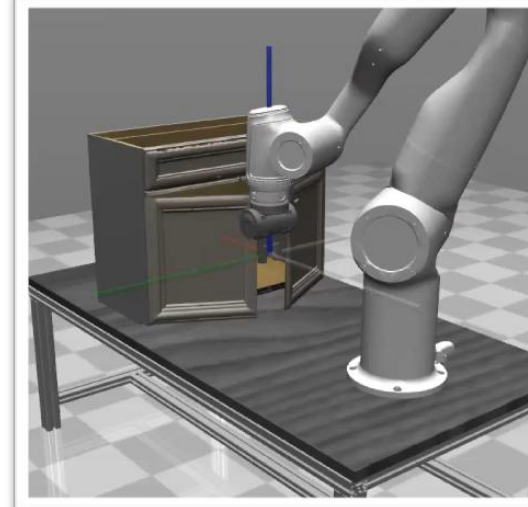
Laptop Closing



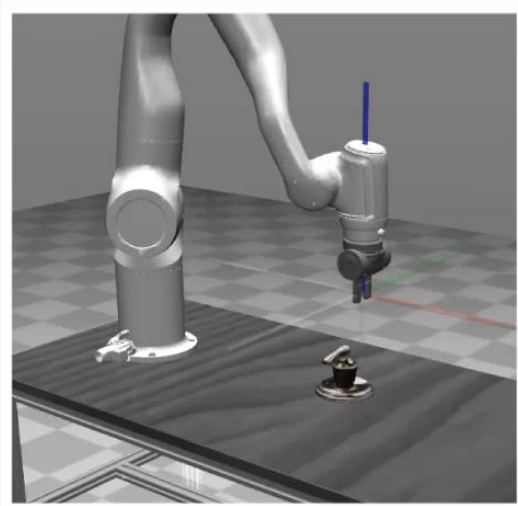
Eyeglass Temple Folding



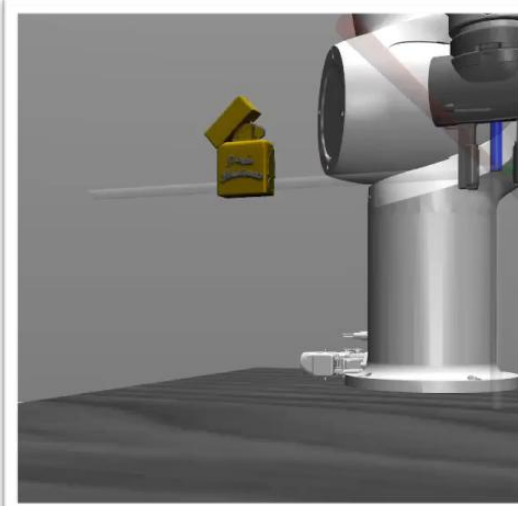
Door Opening and Closing



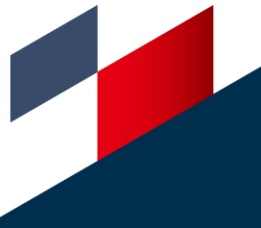
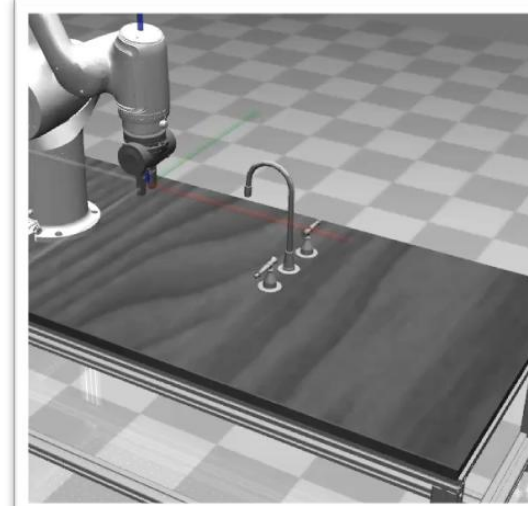
Handle Manipulation



Lighter Snapping Open



Faucet Switch Manipulation



Thanks for listening!



Home page:

<https://physx-3d.github.io/>



Home page:

<https://physx-anything.github.io/>

