



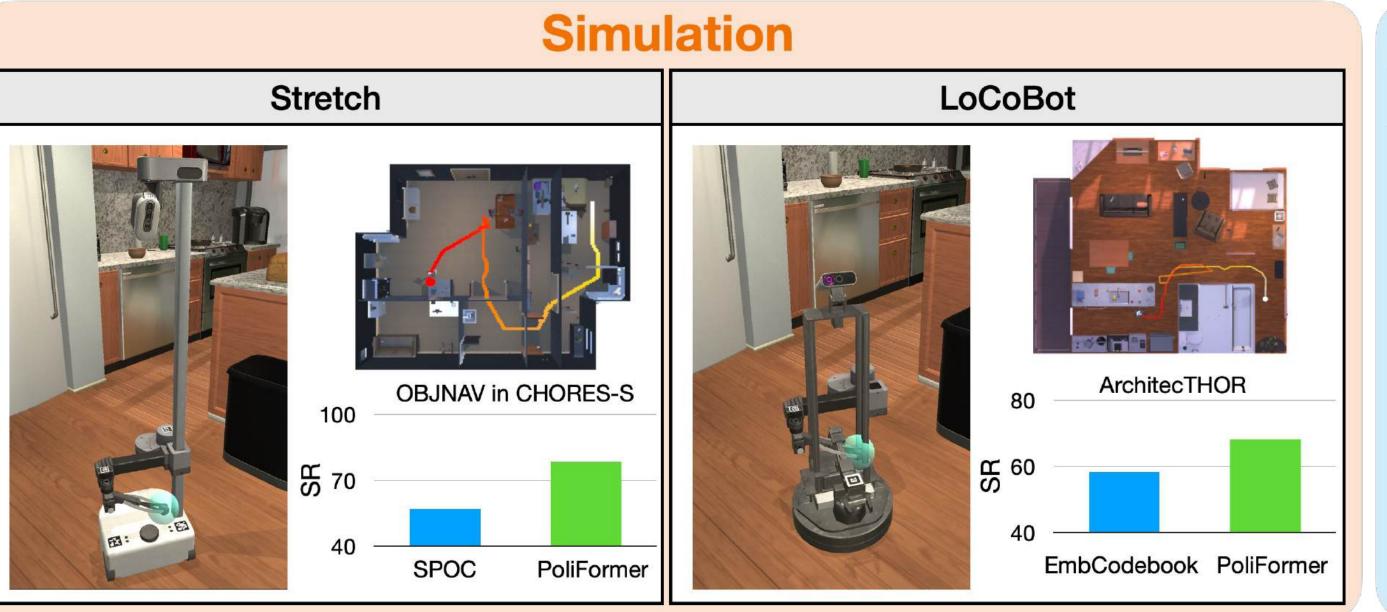
PoliFormer: Scaling On-Policy RL with Transformers Results in Masterful Navigators

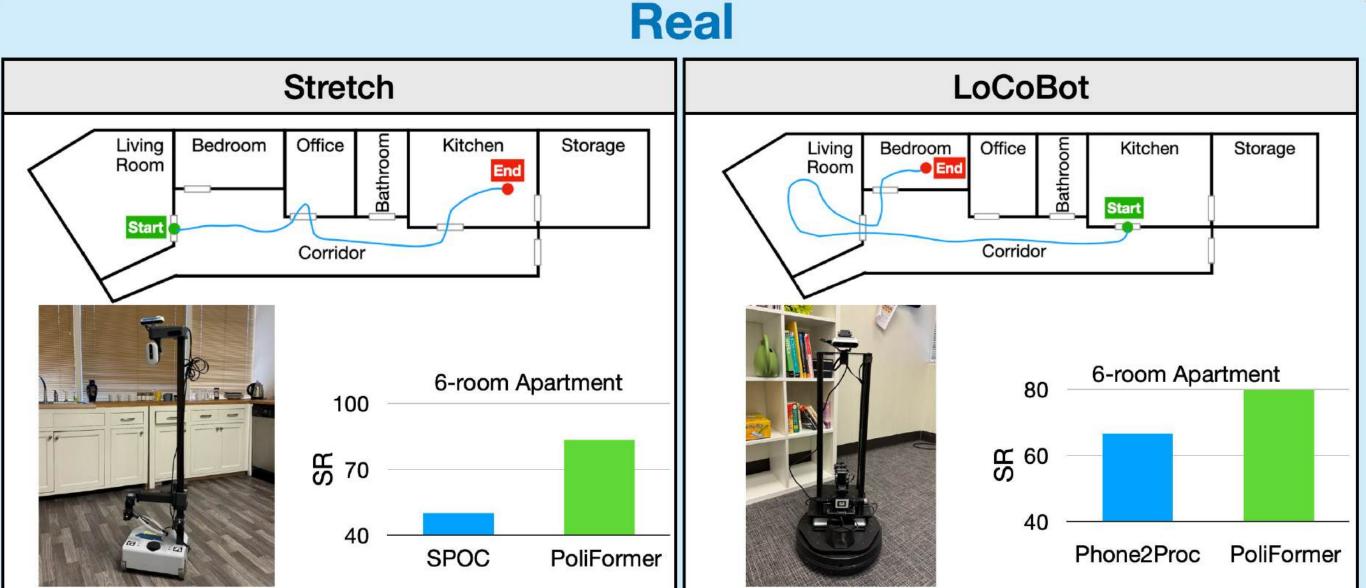


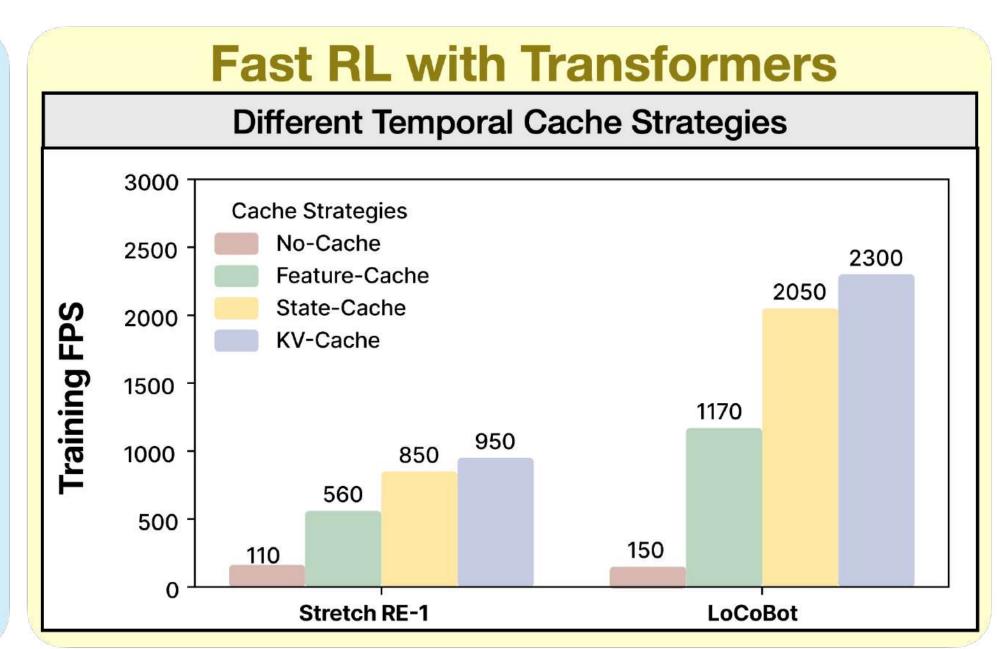
Kuo-Hao Zeng, Zichen "Charles" Zhang, Kiana Ehsani, Rose Hendrix, Jordi Salvador,

Alvaro Herrasti, Ross Girshick, Aniruddha Kembhavi, Luca Weihs

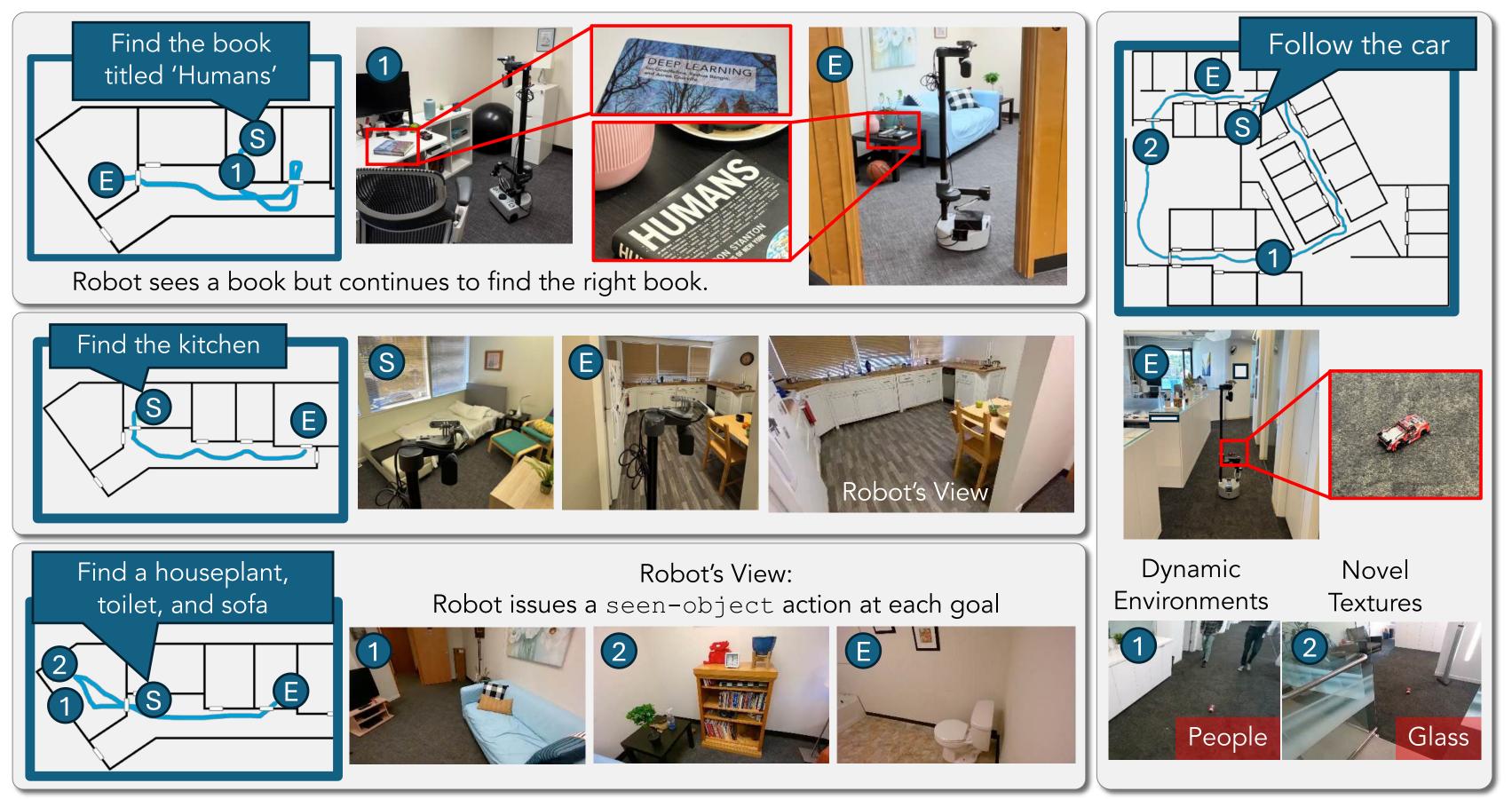
Scaling On-Policy RL Training with Transformers Continual Improvement with Scale **Hundreds of Parallel Rollouts** Transformer-based Policy with Visual Foundation Model CHORES (val) 9 2 8 **Causal Transformer Decoder** STATE TOKEN → Vision KV-cache Transformer 5 50 ⊗ 40 Bate 00 **Goal Encoder** a sofa 700M 100M 10M Hundreds of Millions of Model Parameters Training steps



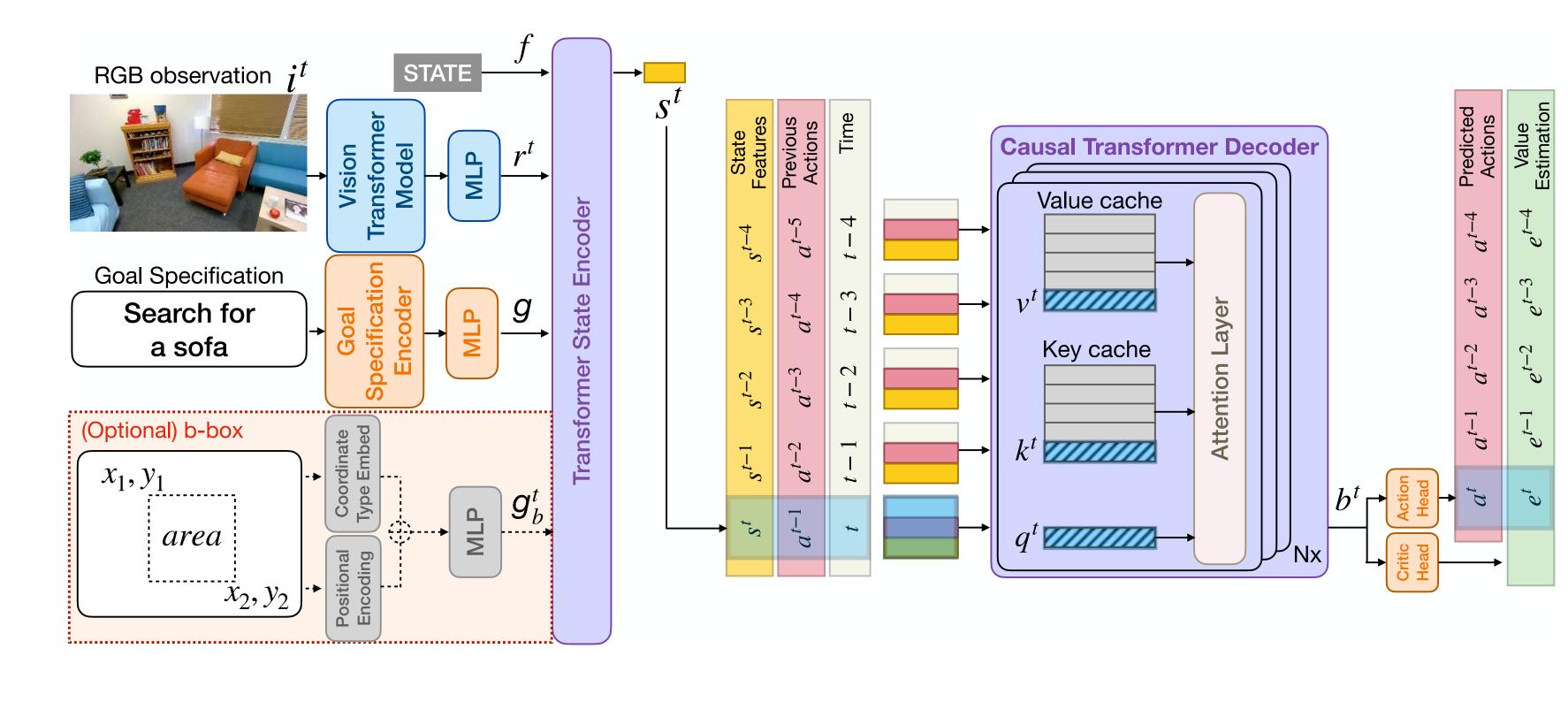




Zero-Shot Downstream Tasks







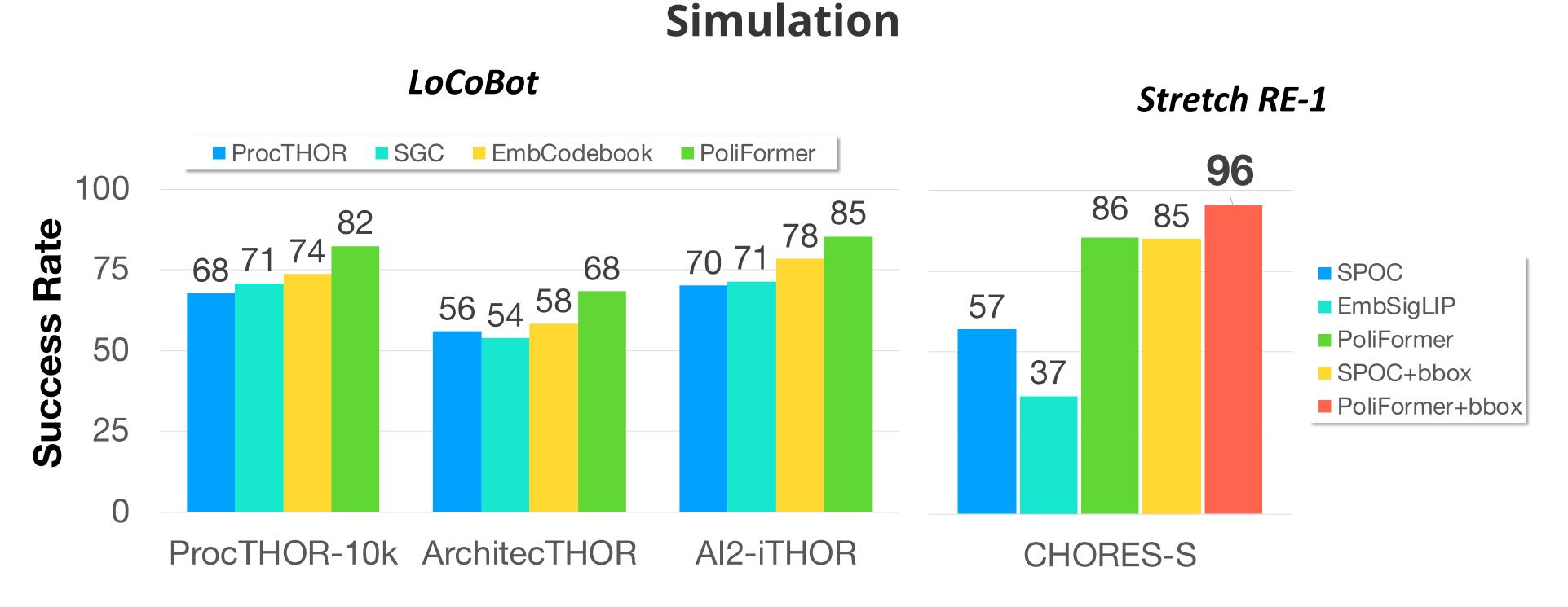
Scale Model and Data

Scale in Architecture

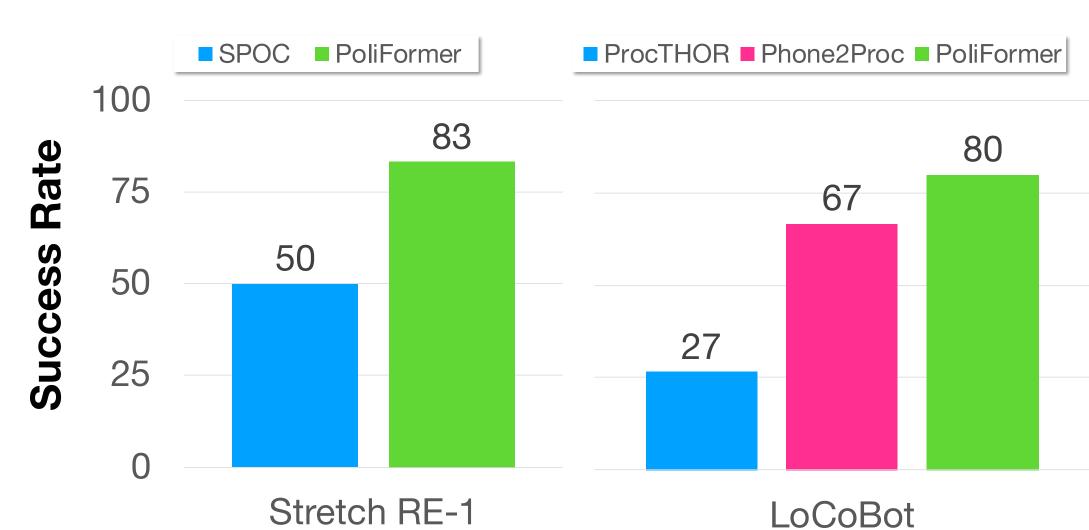
Vision Transformer Model and Transformer State Encoder

Causal Transformer Decoder with KV-Cache

Quantitative Results

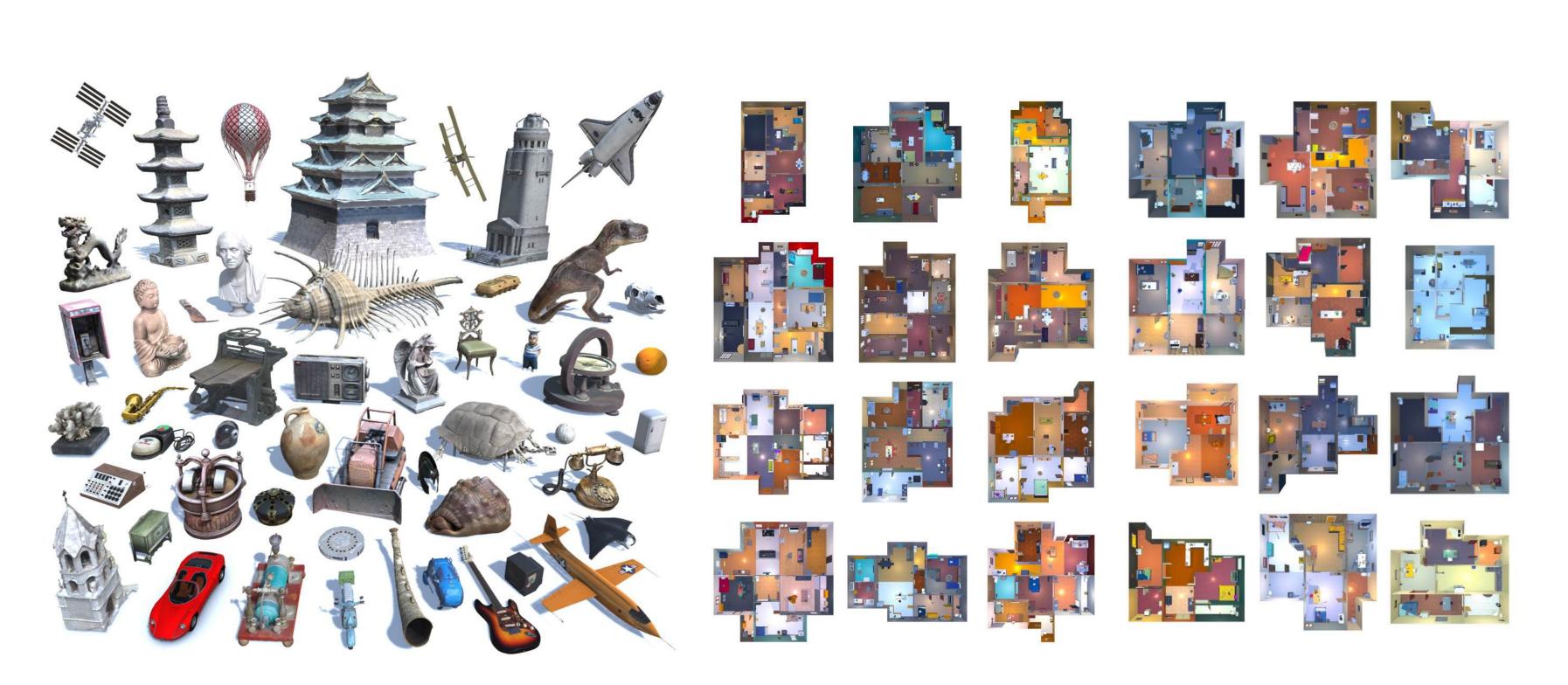


Real-World (zero-shot)



SoTA results in simulation and the real-world, across multiple embodiments.

Scale in Diverse Environment Interactions



Hundreds of parallel rollouts and large batch sizes lead to high training throughput and allow us to train for **700M** training interactions.

150k ProcTHOR houses with ~40k Objaverse assets.