The Data Pyramid for Building Generalist Agents

Yuke Zhu

December 8, 2022

UT Robot Perception & Learning Lab

TEXAS Robotics

TEXAS
The University of Texas at Austin

Building AI for Generalist Agents

household robots

virtual agents

self-driving cars

Open-Ended Objectives

Massively Multitask

World Knowledge

Building AI for Generalist Agents

What data sources shall we use?

How can we best use them?

househo

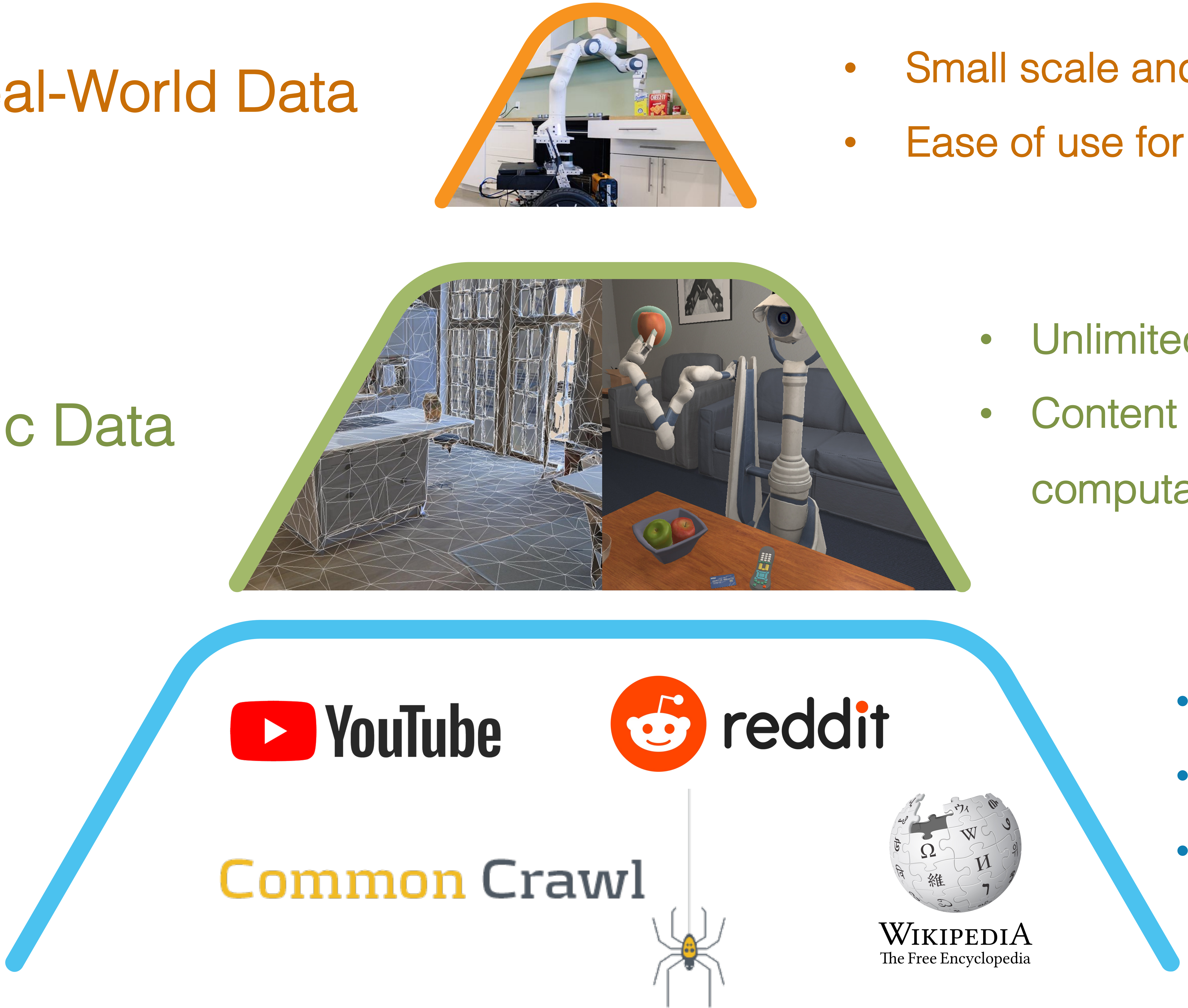
g cars

Open-Ended Objectives

Massively Multitask

World Knowledge

The Data Pyramid for Generalist Agents



Real-World Data

- Small scale and expensive to collect
- Ease of use for agent learning, direct transfer

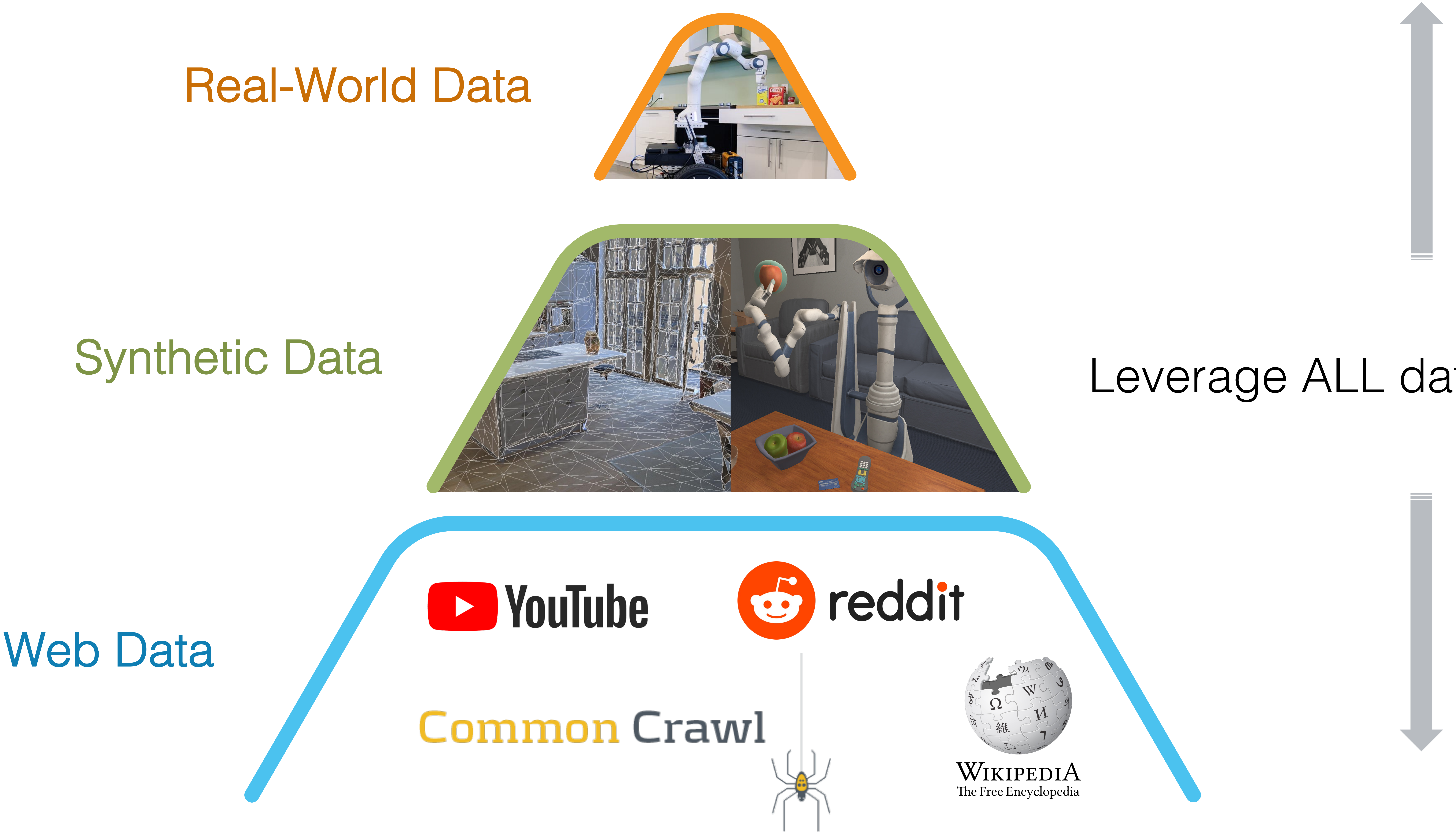
Synthetic Data

- Unlimited simulated data (in theory)
- Content creation challenge, reality gap, computational burden

Web Data

- Massive scale and ever-growing
- Multimodal and unstructured
- Human-centered data

The Data Pyramid for Generalist Agents



Leverage ALL data available!

MineDojo: Framework for Building Generalist Agents

Open-ended Environment | **Generalist Agent** | **Internet-scale Knowledge Base**

“MineDojo: Building Open-Ended Embodied Agents with Internet-Scale Knowledge.” Fan, Linxi et al. NeurIPS 2022

MineDojo: Framework for Building Generalist Agents

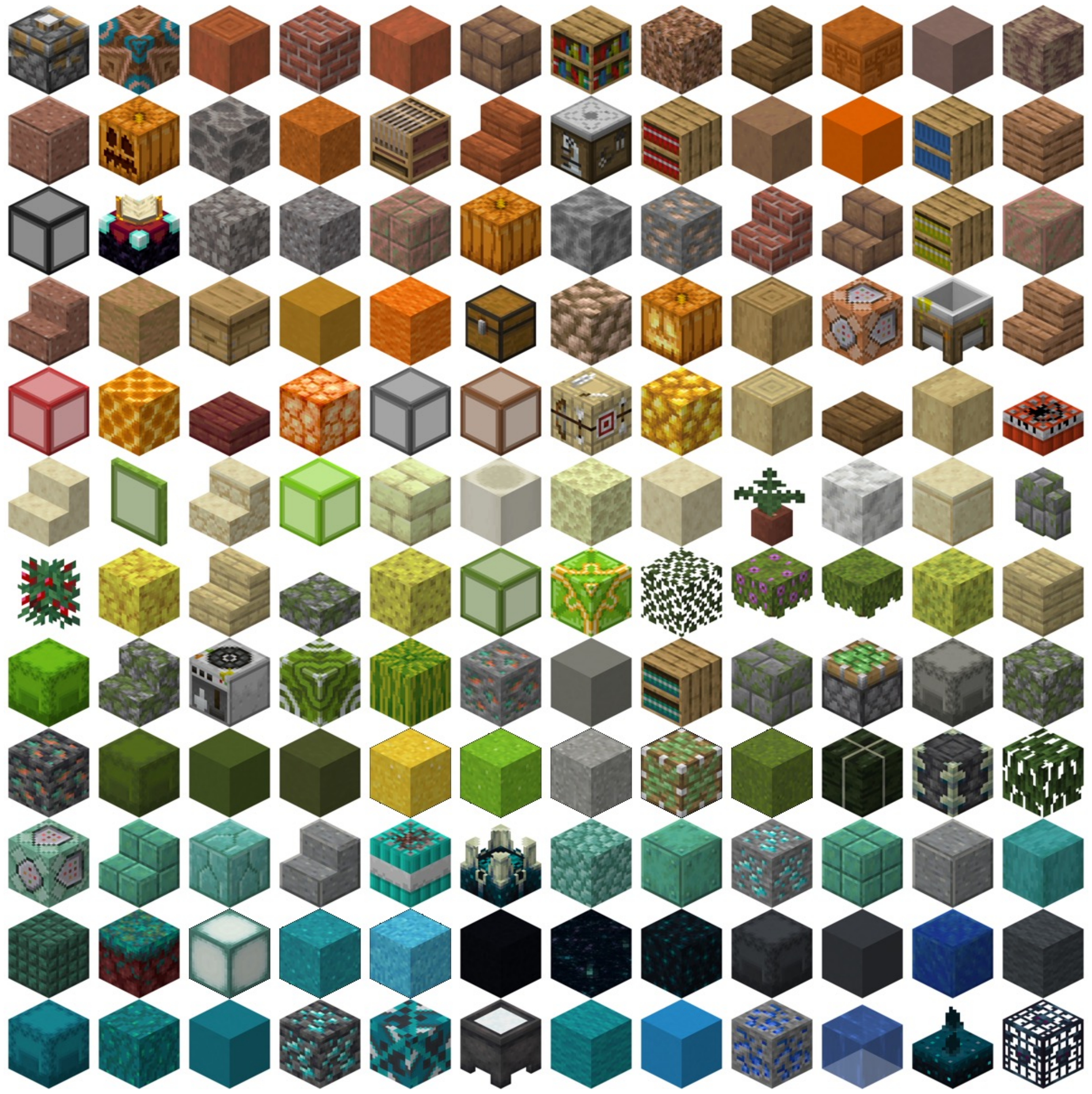
Diverse open-ended environment

- Complex 3D world
- 1500+ programmatic tasks
  - ➢ Success conditions & reward functions are well-defined
  - ➢ Templated language prompts
- 1500+ creative tasks
  - ➢ No well-defined or easily automated success criterion
  - ➢ Reward function is very difficult to formulate

MineDojo: Framework for Building Generalist Agents

Internet-scale knowledge base

YouTube



Wiki

Reddit

"MineDojo: Building Open-Ended Embodied Agents with Internet-Scale Knowledge." Fan, Linxi et al. NeurIPS 2022


MINEDOJO YouTube
730K videos
300K hours
2.2B words

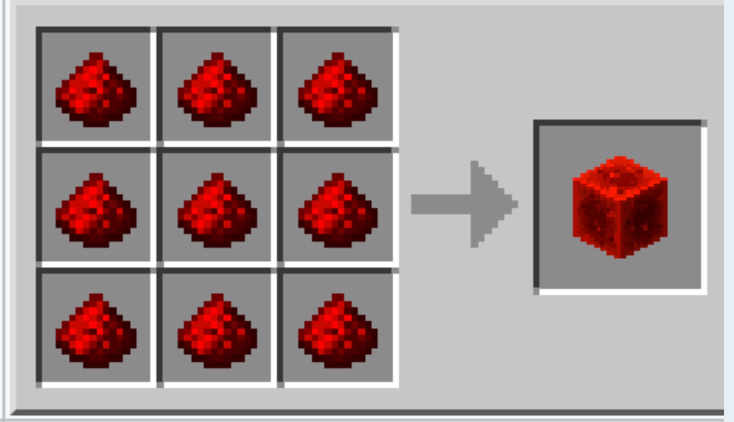
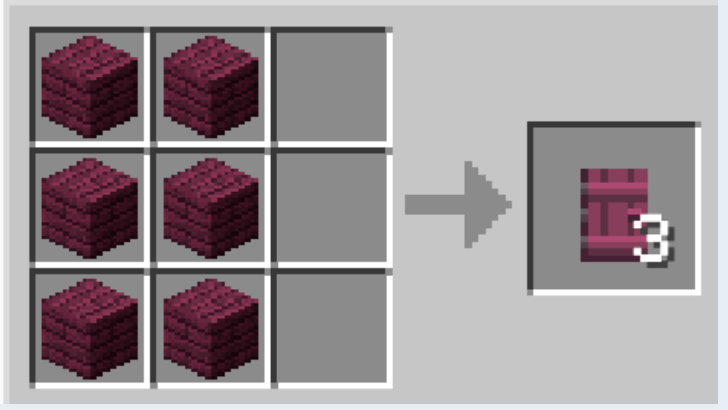
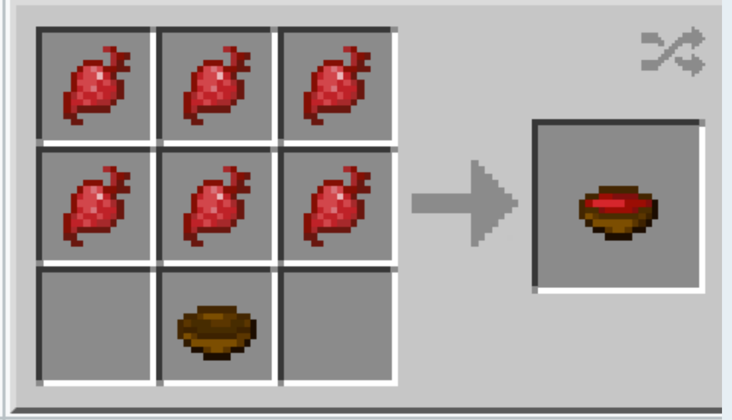
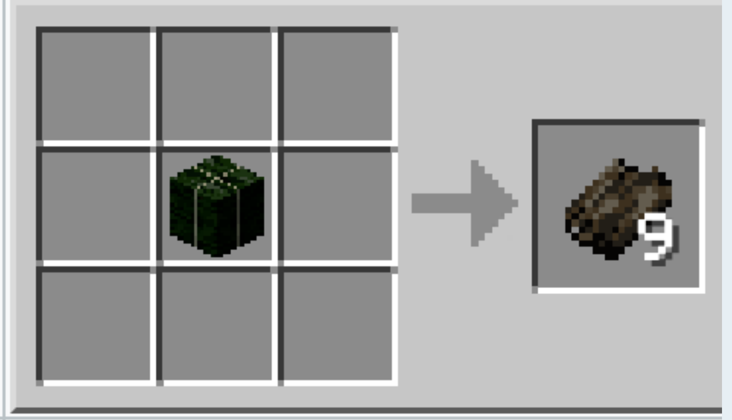
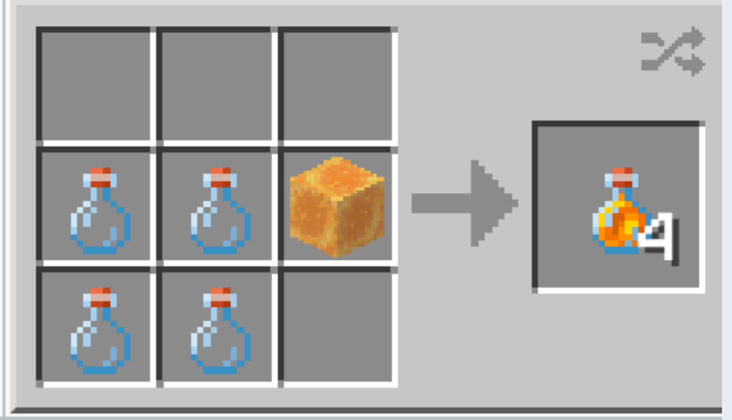
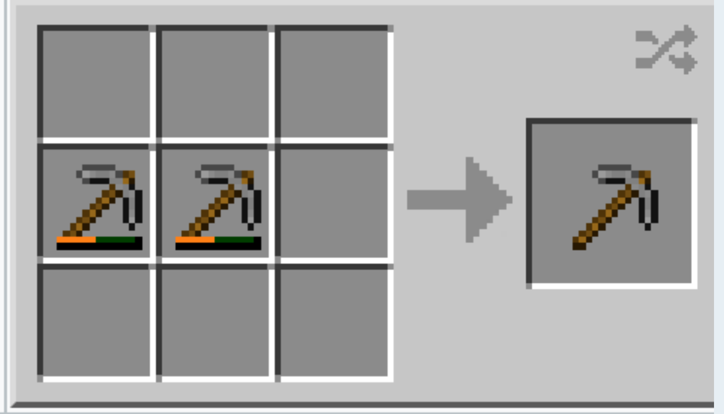

MINEDOJO Wiki
~7K multimodal pages:
texts, images, tables, diagrams

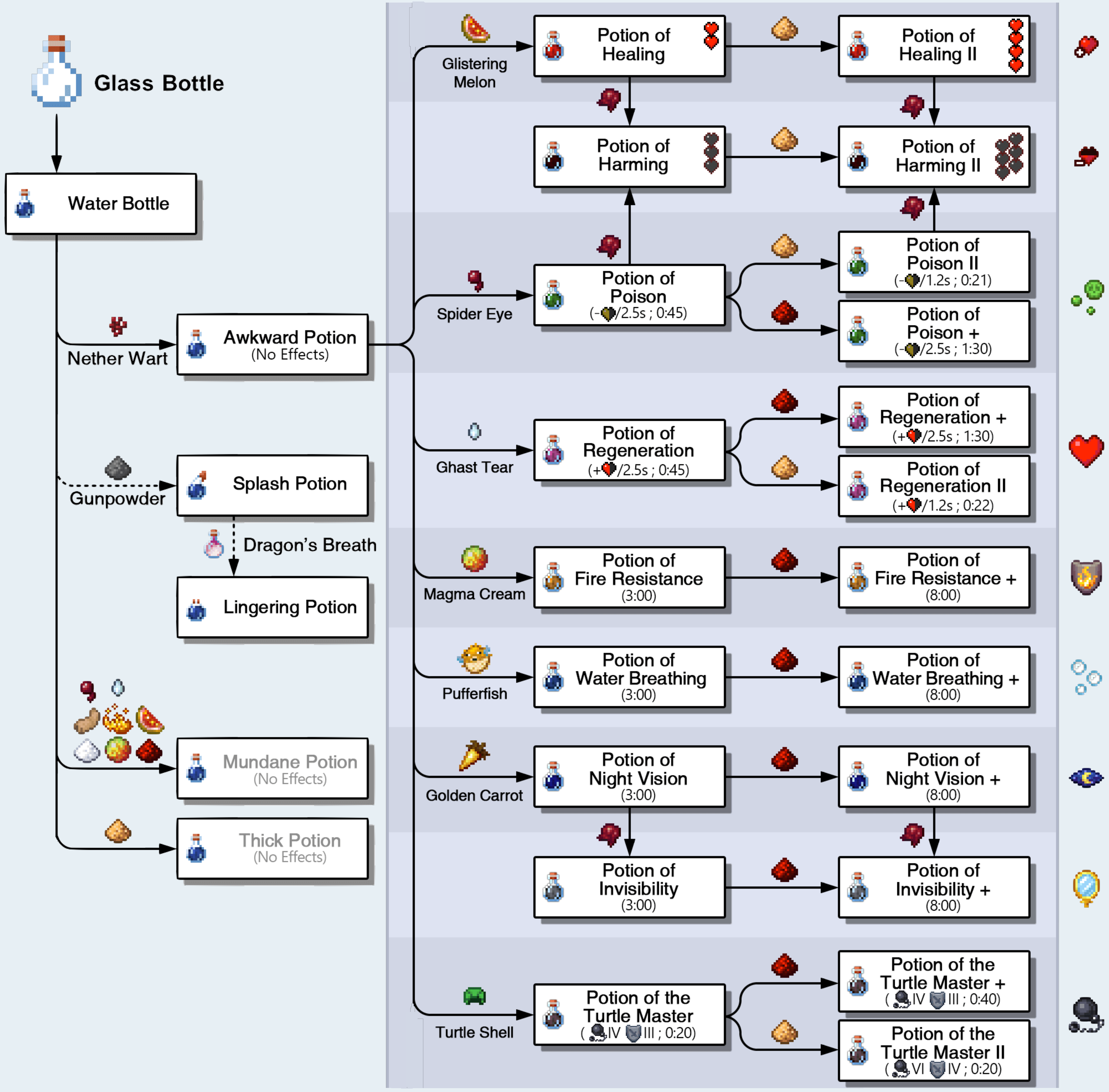
Hostile mobs

Passive mobs

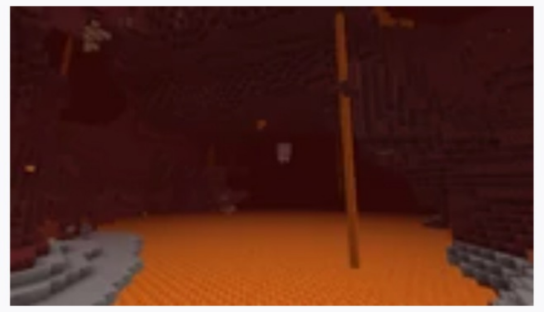
Neutral mobs

Nether Mobs

Name	Ingredients	Crafting recipe
Dropper	Cobblestone + Redstone Dust	
Daylight Detector	Glass + Nether Quartz + Any wood Slab	
Piston	Any Planks + Cobblestone + Iron Ingot + Redstone Dust	
Sticky Piston	Slimeball + Piston	
Block of Redstone	Redstone Dust	
Wood Door	Matching Planks	
Golden Carrot	Gold Nugget + Carrot	
Cake	Milk Bucket + Sugar + Egg + Wheat	
Beetroot Soup	Beetroot + Bowl	
Dried Kelp	Dried Kelp Block	
Honey Bottle	Glass Bottle + Honey Block	
Rabbit Stew	Cooked Rabbit + Carrot + Baked Potato + Any Mushroom + Bowl	
Pickaxe	Any Planks or Iron Ingot or Gold Ingot or Diamond + Stick	
Stone Pickaxe	Any stone-tier block + Stick	
Stone Pickaxe	Cobblestone or Blackstone or Cobbled Deepslate + Stick	
Pickaxe	Matching Damaged Pickaxes	
Recovery Compass	Echo Shard + Compass	
Spyglass	Amethyst Shard + Copper Ingot	



Product	Ingredient	Exp	Description
Copper Ingot	Copper Ore	0.7	Used to craft various items, including spyglasses, lightning rods, and copper blocks.
Iron Ingot	Iron Ore	0.7	Used to craft various items, including blast furnaces, anvils, iron blocks, iron nuggets, rails, buckets, cauldrons, chains, compasses, crossbows, flint-and-steels, heavy-weighted pressure plates, hoppers, iron trapdoors, minecarts, pistons, shears, shields, iron armor, iron tools, stonecutters and tripwire hooks.
Gold Ingot	Gold Ore	1	Used to craft various items, including netherite ingots, gold blocks, golden apples, gold nuggets, clocks, golden armor, golden tools, powered rails and light-weighted pressure plates. Also used as a currency for bartering.
Diamond	Diamond Ore	1	Used to craft various items, enchanting tables, jukeboxes and diamond blocks. When normally mined drops 1 diamond and 3–7.

Biome name	Features	Description	Screenshot [hide]
Nether Wastes	Netherrack, Glowstone, Soul Sand, Nether Quartz Ore, Ghasts, Blazes, Zombified Piglins, Nether Fortresses, Wither Skeletons, Lava, Magma cubes, Gravel, Magma Blocks, Bastion Remnants, Ruined Portals, Piglins,Nether Gold Ore	**Temperature: 2.0. Rainfall: 0.0.** This is one of the biomes used to generate the Nether. Within this biome mobs such as ghasts, packs of piglins, zombified piglins and the occasional magma cubes and endermen spawn. Certain structures, such as Nether quartz ore and glowstone blobs, and Nether fortresses generate only in the Nether.	 Nether Wastes

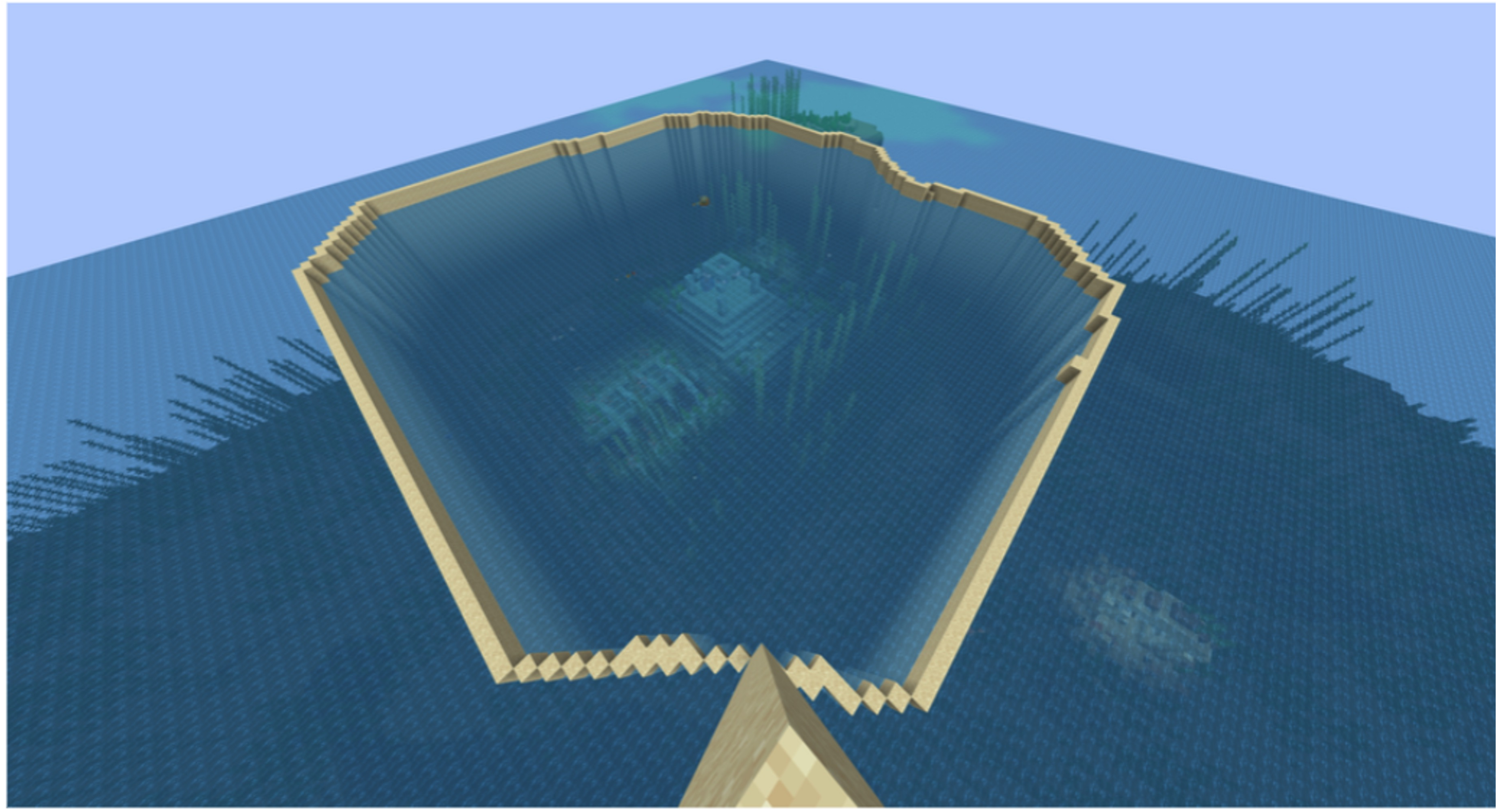

MINECRAFT
MINEDOJO Reddit
340K posts
6.6M comments
560K images
Minecraft on reddit
Join
Posts
Wiki
PINNED BY MODERATORS
Posted by u/kidmania01 3 days ago
Builds
Minecraft Bi-Weekly Build Challenge #152: Fishing Hut
20 Comments
Share
Save
Posted by u/Hugleton 6 hours ago
what do you guys think?
About Community
Minecraft community on reddit.
6.6m
Allay Admirers
5.6k
Copper Button Pushers
Created Jun 11, 2009
r/Minecraft topics
Minecraft
Sandbox game
Open world
Action-adventure game
Gaming
r/Minecraft Rules

55

Posted by u/BlockTrapKing 11 months ago

My first time encasing an Ocean Monument. Help! What is the best way to sponge it!

Help



49 Comments · Award · Share · Save · ...

With sponges

5x5 sections. Sponges only have so much power.

Sponges or a lot of dirt/sand

5x5 grid of sand squars. No joke but it let's you go down in a back and forth without fear of creating more source blocks

685

r/Minecraft · Posted by u/ERROR_173 2 months ago

does anyone know why some of my wheat won't grow?(farming)

Help

121 Comments · Award · Share · Save · ...

Gotta be the light add more torches

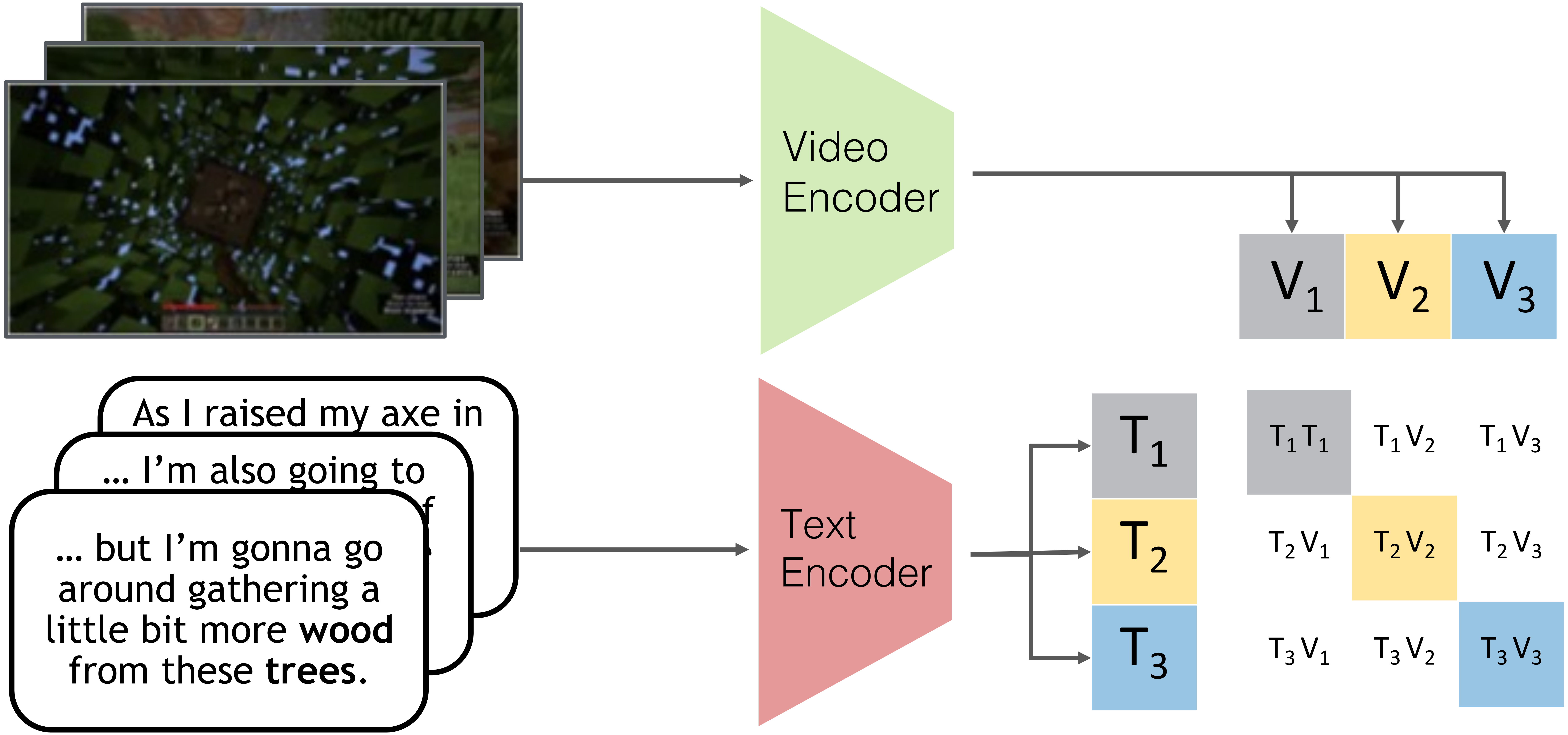
Light

It'll be your light level, as the wheat near the torch has grown. All you need to do is place more torches and you should see a difference.

light lvl too low

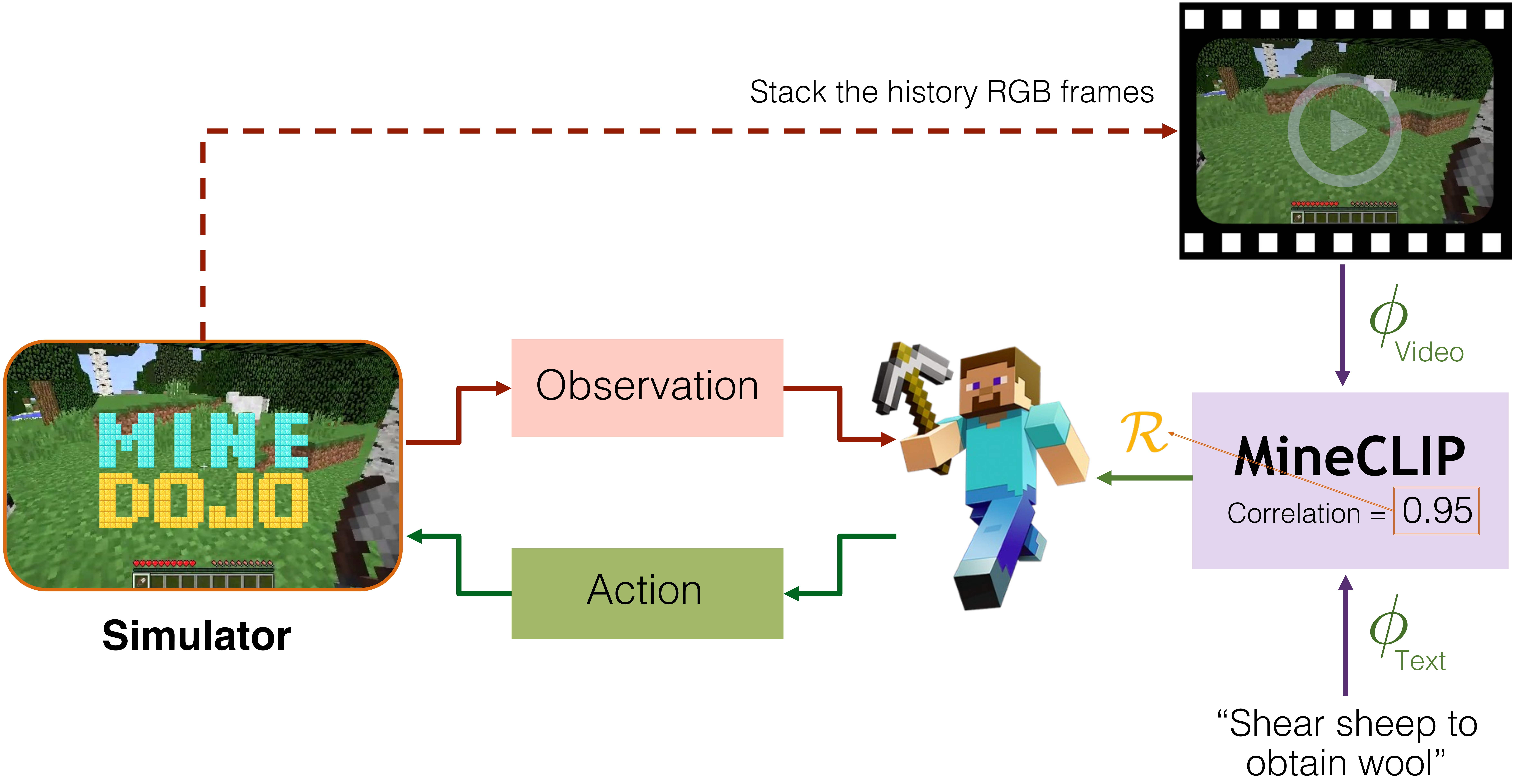
MineDojo: Framework for Building Generalist Agents

MineCLIP: scalable reward learning from YouTube videos



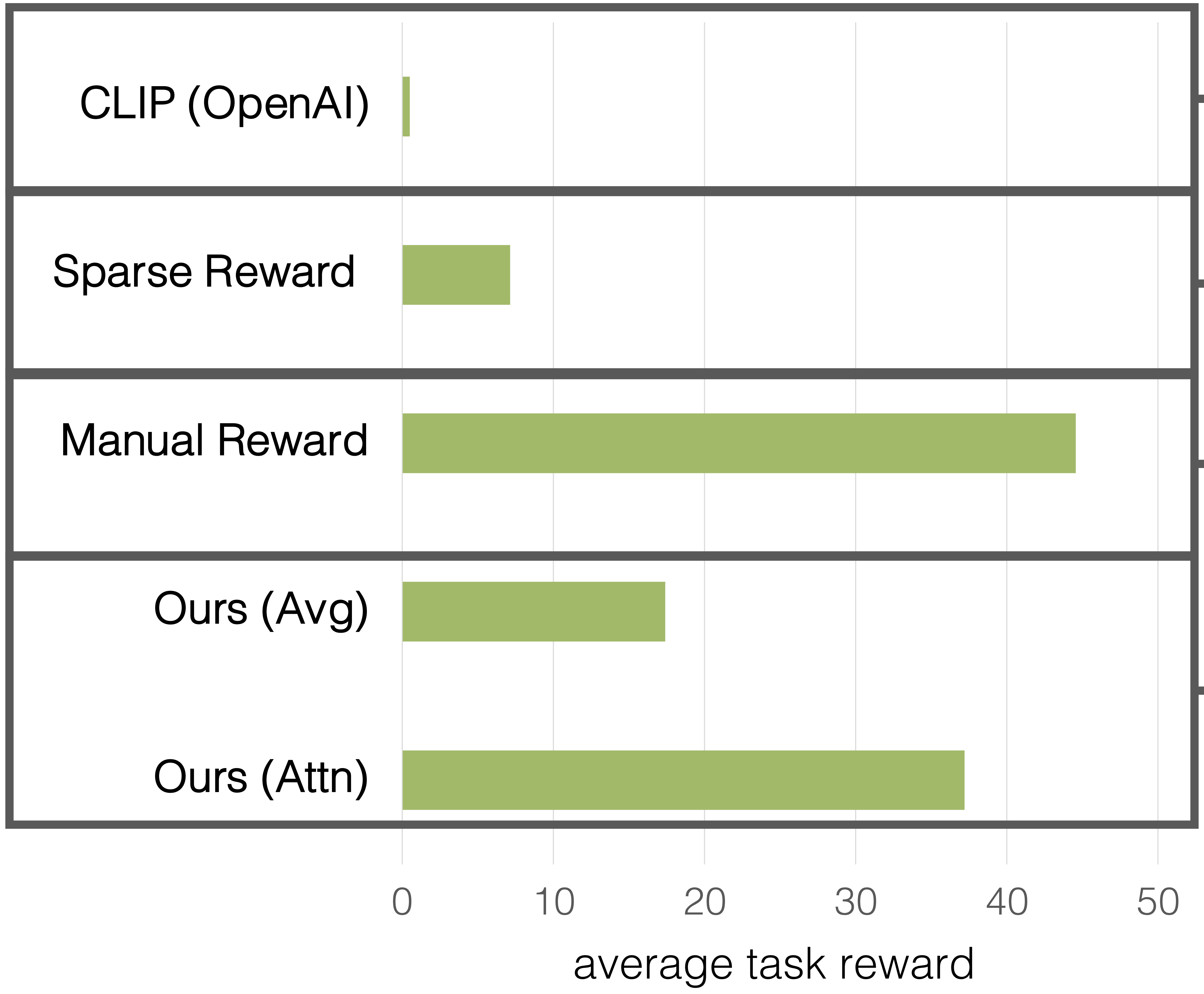
[Xu et al. 2022; Radford et al. 2021]

MineDojo: Framework for Building Generalist Agents




Milk a cow
Hunt a cow
Shear a sheep
Hunt a sheep
Combat a spider
Combat a zombie
Combat a pigman
Combat an enderman
Find a Nether Portal
Find an ocean
Dig a hole
Lay carpet

MineDojo: Framework for Building Generalist Agents



Pre-trained CLIP model from OpenAI fails to generalize.

Binary task-completion rewards lead to exploration challenges.

Dense manually curated rewards show oracle performances.

Two variants of MineCLIP models learned from YouTube videos

Towards Human-Level Generalist Minecraft Agents

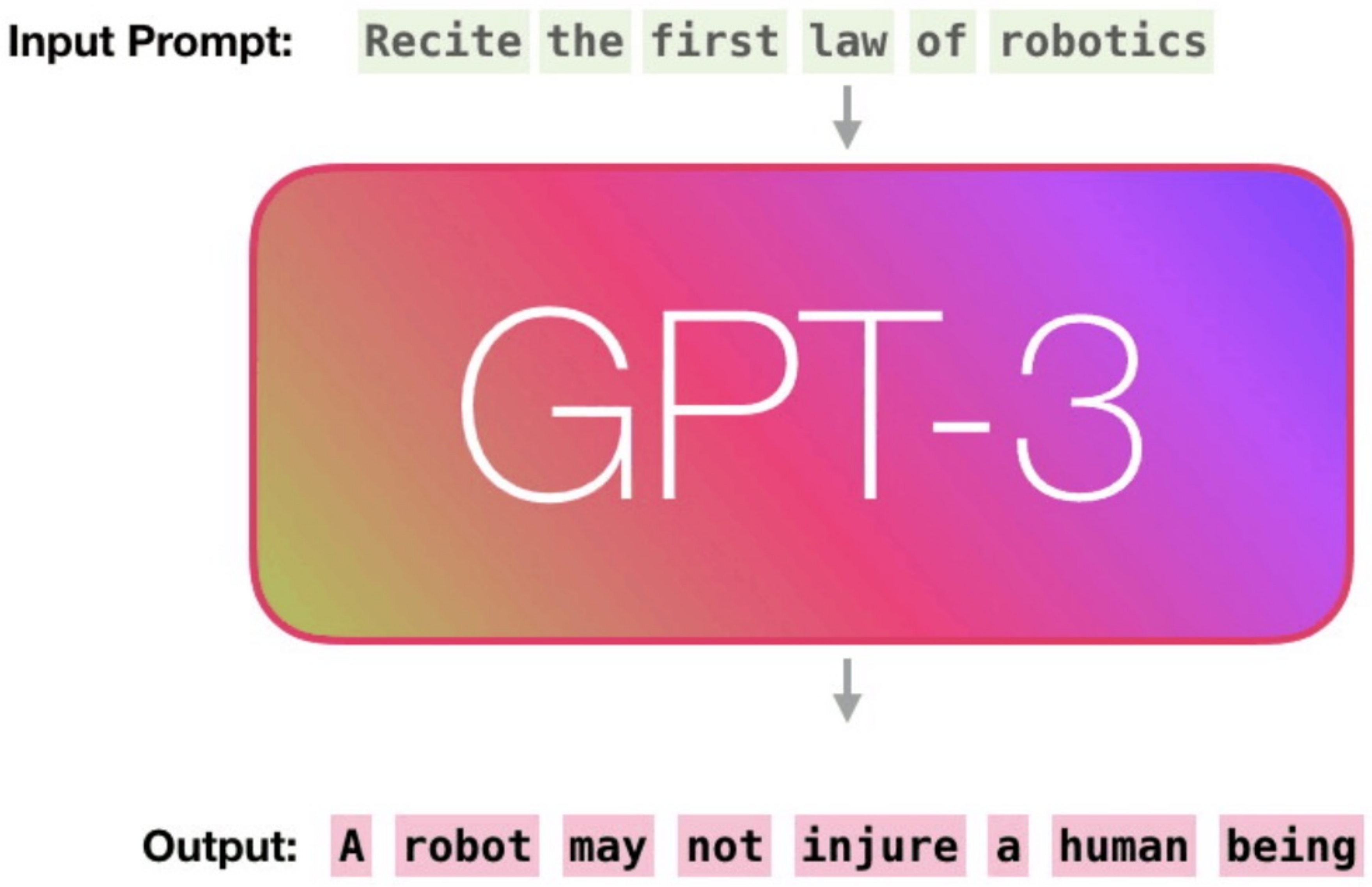
Decorate winter wonderland

Simulate a tiny CPU

Recreate Hogwarts Castle

Fight Ender Dragon

LID: Interactive Task Planning with Large Language Models

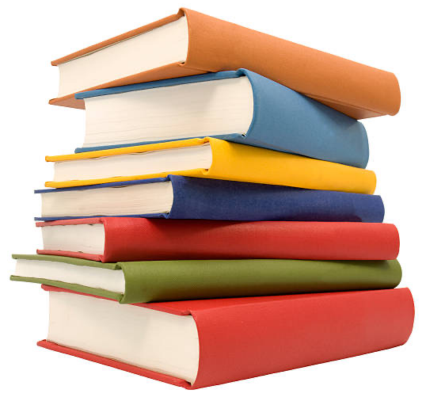


[Credit: Jay Alammar]

Trained on 300 billion tokens of text

Common Crawl

WIKIPEDIA
The Free Encyclopedia



Learning objective: predict the next word

Application Domains:

text summarization, visual question answering, conversational AI, …

“Pre-Trained Language Models for Interactive Decision-Making.” Li, Shuang et al. NeurIPS 2022

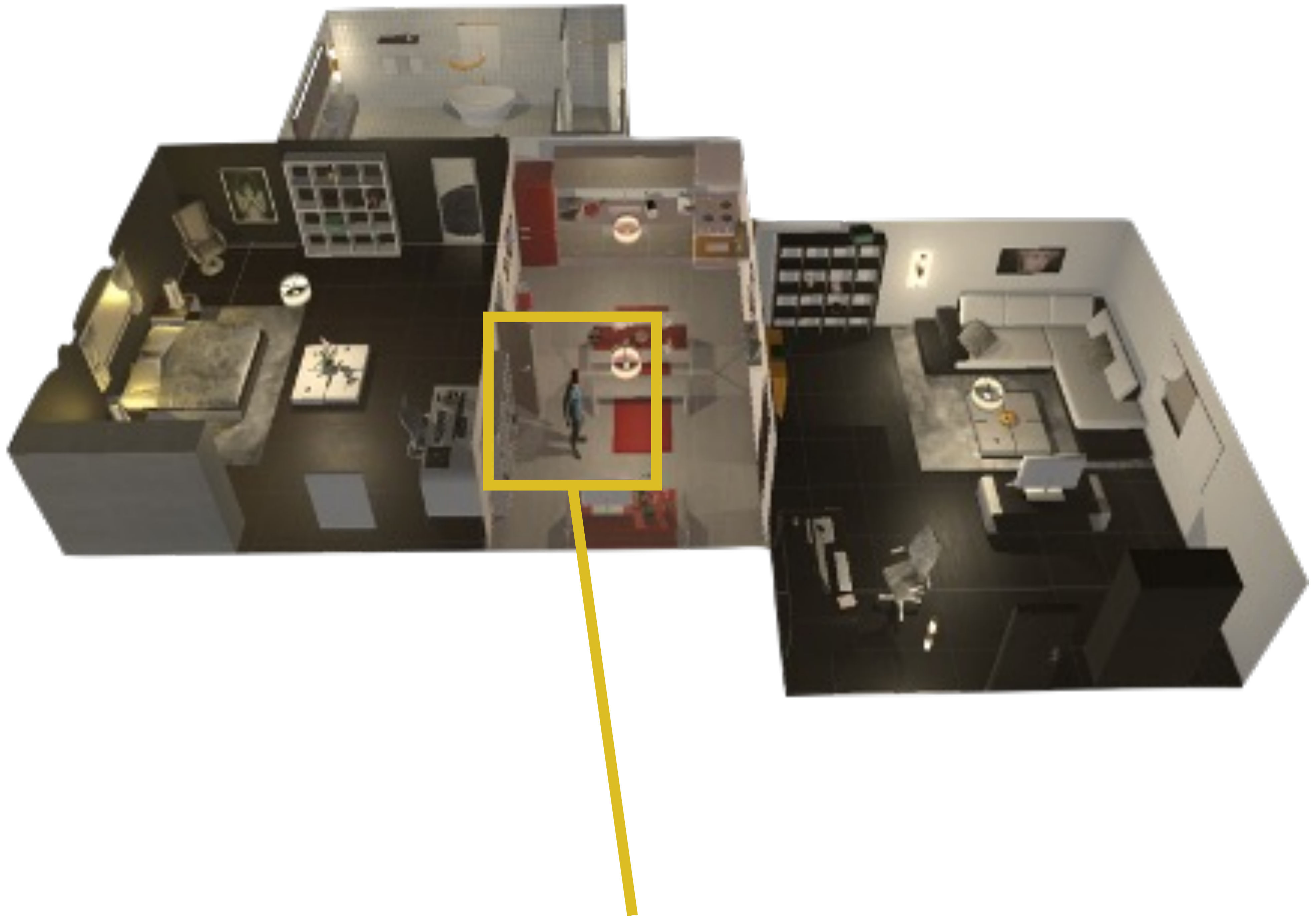
LID: Interactive Task Planning with Large Language Models

Observation LLM as Policy Task Plan

Can we harness the commonsense knowledge in large language models for high-level task planning?

“Pre-Trained Language Models for Interactive Decision-Making.” Li, Shuang et al. NeurIPS 2022

LID: Interactive Task Planning with Large Language Models

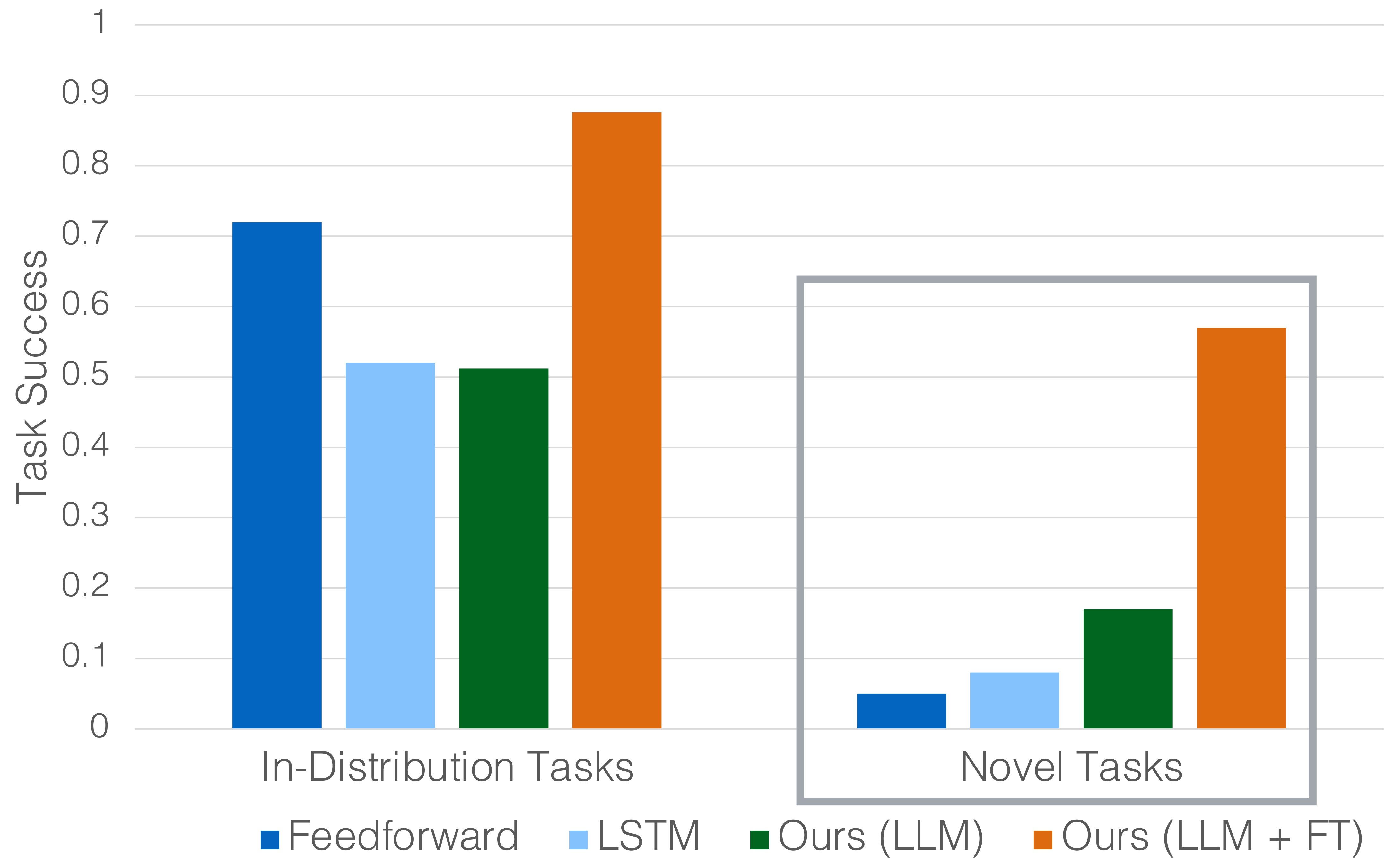


“Pre-Trained Language Models for Interactive Decision-Making.” Li, Shuang et al. NeurIPS 2022

LID: Interactive Task Planning with Large Language Models

"Pre-Trained Language Models for Interactive Decision-Making." Li, Shuang et al. NeurIPS 2022

LID: Interactive Task Planning with Large Language Models



LID-Text (Ours)
on the **In-Distribution** setting

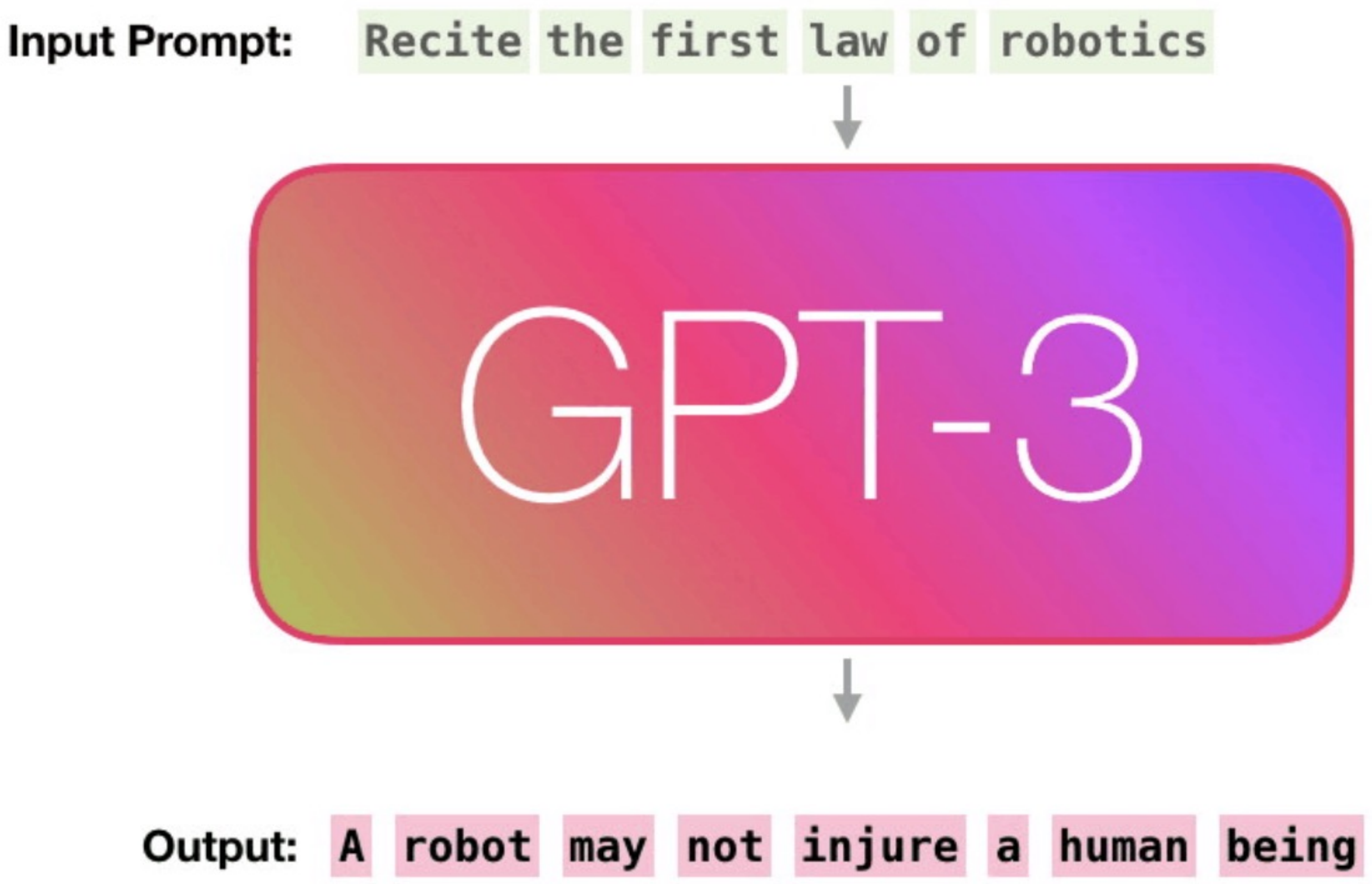
LID-Text (Ours)
on the **Novel Tasks** setting

Speed×2

Significant improvement in combinatorial generalization to novel tasks

"Pre-Trained Language Models for Interactive Decision-Making." Li, Shuang et al. NeurIPS 2022

VIMA: General Robot Manipulation with Multimodal Prompts



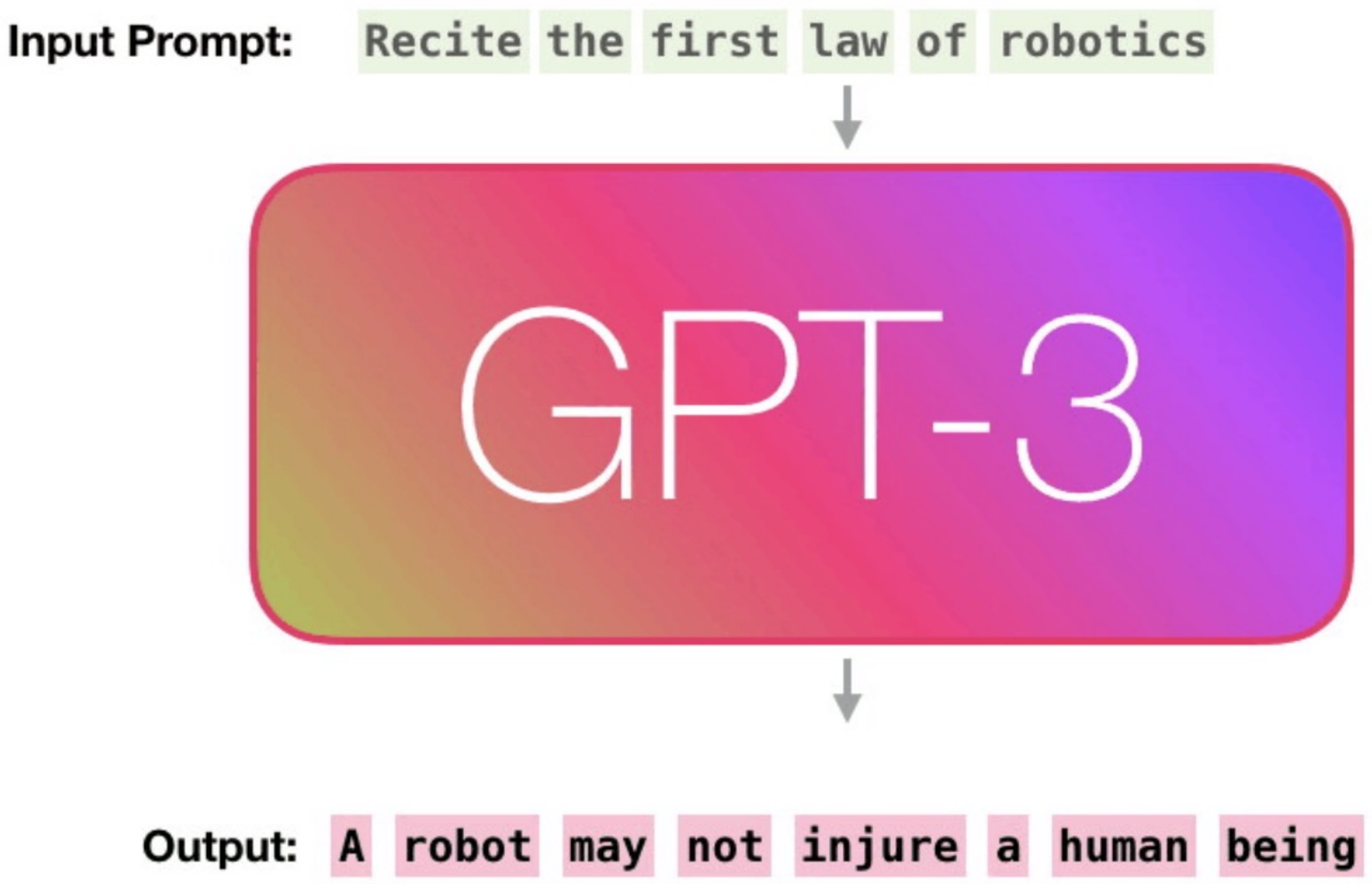
[Credit: Jay Alammar]

Input Prompt:
How to "find material and craft a gold pickaxe" in Minecraft? Let's think step by step.

GPT-3:

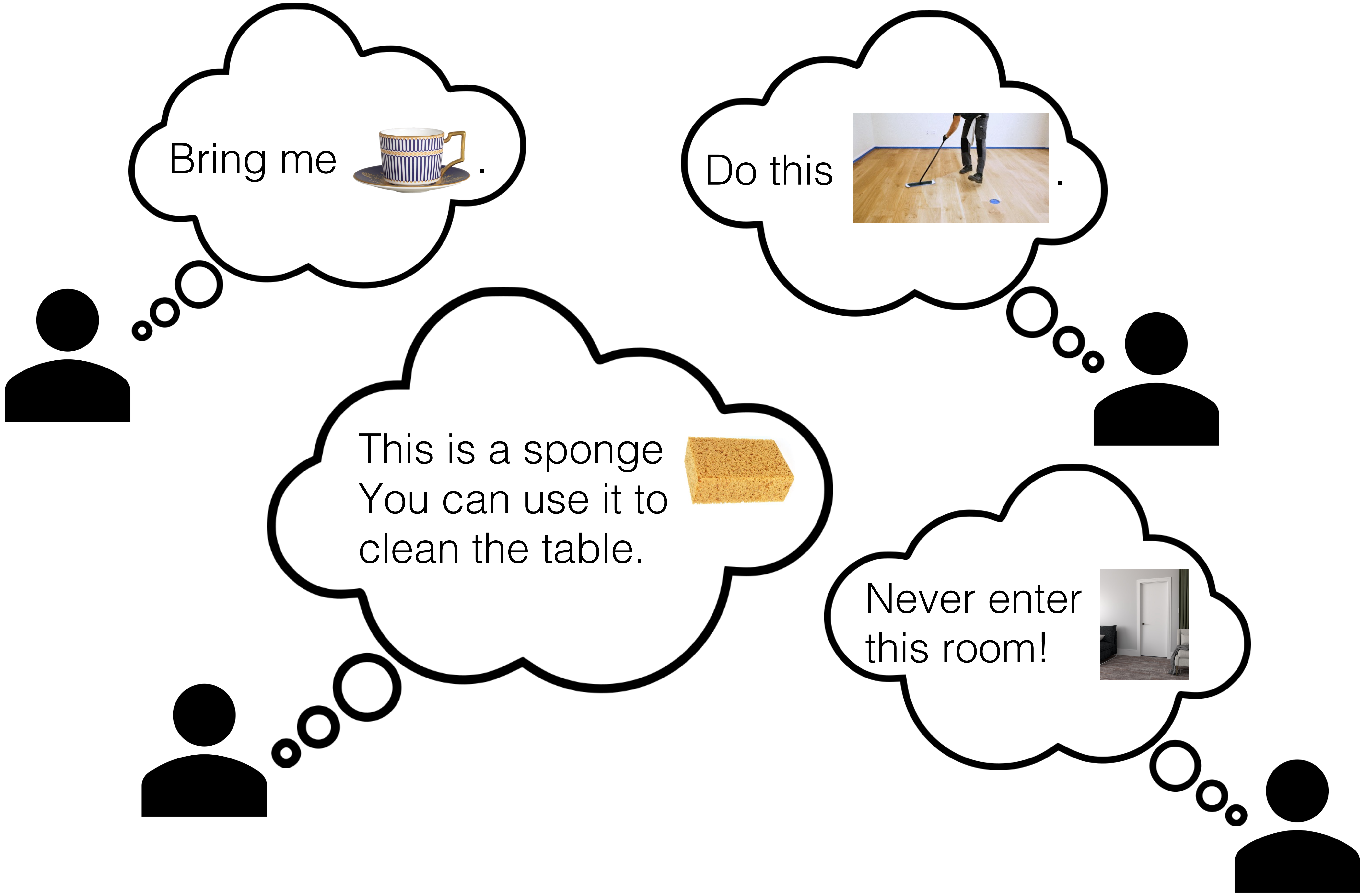
1) Find a place with a lot of trees;
2) Cut down the trees and gather the wood;
3) Find a place with a lot of stone;
4) Mine the stone and gather the cobblestone;
5) Find a place with a lot of iron;
6) Mine the iron and gather the iron ingots;
7) Find a place with a lot of gold;
8) Mine the gold and gather the gold ingots;
9) Craft a gold pickaxe.

VIMA: General Robot Manipulation with Multimodal Prompts

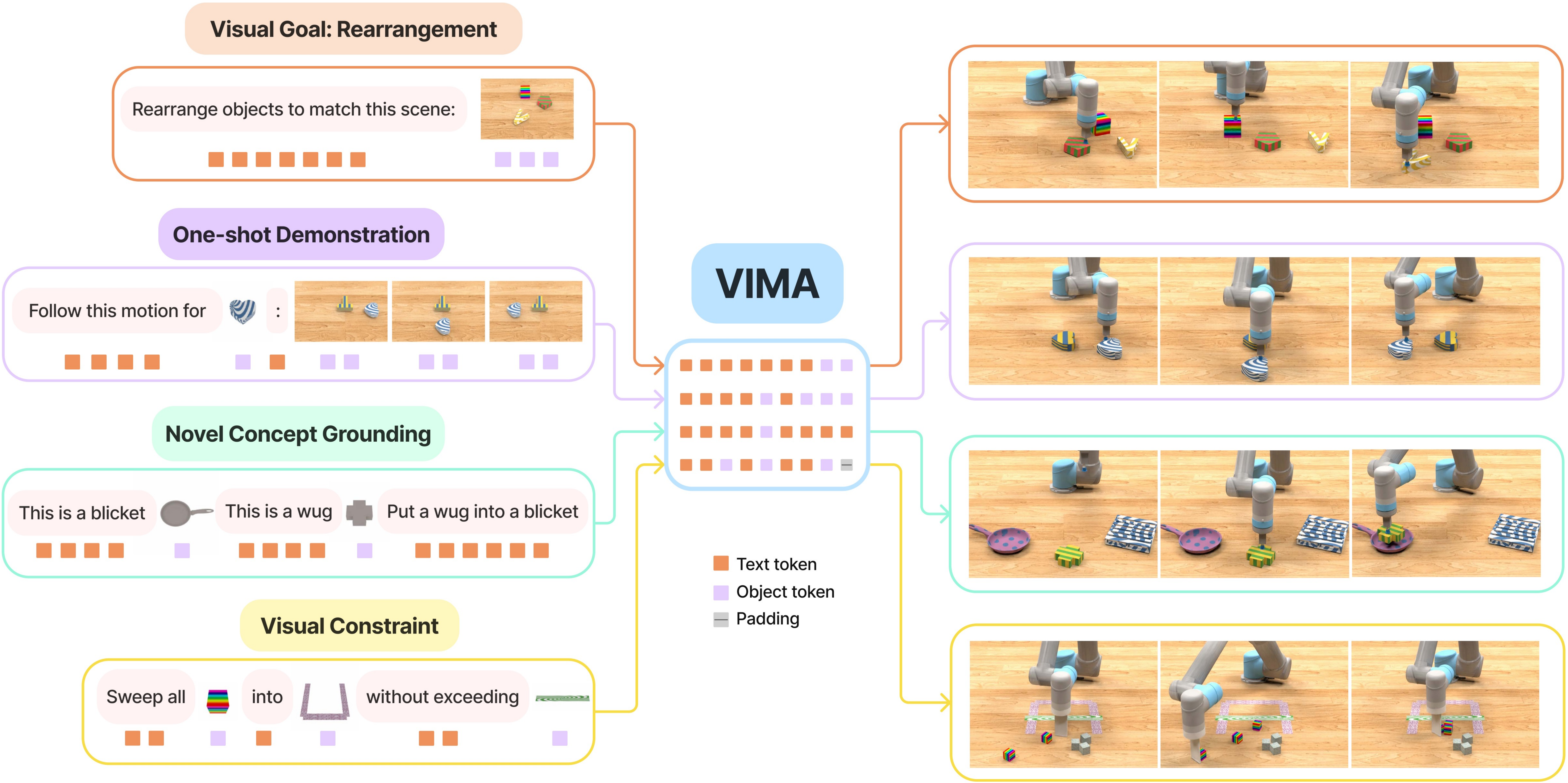


[Credit: Jay Alammar]

What if we can prompt a household robot to …



VIMA: General Robot Manipulation with Multimodal Prompts



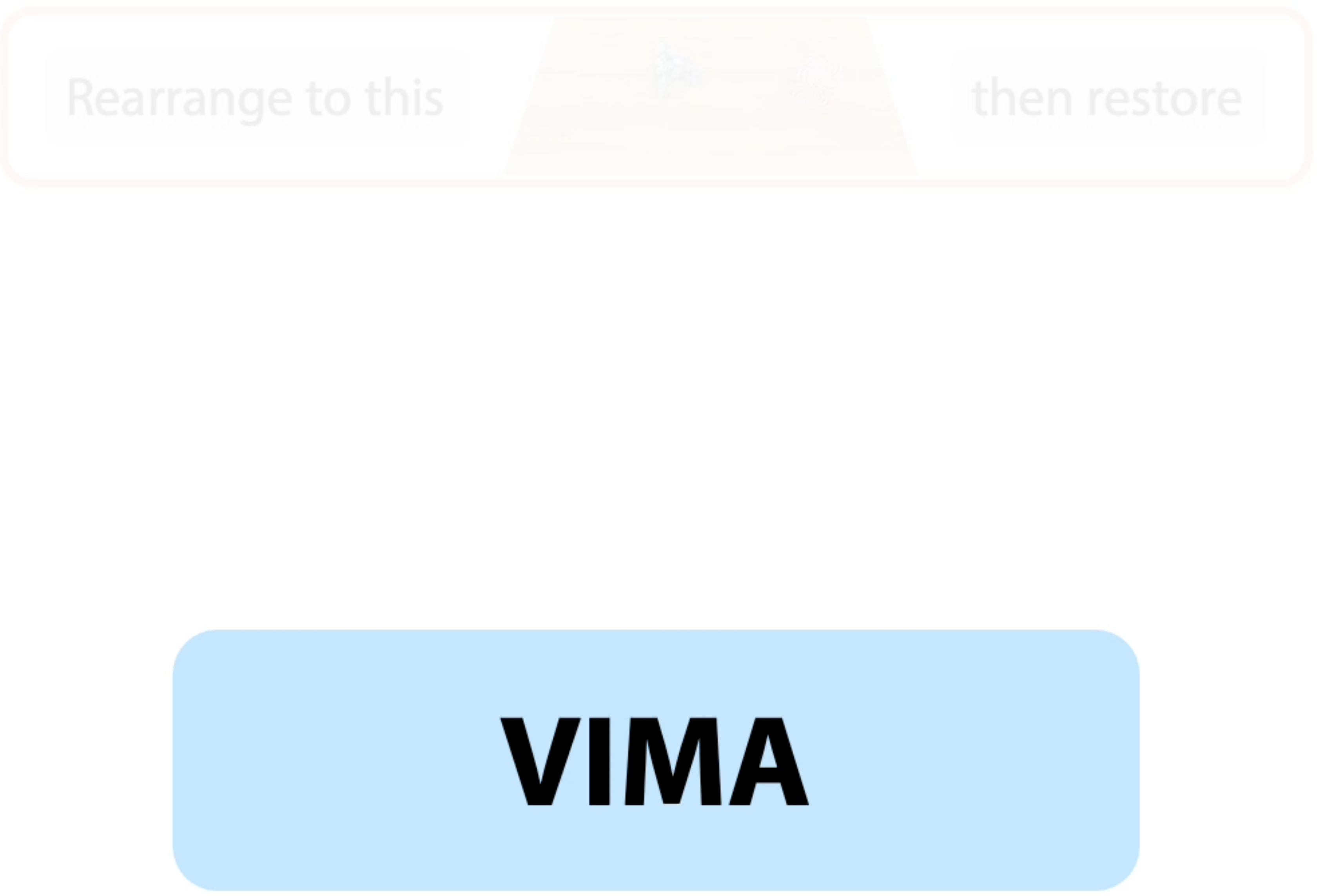
generalist robot agent for multi-task learning and zero-shot generalization

"VIMA: General Robot Manipulation with Multimodal Prompts." Jiang, Yunfan et al. ArXiv 2022

VIMA: General Robot Manipulation with Multimodal Prompts

VIMA: Visuo-Motor Attention model

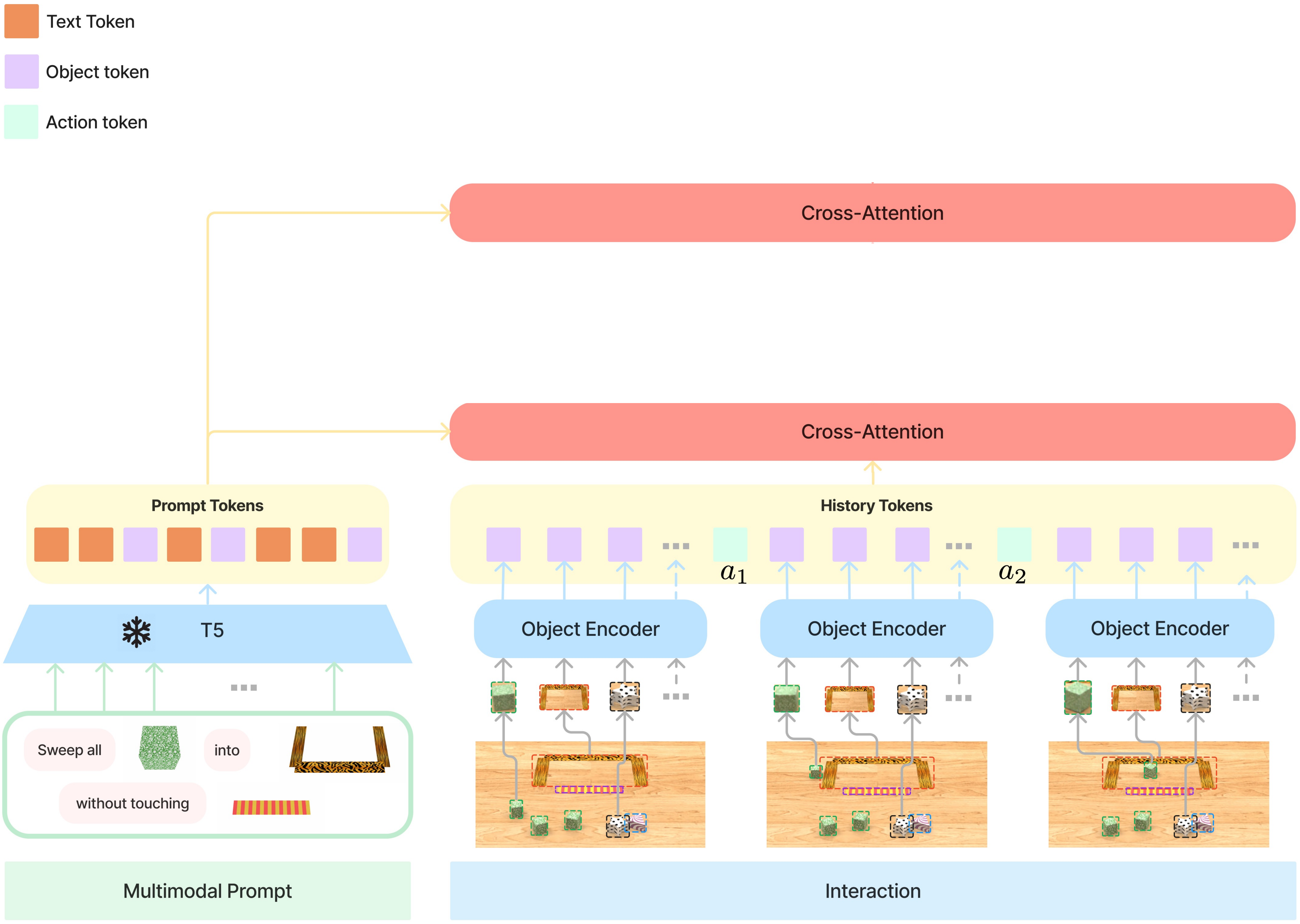
- Transformer encoder-decoder;
- Encode multimodal prompts with a frozen LM (Google T5);
- Decode robot arm actions given the prompt and interaction history.



“VIMA: General Robot Manipulation with Multimodal Prompts.” Jiang, Yunfan et al. ArXiv 2022

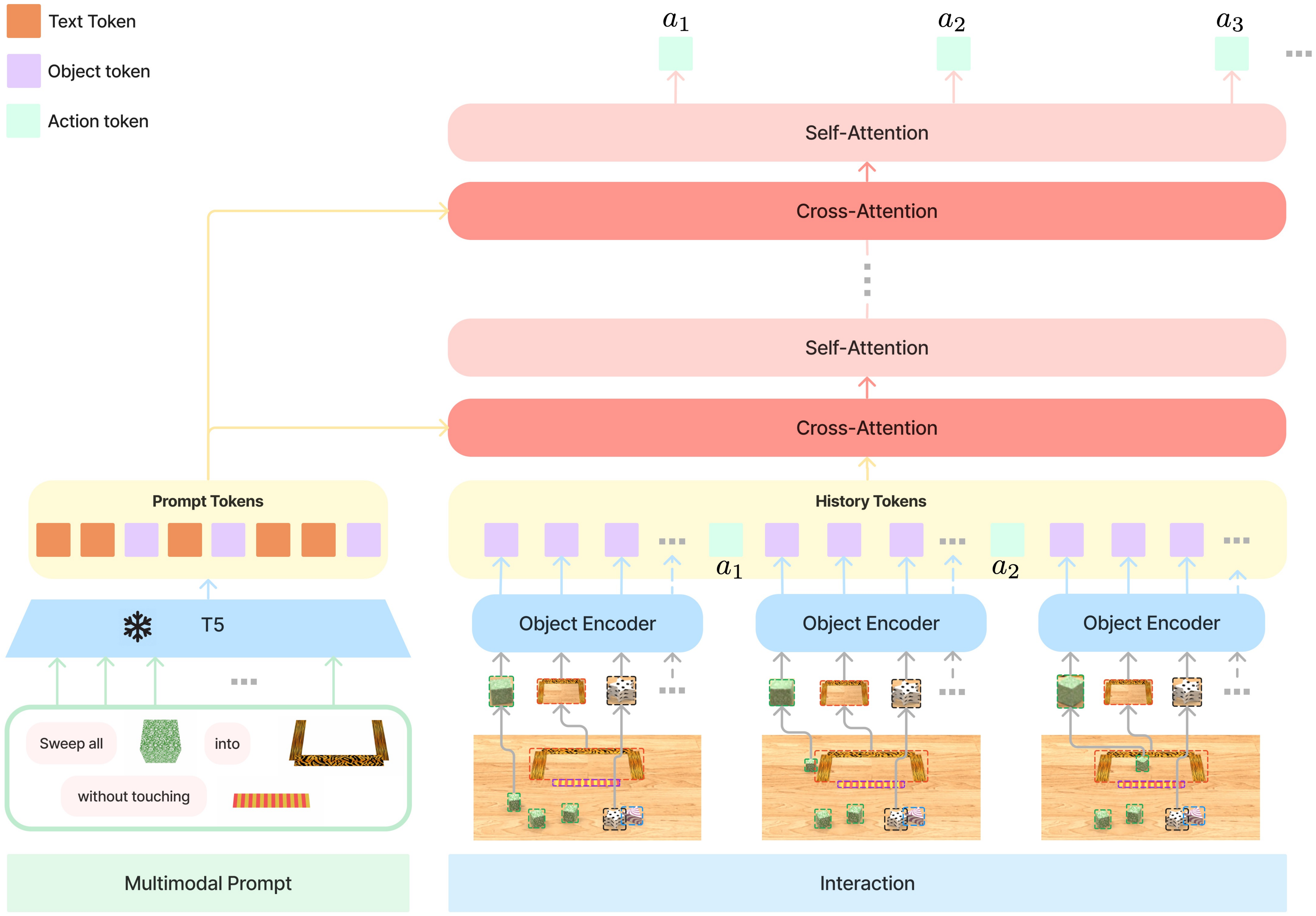
VIMA: General Robot Manipulation with Multimodal Prompts

- Cross-attention to condition history on prompt;



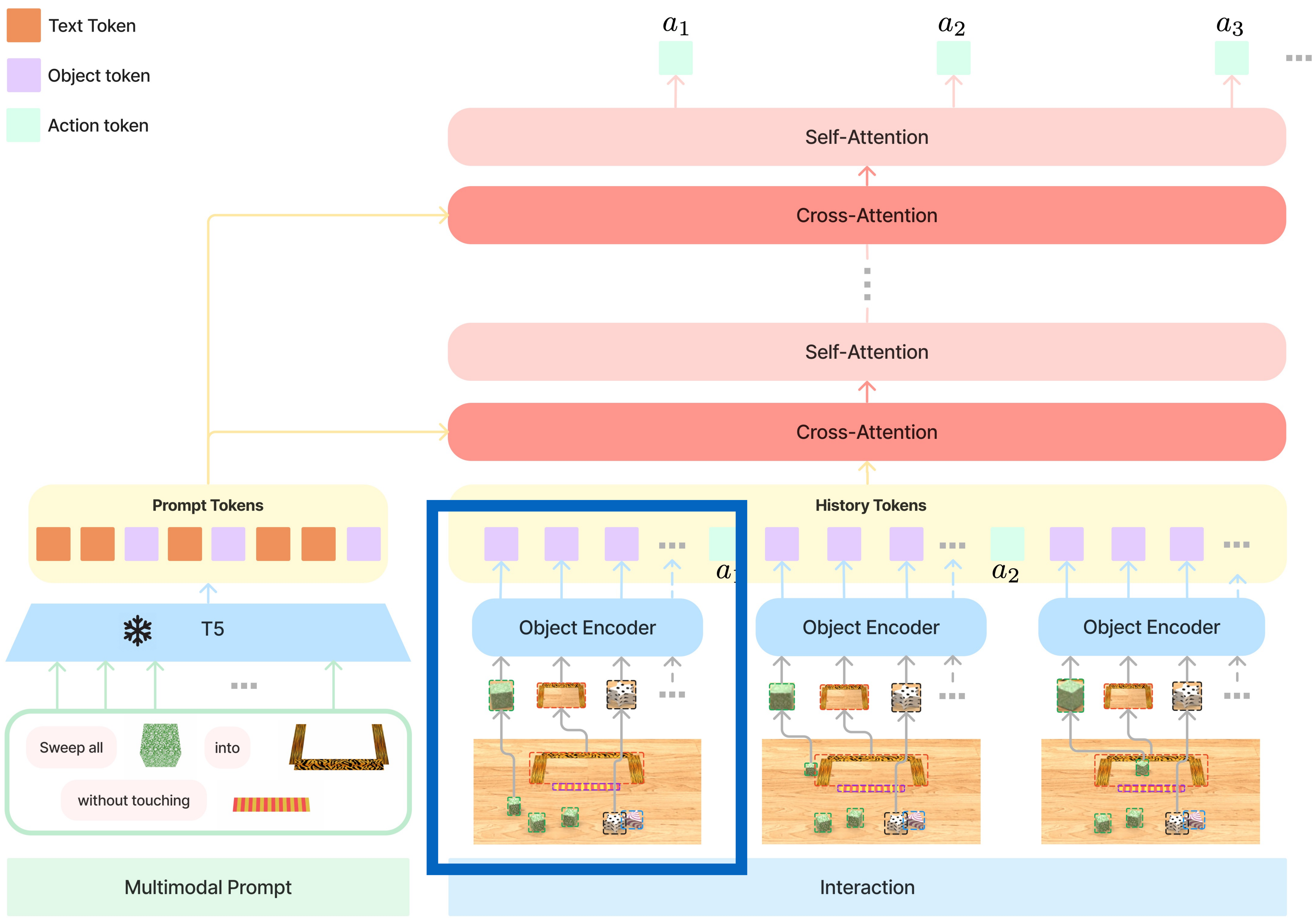
VIMA: General Robot Manipulation with Multimodal Prompts

- Cross-attention to condition history on prompt;
- Alternate cross-attention and causal self-attention to decode actions;



VIMA: General Robot Manipulation with Multimodal Prompts

- Cross-attention to condition history on prompt;
- Alternate cross-attention and causal self-attention to decode actions;
- Objects as tokens.



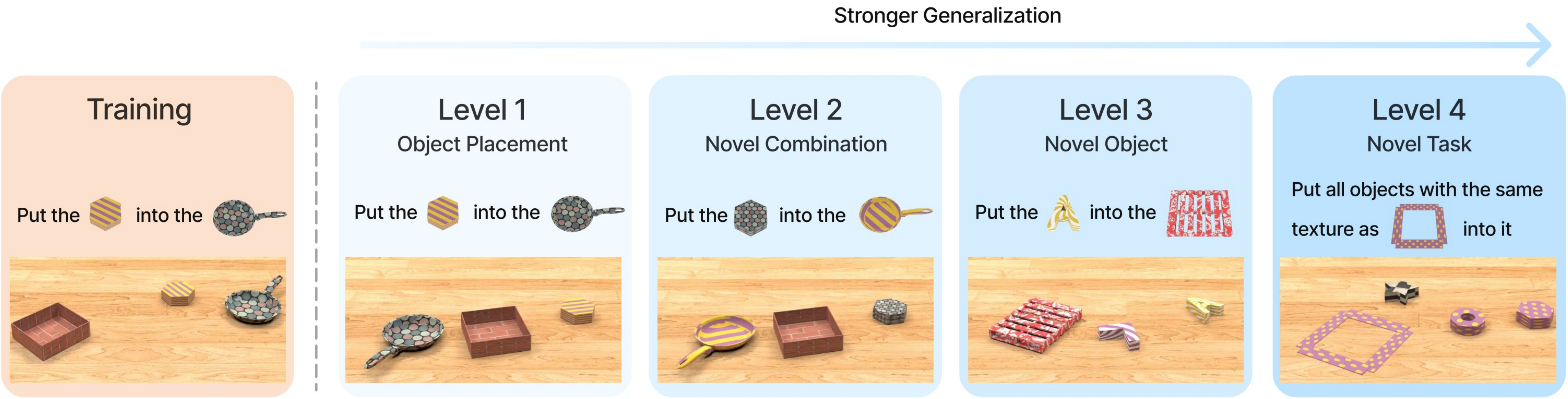
VIMA: General Robot Manipulation with Multimodal Prompts

Open-source large-scale simulation benchmark

- 6 task categories unified by the sequence model;
- 1000s of procedurally generated tasks, paired with multimodal prompts;
- Scripted oracles to generate expert demonstrations.

VIMA: General Robot Manipulation with Multimodal Prompts

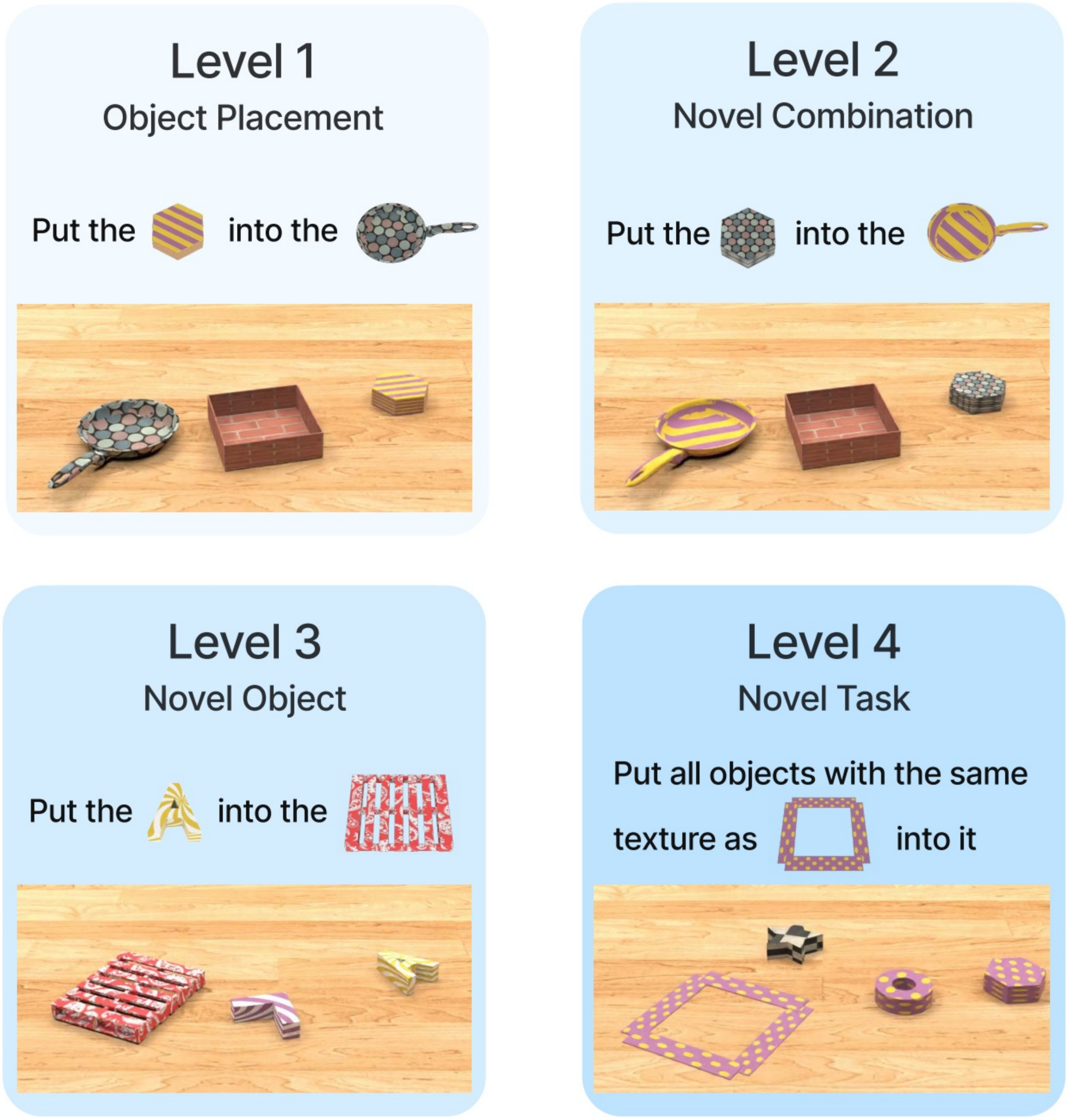
Four-level evaluation protocol for systematic generalization

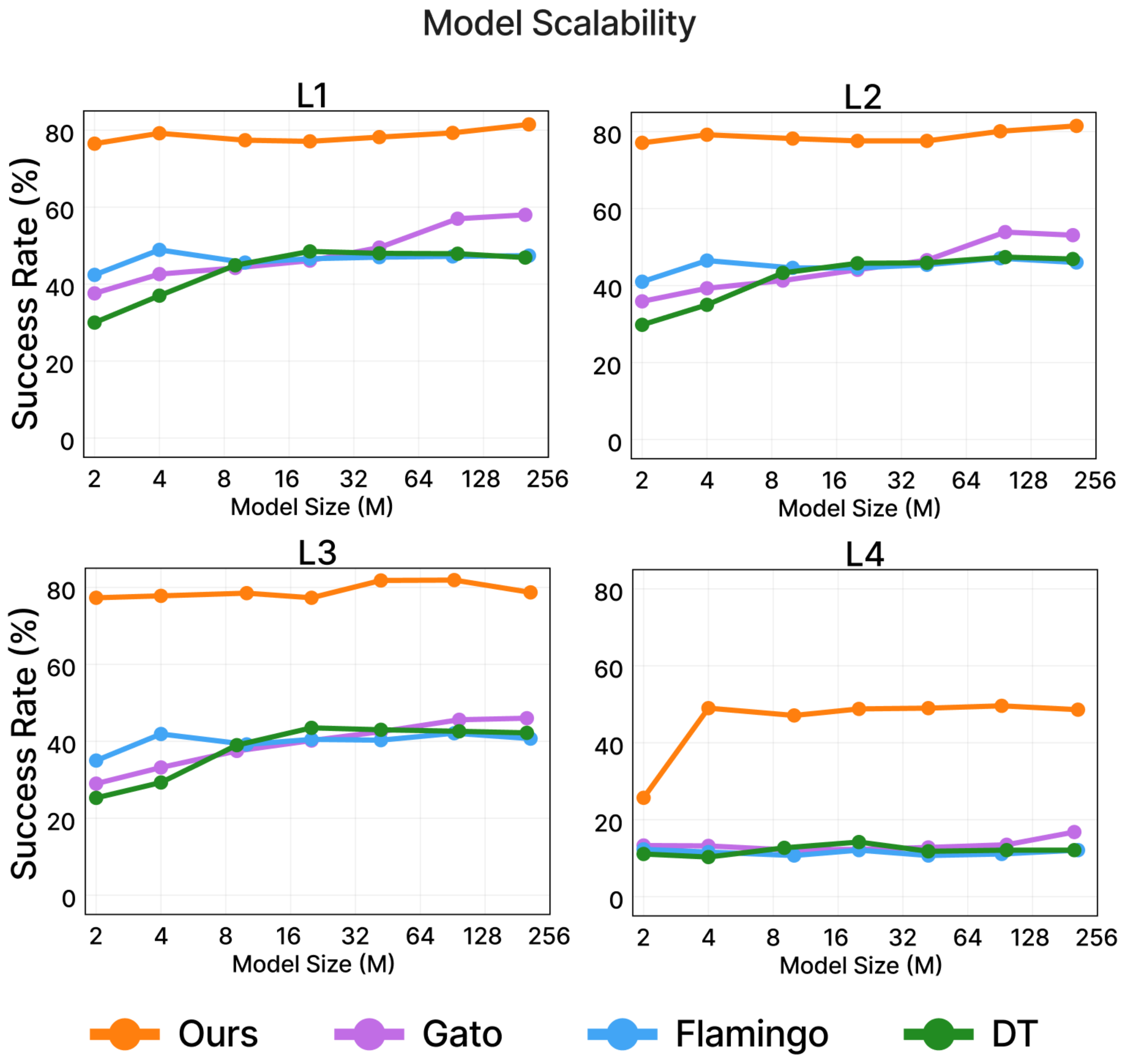


- systematically measuring the zero-shot generalization capability of trained agents.
- Each level deviates more from the training distribution and is strictly more challenging.

VIMA: General Robot Manipulation with Multimodal Prompts

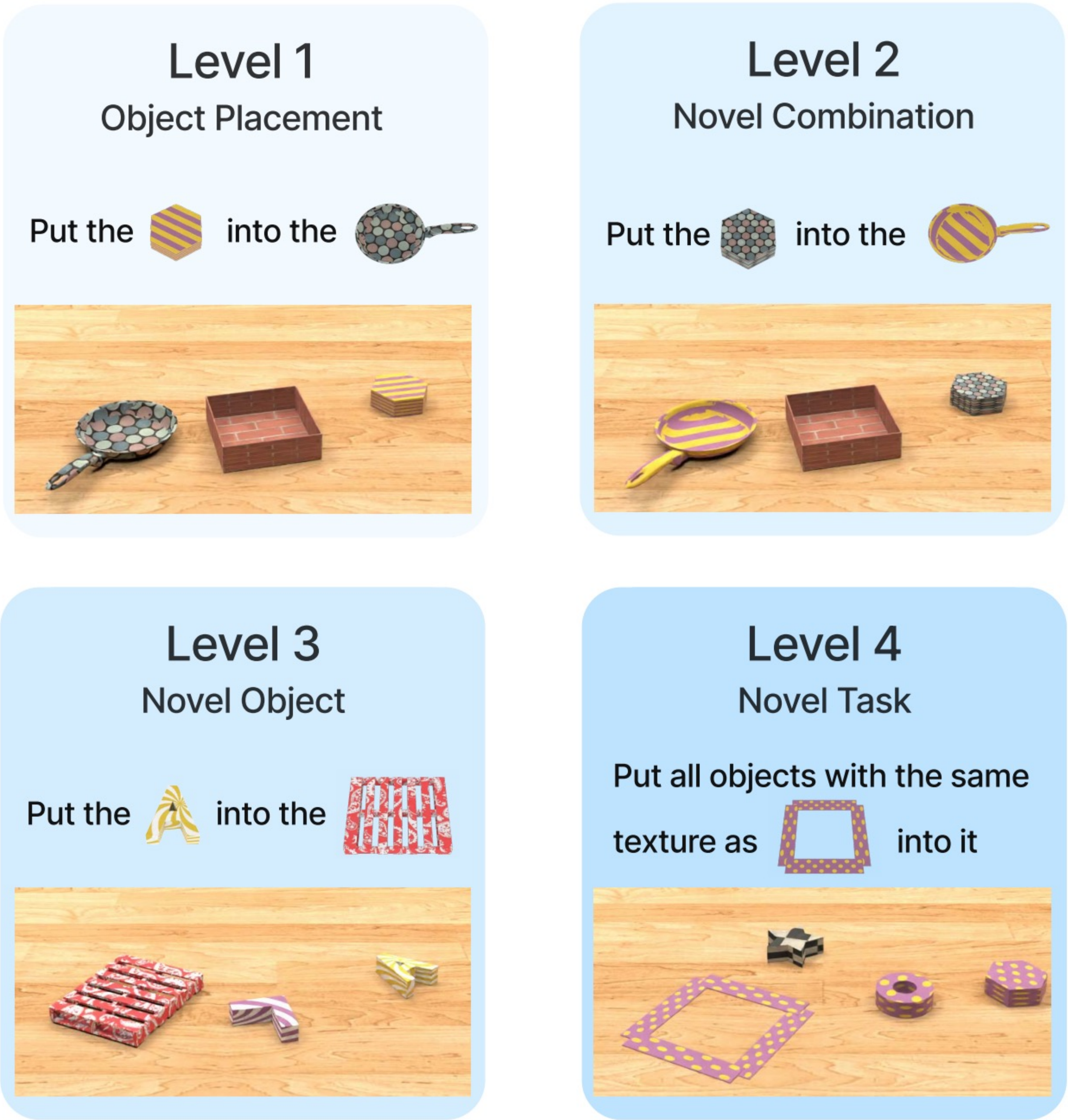
Model scalability from 2M to 200M parameters

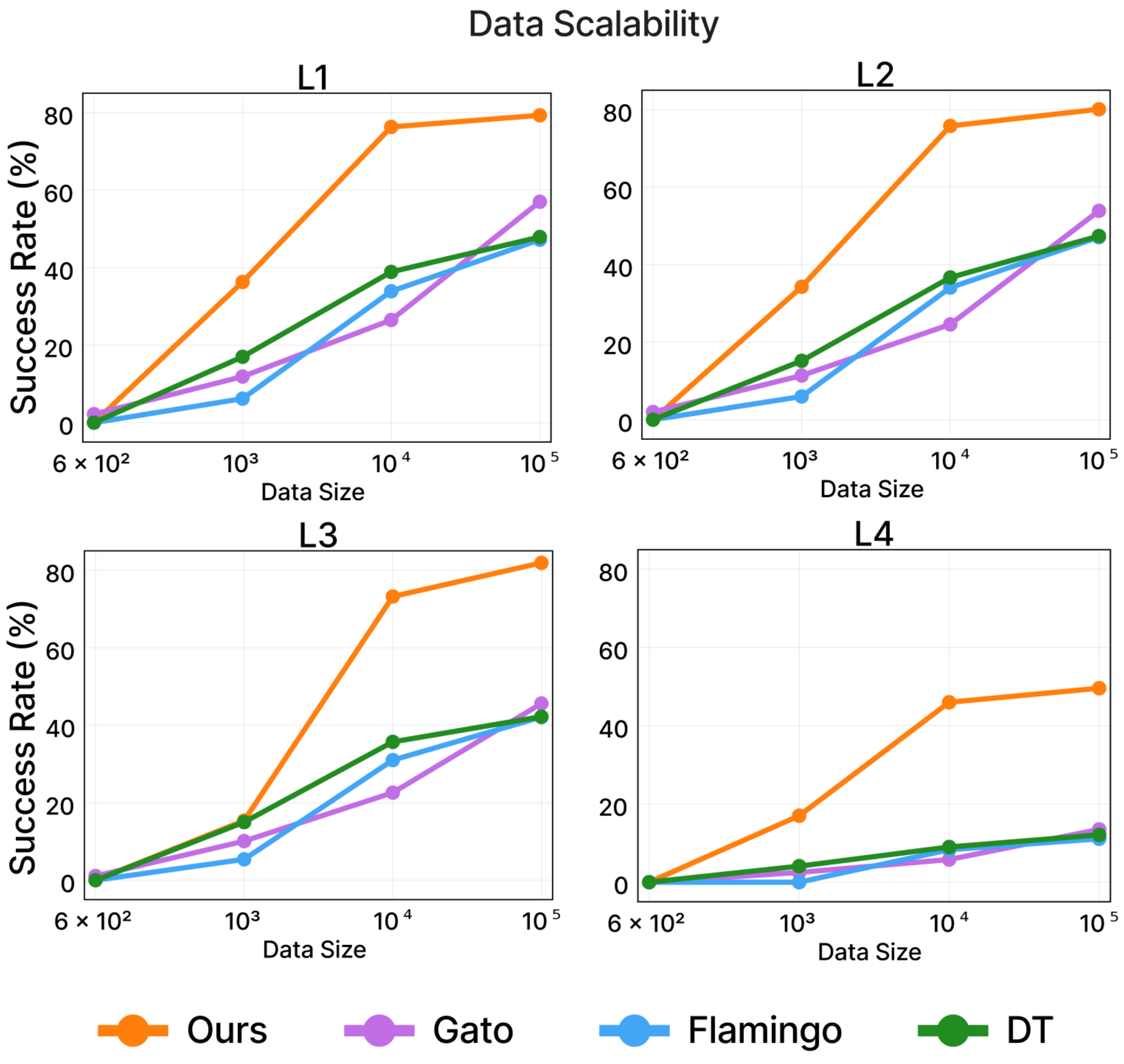




VIMA: General Robot Manipulation with Multimodal Prompts

Data scalability from 0.1% to full dataset





vimalabs.github.io

SAILOR: Data-Efficient Learning from Diverse Prior Data

Real-world data is expensive. Can we gather all prior data to accelerate learning of target tasks?

Prior Robotic Dataset: related tasks, task-agnostic data, etc.

+

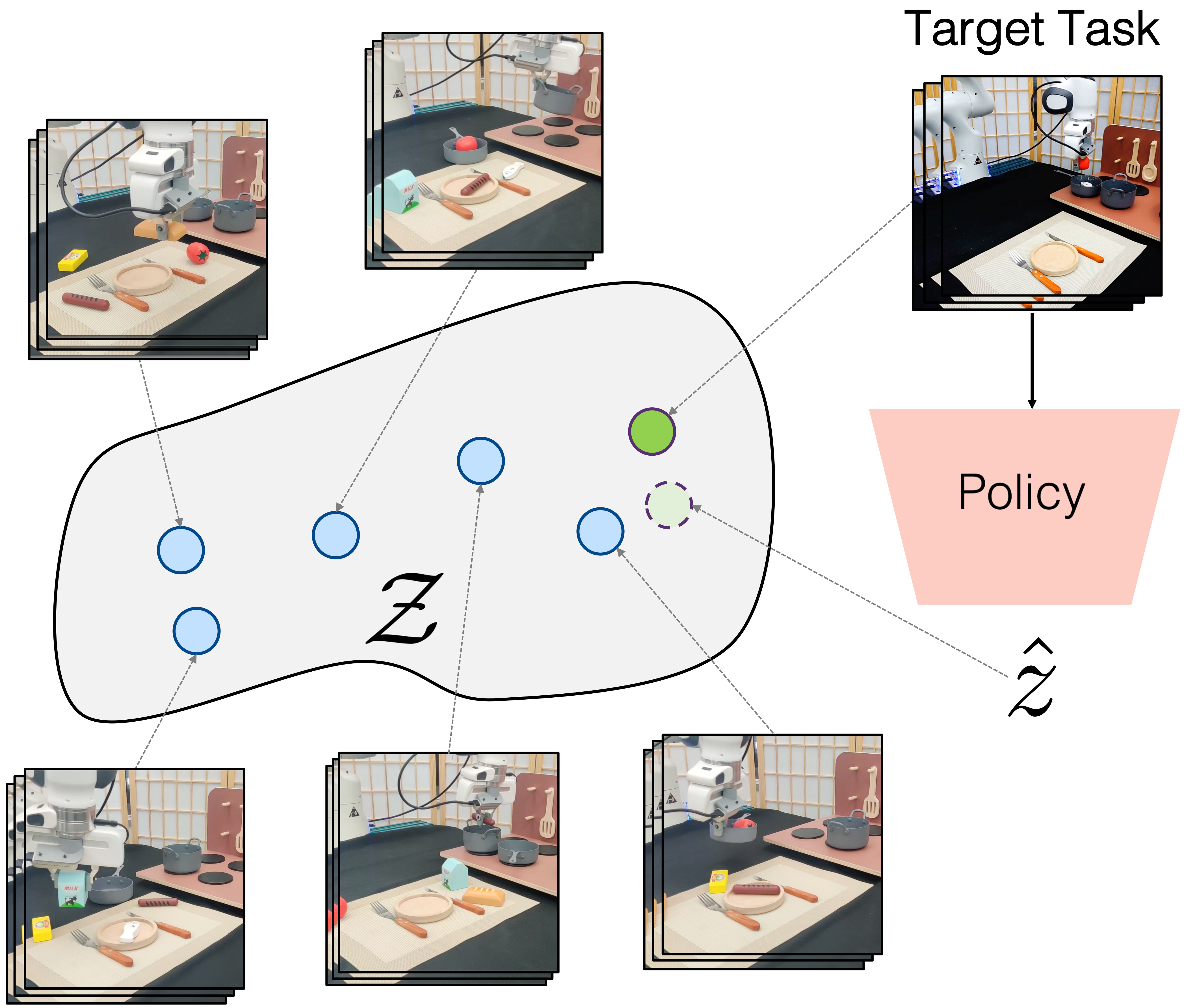
Target Task

large-scale prior robotic data collected in diverse settings

small amount of target task data

SAILOR: Data-Efficient Learning from Diverse Prior Data

Mapping past experiences into latent skills



Extracts reusable behaviors from rich, diverse data

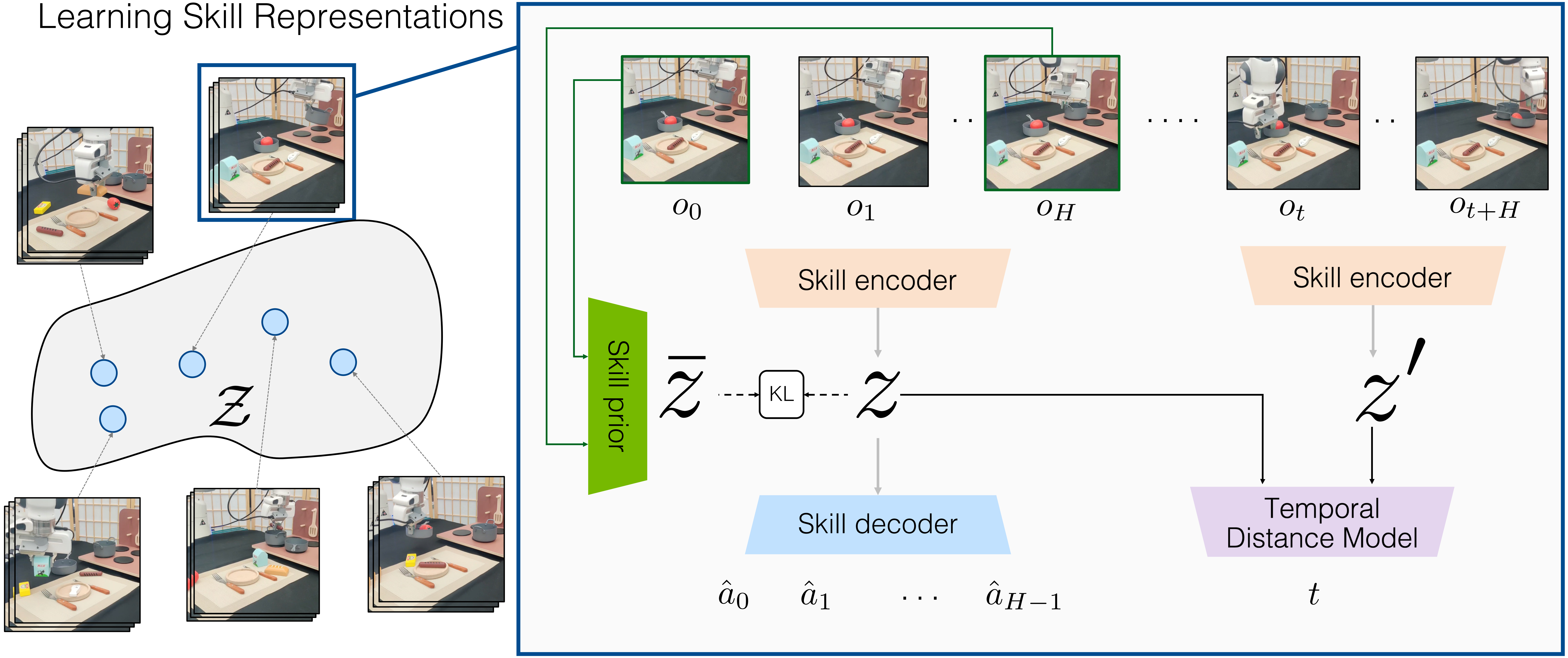
How to learn skill representations that enable effective **downstream learning**?

How to learn **robust policy** given limited target task demonstrations?

"Learning and Retrieval from Prior Data for Skill-based Imitation Learning." Nasiriany, Soroush et al. CoRL 2022

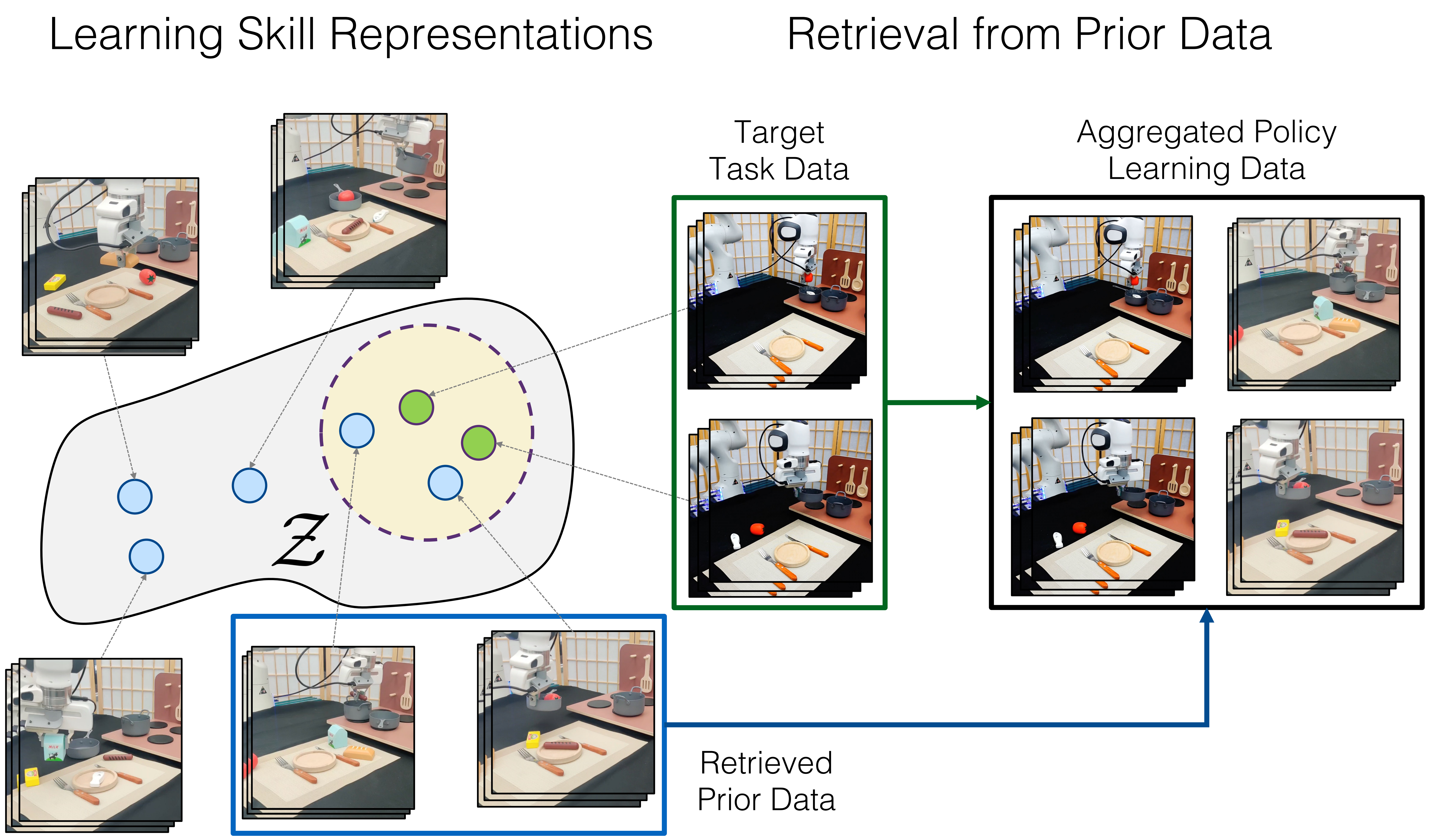
SAILOR: Data-Efficient Learning from Diverse Prior Data

Learning Skill Representations



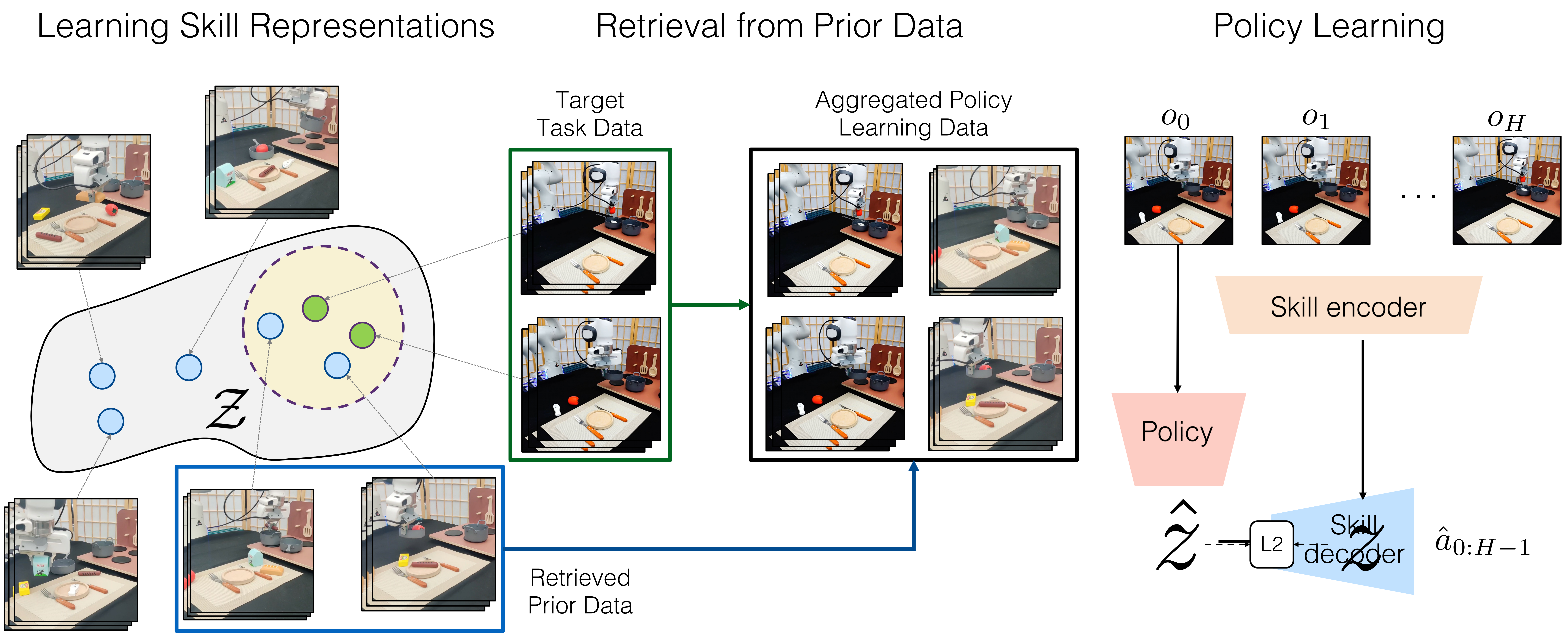
"Learning and Retrieval from Prior Data for Skill-based Imitation Learning." Nasiriany, Soroush et al. CoRL 2022

SAILOR: Data-Efficient Learning from Diverse Prior Data



"Learning and Retrieval from Prior Data for Skill-based Imitation Learning." Nasiriany, Soroush et al. CoRL 2022

SAILOR: Data-Efficient Learning from Diverse Prior Data

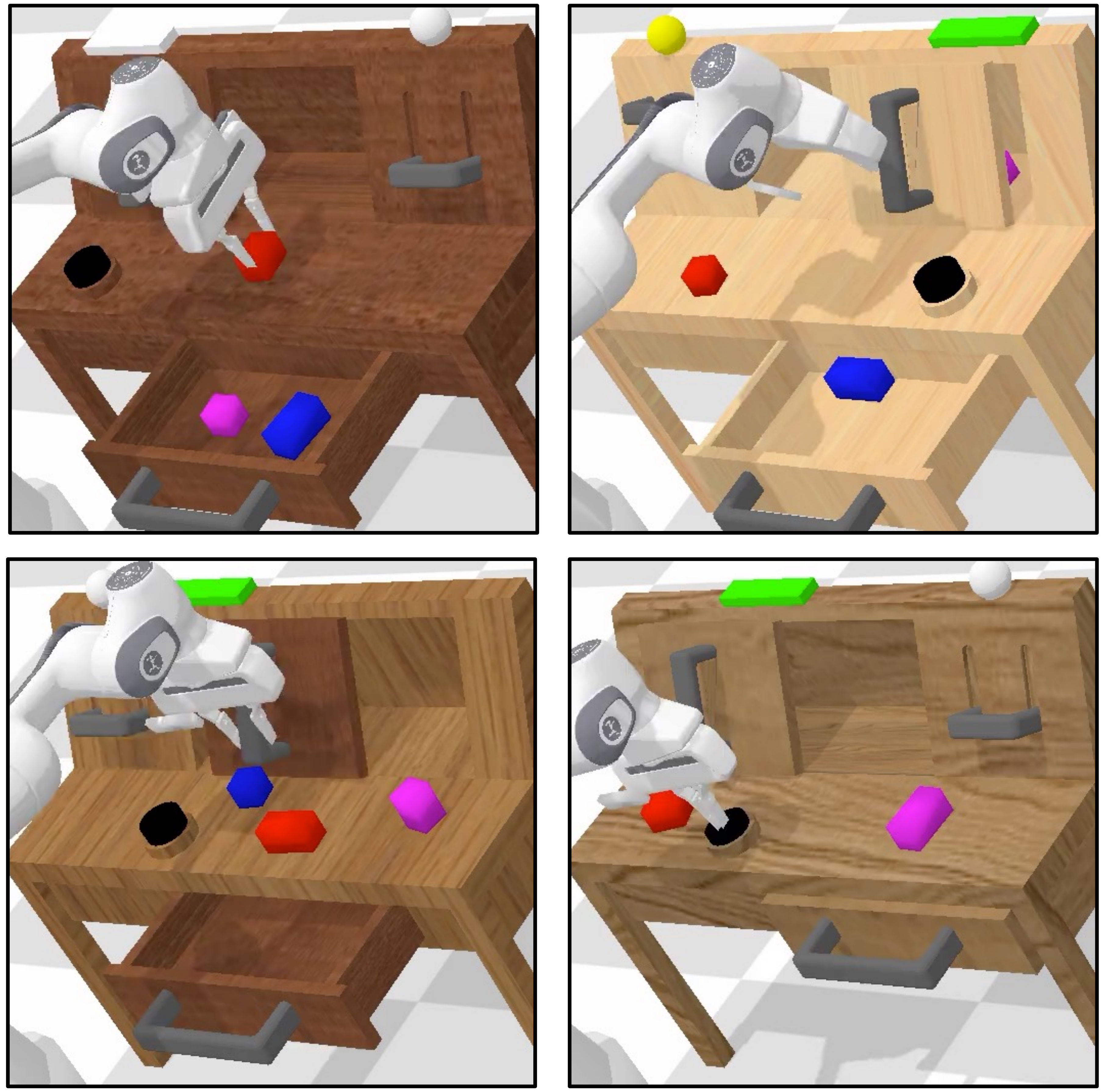


"Learning and Retrieval from Prior Data for Skill-based Imitation Learning." Nasiriany, Soroush et al. CoRL 2022

SAILOR: Data-Efficient Learning from Diverse Prior Data

Prior Data

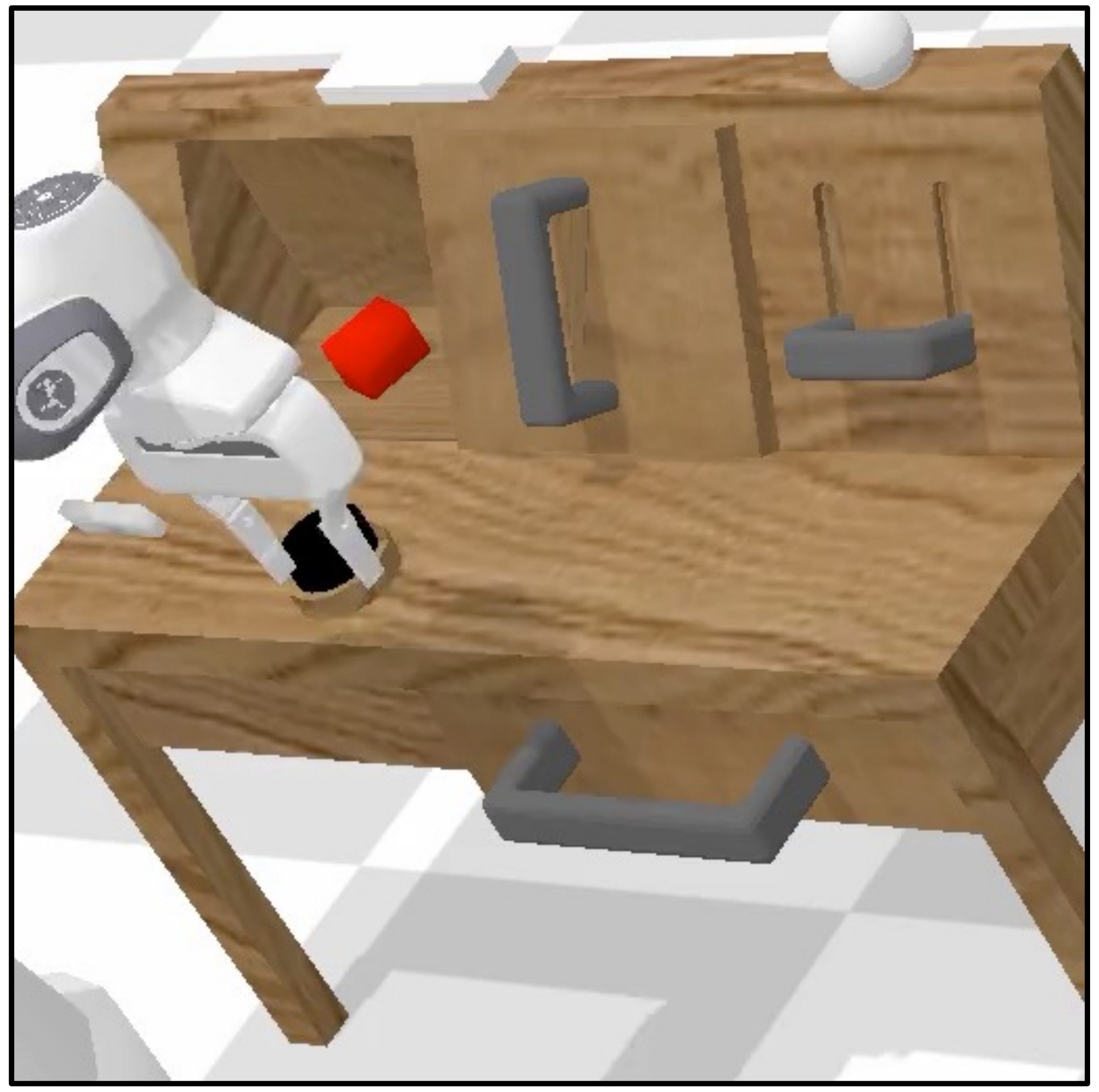
~24 hours play data across four environments



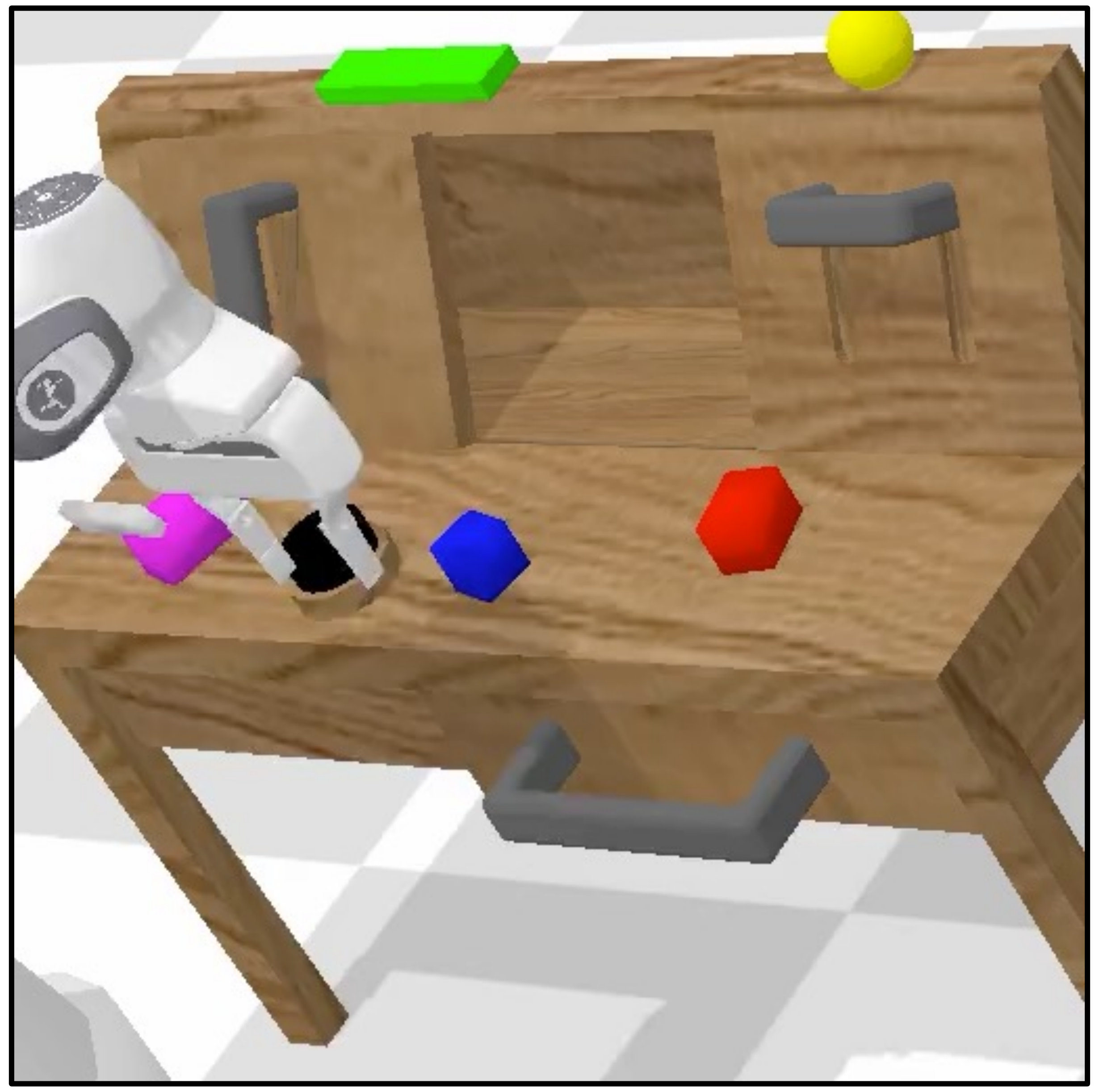
Target Tasks

30 human demonstrations collected in ~30 minutes

Setting Up

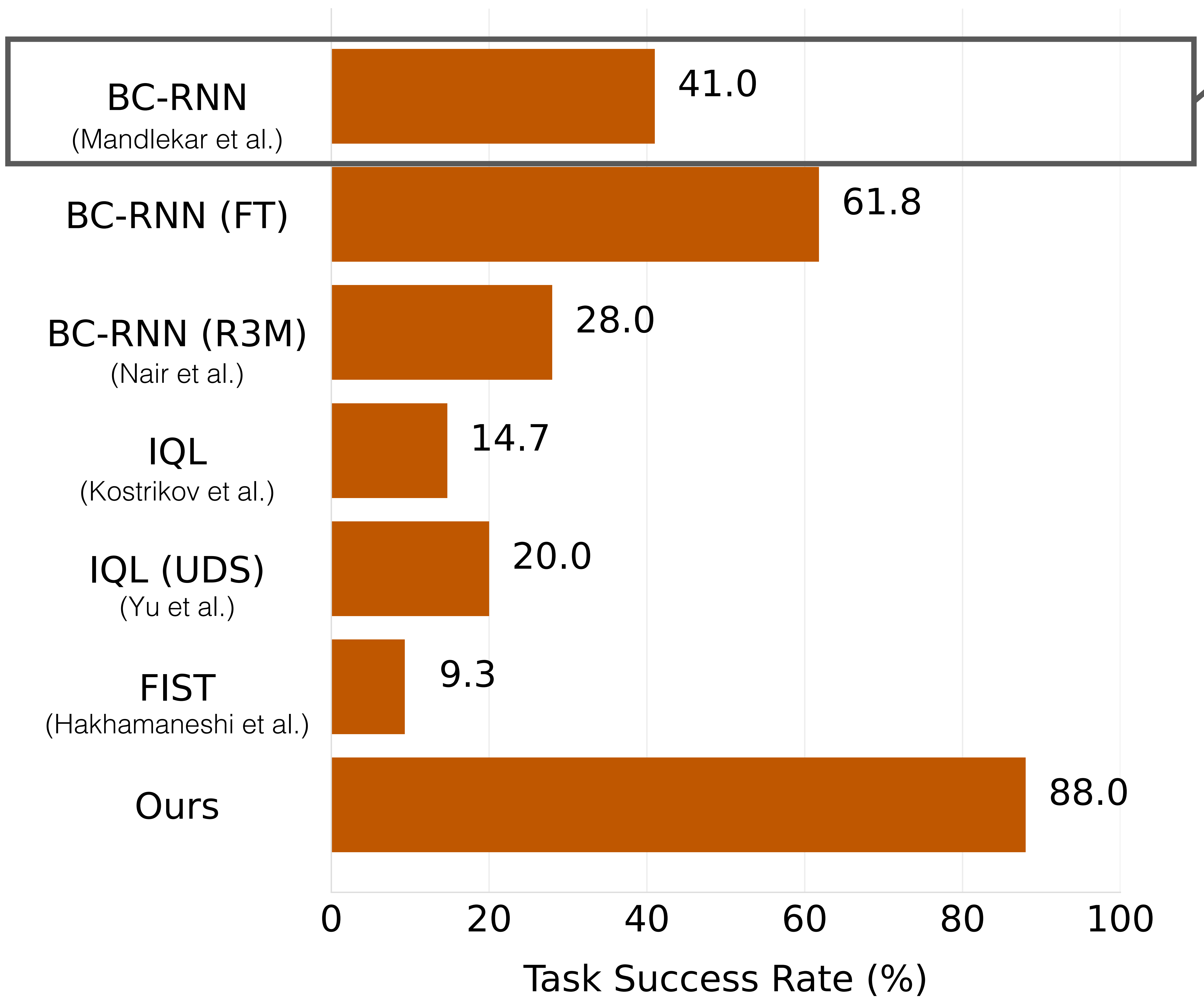


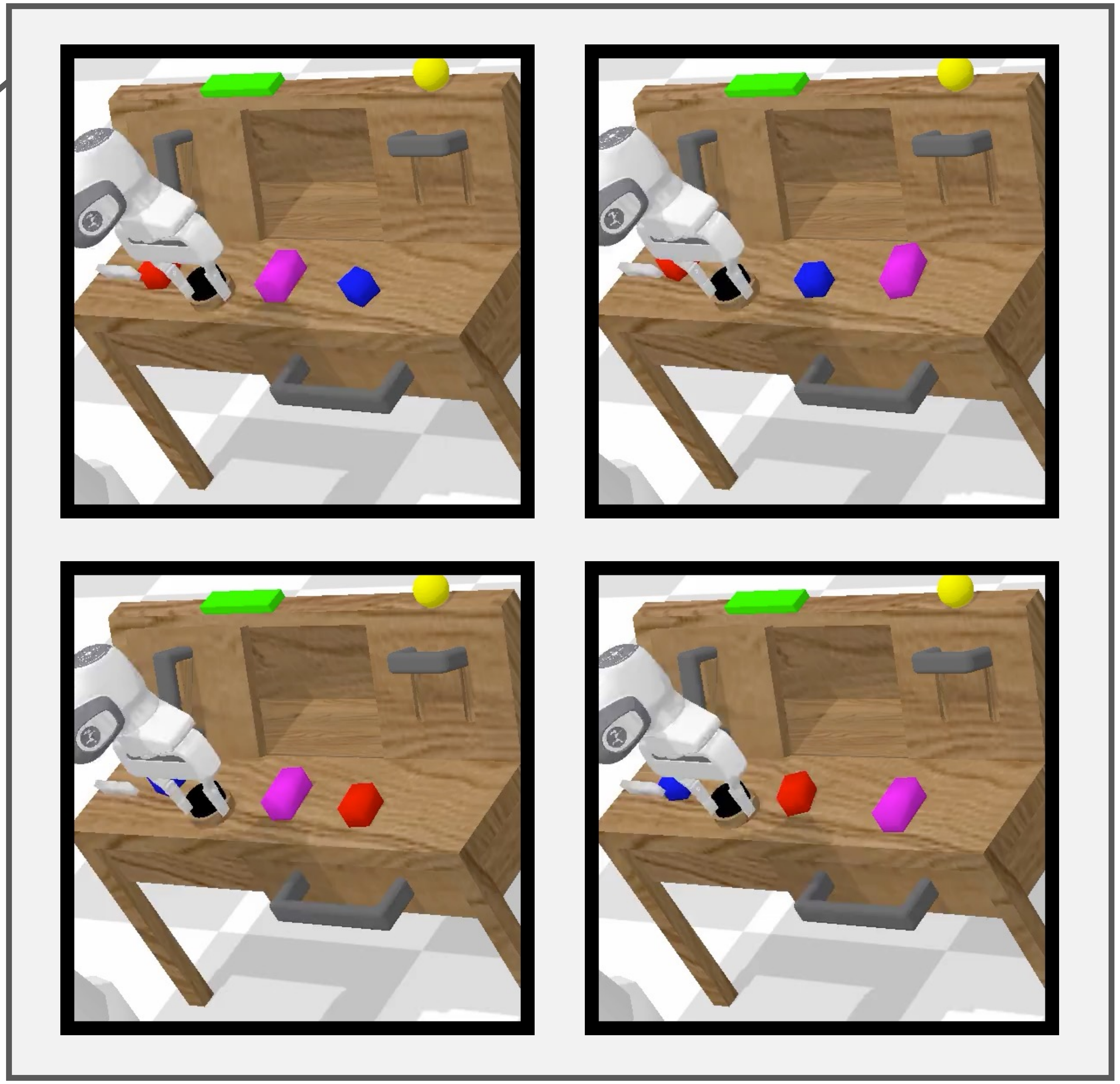
Cleaning Up



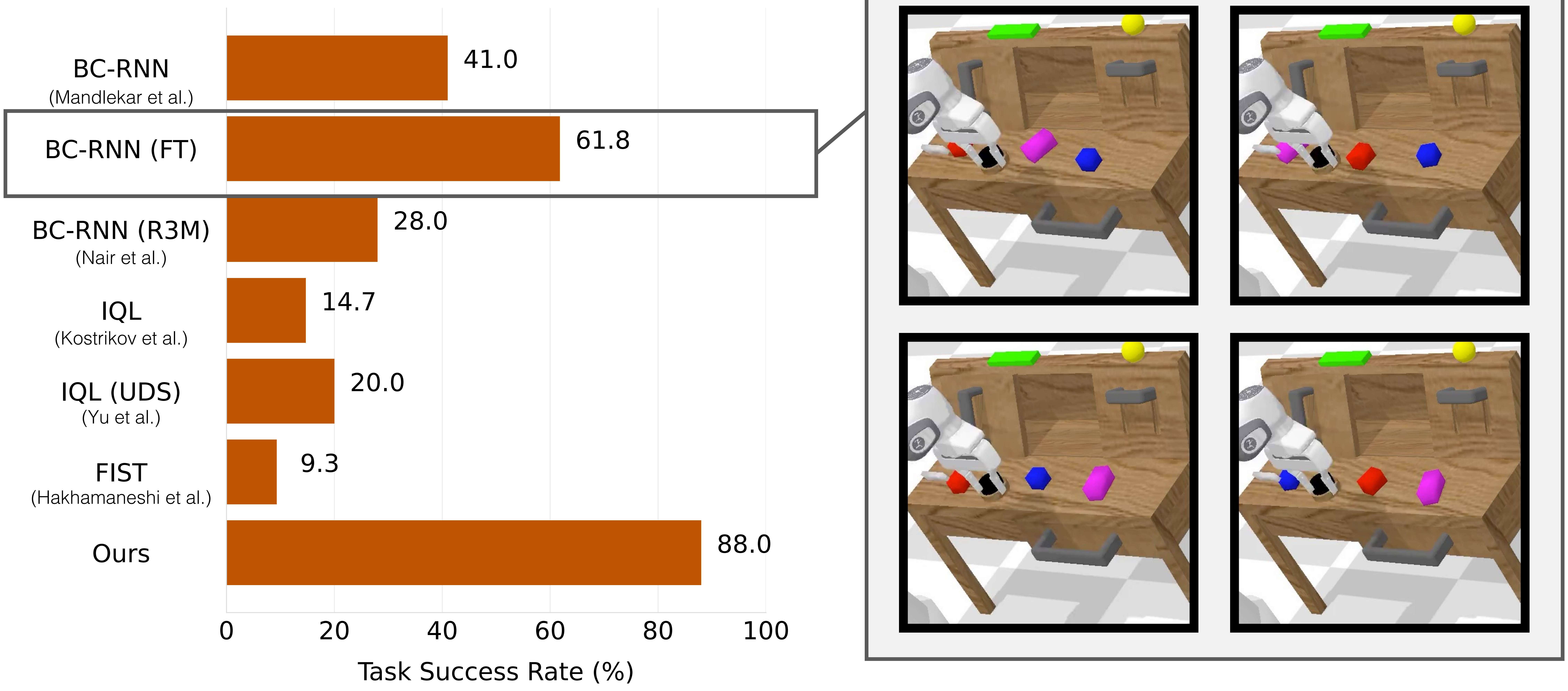
CALVIN: A benchmark for Language-Conditioned Policy Learning for Long-Horizon Robot Manipulation Tasks, Mees et al., IEEE-RAL 2022

SAILOR: Data-Efficient Learning from Diverse Prior Data

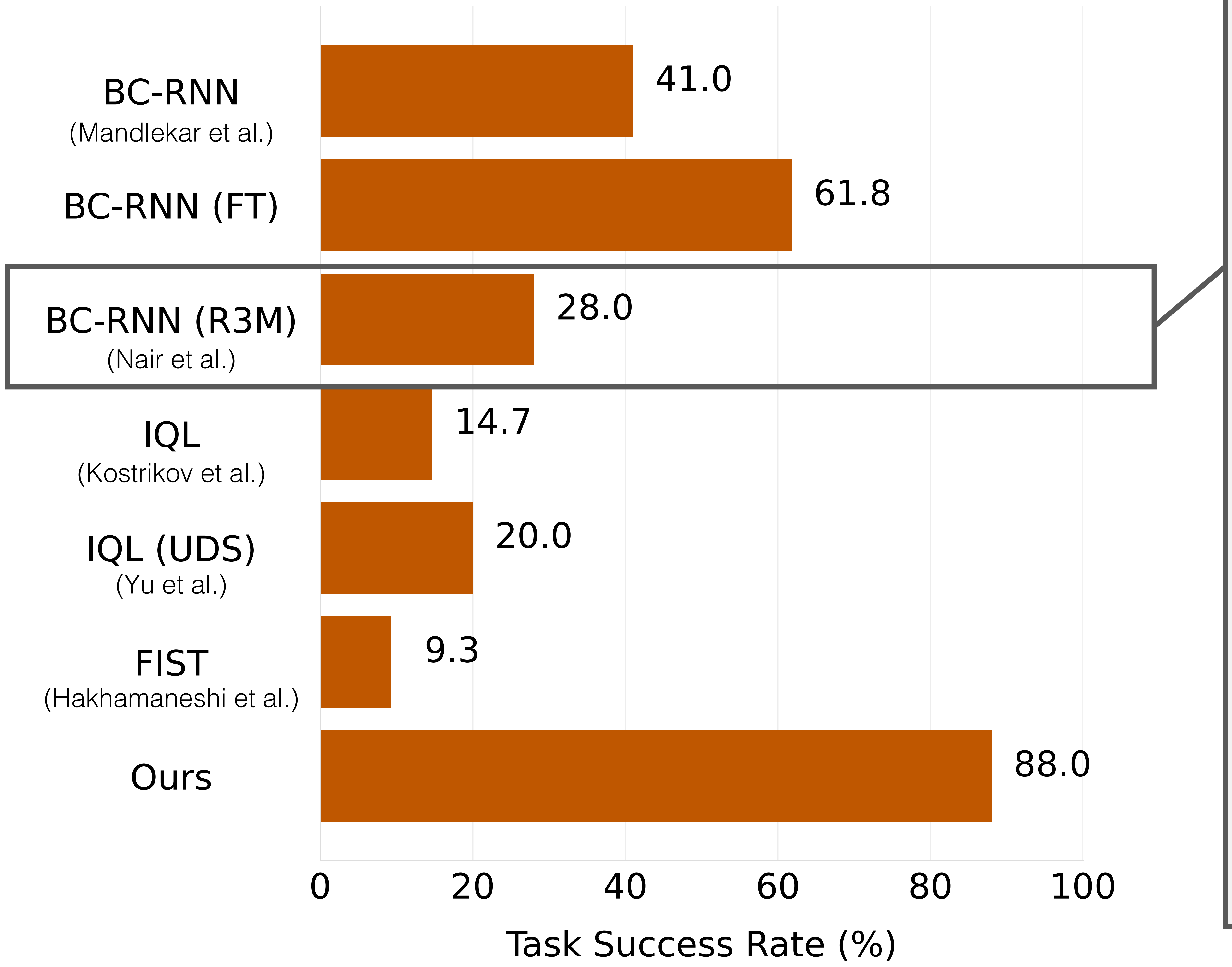


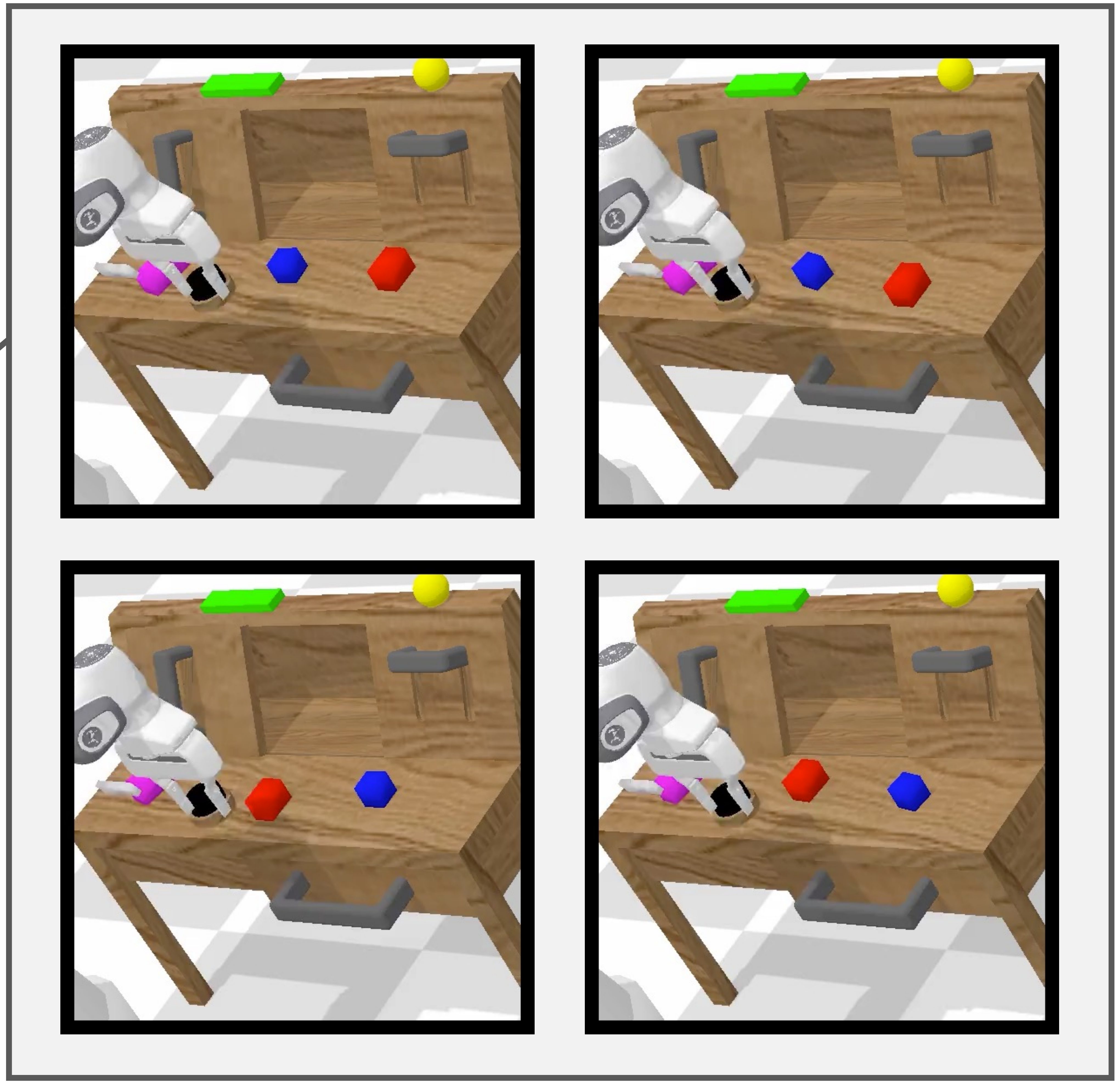


SAILOR: Data-Efficient Learning from Diverse Prior Data

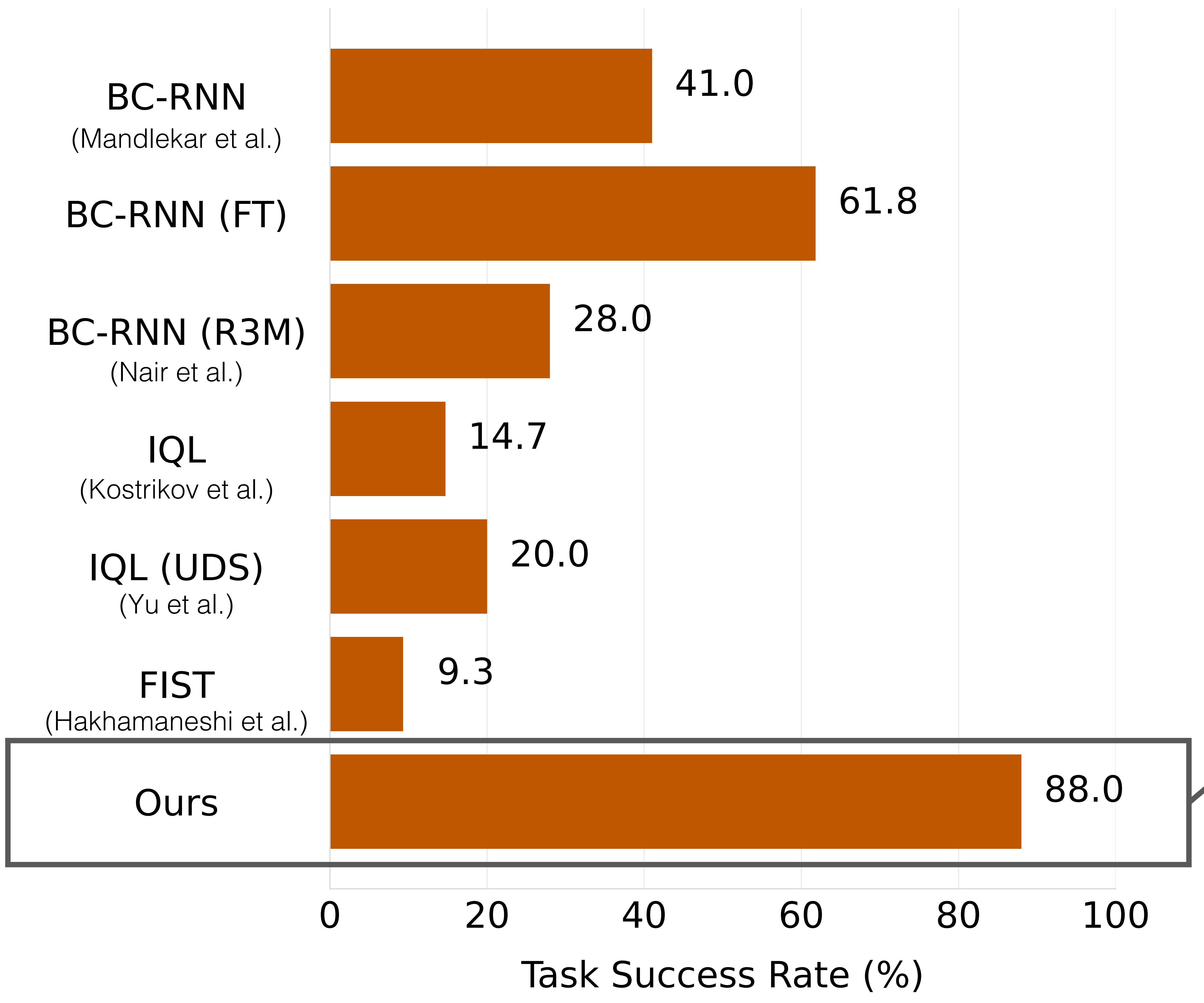


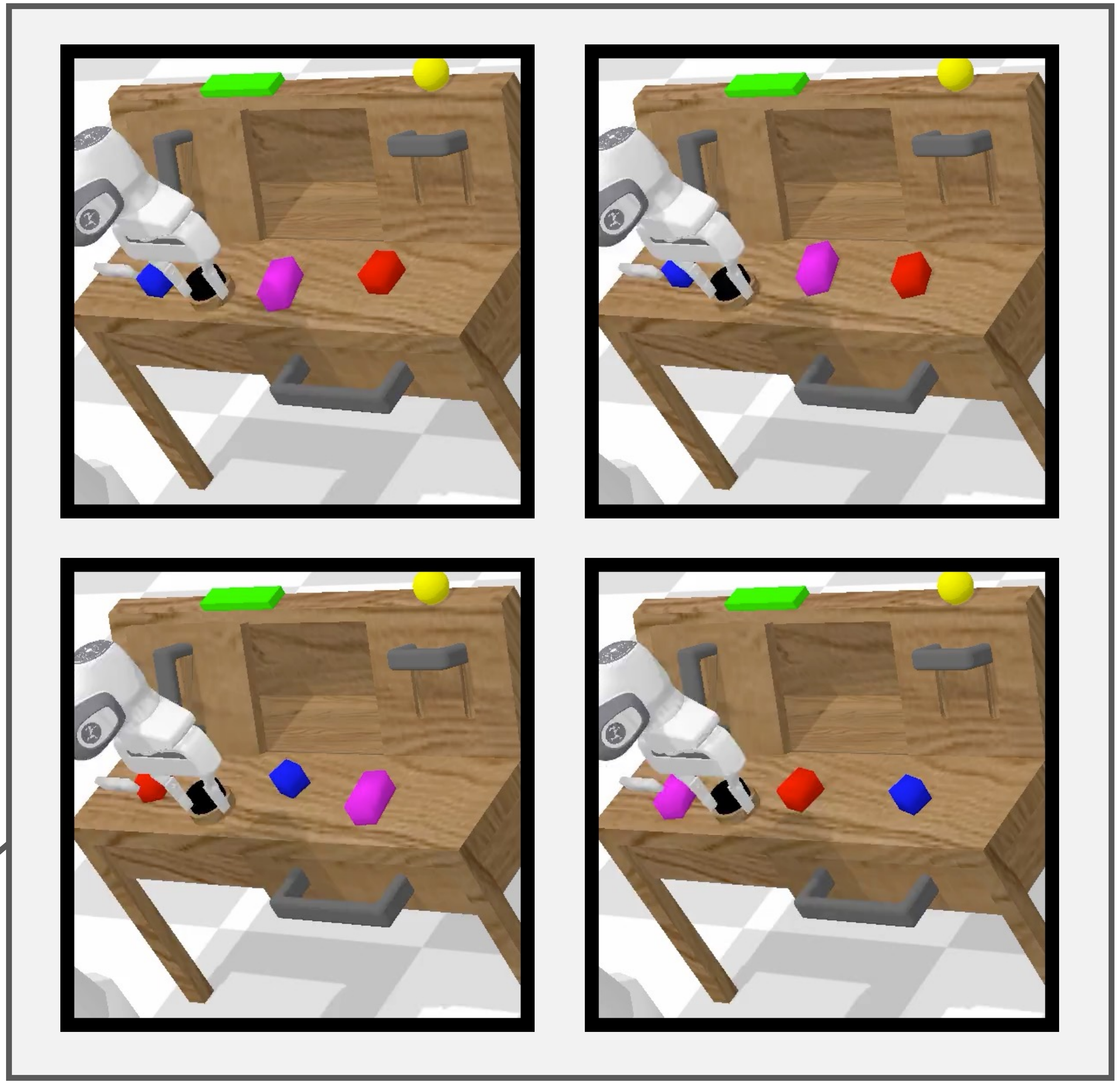
SAILOR: Data-Efficient Learning from Diverse Prior Data





SAILOR: Data-Efficient Learning from Diverse Prior Data





SAILOR: Data-Efficient Learning from Diverse Prior Data

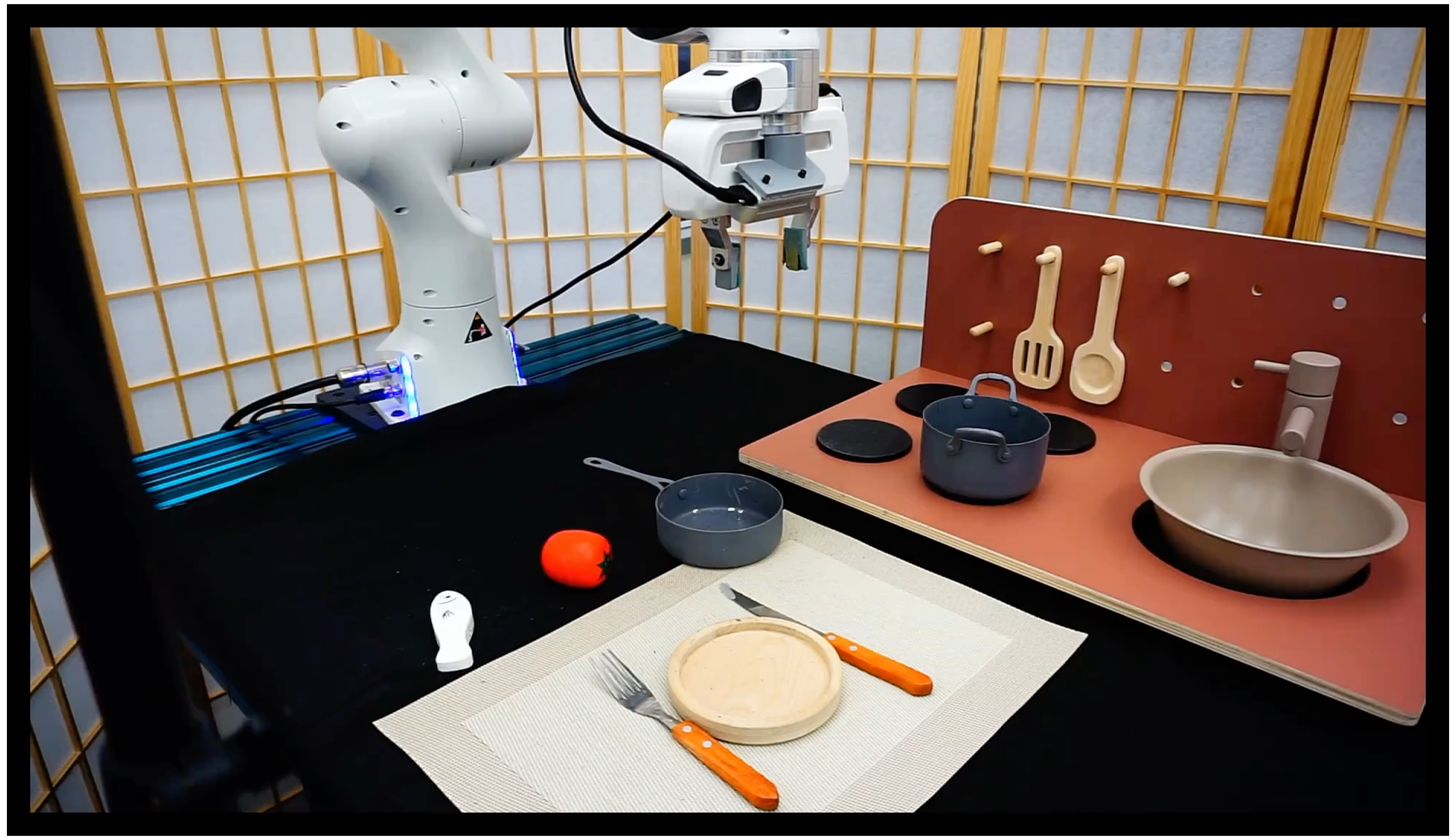
Large-scale prior data collection in real world

- Diverse kitchen environment
- 8 food items, one pot, one pan
- Unstructured interaction data: 300,000 timesteps.

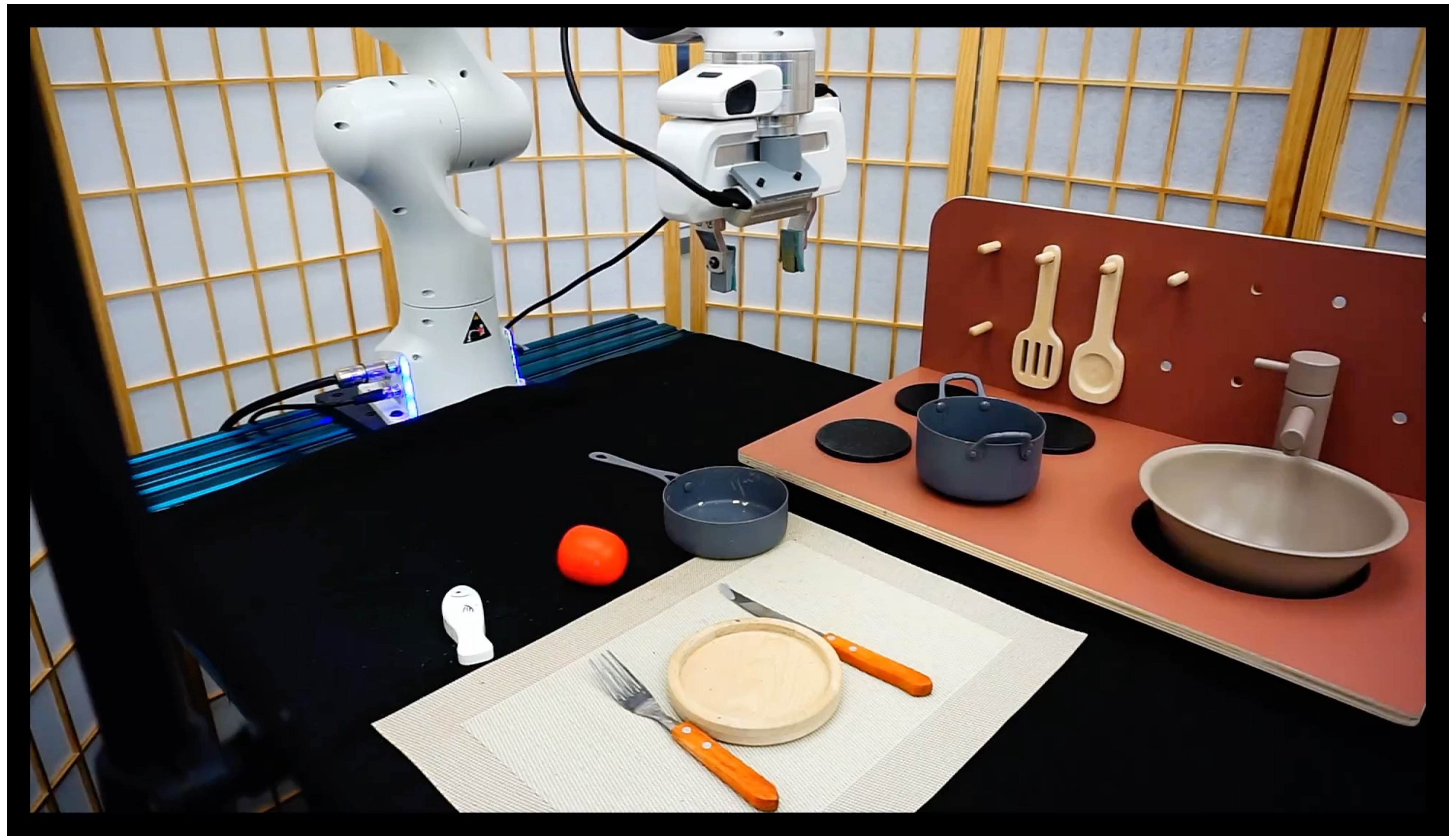
"Learning and Retrieval from Prior Data for Skill-based Imitation Learning." Nasiriany, Soroush et al. CoRL 2022

SAILOR: Data-Efficient Learning from Diverse Prior Data

Effective imitation learning for vision-based manipulation



BC-RNN (FT): 46.7% success rate



Ours: 76.7% success rate

Vision-based policies trained on 30 demonstrations for the target task

ut-austin-rpl.github.io/sailor

VIOLA: Imitation Learning with Object-Centric Priors

Raw visual observation

"VIOLA: Imitation Learning for Robot Manipulation with Object Proposal Priors." Zhu, Yifeng et al. CoRL 2022

VIOLA: Imitation Learning with Object-Centric Priors

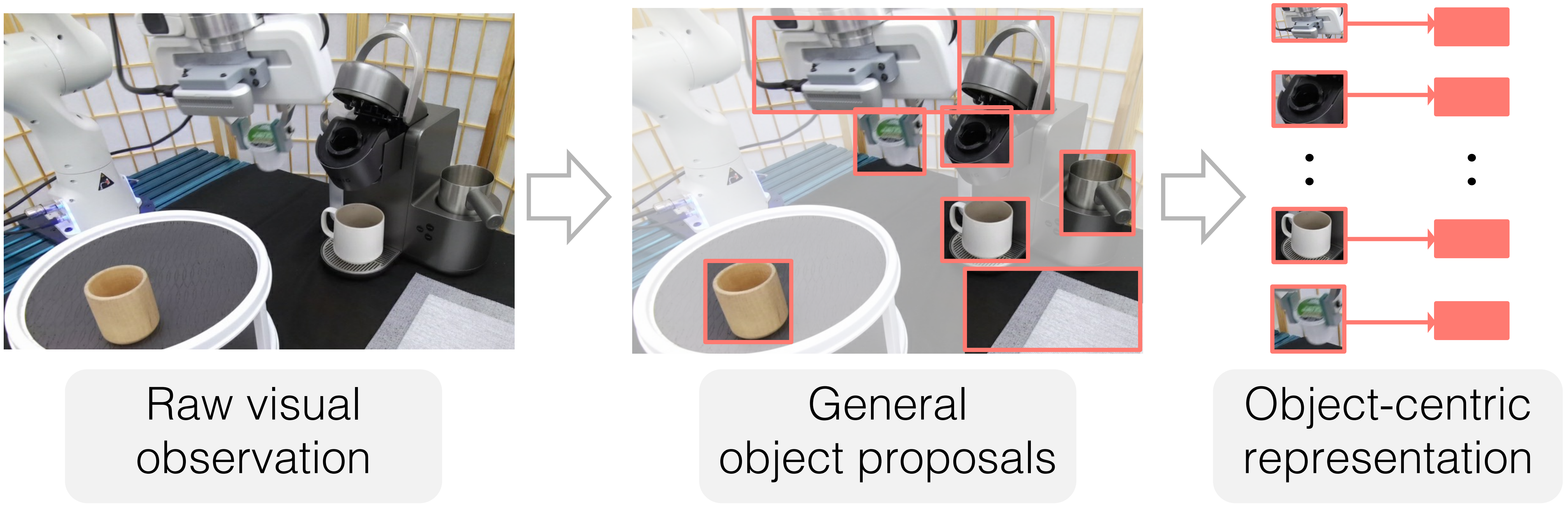
Raw visual observation

General object proposals

Pre-trained region proposal network on large image datasets

"VIOLA: Imitation Learning for Robot Manipulation with Object Proposal Priors." Zhu, Yifeng et al. CoRL 2022

VIOLA: Imitation Learning with Object-Centric Priors



Encoding object visual appearances and their spatial locations

“VIOLA: Imitation Learning for Robot Manipulation with Object Proposal Priors.” Zhu, Yifeng et al. CoRL 2022

VIOLA: Imitation Learning with Object-Centric Priors

Using transformers to select task-relevant objects and reason about their relations

"VIOLA: Imitation Learning for Robot Manipulation with Object Proposal Priors." Zhu, Yifeng et al. CoRL 2022


Workspace Image
VIOLA can reward a hard-working researcher
with a cup of good coffee!

The Data Pyramid for Generalist Agents

Real-World Data

SAILOR & VIOLA [Nasiriany et al. & Zhu et al. CoRL 2022]: Data-efficient and robust imitation learning in the real world with diverse prior robot data

Synthetic Data

VIMA [Jiang et al. NeurIPS FMDM 2022]: Building generalist robotic agents with multimodal prompts in large-scale simulation

Web Data

YouTube

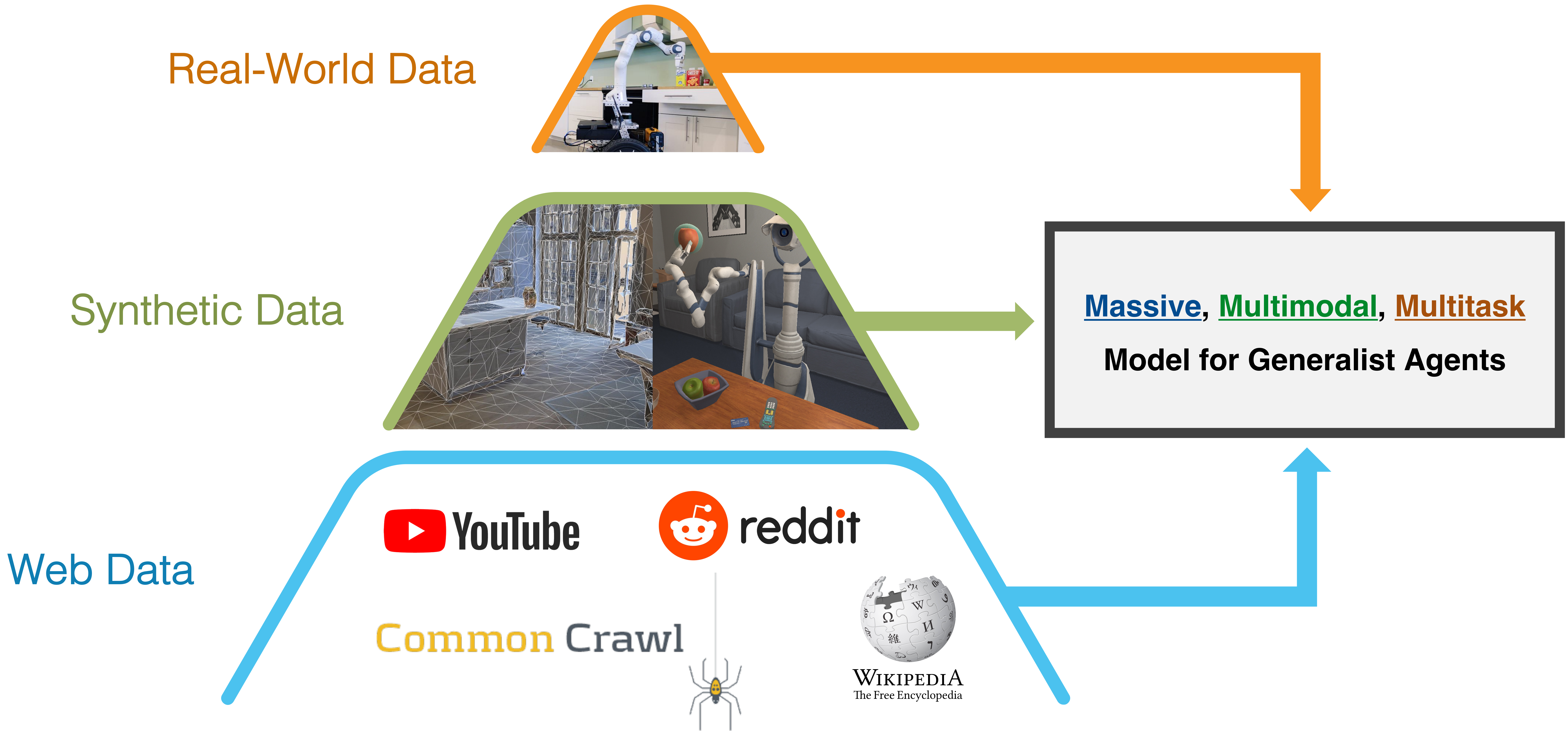
reddit

Common Crawl

Wikipedia
The Free Encyclopedia

MineDojo, LID [Fan et al. & Li et al. NeurIPS 2022]: Building generally capable agents with Internet-scale knowledge

The Data Pyramid for Generalist Agents



Acknowledgement


Soroush Nasiriany

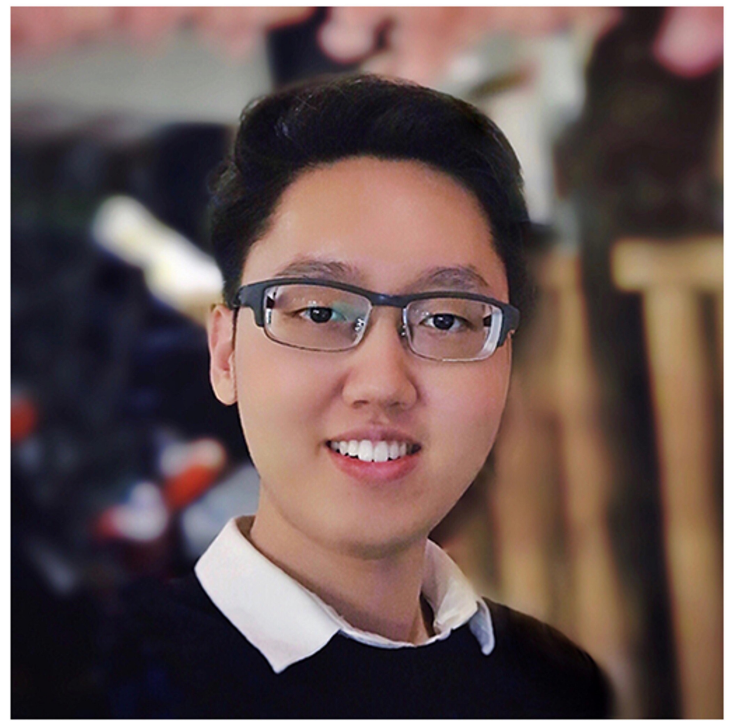

Yifeng Zhu


Shuang Li


Ajay Mandlekar


Linxi “Jim” Fan


Yunfan Jiang


Guanzhi Wang


Anima Anandkumar

rpl.cs.utexas.edu