Using Both Demonstrations and Language Instructions to Efficiently Learn Robotic Tasks

The University of Texas at Austin

https://deltaco-robot.github.io/ (PDF + Code + Datasets)
Albert Yu, Ray Mooney
{albertyu, mooney}@utexas.edu

PDF   Website

Why Condition Robotic Policies on Demonstrations + Language?

Learn Efficiently with Multimodality

How to make noodles — Visual demonstration → How to make a tent
Language instruction
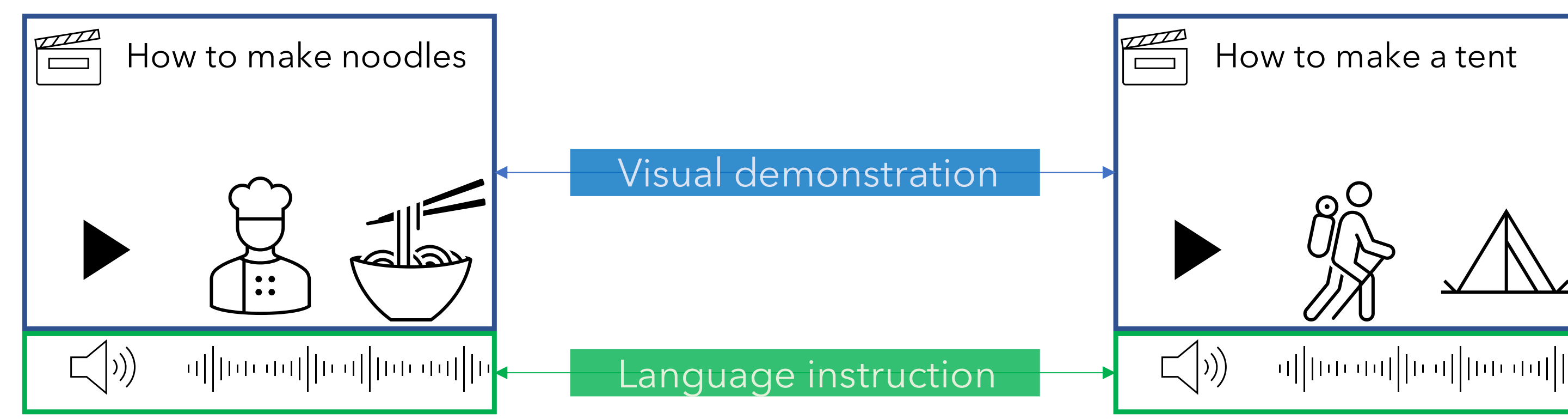
Humans teach/learn with demos + language (e.g. instructional videos).

Resolve Ambiguities

Demo-only Task Specification

Language-only Task Specification
Instruction: *"Grab the left* pointy* gold* object."*
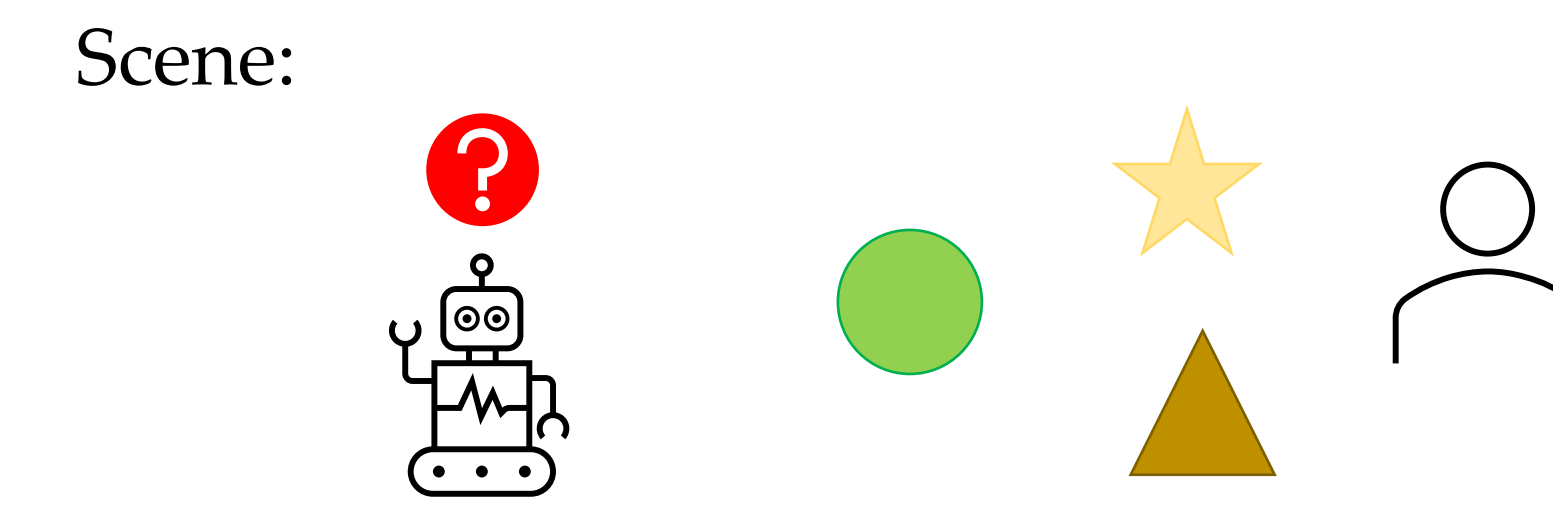
Scene:
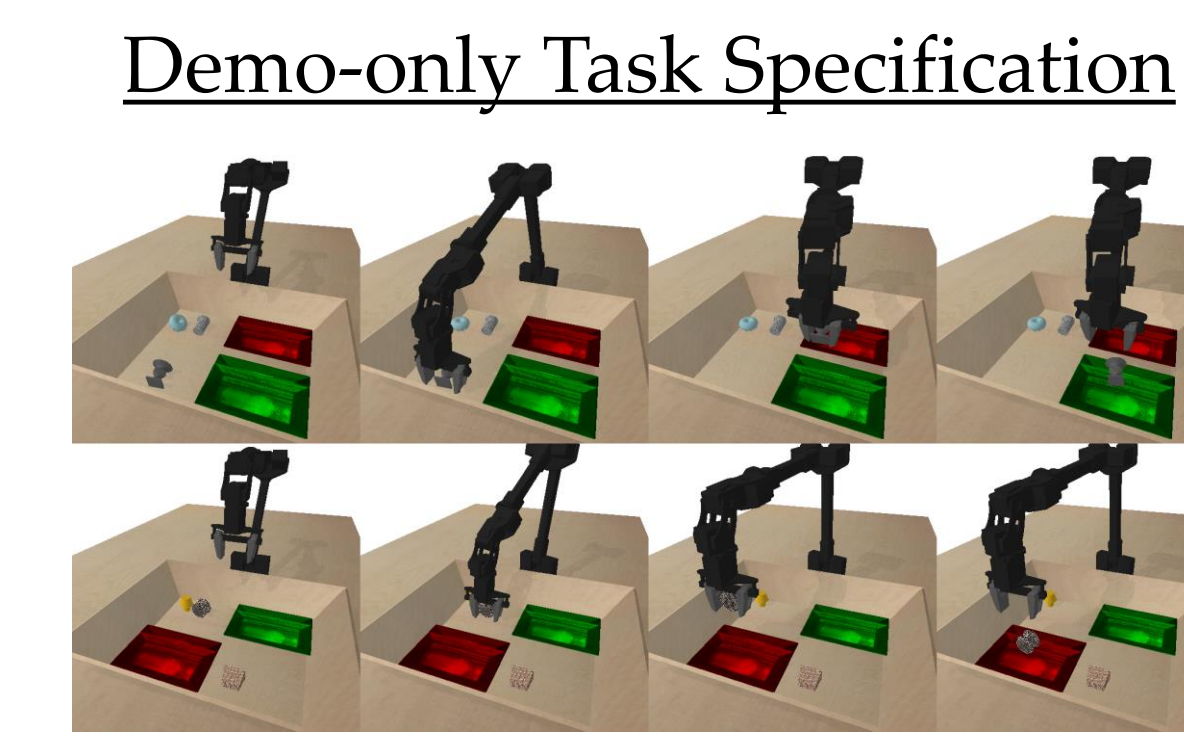
Hard to infer a task from multiple demos.   Instruction contains many ambiguities (*).

Specifying tasks with demos + language helps resolve ambiguities.

Summary

Contributions

1. 1st work (we know of) showing that learning with demos + language attains better generalization and sample efficiency than prior (unimodal) methods
2. Framework for learning from both demos + language in training + testing
3. Benchmark of 300 highly-randomized pick-and-place tasks

Main Takeaways

1. **Generalization**: Learning from both demonstrations and language with DeL-TaCo leads to better generalization than learning from one modality
2. **Sample-Efficiency**: DeL-TaCo reduces the need for ~50 demonstrations over policies conditioned only on demonstrations
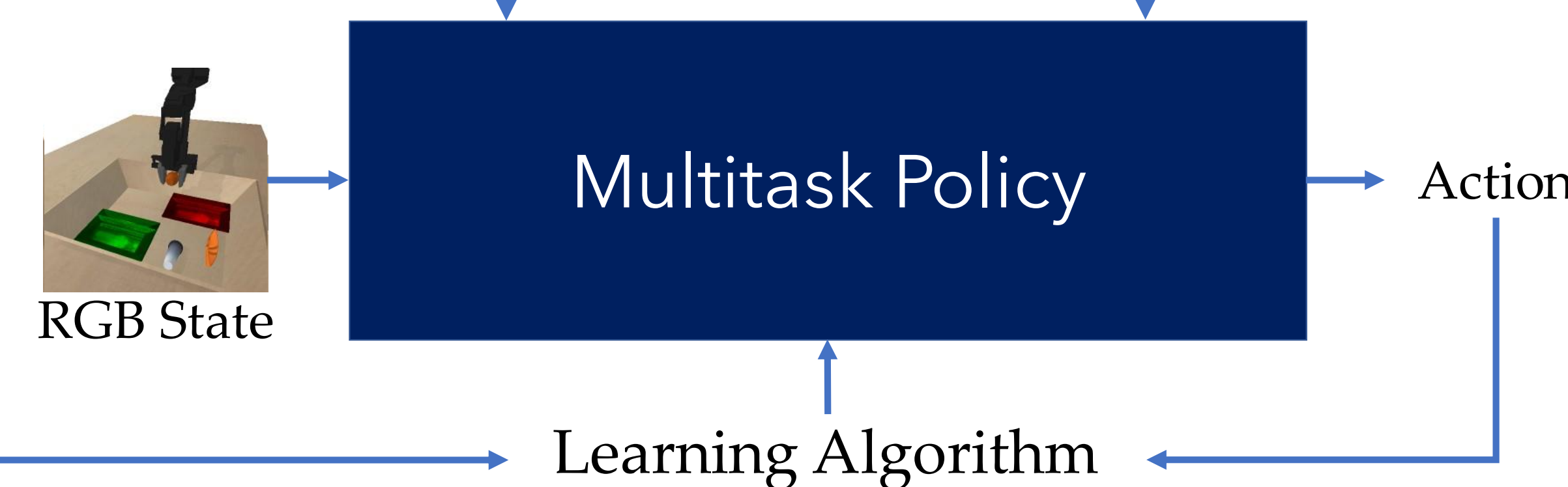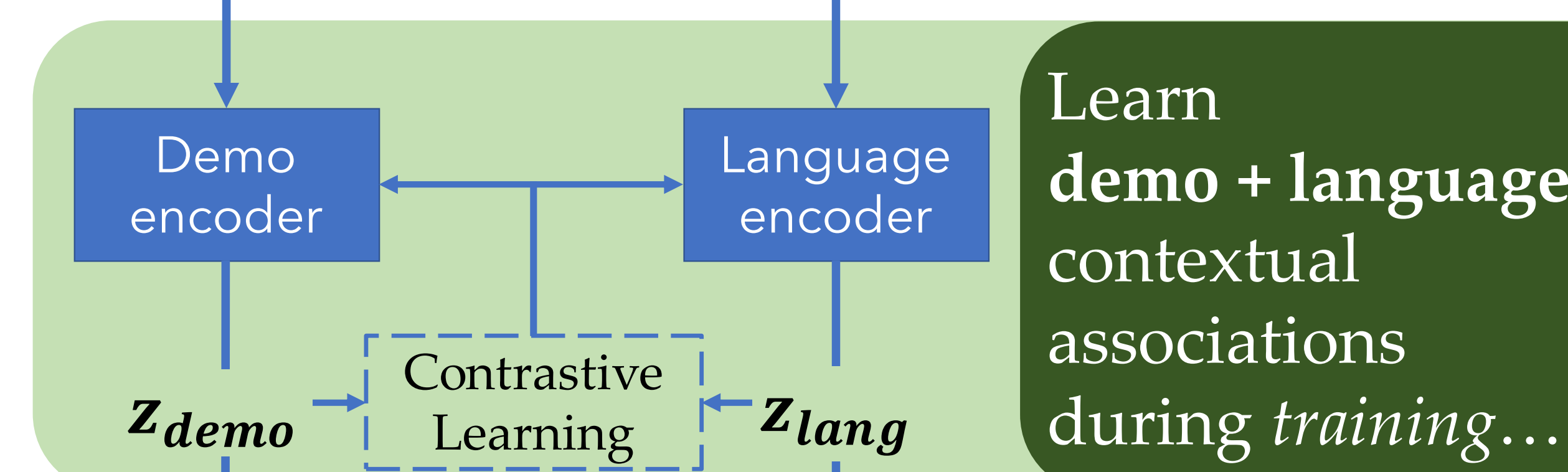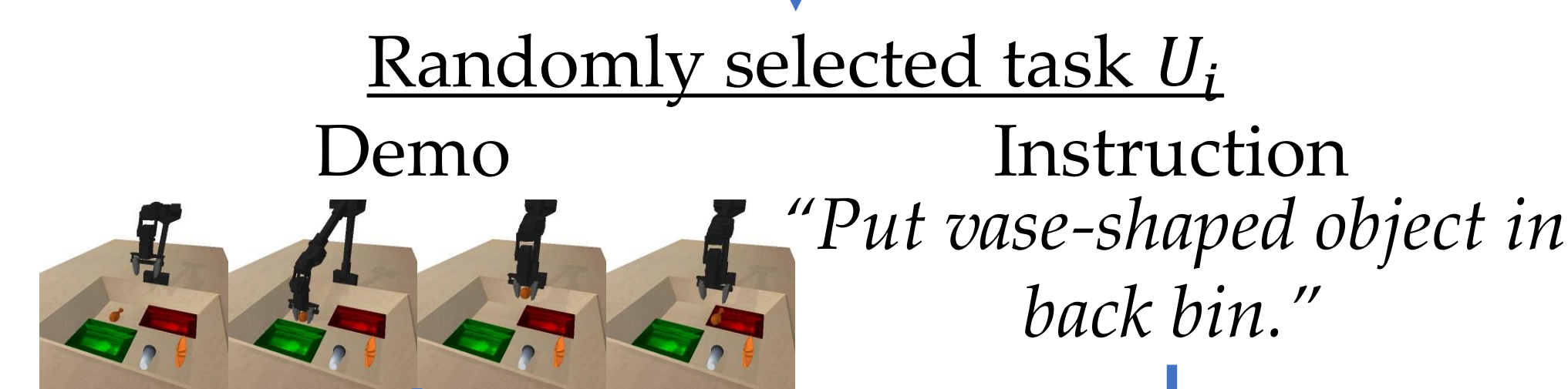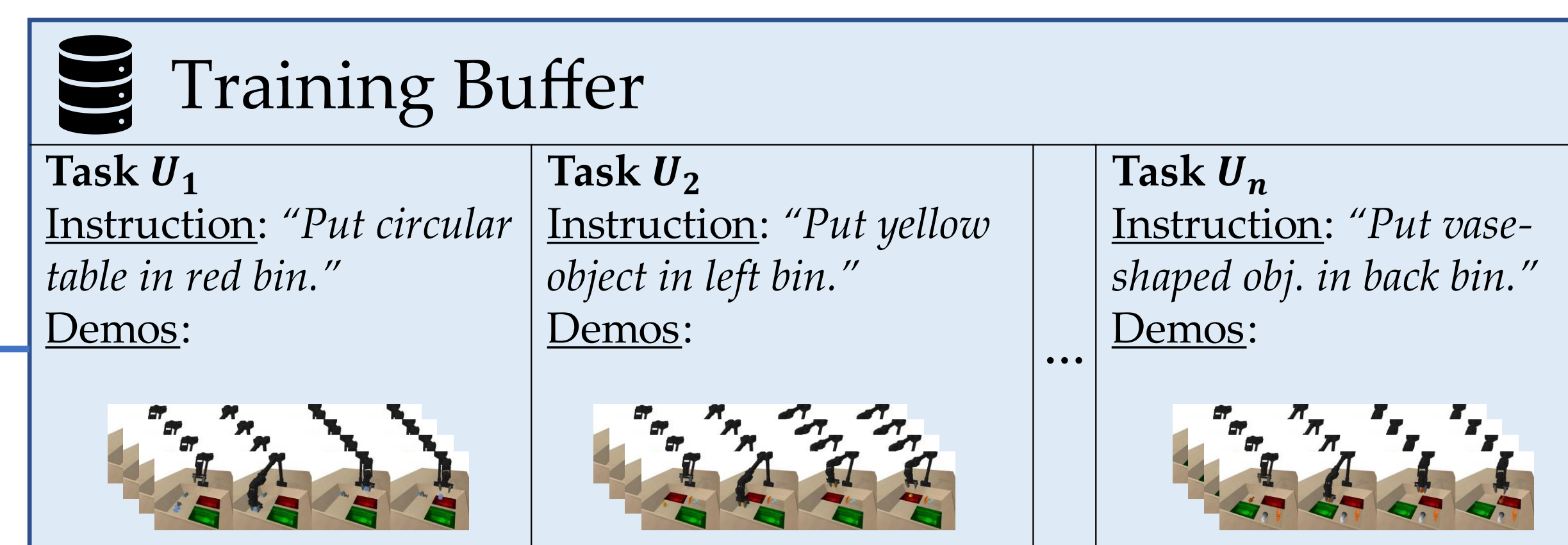
Overview

DeL-TaCo (Demo-Language Task Conditioning)

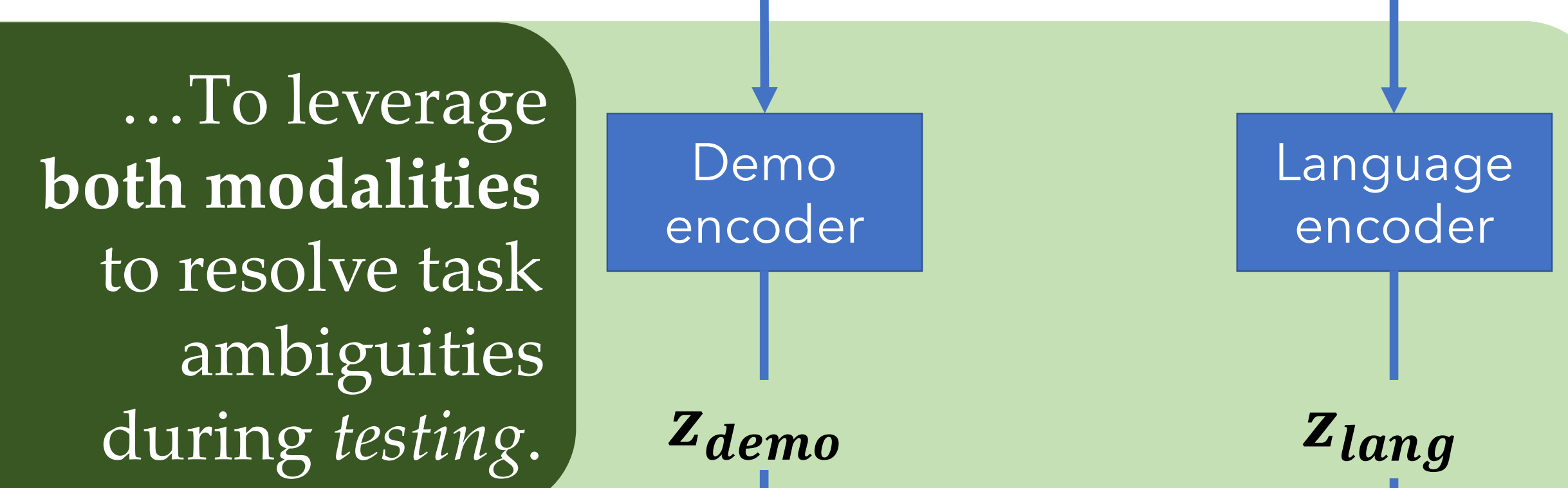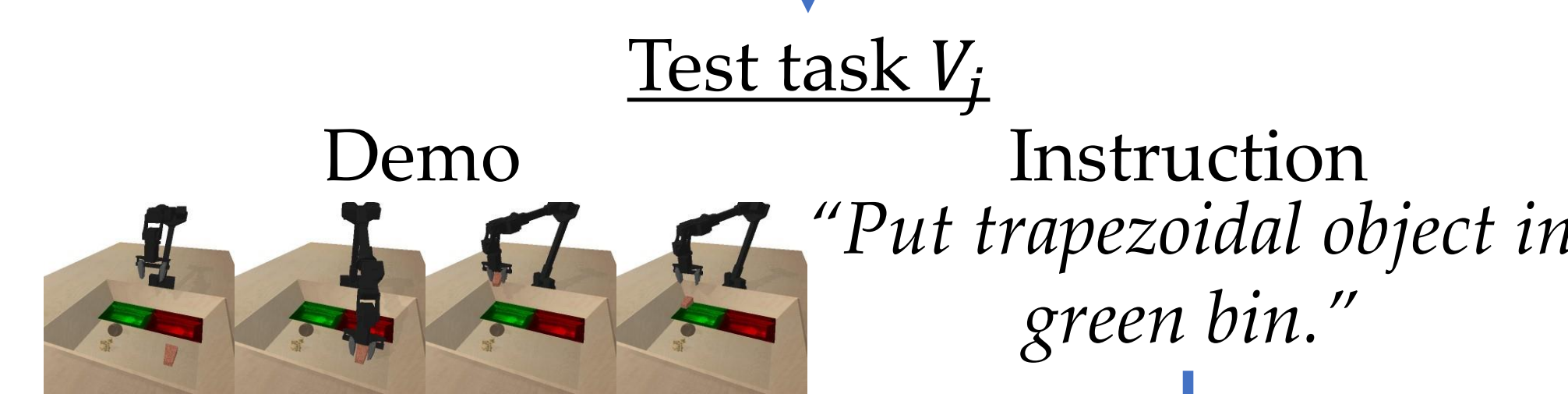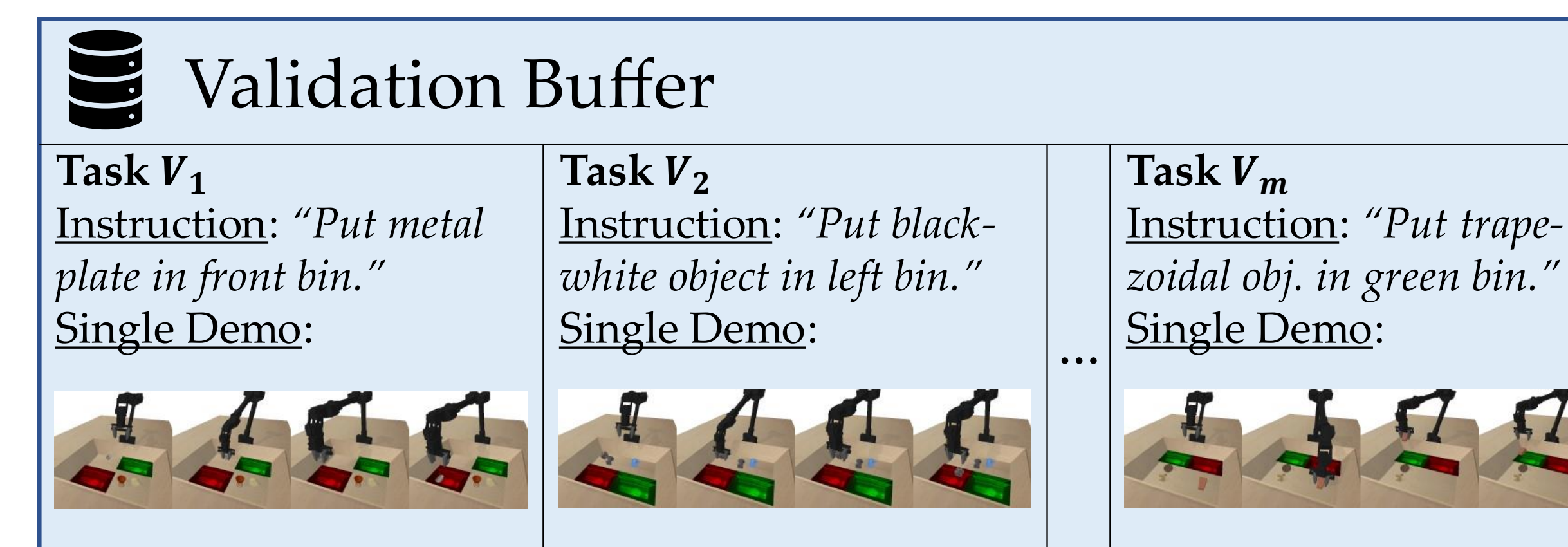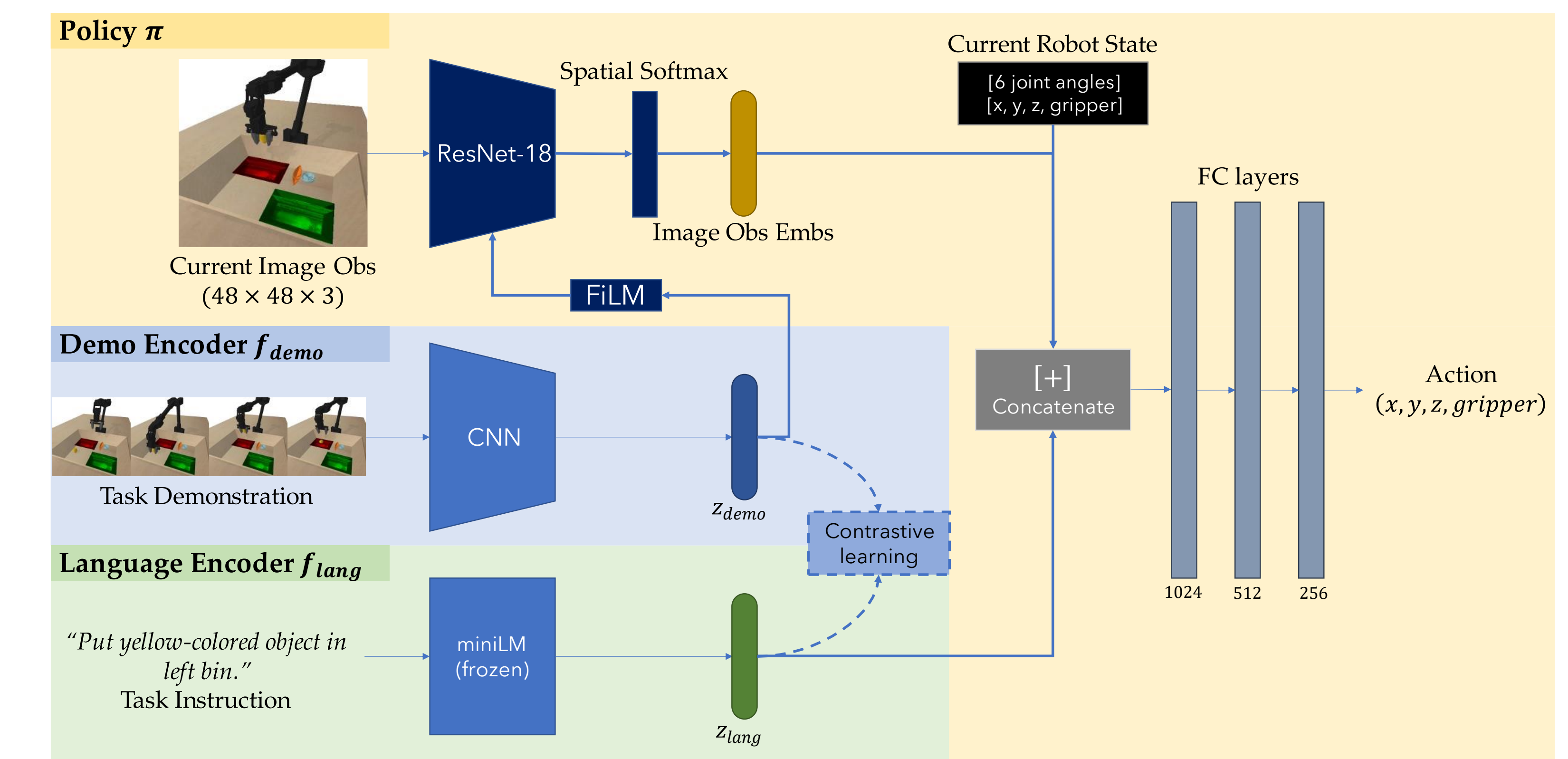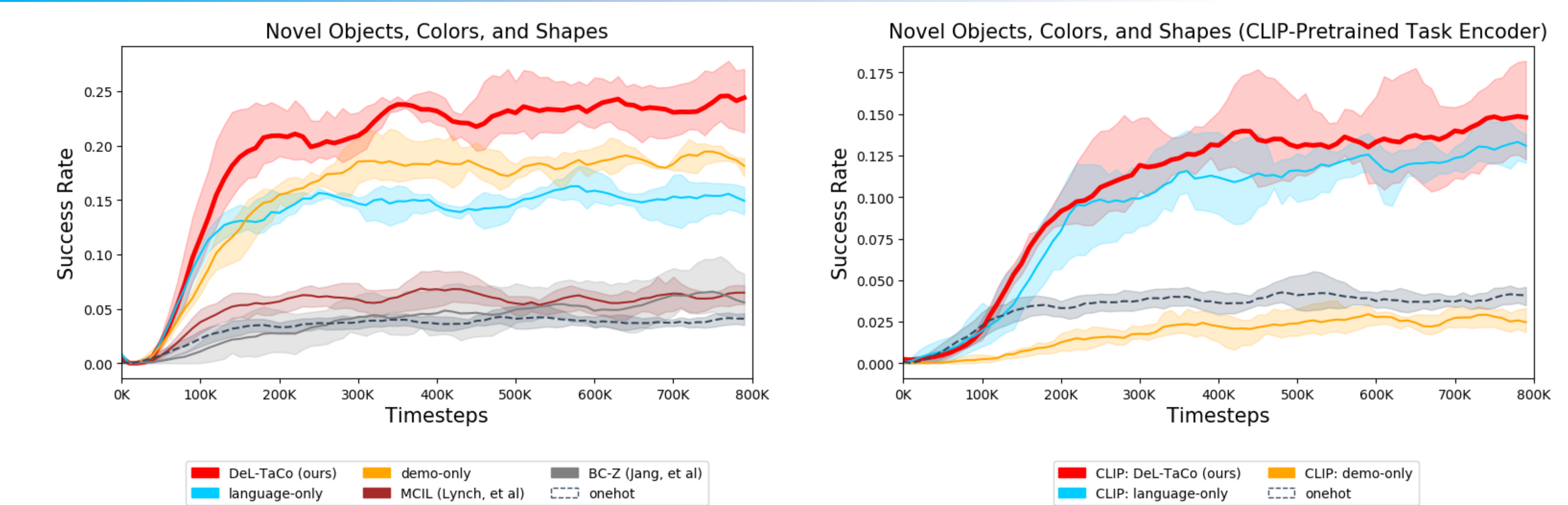Simultaneously conditions on *both* demos *and* language to teach robots new tasks, unlike prior work.

Training

Train a single multi-task policy on hundreds of tasks.

Training Buffer

Task U_1
Instruction: *"Put circular table in red bin."*
Demos:

Task U_2
Instruction: *"Put yellow object in left bin."*
Demos:

Task U_n
Instruction: *"Put vase-shaped obj. in back bin."*
Demos:

Randomly selected task U_i

Demo   Instruction
"Put vase-shaped object in back bin."

Demo encoder → Z_{demo} — Contrastive Learning — Z_{lang} ← Language encoder

Learn **demo + language** contextual associations during *training*…

RGB State → Multitask Policy → Action

Learning Algorithm

Testing

One-shot generalization to ~100 new tasks (new objects, colors, shapes).

Validation Buffer

Task V_1
Instruction: *"Put metal plate in front bin."*
Single Demo:

Task V_2
Instruction: *"Put black-white object in left bin."*
Single Demo:

Task V_m
Instruction: *"Put trapezoidal obj. in green bin."*
Single Demo:

Test task V_j

Demo   Instruction
"Put trapezoidal object in green bin."

Demo encoder → Z_{demo} Z_{lang} ← Language encoder

…To leverage **both modalities** to resolve task ambiguities during *testing*.

RGB State → Multitask Policy → Action

Architecture

Policy π

Current Image Obs (48 × 48 × 3) → ResNet-18 → Spatial Softmax → Image Obs Embs

Current Robot State [6 joint angles] [x, y, z, gripper]

FiLM

Demo Encoder f_{demo}
Task Demonstration → CNN → z_{demo} — Contrastive learning

Language Encoder f_{lang}
"Put yellow-colored object in left bin." Task Instruction → miniLM (frozen) → z_{lang}

[+] Concatenate → FC layers (1024, 512, 256) → Action $(x, y, z, gripper)$

Results

Generalization to Novel Objects, Colors, and Shapes

Novel Objects, Colors, and Shapes

Novel Objects, Colors, and Shapes (CLIP-Pretrained Task Encoder)

Legend: DeL-TaCo (ours); language-only; BC-Z (Jang, et al); MCIL (Lynch, et al); onehot
CLIP: DeL-TaCo (ours); CLIP: language-only; CLIP: demo-only; onehot

Demo Encoder	Language Encoder	Task Conditioning	Success Rate ± SD (%)
−	−	One-hot (lower bound)	4.9 ± 1.7
−	−	One-hot Oracle (upper bound)	69.3 ± 7.4
CLIP (frozen)		Language-only	14.4 ± 2.4
		Demo-only	3.5 ± 1.0
		DeL-TaCo (ours)	**15.8 ± 2.2**
−	miniLM (frozen)	Language-only	17.1 ± 2.2
CNN	−	Demo-only	20.8 ± 2.4
CNN	−	BC-Z [1] (Demo-only)	6.7 ± 2.3
CNN	miniLM (frozen)	MCIL [2] (Demo-only or Lang-only)	7.5 ± 1.2
CNN	miniLM (frozen)	**DeL-TaCo (ours)**	**25.8 ± 3.4**

[1] E. Jang, A. Irpan, M. Khansari, D. Kappler, F. Ebert, C. Lynch, S. Levine, C. Finn. "BC-Z: Zero-Shot Task Generalization with Robotic Imitation Learning." CoRL, 2021.
[2] Corey Lynch and Pierre Sermanet. "Language Conditioned Imitation Learning Over Unstructured Data." RSS, 2021.

How many Demonstrations is Language Worth?

Task Conditioning	Demo-only					DeL-TaCo (ours)
# demos per test-task finetuned on	0	10	25	50	100	0
Success Rate (%)	20.8	23.4	24.6	26.1	32.9	25.8
± SD (%)	±2.4	±1.8	±2.5	±2.6	±2.5	±3.4