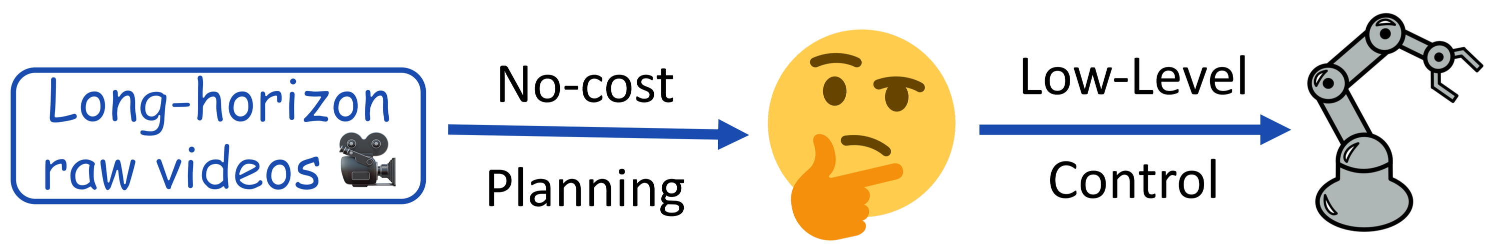
Universal Visual Decomposer: Long-Horizon Manipulation Made Easy

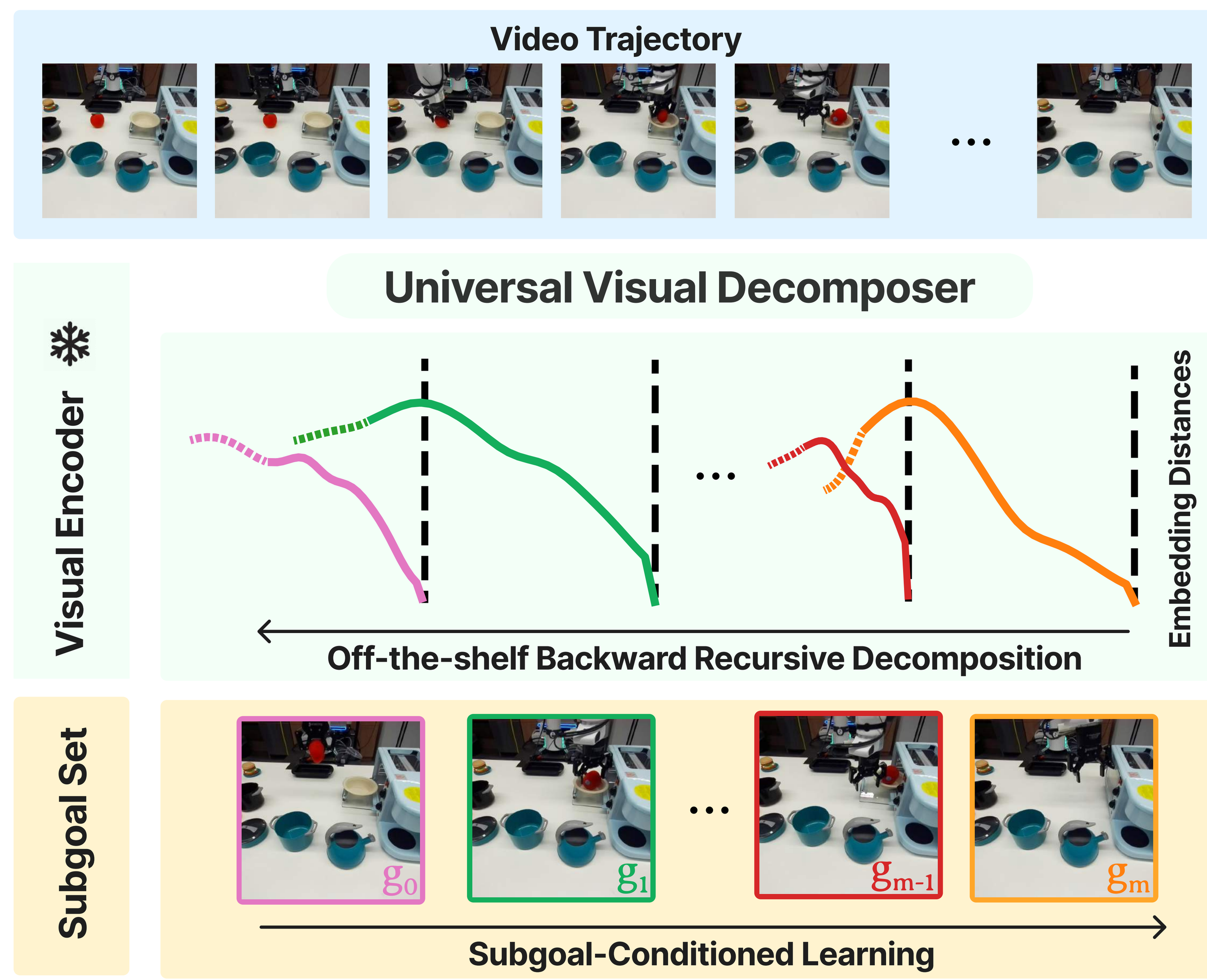
Zichen “Charles” Zhang*, Yunshuang Li*, Osbert Bastani, Abhishek Gupta, Dinesh Jayaraman, Yecheng Jason Ma†, Luca Weihs†

Motivation

- Multi-stage long horizon manipulation is challenging.
- Previous works may use LLM or Generation model for task or goal decompositions.

But can we have an *off-the-shelf* method - *NO extra data, training, cost, or task knowledge required*? **YES!**





- *Decomposes* long-horizon tasks into meaningful sub-stages
- Enables *OOD generalization* in Sim & Real
- Solves long-horizon multi-stage manipulation using RL from vision *without* reward engineering
- Applicable to **ANY** visuomotor policy training

Universal Visual Decomposer (UVD)

UVD discovers subgoals by detecting phase shifts in the embedding space of the pre-trained representation.

- Pseudocode

Algorithm: Universal Visual Decomposer

Init: frozen visual encoder $\phi, \tau = \{o_0, \dots, o_T\}$

Init: set of subgoals $\tau_{goal} = \{\}, t = T$

While t *not small enough:*

$\tau_{goal} = \tau_{goal} \cup \{o_t\}$

$o_{t-n-1} := \text{argmax}_{o_h} d_\phi(o_h; o_t) < d_\phi(o_{h+1}; o_t), h < t$ (Eq. 3)

$t = t - n$

End

- UVD - Policy Learning

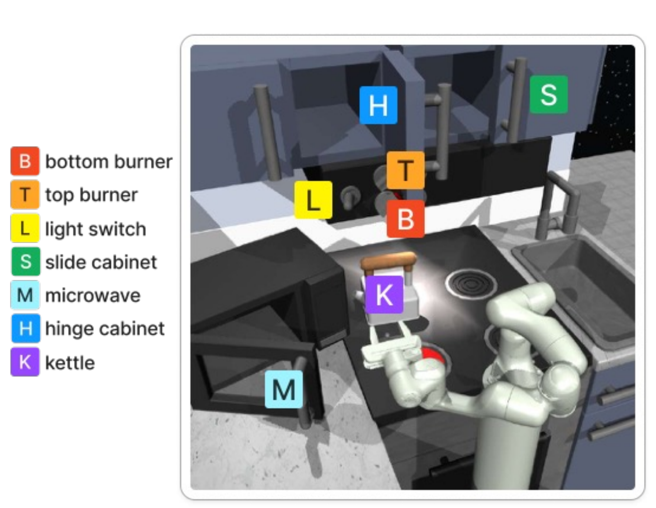
With UVD's recursive decomposition ⏪, it can build upon any standard *goal-conditioned* visuomotor policy training ⏩

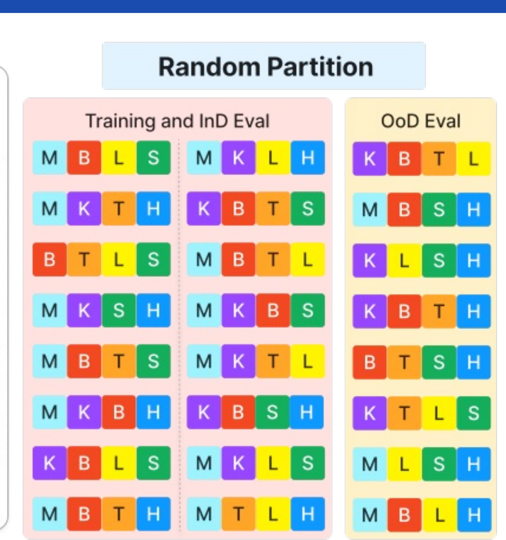
- UVD - RL Rewards

Progressive and optimally monotonic *goal-embedding distance difference* using UVD subgoals.

$$R(o_t, o_{t+1}; \phi, g_i) := d_\phi(o_t; g_i) - d_\phi(o_{t+1}; g_i)$$

Experiments





FrankaKitchen Simulation

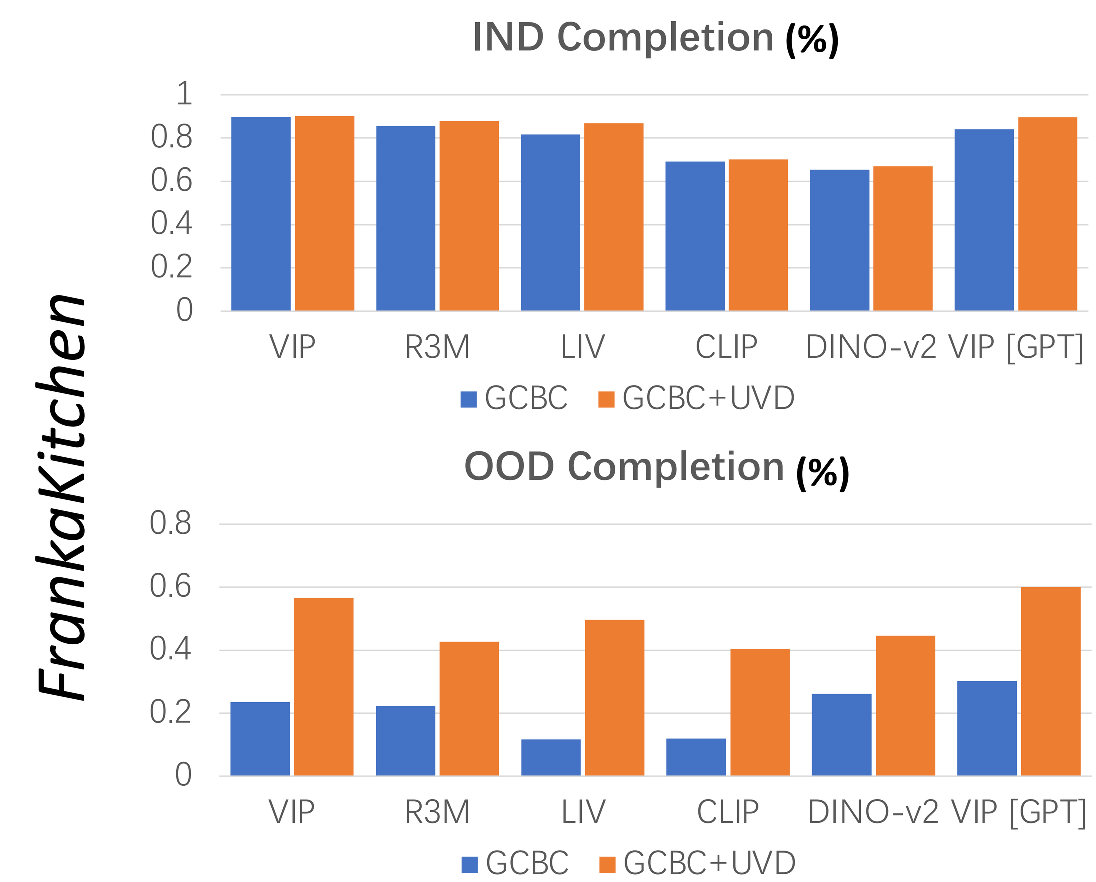
- *4 out of 7 objects are manipulated in an arbitrary order*
- **Random train-eval partitions:** While training on 16 sub-task seqs, the reset of 8 seqs are for compositional OOD evaluation.

Real-world experiments

- *Apple-in-Oven*: picking apple, placing apple in the bowl, pushing the bowl into the oven, closing the oven.
- *Fries-and-Rack:* picking a bowl, pour fries out of the bowl, placing the bowl on the rack
- *Fold-Cloth:* diagonal fold, quarter fold, eighth fold, etc
- Initial/Intermediate novel states for OOD evaluation.

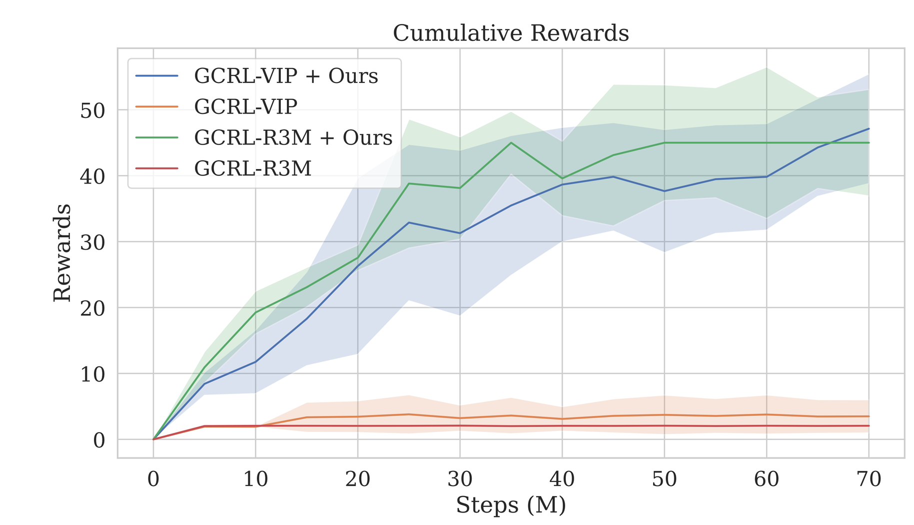
Results

• UVD for Imitation Learning



- Plug-and-play UVD significantly boosts *compositional OOD performance* across *various* visual backbones and policies.

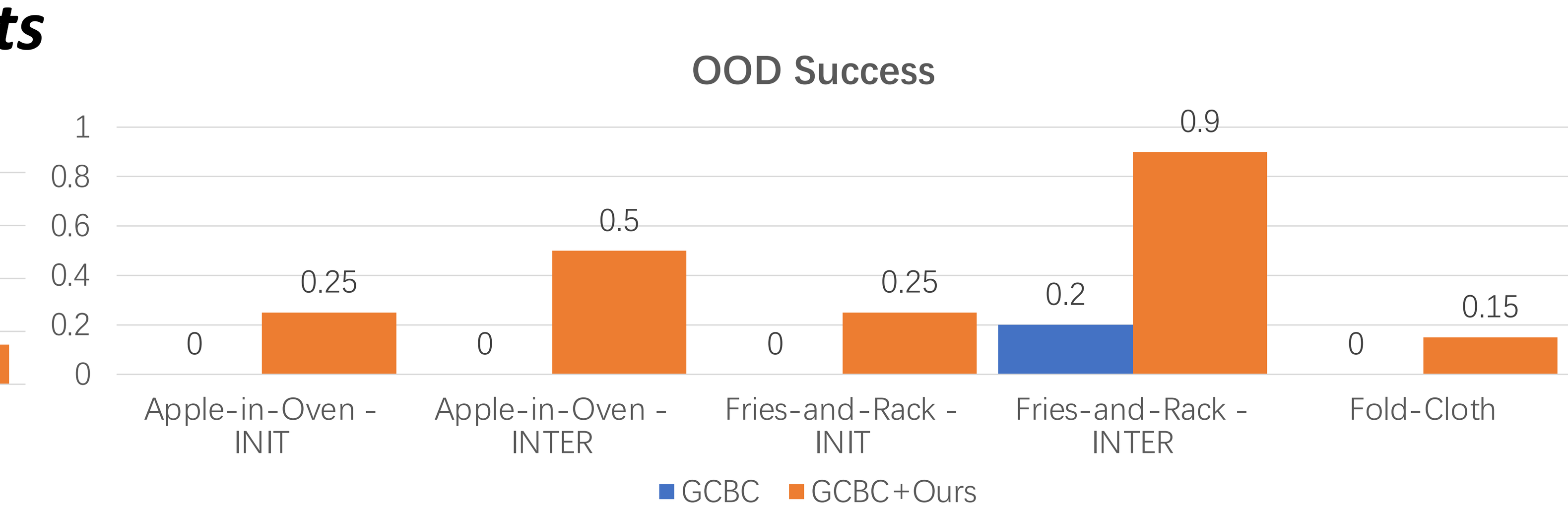
• UVD for Reinforcement Learning



Method	Success	Completion
GCRL-VIP	0.0 / 0.0	0.09 / 0.25
GCRL-VIP + Ours	**0.65 / 1.0**	**0.75 / 1.0**
GCRL-R3M	0.0 / 0.0	0.09 / 0.25
GCRL-R3M + Ours	**0.649 / 1.0**	**0.82 / 1.0**

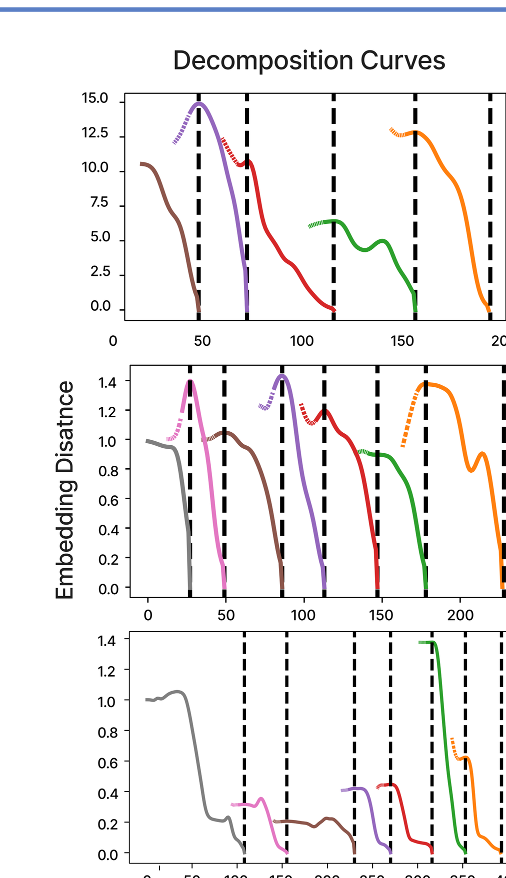
- UVD's detection of monotone trends in feature space allows us to provide a naturally progressive reward signal

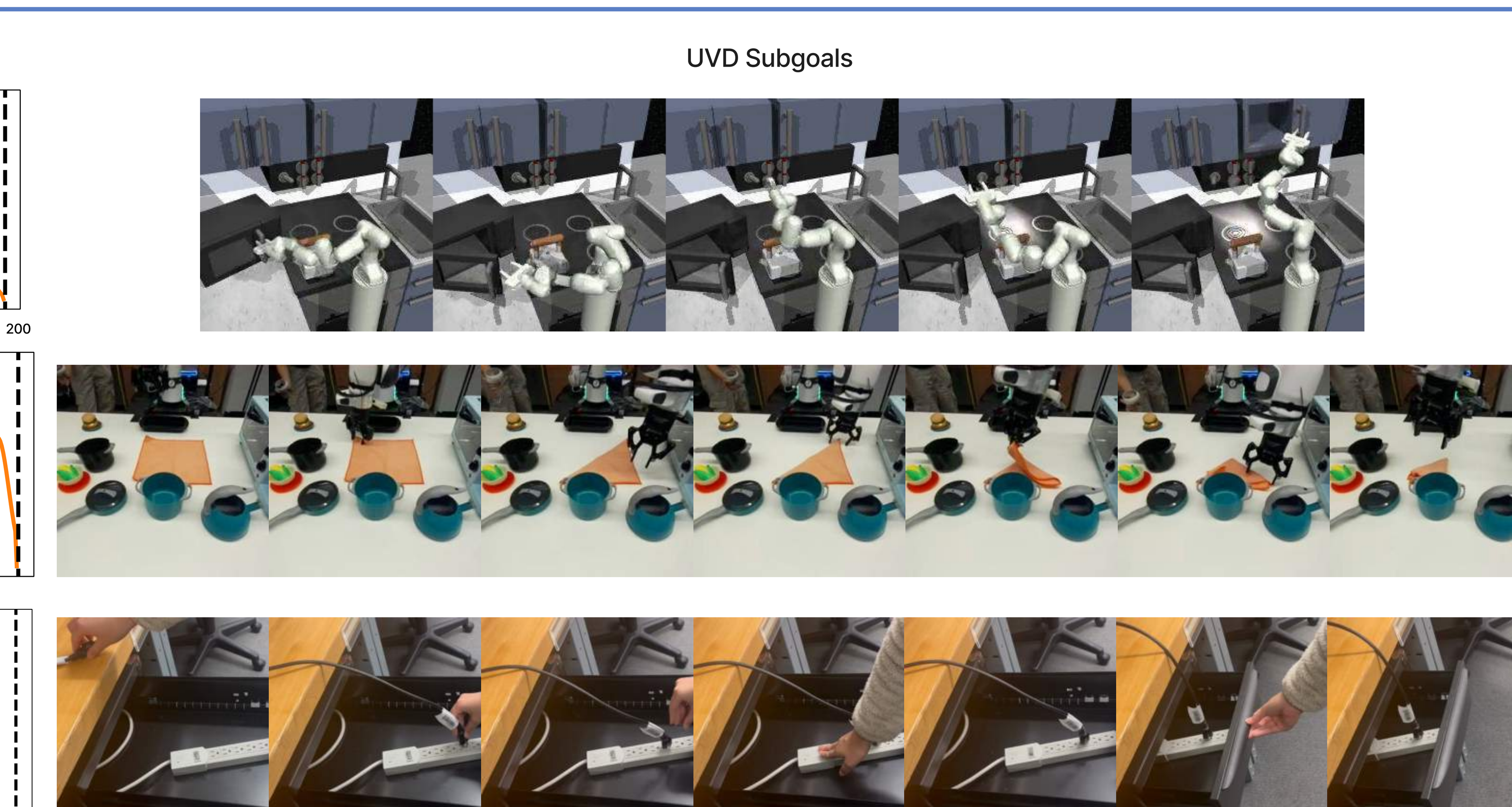
Real-World Experiments



- Baseline w/o UVD all *failed* for the initial (INIT) and intermediate (INTER) OOD states.
- UVD can also enable agents to auto-skip sub-stages preemptively finished by humans (robust to the intervention) and can reset to redo certain stages during deployment (recovery).

Visualizations





Millisecond level runtime for UVD!

	# frames	Load	Preprocess	UVD
FrankaKitchen	226.9	0.023	0.155	0.0023
In the wild	698	1.011	0.450	0.011

Check more visualizations, and online hosted demo at

zcczhang.github.io/UVD