POSTECH

INDUSTRIAL AND MANAGEMENT ENGINEERING, POSTECH

Towards Personalized Household Robots

Sangwoo Mo

Assistant Professor
POSTECH IME

Today: The Same Robot for Everyone

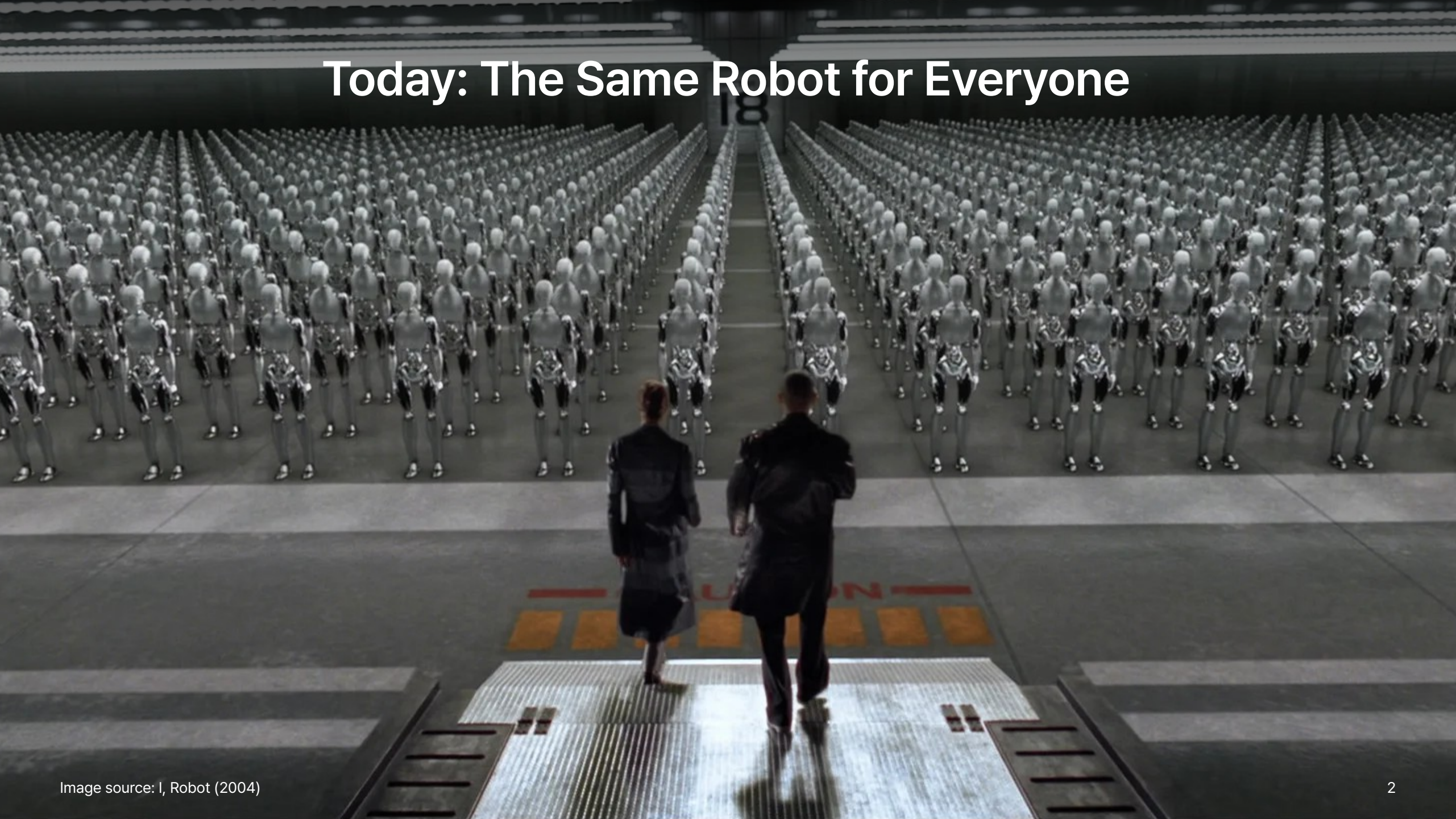


Image source: I, Robot (2004)

Tomorrow: Personalized Robots



Image source: I, Robot (2004)

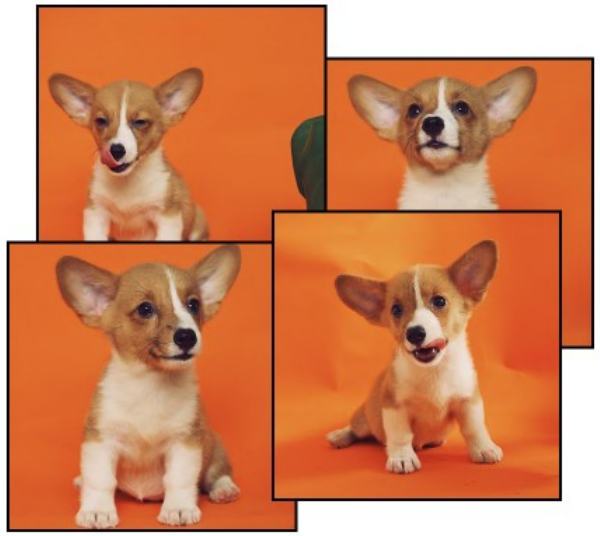
First Step: Personalized Object Manipulation



Image source: AI generated

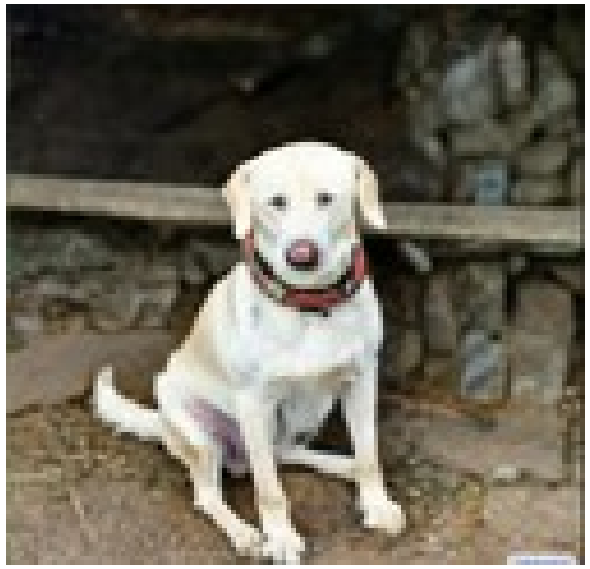
Prior Work: Personalized Content Creation

Reference photos



My dog

Generic model



A dog

Personalized model



My dog

Image source: Ruiz et al., DreamBooth: Fine Tuning Text-to-Image Diffusion Models for Subject-Driven Generation, CVPR 2023

Prior Work: Personalized VLMs

Is <bo> in this photo?

Reference photo



<bo>: my dog

Generic model

“I’m not sure what you’re referring to.”

Personalized model

“Yes, <bo> is in this photo.”

Image source: Nguyen et al., Yo’LLaVA: Your Personalized Language and Vision Assistant, NeurIPS 2024

Question: Can We Personalize VLAs?

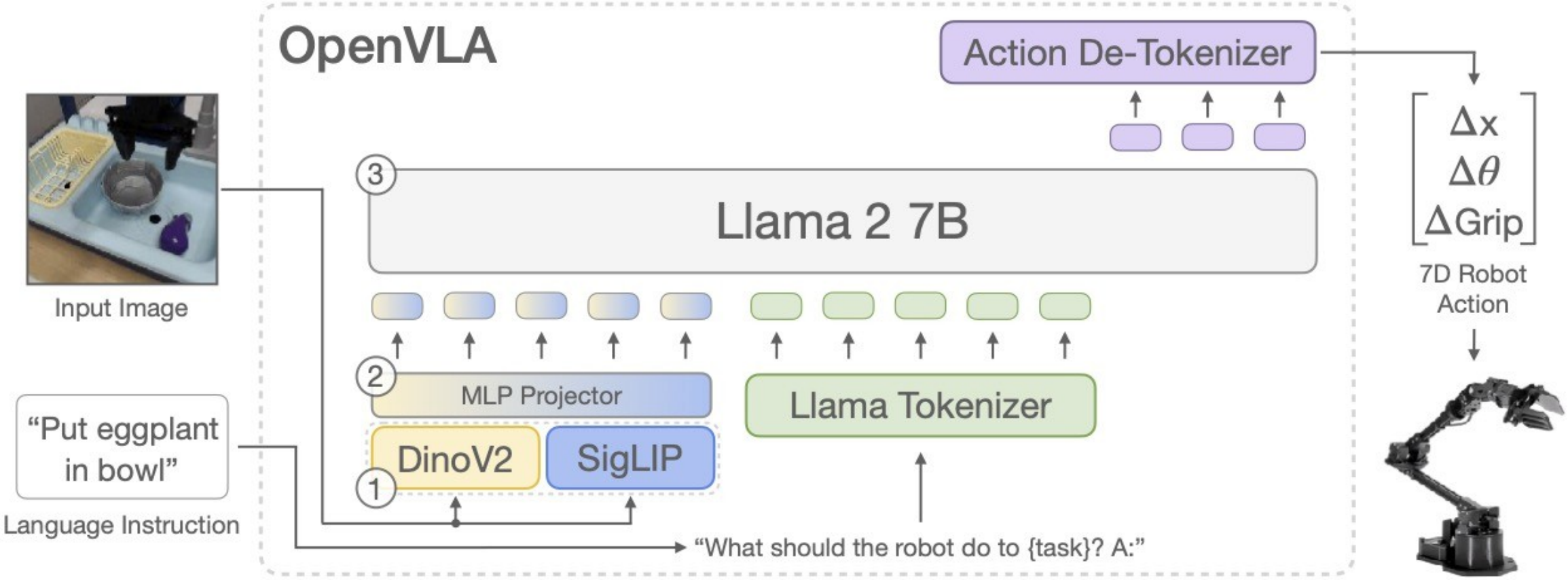


Image source: Kim et al., OpenVLA: An Open-Source Vision-Language-Action Model, CoRL 2024

Bring My Cup! Personalizing Vision-Language-Action Models with Visual Attentive Prompting



Sangoh Lee



Sangwoo Mo†



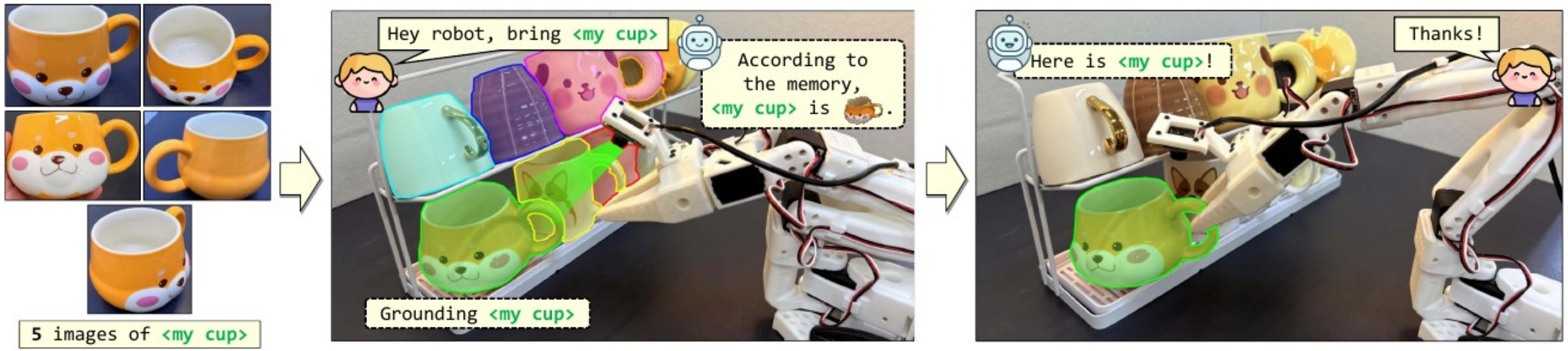
Wook-Shin Han†

† Equal advising

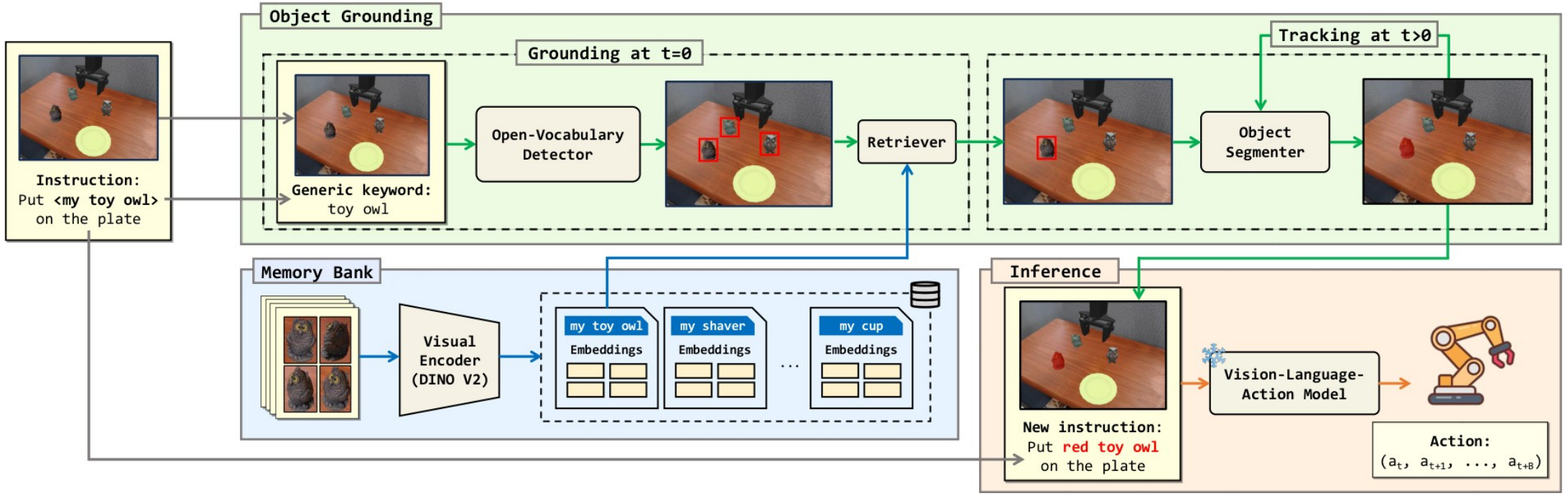
ICML 2026

Image source: Authors' profile pages / POSTECH

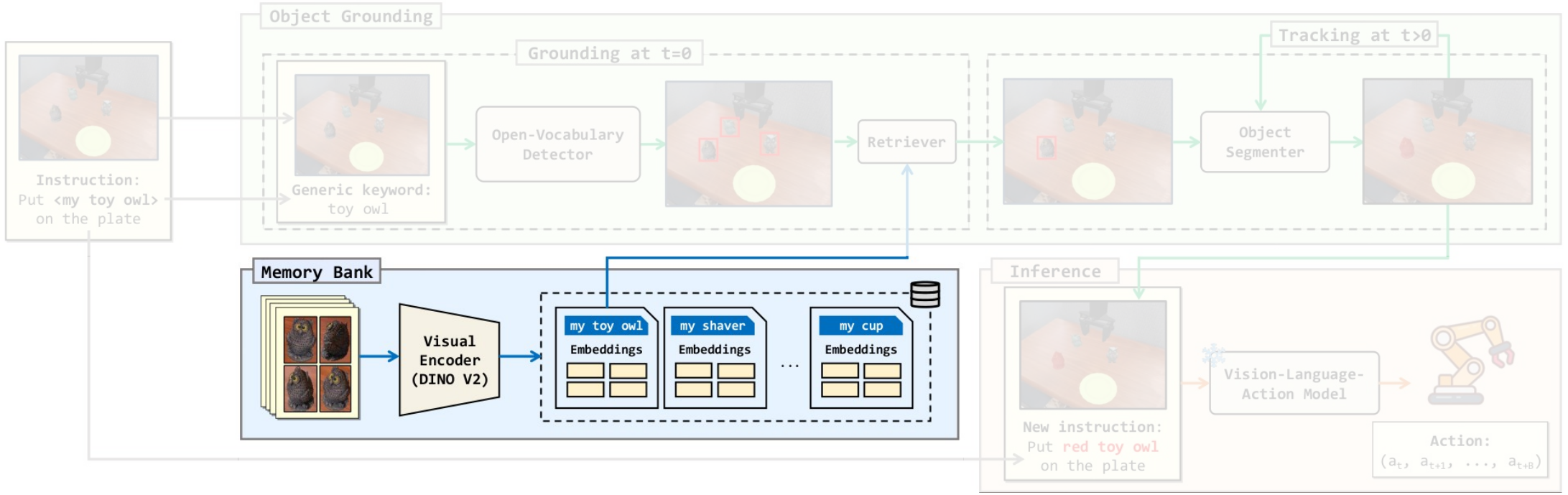
Bring a Cup vs. Bring My Cup



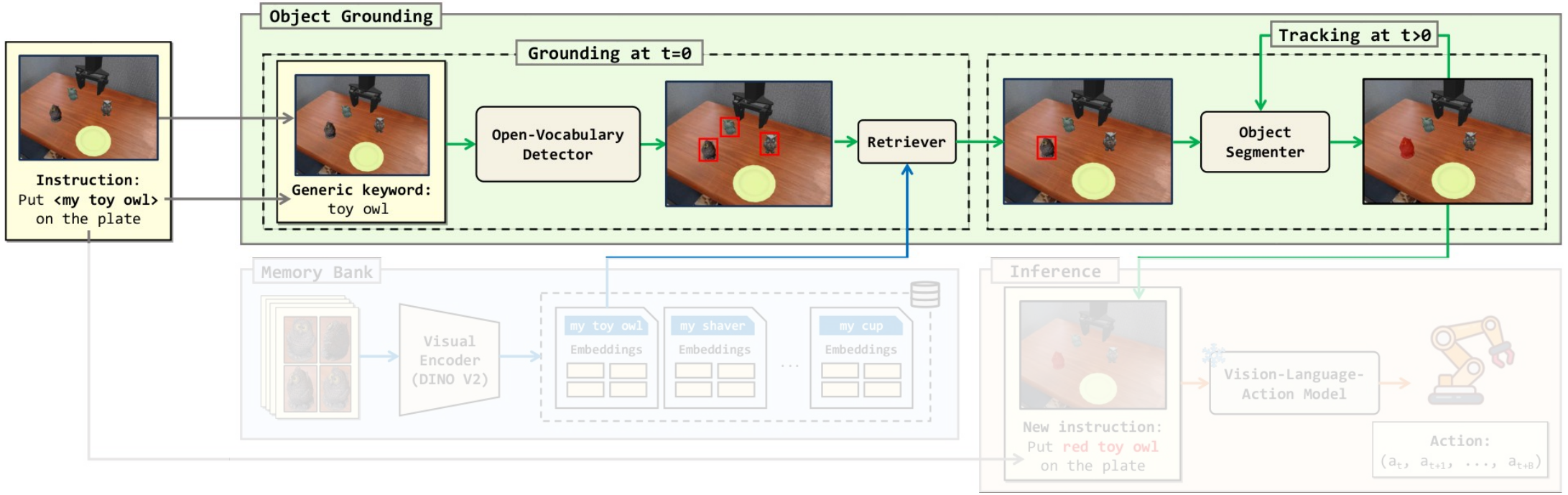
Method: Visual Attentive Prompting (VAP)



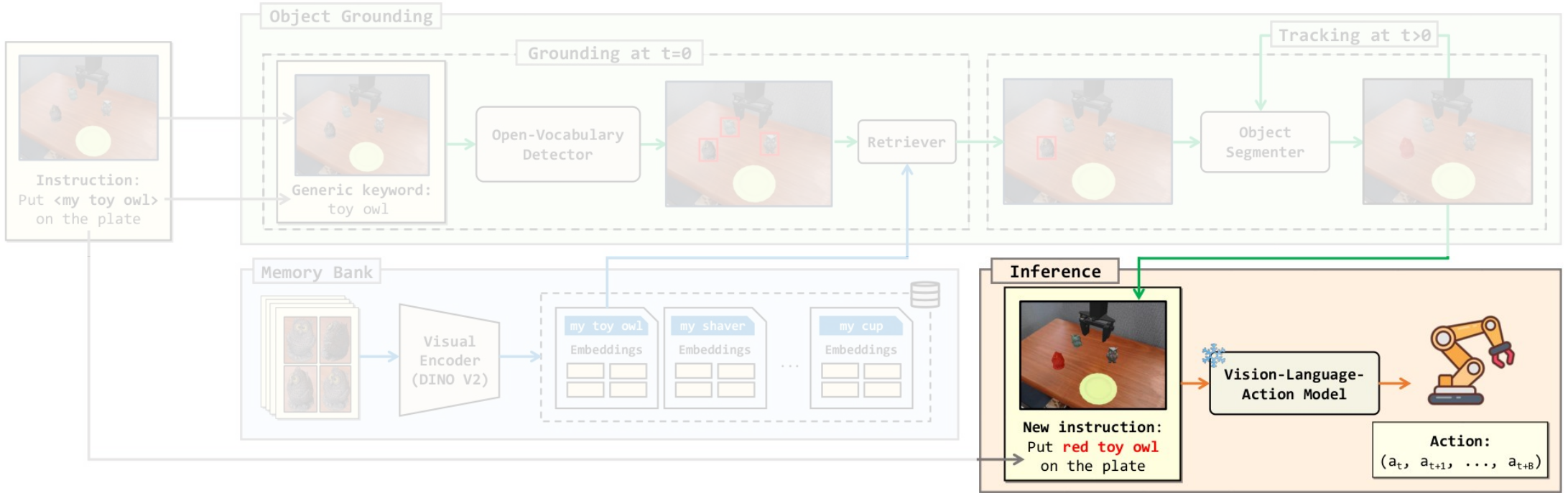
Step 1: Memory Bank



Step 2: Object Grounding



Step 3: Visual Prompting



Baselines

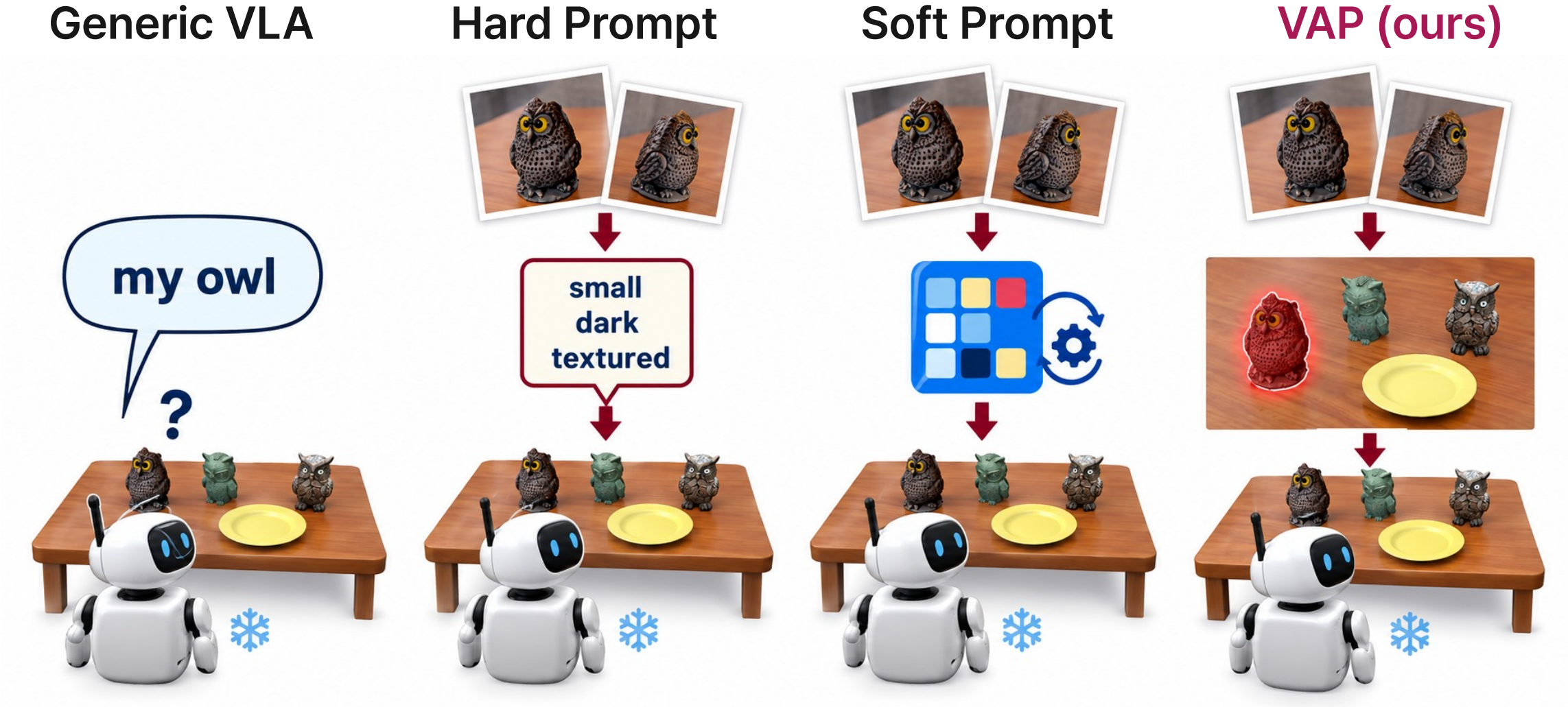


Image source: AI generated

Simulation Results

“Pick my pen holder”



Simulation Results

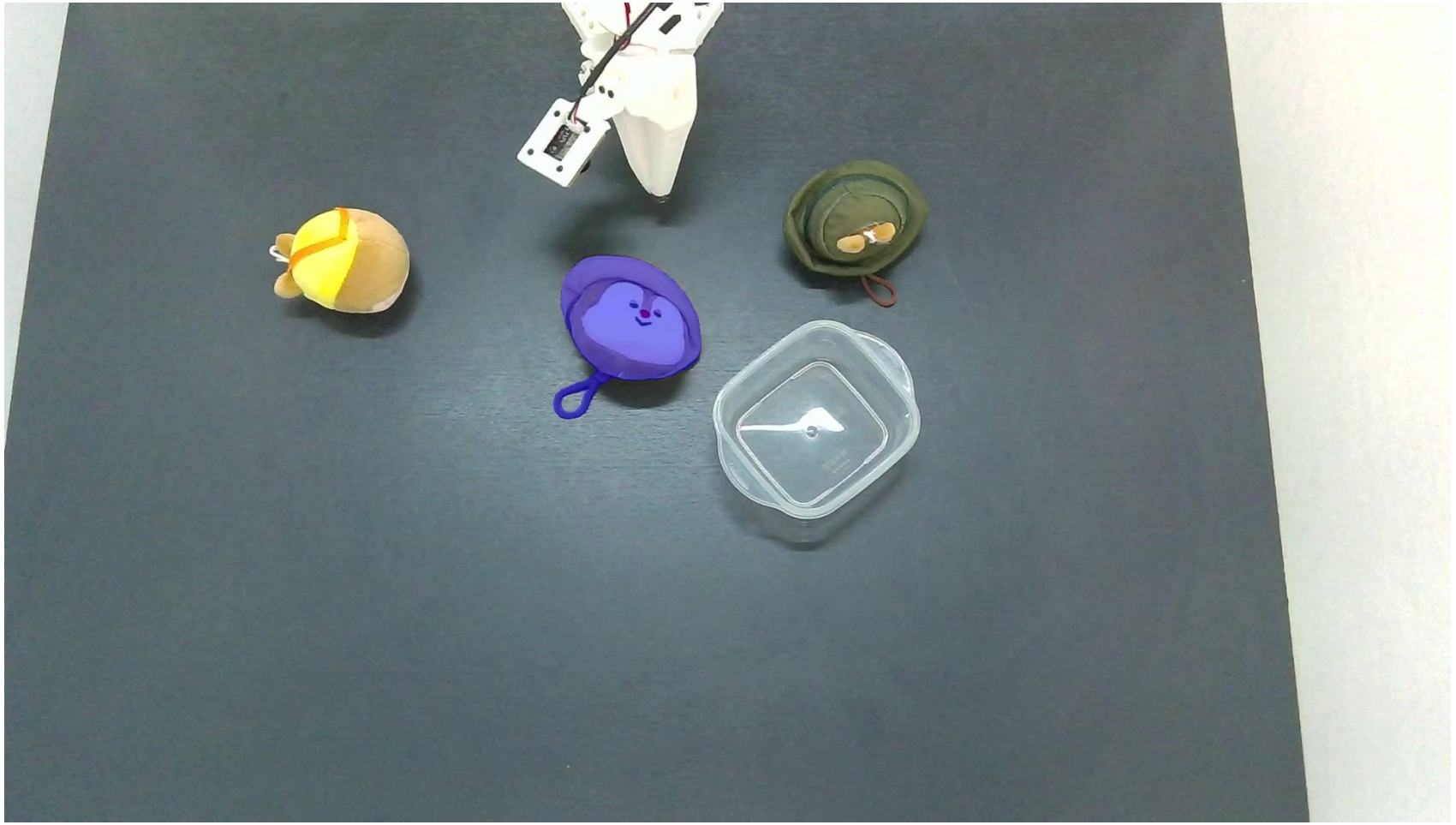
Method	Task 1: Pick my pen holder		Task 2: Move my bottle near coke can	
	CMR (%)	SR (%)	CMR (%)	SR (%)
Track 1: Visual Matching with Unseen Personalization Objects				
π_0	10.5	8.5	50.4	48.6
π_0 + Soft Prompt	33.1	24.0	60.0	51.7
π_0 + Hard Prompt (Short)	11.0	8.9	66.7	57.5
π_0 + Hard Prompt (Long)	11.2	9.1	64.5	58.8
π_0 + VAP	**89.2**	**60.3**	**83.2**	**75.0**
Track 2: Variant Aggregation with Unseen Personalization Objects				
π_0	17.4	10.2	49.2	41.3
π_0 + Soft Prompt	31.8	26.4	57.7	44.7
π_0 + Hard Prompt (Short)	21.6	14.4	71.5	54.2
π_0 + Hard Prompt (Long)	17.3	12.6	67.7	51.3
π_0 + VAP	**87.3**	**58.2**	**72.4**	**62.5**

Pen holder success

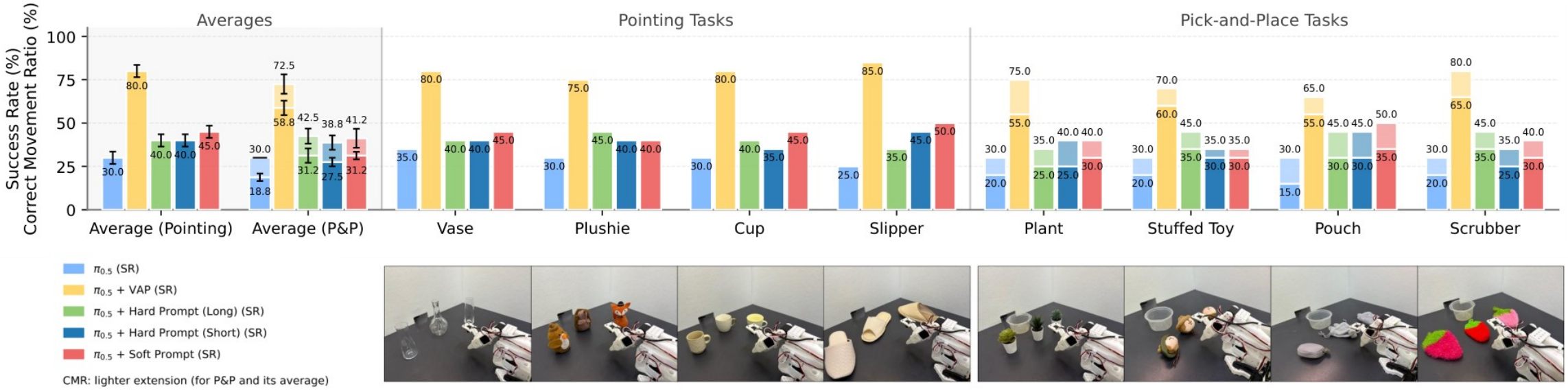
8.5% → 60.3%

Real-world Results

“Put my stuffed toy into the plastic bowl”



Real-world Results



Selection success

30.0% → 80.0%

Pick-and-place success

18.8% → 58.8%

Future Direction I: Towards Whole-body Control







Manipulation → **Mobile Manipulation** → **Whole-body Control**

Image source: Clearpath Robotics, Hello Robot, Figure

Future Direction II: Personalized Perception → Action



Which one is mine?

How should I handle mine?

Image source: AI generated

Future Direction III: Personalization → Contextualization



Image source: RoboCasa

One More Thing: Accessibility

Cannot type. Cannot speak.



How can they ask a robot for help?

Image source: AI generated

EOD: Towards Personalized Household Robots



Image source: AI generated