Active Learning ⋆

Zhifeng Han [*] **Claire Walton** [**]

[*] *University of Texas at San Antonio, San Antonio, TX 78249 USA (e-mail: zhifeng.han@utsa.edu).*
[**] *University of Texas at San Antonio, San Antonio, TX 78249 USA (e-mail: claire.walton@utsa.edu).*

Abstract: As AI agents (LLMs and Robots) accumulate "crossed paths" of knowledge, internal world models become prone to "pre-convicted" illusions—internal biases that override reality. This framework proposes a Focus Mode that uses Active Learning to dynamically balance internal predictions with direct environmental interactions, ensuring the agent remains grounded in the "Working Sense."

Keywords: LLM, robotic, Reinforcement Learning, world model, Active Learning, Focus Mode, Working Sense

1. INTRODUCTION

lo We use world model to predict the future based on historical data.Hafner et al. (2019) world models allow for planning and behavior learning given only small amounts of real world interaction. In the past the robot research focussed on the dynamic of world or the dynamic of the robot itself. However, in this paper, we focus on the dynamic of the environment and dynamic of robot body framework. We emphasize the importance of understanding the environment's dynamics and the robot's body dynamics for effective planning and behavior learning. This approach can lead to more efficient and adaptive robots that can handle a wide range of tasks in various environments.

2. BACKGROUND

2.1 Main challenges

3. ARCHITECTURE

We propose a novel architecture called Active Learning and Teaching Architecture (ALTA) that integrates three key components:

- **Active Learning Module**: This module enables the robot to actively select informative samples for learning, reducing the need for extensive labeled data.
- **Teaching Module**: This component facilitates the robot's ability to teach and transfer knowledge to other agents or systems.
- **Adaptive Planning Module**: This module allows the robot to adapt its planning strategies based on new information and changing environments.

The overall architecture can be represented as follows:

human demonstrate task as input to Active Learning Module output to Teaching Module then output to Adaptive Planning Module finally robot performs task. As human pick and place object the robot learn from human demonstration through Active Learning Module, then Teaching Module help robot to generalize the task to different objects, finally Adaptive Planning Module enable robot to plan the pick and place task efficiently.

The teaching module leverages techniques from Schaal et al. (2003) to facilitate effective knowledge transfer. we utilize dreamer to learn the world model and plan the task through adaptive planning module. The active learning module is designed to minimize human intervention by selecting the most informative samples for learning, as discussed in Ebert et al. (2018).

the integration of these components allows ALTA to effectively learn from limited data,Hafner et al. (2019) teach other agents, and adapt to dynamic environments,

3.1 XArm Visual Pick and Place

While the UR5 robot is a high performance industrial robot, the XArm is an accessible low-cost 7 DOF manipulation, which we control at approximately 0.5 Hz. Similar to Section 3.2, the task requires localizing and grasping a soft object and moving it from one bin to another and back, shown in Figure 6. We connect the object to the gripper with a string, which makes it less likely for the object to get stuck in corners at the cost of more complex dynamics. The sparse reward, discrete action space, and observation space match the UR5 setup except for the addition of depth image observations.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work the author(s) used NotbookLM and ChatGPT in order to help English as secondary language speaker. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

⋆ This work was supported in part by the U.S. National Science Foundation (NSF) under Award No. 1826086.

REFERENCES

Ebert, D.D., Van Daele, T., Nordgreen, T., Karekla, M., Compare, A., Zarbo, C., Brugnera, A., Øverland, S., Trebbi, G., Jensen, K.L., et al. (2018). Internet-and mobile-based psychological interventions: applications, efficacy, and potential for improving mental health. *European Psychologist*.

Hafner, D., Lillicrap, T., Ba, J., and Norouzi, M. (2019). Dream to control: Learning behaviors by latent imagination. *arXiv preprint arXiv:1912.01603*.

Schaal, S. et al. (2003). Dynamic movement primitives: Learning attractor models for motor behavior. In *IEEE Robotics and Automation*.