A Robot Learning System for Viewpoint-aware Legible Motion Planning

Carnegie Mellon University

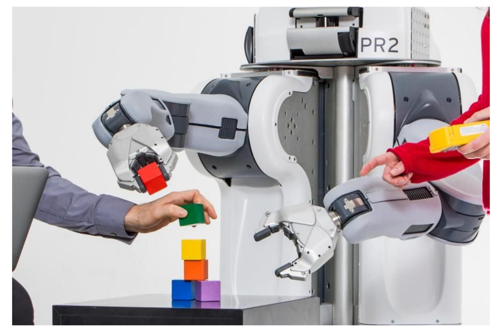
Workshop on Learning for Assistive Robotics

Khai Nguyen*, Yi-Hung Chiu*, Prakrit Tyagi*, Saurav Kambil* and Inseung Kang

Department of Mechanical Engineering, Carnegie Mellon University

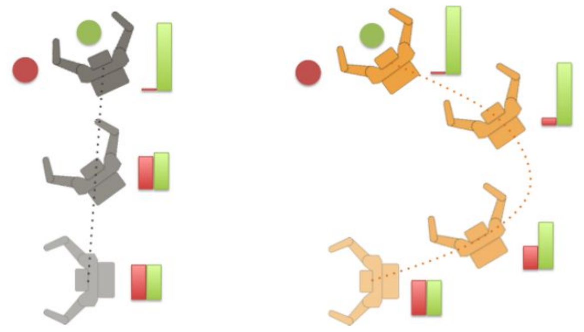
ROBOTICS SCIENCE AND SYSTEMS

Legibility: the key to effective assistive robotics and shared autonomy

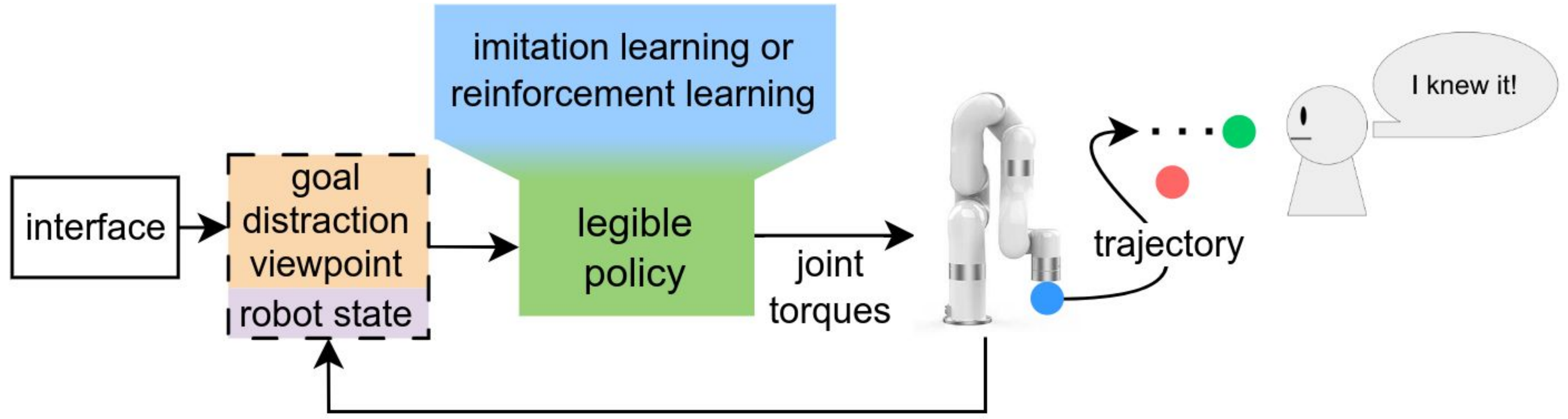


Legibility is crucial for efficiency and trust in environments where robots collaborate closely with humans, e.g. manufacturing and healthcare.

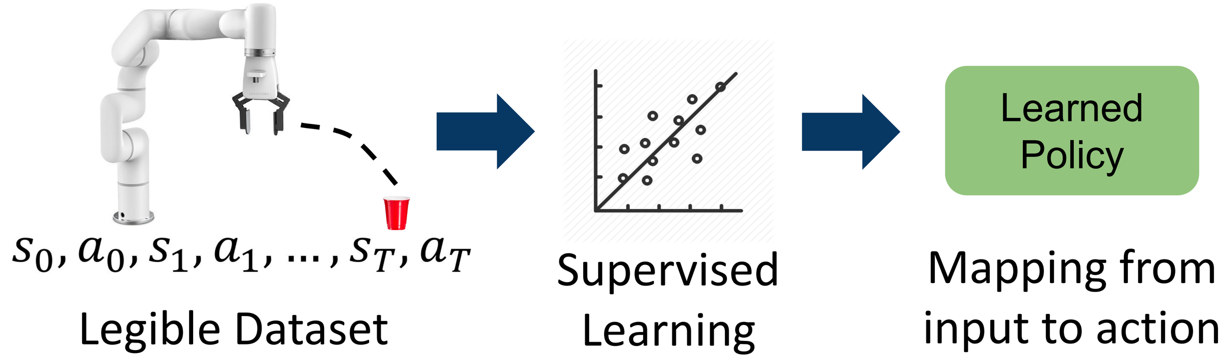
It involves designing robotic motions that are legible or *intent-expressive*, allowing humans to predict the robot's goal.



A novel robot learning system that integrates observer's viewpoint to generate adaptive, legible motions



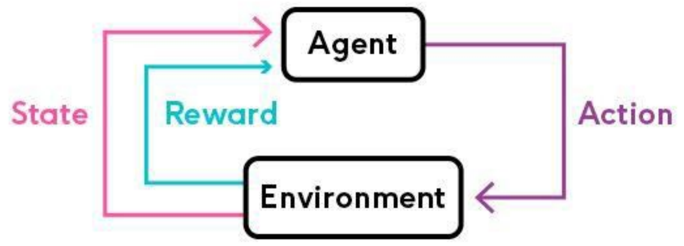
Imitation learning (IL) performs supervised learning on real legible human demonstrations (images A & B)



Legible Human Demonstrations/Dataset. A single human operator manually controls the robot to reach the goal in a way they consider legible.

Training Procedure. We perform standard supervised learning on the collected legible observation-action data. The policy converges within 1000 iterations.

Reinforcement learning (RL) optimizes with a legibility reward by interacting with the environment (see C & D)

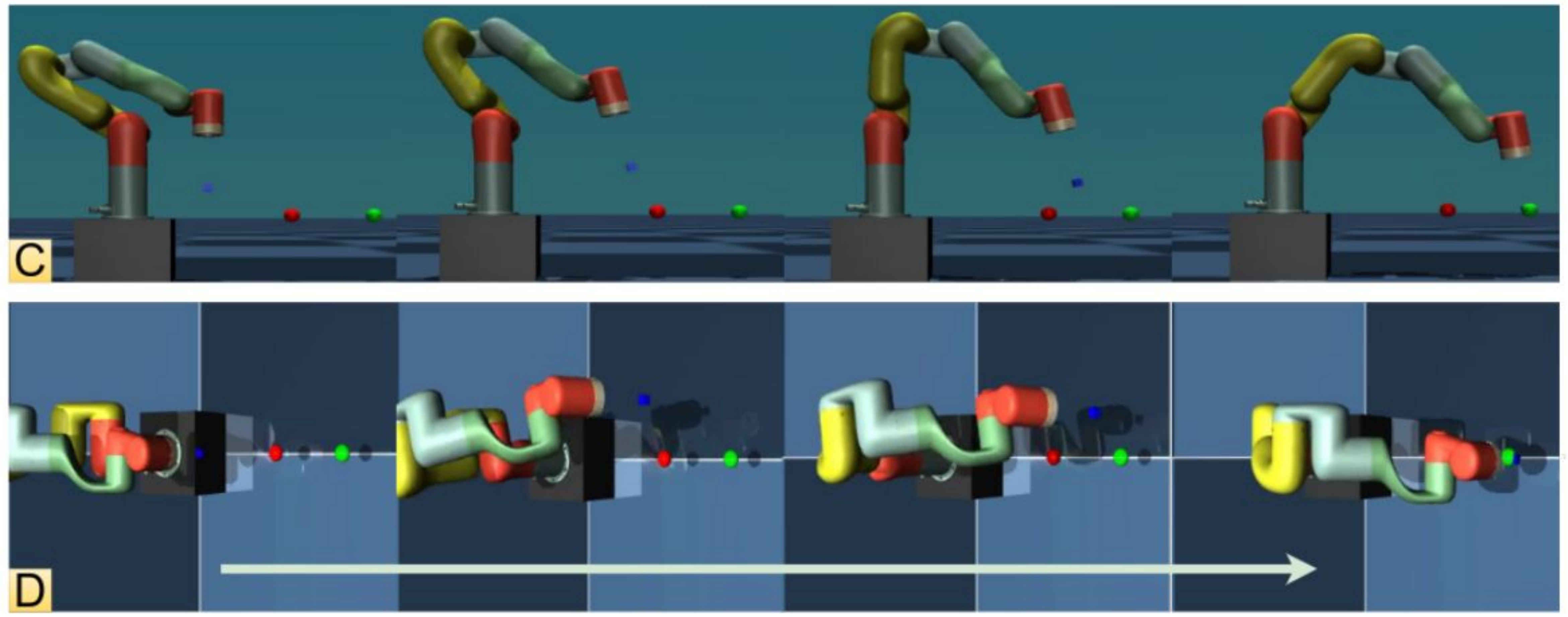


Legibility Reward. We define a reward function:

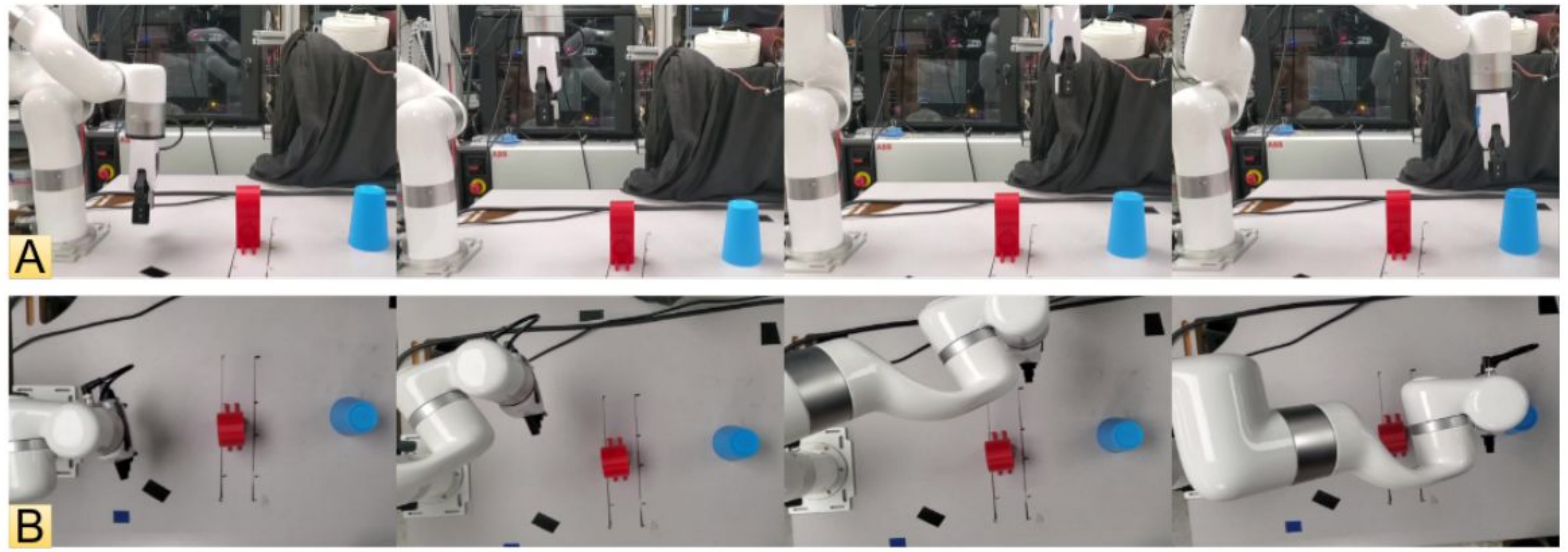
$$r = -\alpha_1 d_{\text{goal}} + \alpha_2 g(d_{\text{distract}}, \beta_1) - \alpha_3 e_{\text{rot}} + \alpha_4 g(z, \beta_2)$$

$$g(x, y) = -1 \ \text{if} \ x \leq y, \ \ \text{otherwise} \ 0$$

Training Procedure. We use MuJoCo as the physics simulator and the soft actor-critic (SAC) algorithm to optimize the network's parameters. Policies converge after roughly 200K training steps.

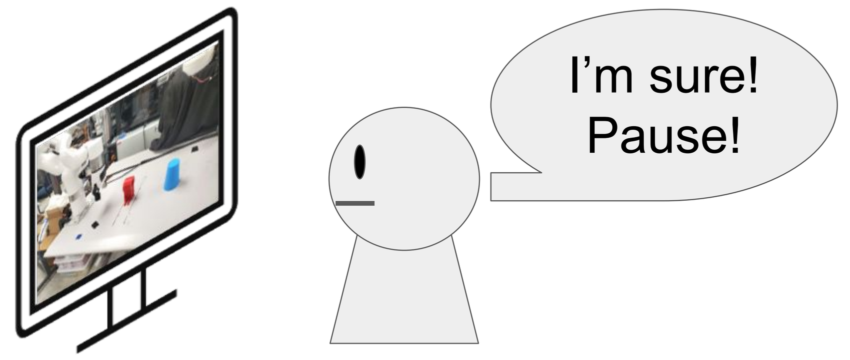


Both RL in simulation (C, D) and IL in real world (A, B) exhibit expected behaviors.



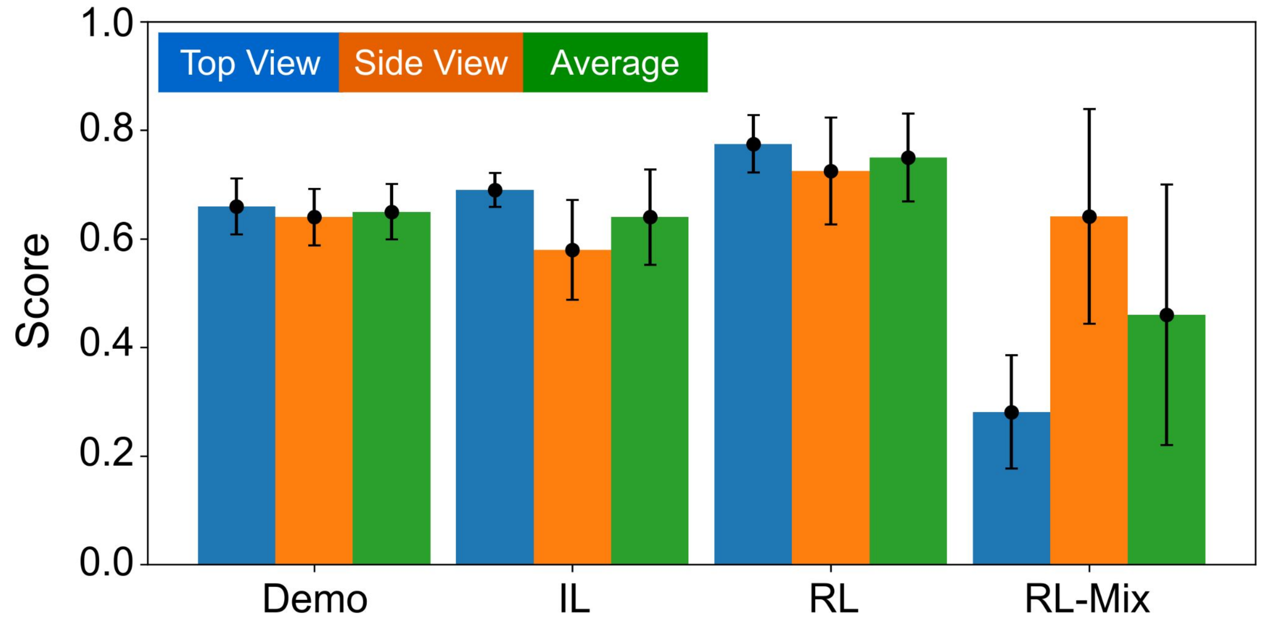
We conduct a human study with a legibility score system to objectively assess the performance

Ten participants watched eight videos and paused when confident about the robot's intent and reported which object the robot was targeting.



Legibility score:

$$\text{correct} \cdot \left(1 - \frac{t_{\text{pause}}}{t_{\text{total}}}\right)$$



- The IL agent achieved nearly the same average score as its demos ⇒ IL learns implicit objectives from data but limited by its quality.
- The RL policy scored approximately 15% higher than demos ⇒ RL optimizes behaviors based on carefully engineered rewards, outperforming human demonstrations on average.
- The RL-Mix scenario had high variability, the top view video scored poorly at 0.281 ⇒ The legibility of trajectories can vary significantly depending on the observer's viewpoint.

Discussion and Future Work

✅ Environment understanding boosts legibility.

✅ Legibility is subjective and requires proper evaluation.

🦾 Next Step: Bridge the sim-to-real gap for RL and integrate richer sensory inputs to enable full autonomy.