快速訓練RL模型夾取HSKs經驗分享

台達電子(江蘇)

Summer intern 葉行遠

Outlines

1. 問題簡介
2. 技術難處
3. 技術亮點
4. 未來展望

Task

- Pick HSKs while avoiding with pins.
- Place them with the correct pose in the proper position on the tray.

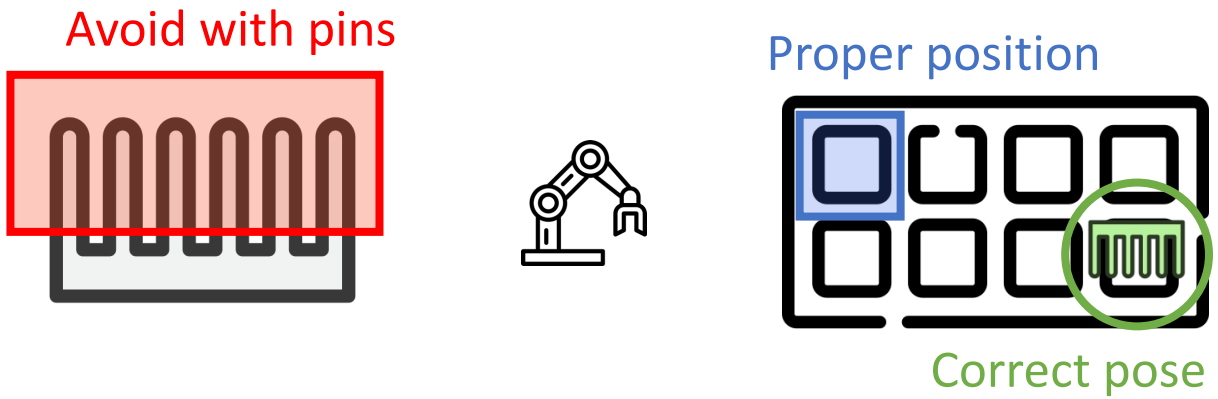
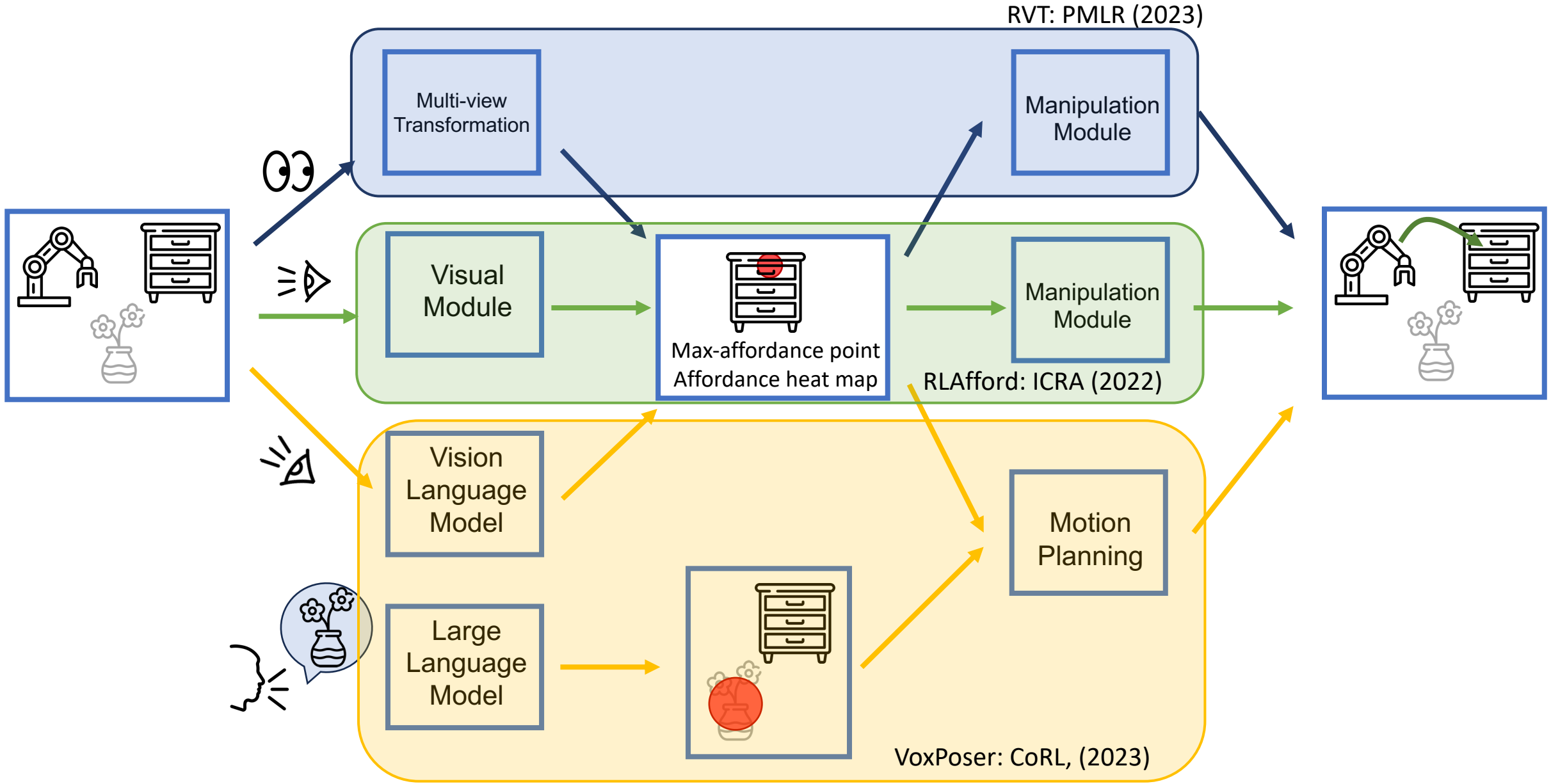


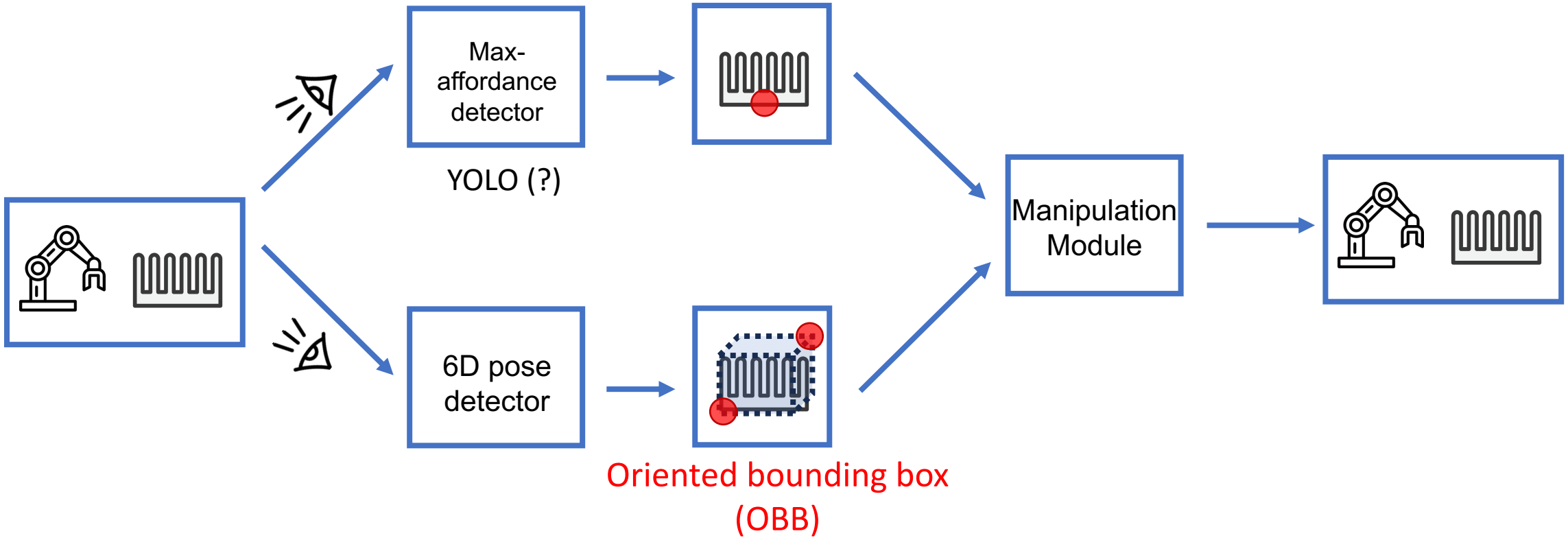
Illustration captured from 上海工博會

The models listed below, RVT, RLAfford, VoxPoser, are all End-to-End.

Literature Review

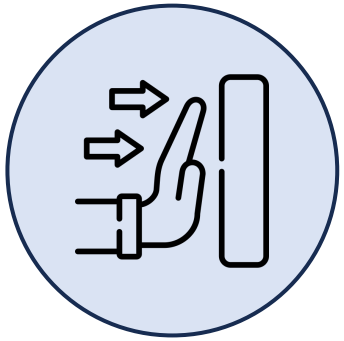


Model architecture



Delta collaborative robots: D-Bot

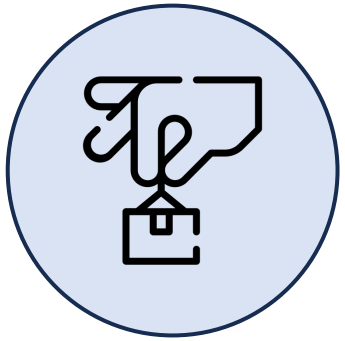
Done



Push

- Toy model

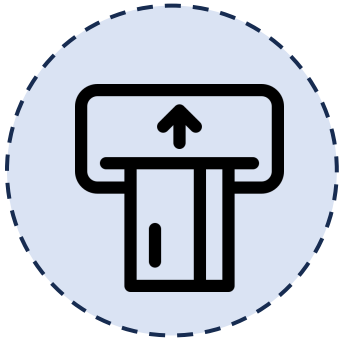
Doing



Pick & Place

- Avoid touch pins
- Place with correct poses

To-Do



Insert

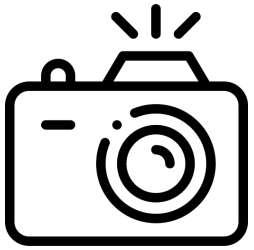
- Insert HSKs

Captured from Delta website

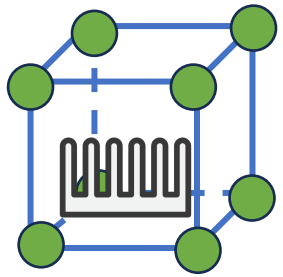
Computation inefficient (Main modification)

By using pose information without using image, we can capture the pin locations.

	GPU	Training time
RLAfford	x	x
VoxPoser	x	x
RVT	8 * V100	1 days

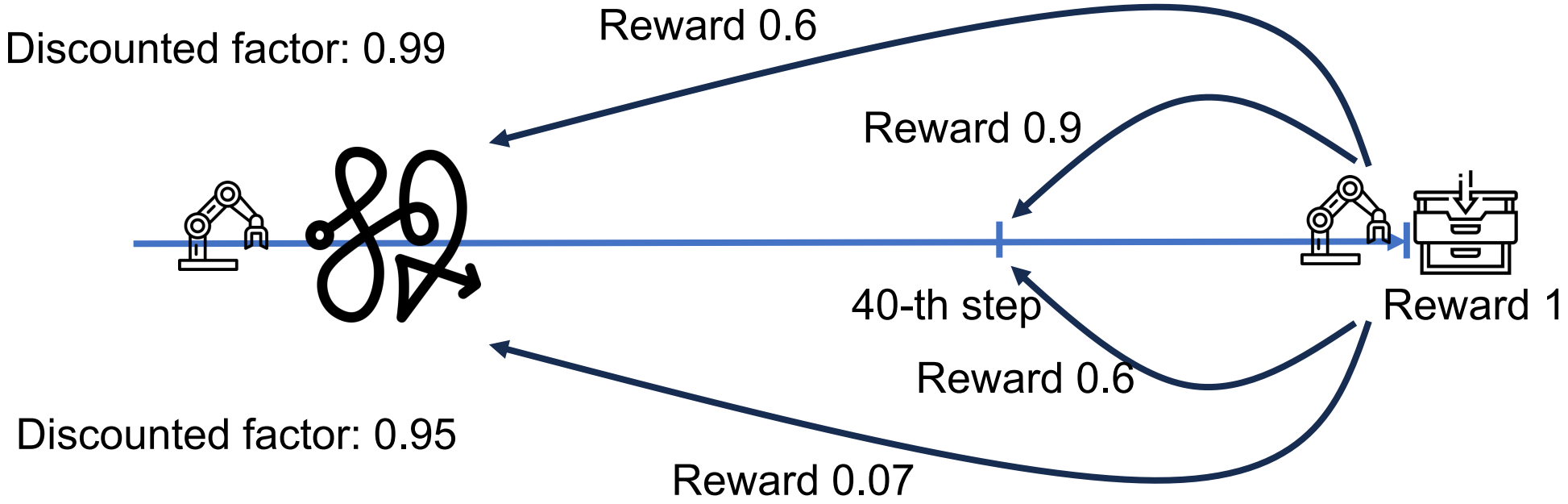


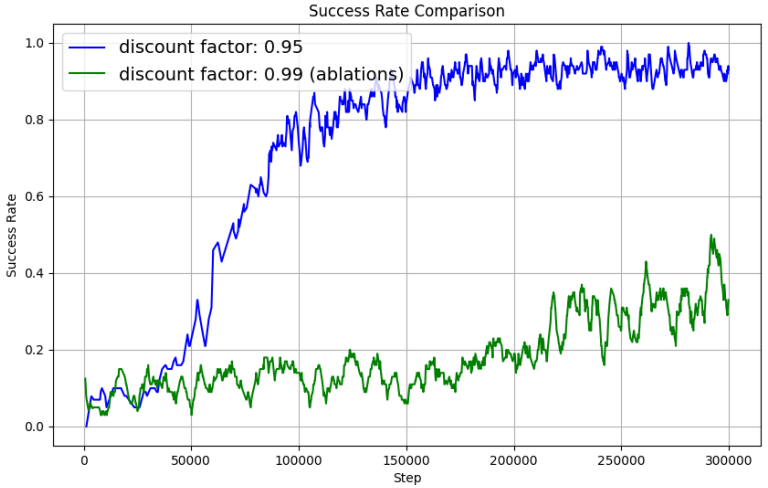
	GPU	Training time
Ours	1 * A1000	8 hours



Sample inefficient

$$n \propto \frac{1}{\varepsilon^2(1-\gamma)^7}$$





Yang, Jin, Wang, Wang, Jordan, *NeurIPS*, (2020)
Yeh, Chang, Yueh, Wu, Bernacchia, Vakilli, *AISTATS*, (2023)

Andrychowicz et al., *NeurIPS*, (2017)

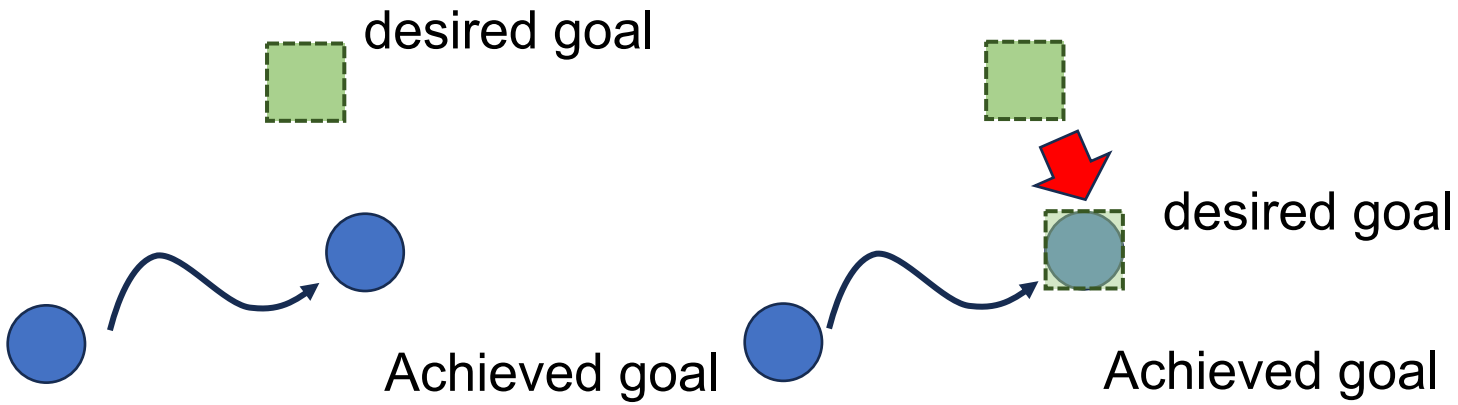
Hindsight Experience Replay (HER)

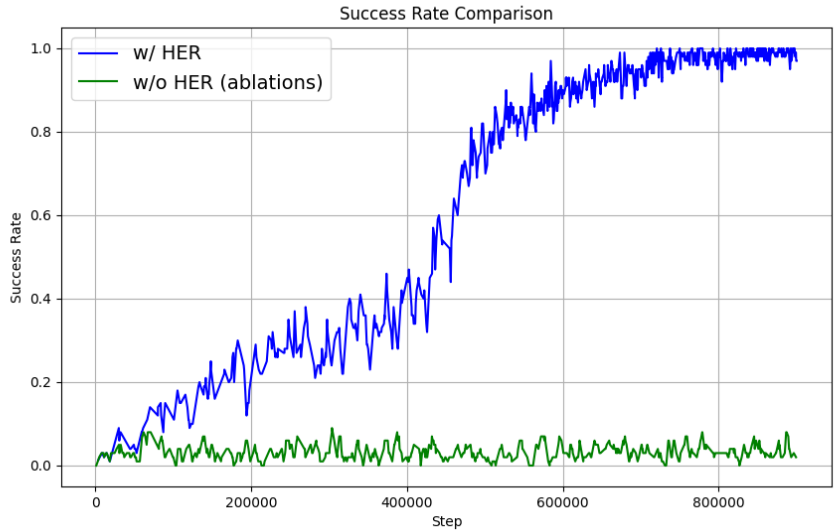
Advantages:

- **Improved Sample Efficiency**: HER allows agents to learn from past experiences, even when they **fail** to achieve their original goals, making it particularly effective in sparse reward environments.
- **Versatility**: HER is applicable to a wide range of tasks, especially those involving long-term planning and exploration, such as robotics and complex decision-making problems.

Disadvantages:

- **Limited Effectiveness in Certain Scenarios**: While HER is beneficial in sparse reward settings, its advantages may diminish in environments with dense rewards.





Dataset Aggregation (DAgger)

Not for disclosure!

Results

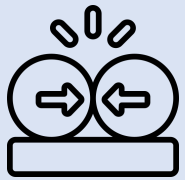
	Push	Pick & Place	Pick & Place HSKs
RLAfford	X	46.5%	X
VoxPoser	X	90%	X
RVT	100.0 ±0.0 %	88.0 ±5.7 %	X
Ours	X	97.6 ±0.6 %	83.2 ±1.6 %

Zhao et al., SSCI, (2024)

Sim-to-Real

Randomness Physical Constant

- Randomness friction constant
- Randomness elasticity constant



Simulation Environment

- Capture images in real world to create a comprehensive simulation model for training.
- Use domain adaptation to align real world image with simulation characteristics.

Transfer in real-world

- Collect the real-world data and transfer our model in such data.

References

1. A. Goyal, J. Xu, Y. Guo, V. Blukis, Y.-W. Chao, D. Fox, (2023), RVT: Robotic View Transformer for 3D Object Manipulation, *PMLR*.
2. Y. Geng, B. An, H. Geng, Y. Chen; Y. Yang, H. Dong, (2023), RLAfford: End-to-End Affordance Learning for Robotic Manipulation, *ICRA*.
3. W. Huang, C. Wang, R. Zhang, Y. Li, J. Wu, Li F.-F., (2023), VoxPoser: Composable 3D Value Maps for Robotic Manipulation with Language Models, *CoRL*.