Learning Robotic Manipulation from Videos Priors via Task-Agnostic Reward Function

By Ahad Jawaid, Daniel George, and Mason Cushing

Problem

Object manipulation is hard

Another Problem

Teleoperation takes a lot of **time** or use **custom hardware**

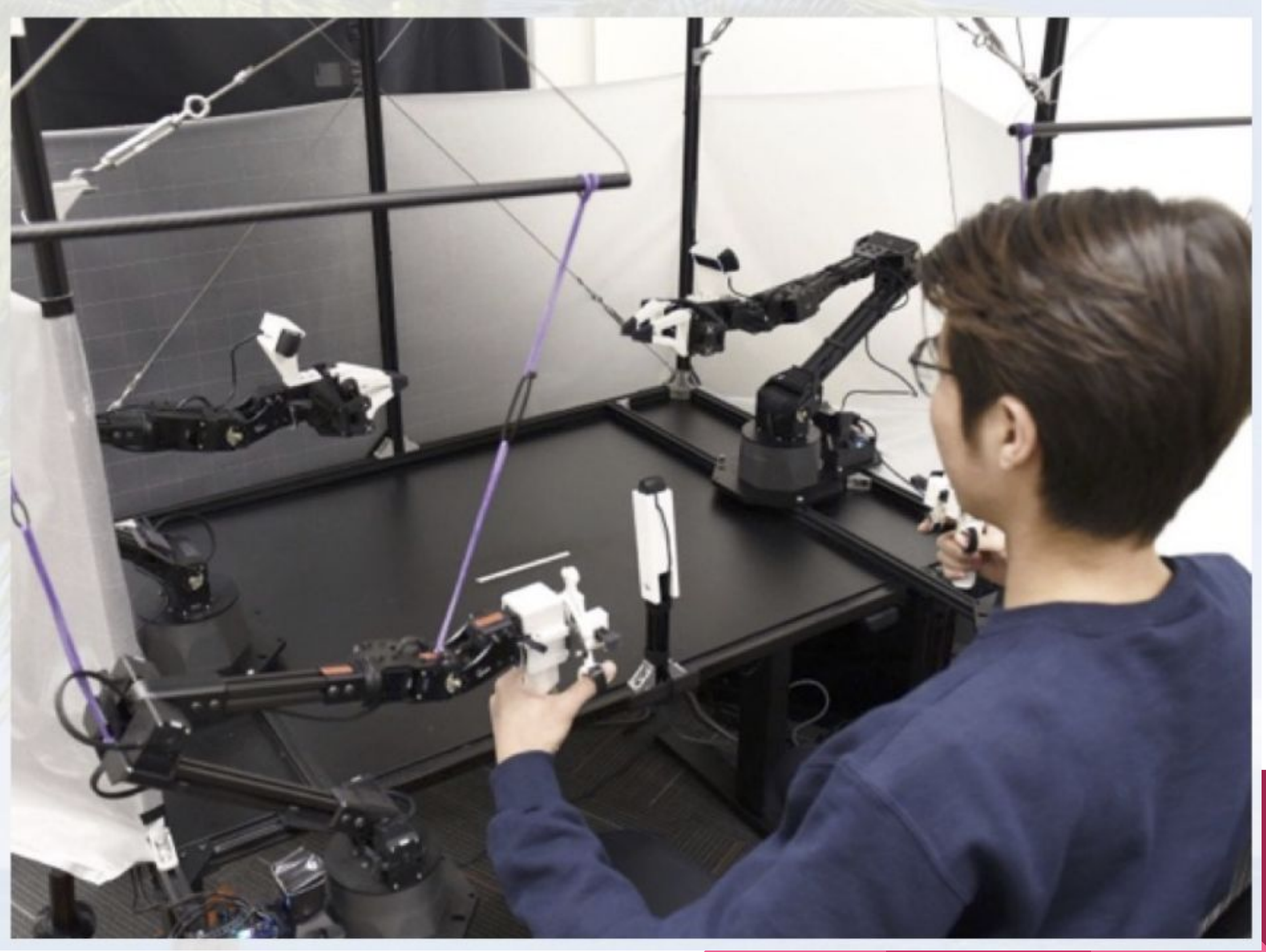
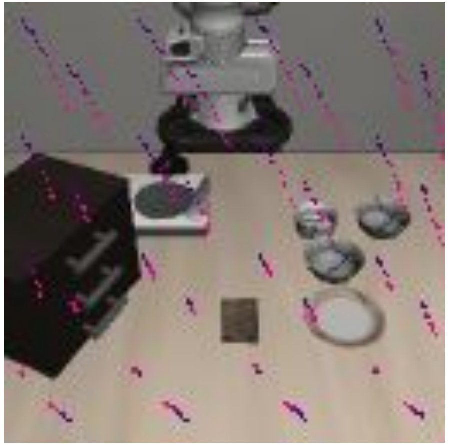
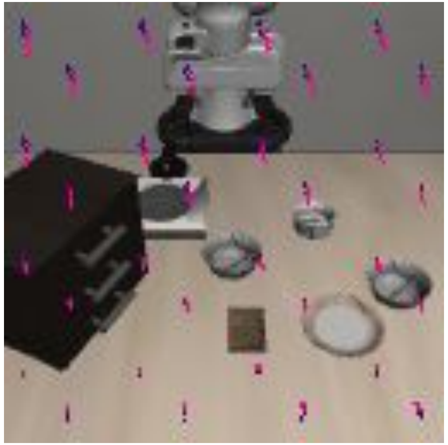


Image from [1]

Aims

- Learn to manipulate objects according to a goal specified in language without teleoperated demonstrations.



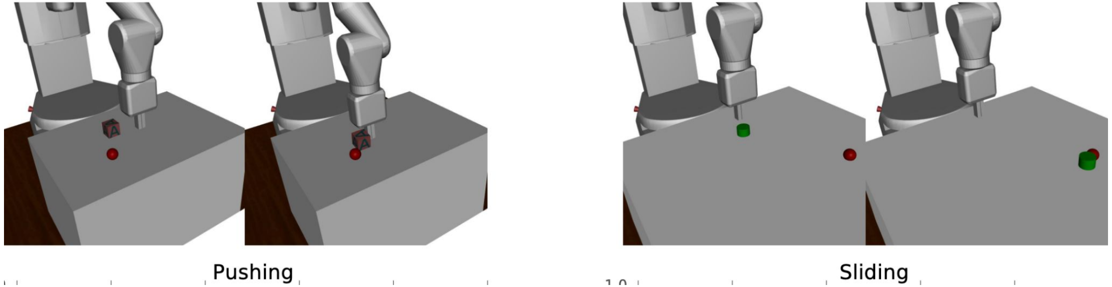


Task: pick up the black bowl between the plate and the ramekin and place it on the plate

Task: pick up the black bowl from table center and place it on the plate

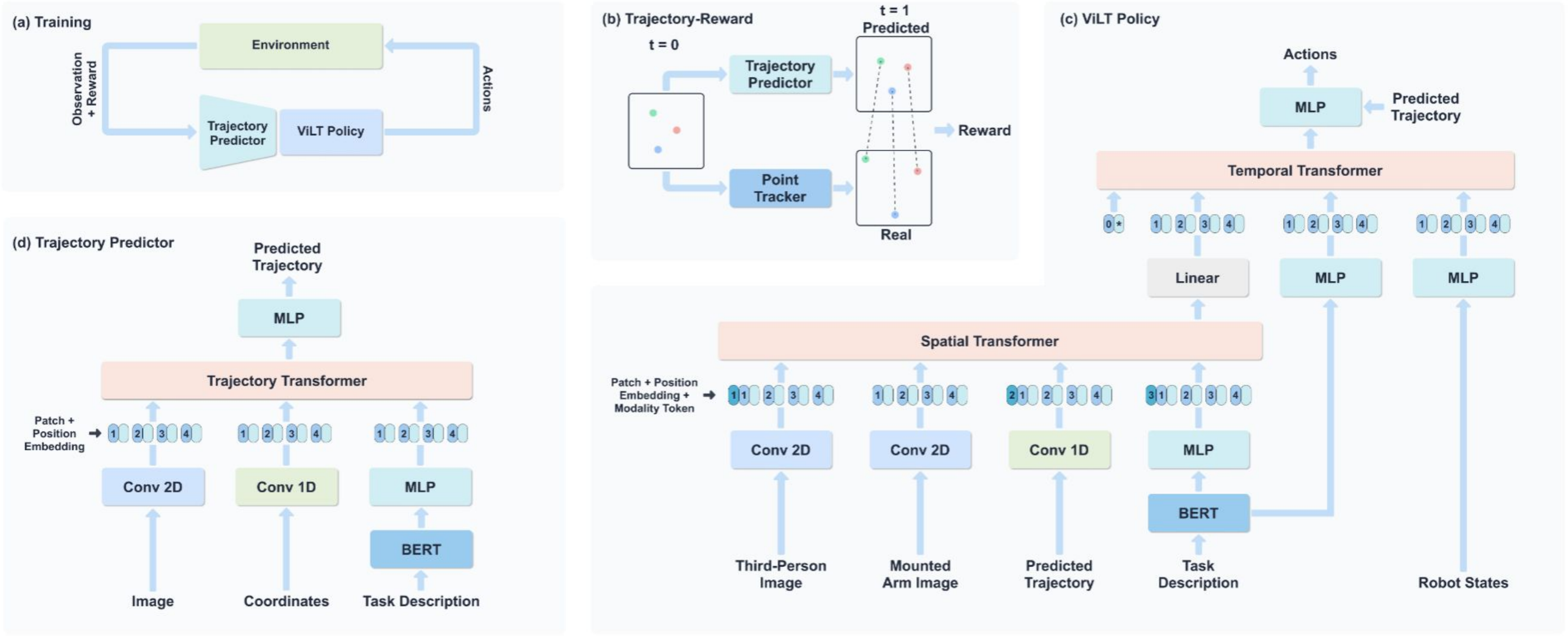
Prior Works

- Sparse Reward
  - Initializing closer to the goal to allow for better exploration [2]



- Dense Task Specific Rewards
  - Create task specific rewards that utilize some attribute that is task specific or robot specific to compute a distance such as tool distance or hand distance

Framework



(a) Training
Environment
Observation + Reward
Actions
Trajectory Predictor
ViLT Policy
(b) Trajectory-Reward
t = 0
t = 1
Predicted
Trajectory Predictor
Point Tracker
Reward
Real
(c) ViLT Policy
Actions
MLP
Predicted Trajectory
Temporal Transformer
Linear
MLP
MLP
Spatial Transformer
Patch + Position Embedding + Modality Token
Conv 2D
Conv 2D
Conv 1D
MLP
BERT
Third-Person Image
Mounted Arm Image
Predicted Trajectory
Task Description
Robot States
(d) Trajectory Predictor
Predicted Trajectory
MLP
Trajectory Transformer
Patch + Position Embedding
Conv 2D
Conv 1D
MLP
BERT
Image
Coordinates
Task Description

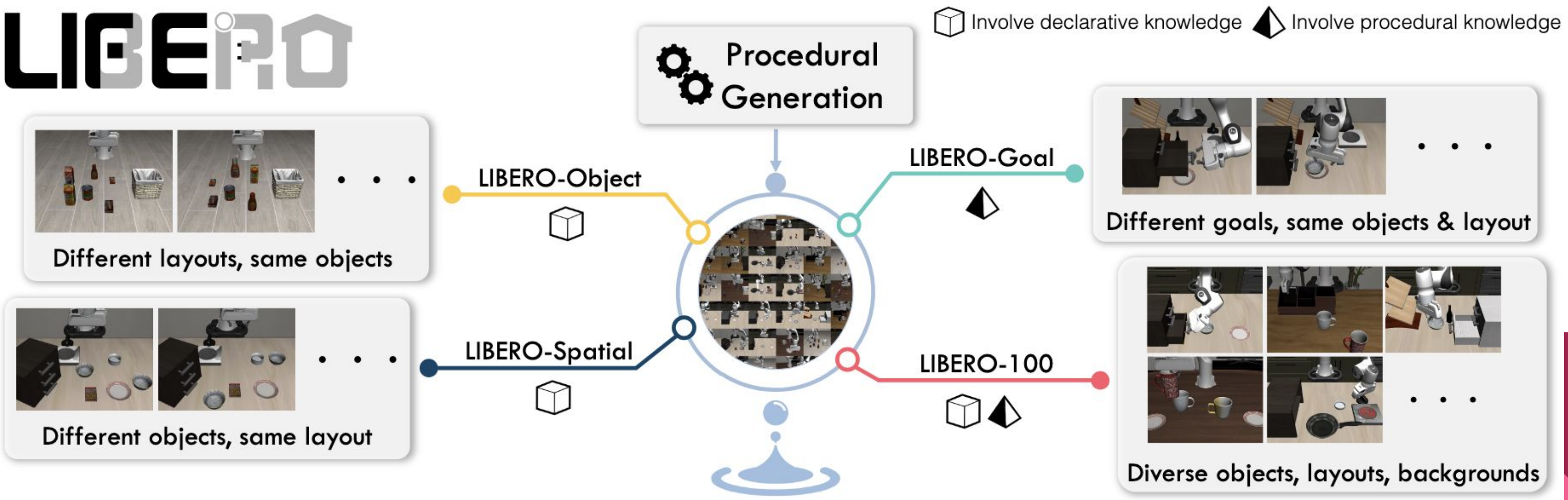
Reward Formulation

$$r(s, a) = c_1 \cdot r_{\text{trajectory}} + c_2 \cdot r_{\text{success}} \tag{2}$$

$$r_{\text{trajectory}} = \exp\left(-\|\mathbf{p}_{\text{real}}(t+1) - \mathbf{p}_{\text{pred}}(t+1)\|_2\right) \tag{3}$$

Training Environments

- Our policy is trained with the LIBERO Environment [3] in the MuJuCo Simulator



Demo of RL

Next Steps

- Use segmentation mask to determine best points to predict
- Try larger policy architecture
- Use an additional unsupervised exploration technique

References

[1] T. Z. Zhao, V. Kumar, S. Levine, and C. Finn, ‘Learning Fine-Grained Bimanual Manipulation with Low-Cost Hardware’, *arXiv [cs.RO]*. 2023.

[2] A. Nair, B. McGrew, M. Andrychowicz, W. Zaremba, and P. Abbeel, ‘Overcoming Exploration in Reinforcement Learning with Demonstrations’, *arXiv [cs.LG]*. 2018.

[3] B. Liu *et al.*, ‘LIBERO: Benchmarking Knowledge Transfer for Lifelong Robot Learning’, *arXiv [cs.AI]*. 2023.