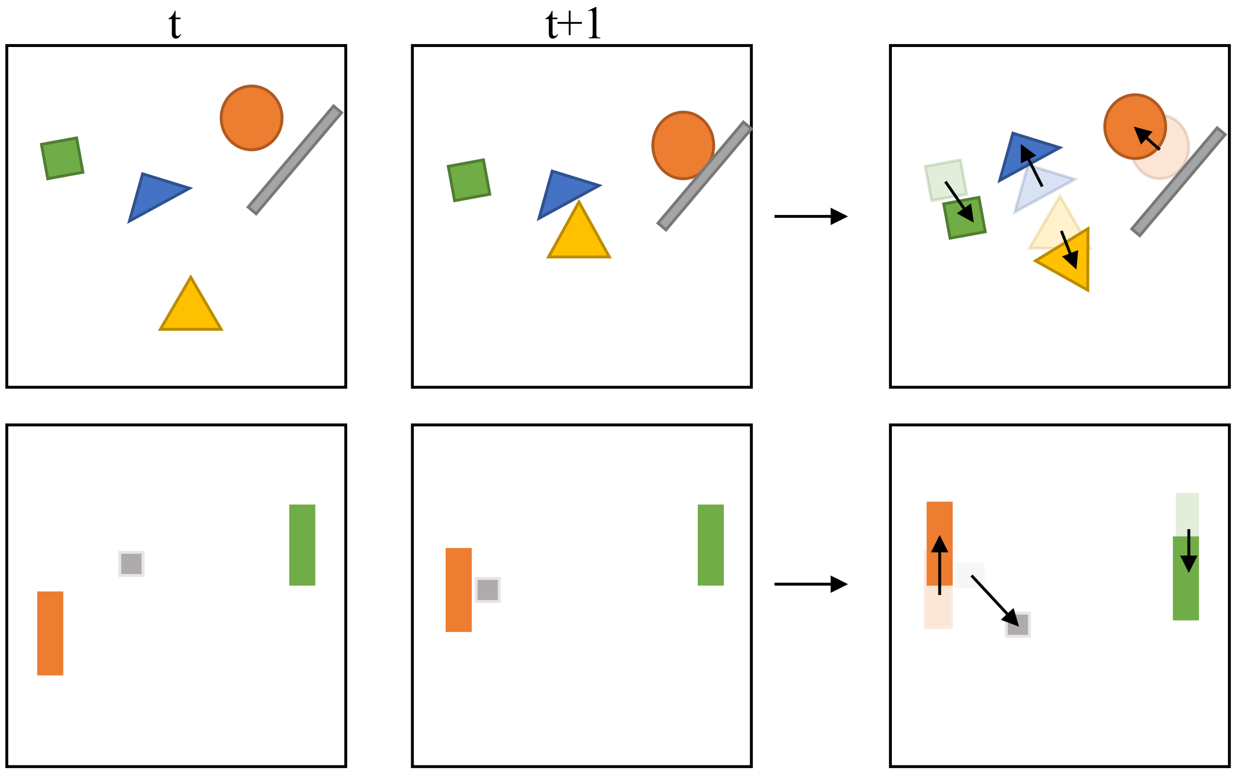
Task-Agnostic Dynamics Priors for Deep Reinforcement Learning

Yilun Du[1], Karthik Narasimhan[2]

[1] MIT, [2] Princeton

Key Questions

- Can we learn physics in a task-agnostic fashion?
- Does it help sample efficiency of RL?
- Can we transfer the learned physics from one environment to other?

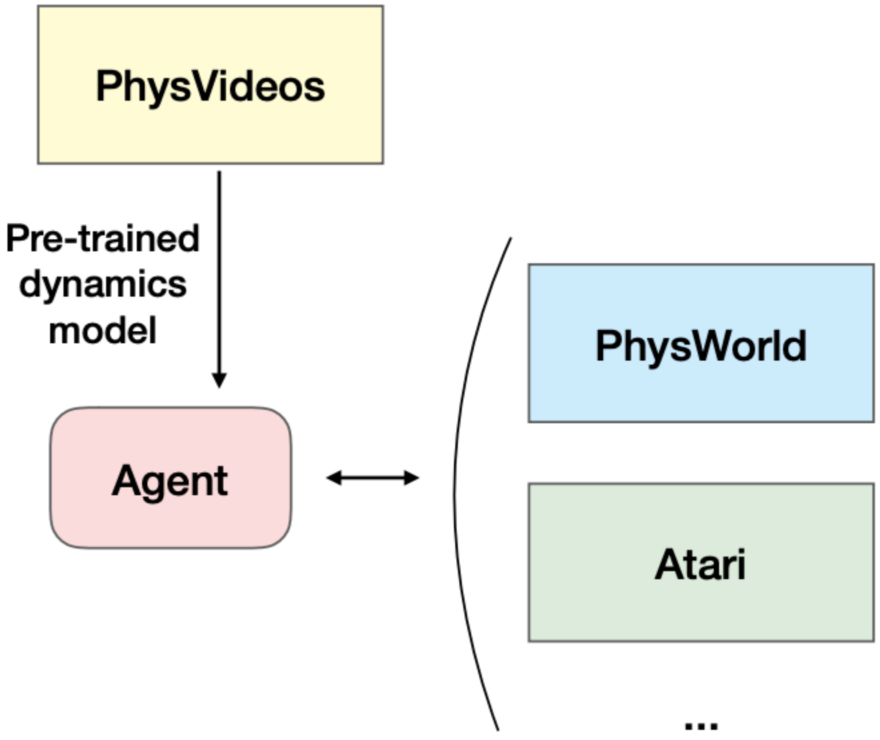


Dynamics Model in RL

- Frame Prediction (Oh et al.(2015), Finn et al.(2016), Weber et al. (2017), …)
  - Action conditional and not easily transferable across environments
- Parameterized physics models (Cutler et al. (2014), Scholz et al.(2014), Zhu et al. (2018), …)
  - Requires manual specification
- Our method: learn physics priors through task-independent data
  - Action unconditional modeling of data
  - Inductive local biases in architecture to reflect local nature of physics

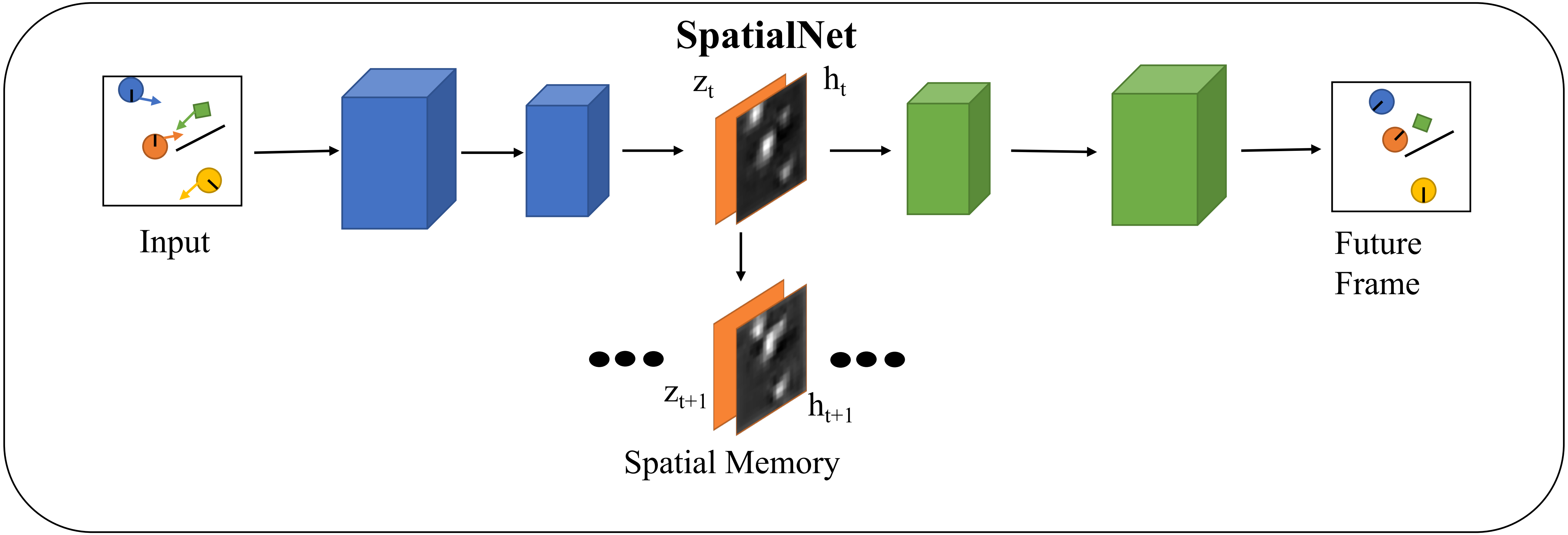
Overall Approach

- Pre-train a frame predictor on physics videos
- Initialize dynamics model and use it to train a policy
- Simultaneously fine-tune dynamics model on target environment.



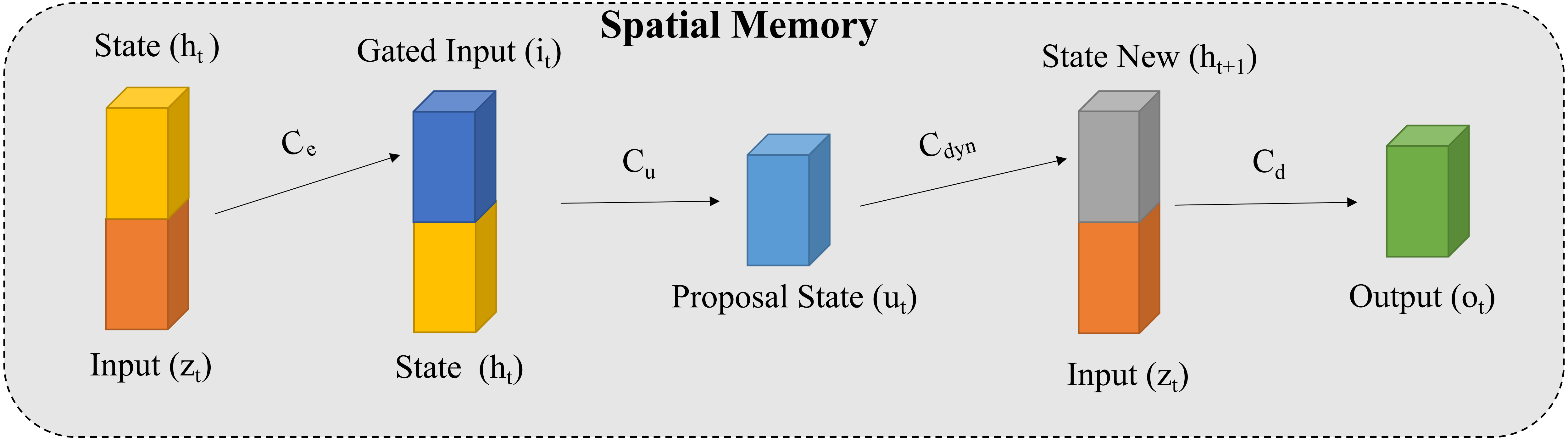
SpatialNet

- Two key operations:
  - Isolation of dynamics of each entity
  - Accurate modeling of dynamic interactions of local spaces around each entity



Spatial Memory

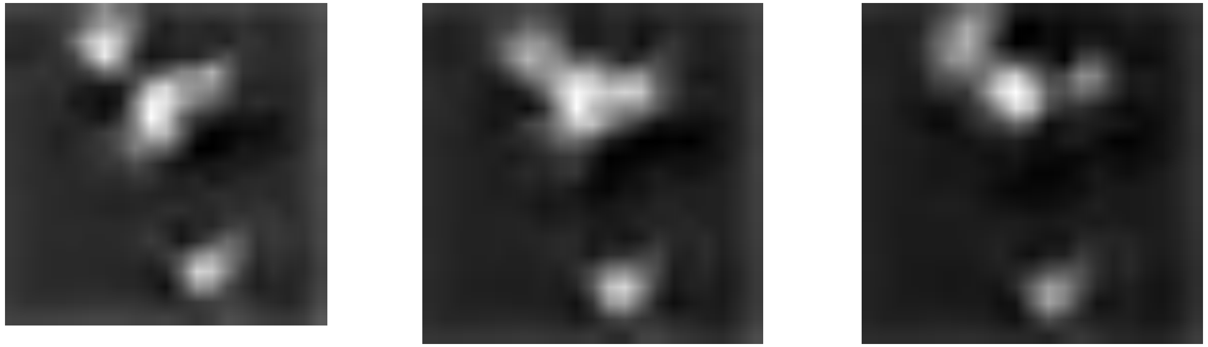
- Use 2D grid memory to locally store dynamic state of each object
- Use convolutions and residual connections to better model dynamics (instead of additive updates in the ConvLSTM model (Xingjian et al., 2015))



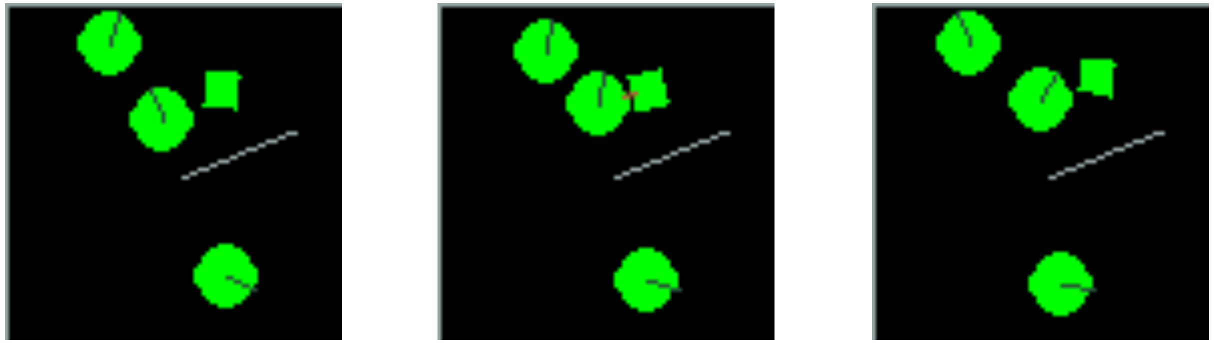
Spatial Memory

- Use 2D grid memory to locally store dynamic state of each object
- Use convolutions and residual connections to better model dynamics (instead of additive updates in the ConvLSTM model (Xingjian et al., 2015))

Spatial Memory State



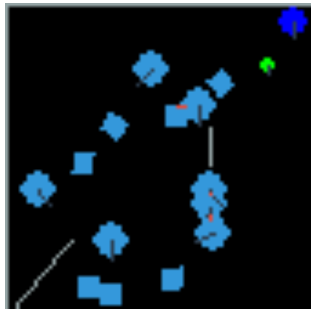
Input Frames



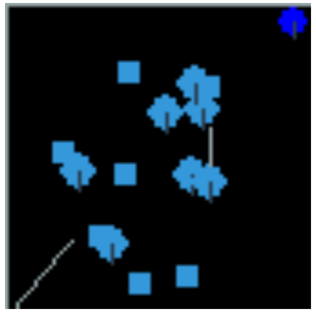
Experimental Setup

- *PhysVideos*: 625k frames of video containing moving objects of various shapes and sizes
- *PhysWorld*: Collection of 2D/3D Physics-centric games
- *Atari*: Stochastic version with sticky actions
- RL agent: Predicted frames stack with observation frames as joint input into a policy
- **Same prior for all tasks**

PhysGoal



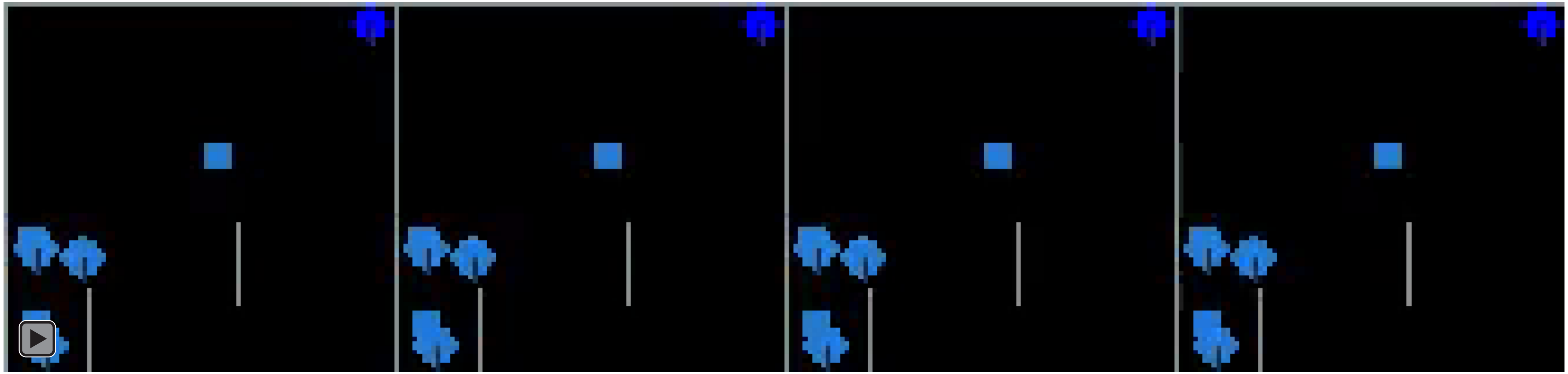
PhysShooter



PhysForage

Phys3D

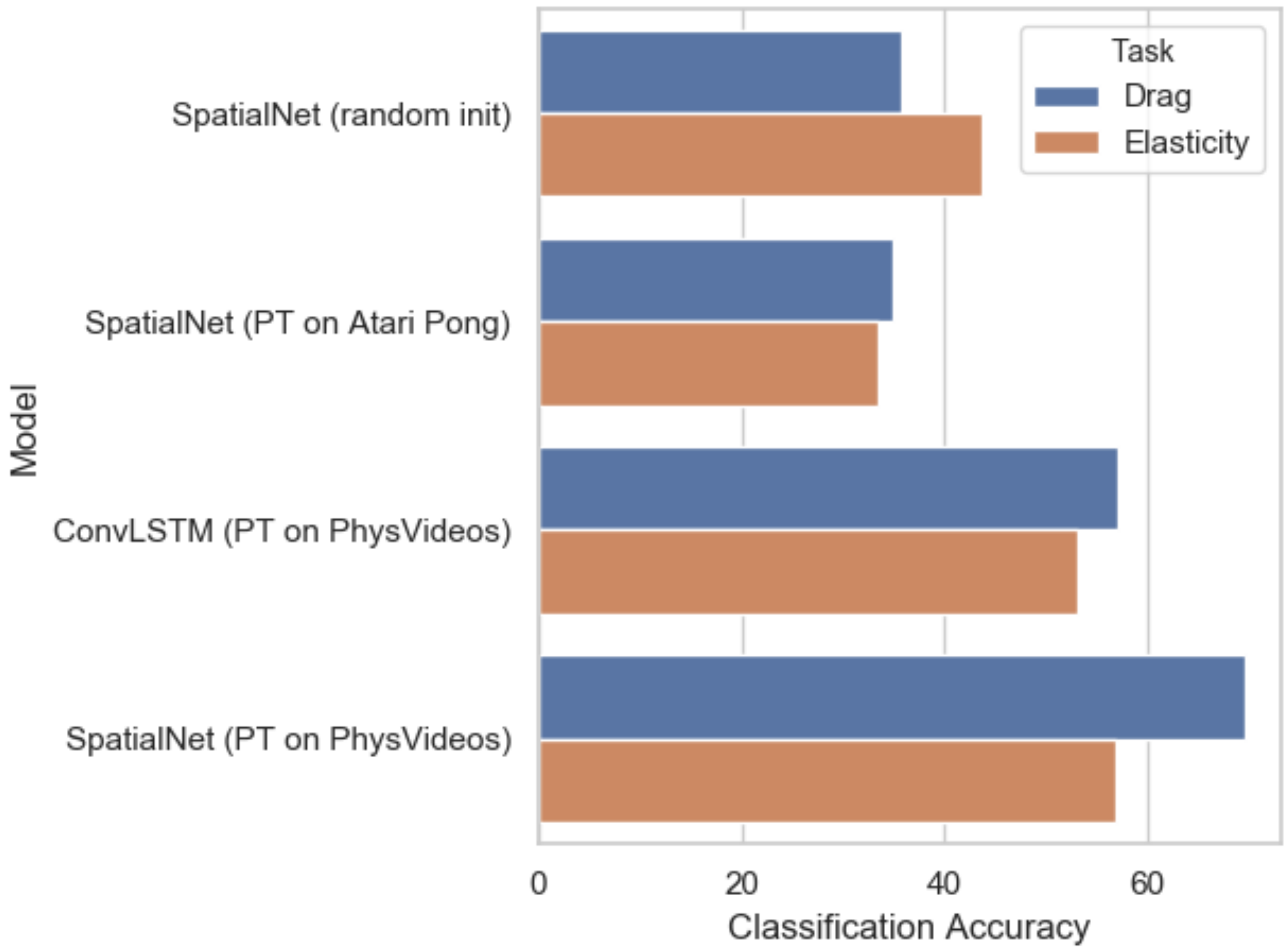
Model Predictions



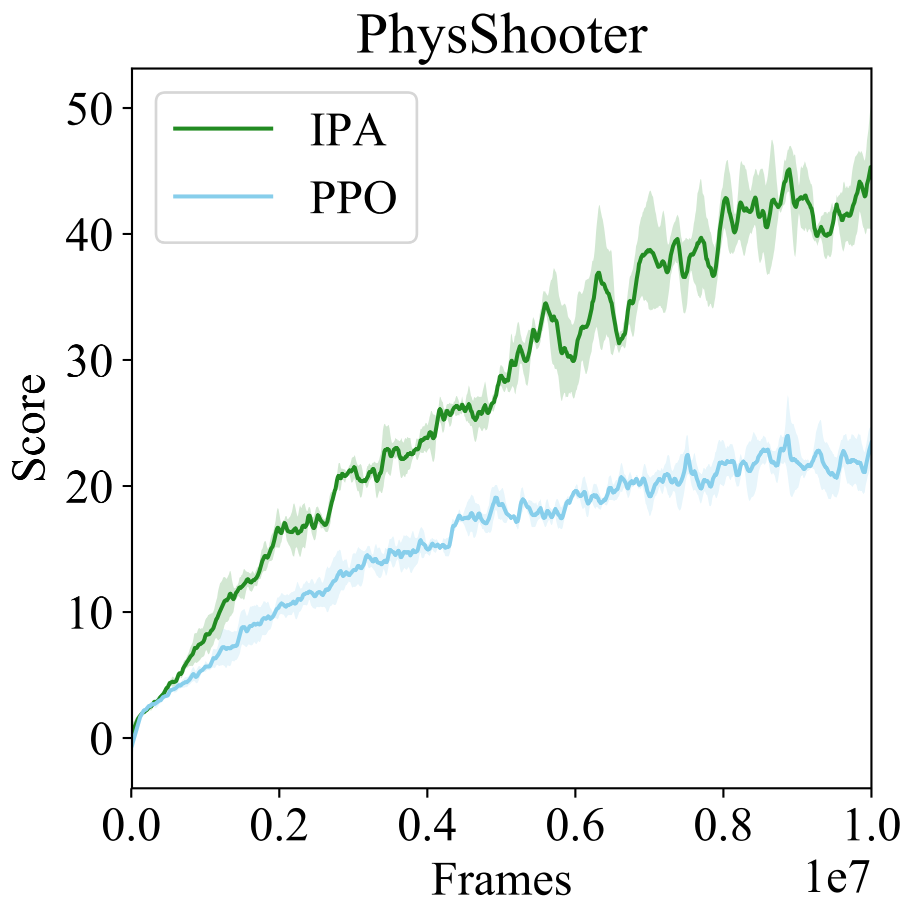
Model	10 step	Objects Lost
RCNet (Oh et al., 2015)	0.0268	1.0
ConvLSTM (Xingjian et al., 2015)	0.0503	0.4
ConvLSTM + Residual	0.0210	0.45
SpatialNet	**0.0176**	**0.13**

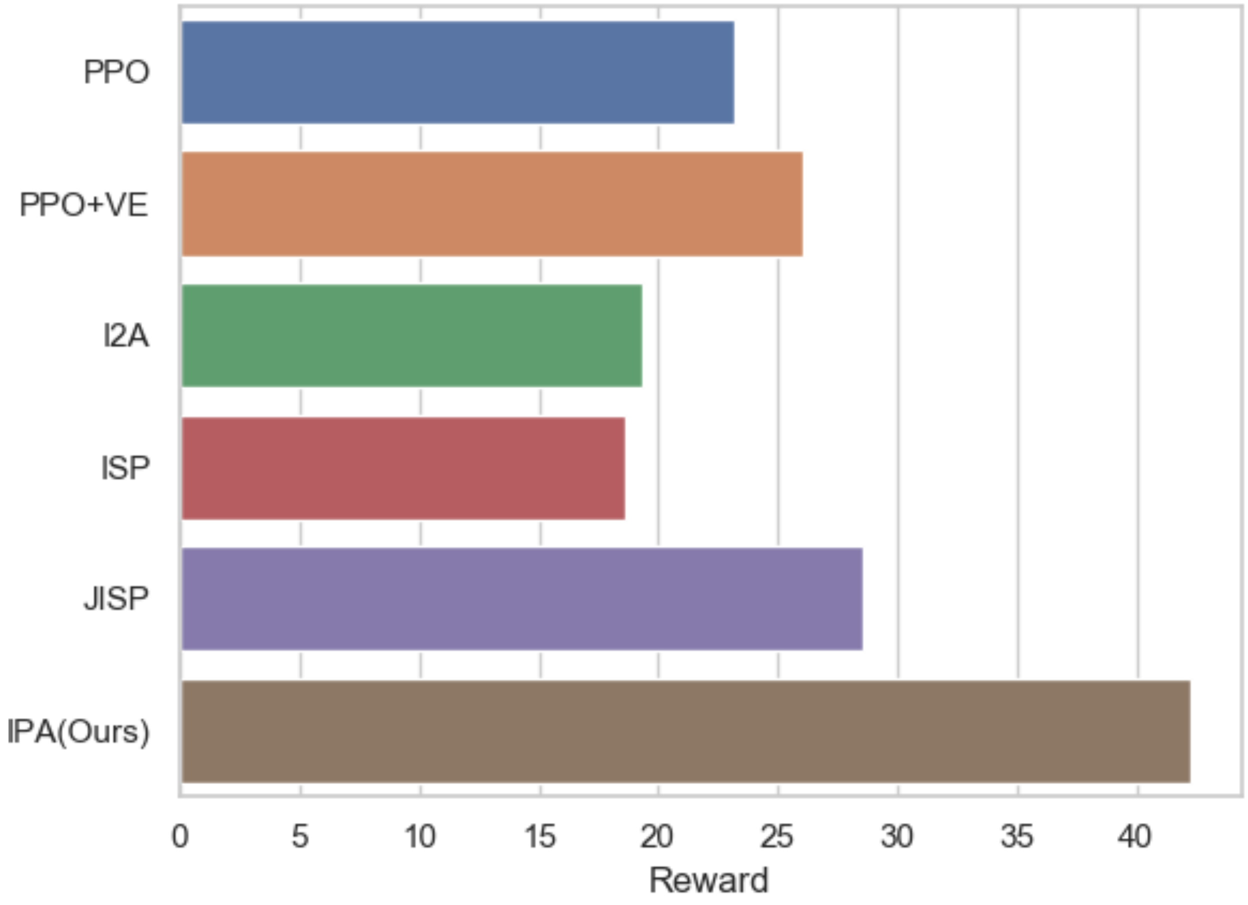
Pixel Prediction Accuracy

Predicting Physical Parameters

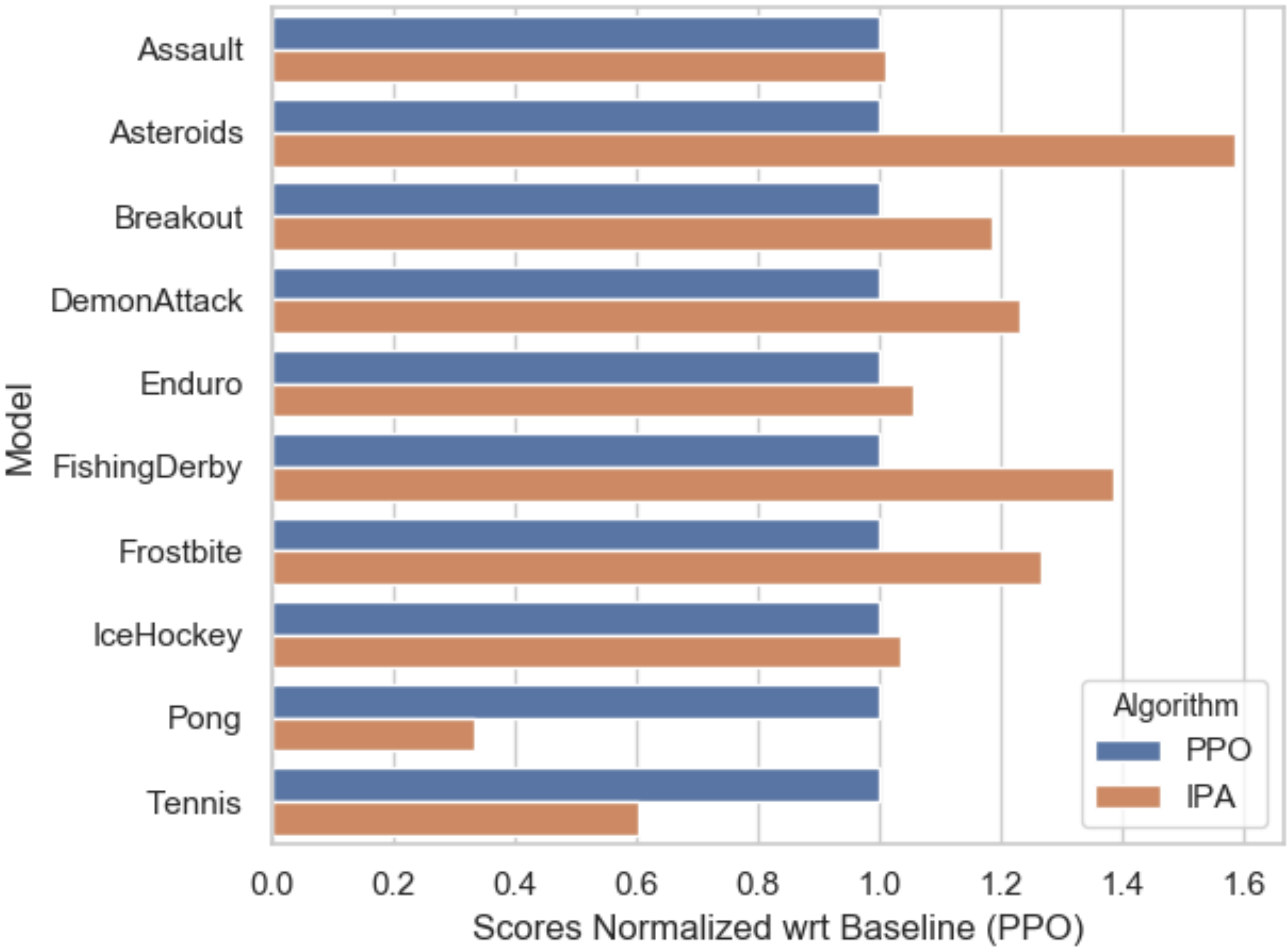


Policy Learning: PhysShooter

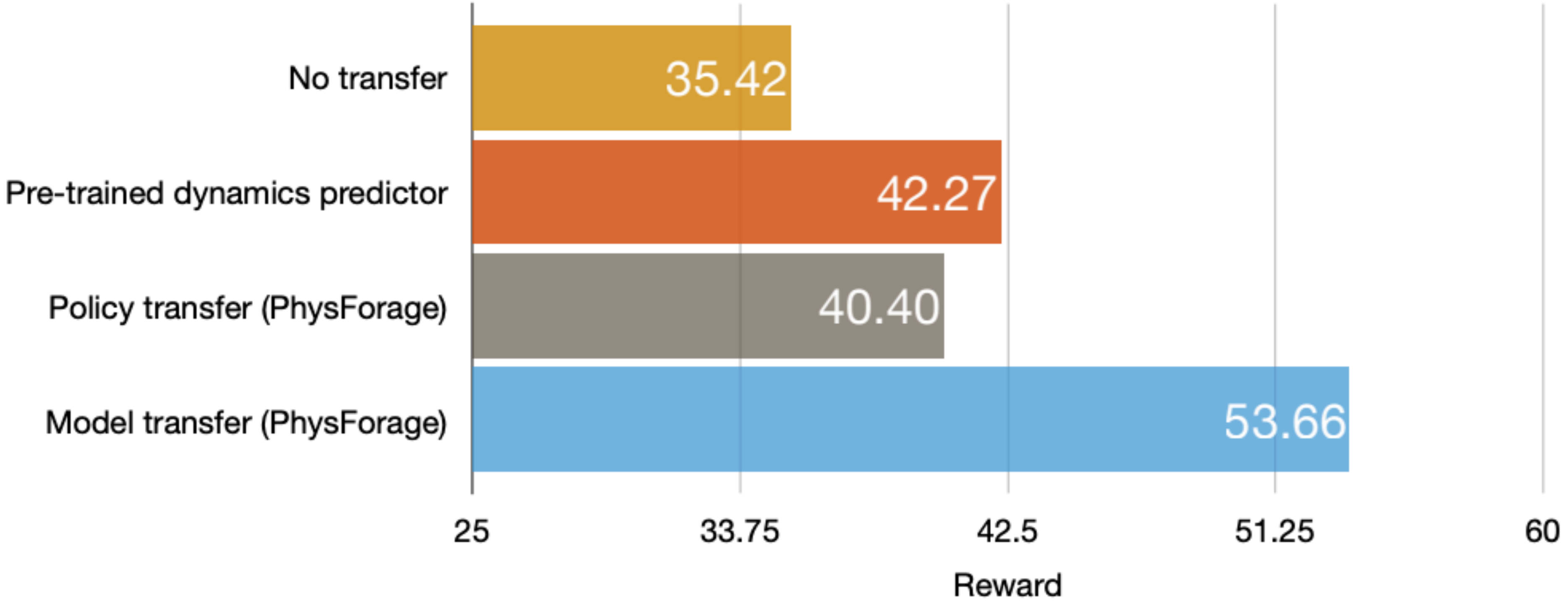




Policy Learning: Atari



Transfer Learning



Target env: PhysShooter

Model Transfer > Model + Policy Transfer > No Transfer

Conclusion

- Task-agnostic priors over models provide a potential solution for improving sample efficiency for RL
- Being task-agnostic allows us to pre-train priors without access to the target task
- Such priors also generalize well to a wide variety of tasks and show good transfer performance