Learning Representations for Versatile Behavior

Chelsea Finn

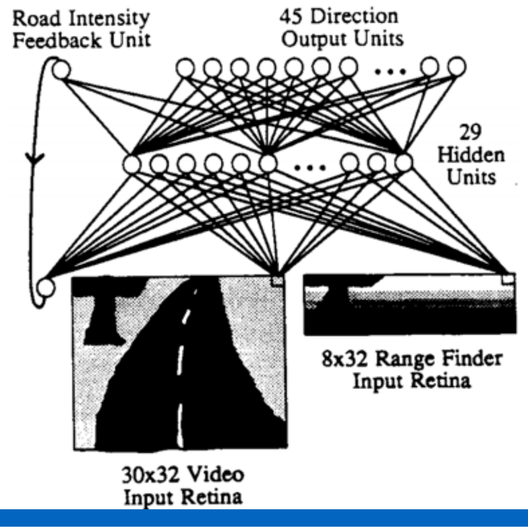
BAIR
BERKELEY ARTIFICIAL INTELLIGENCE RESEARCH

If you gave the parts to a 9 month old…

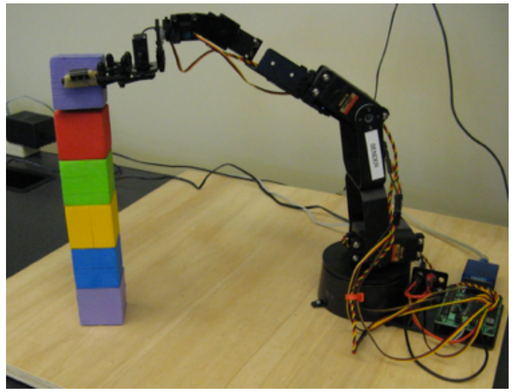
… vs. older child or adult

Humans reuse prior experience.

Robot learning

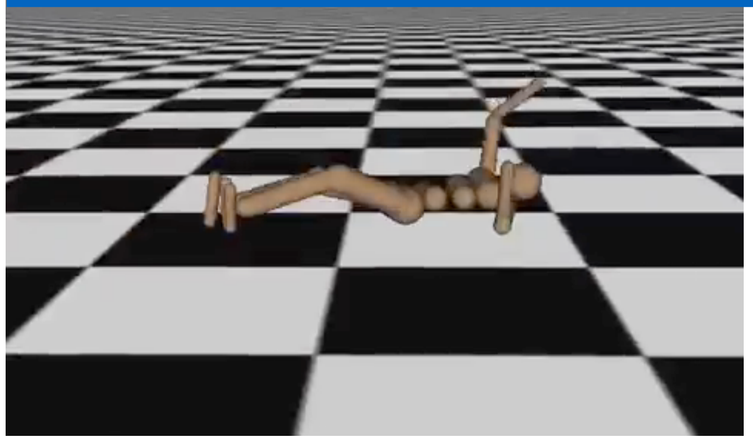


Kohl & Stone '04

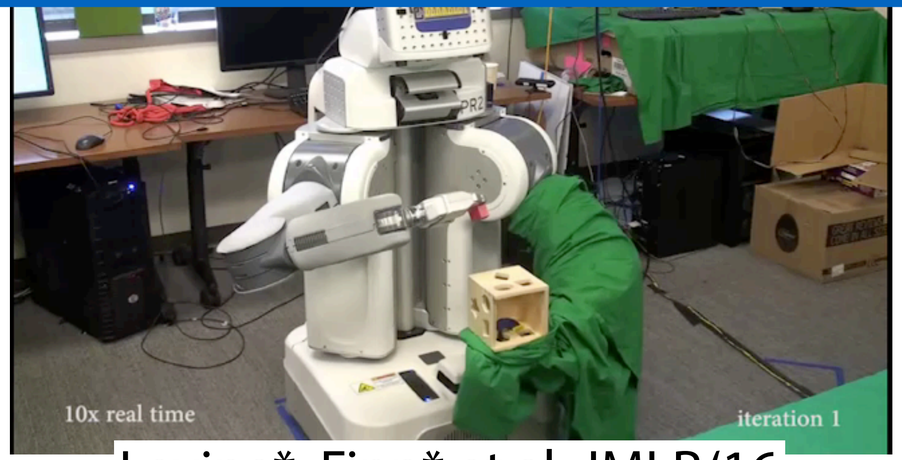


Deisenroth et al. '11

paradigm: train/test on 1 task in 1 environment, starting from scratch



Schulman et al. '16



Levine*, Finn* et al. JMLR '16

Chebotar et al. '17

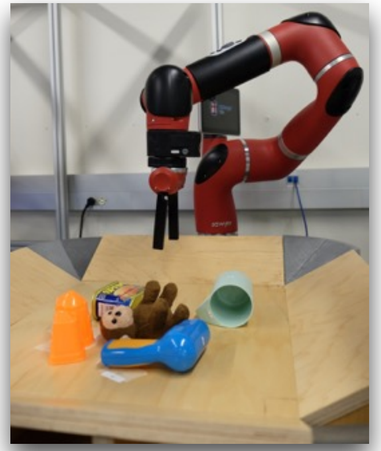
How can robots use past experience?

1. Learn about the physical world

2. Learn to learn

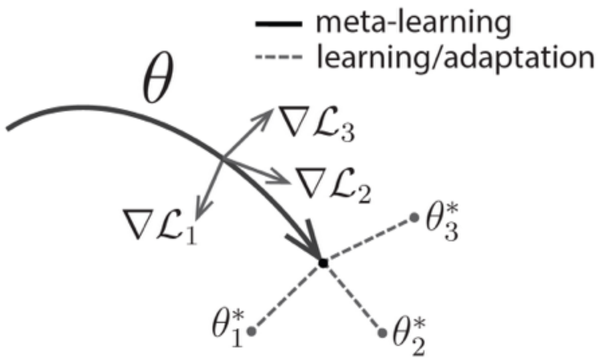
How can robots use past experience?

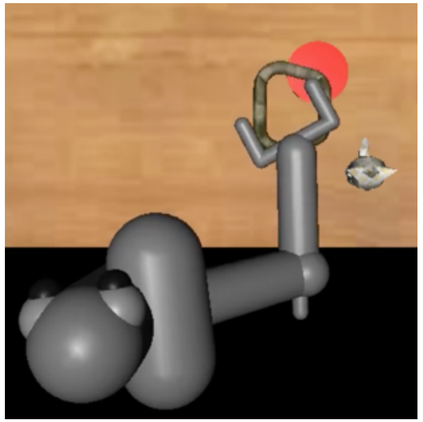
1. Learn about the physical world



Finn et al. NIPS '16, *Finn* & Levine ICRA '17
Ebert et al. '17 (under review)

2. Learn to learn

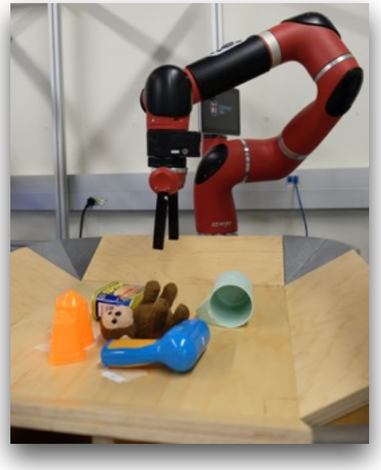




Finn et al. ICML '17
Finn*, Yu* et al. '17 (under review)

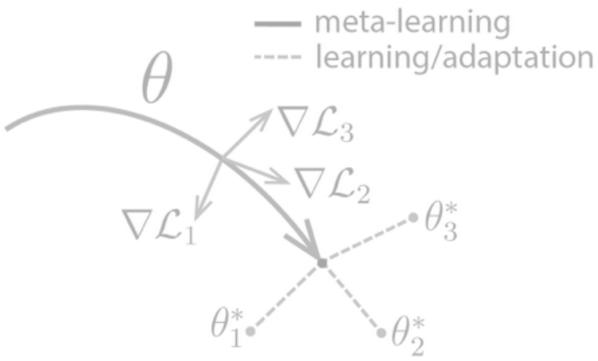
How can robots use past experience?

1. Learn about the physical world



Finn et al. NIPS '16, *Finn* & Levine ICRA '17
Ebert et al. '17 (under review)

2. Learn to learn



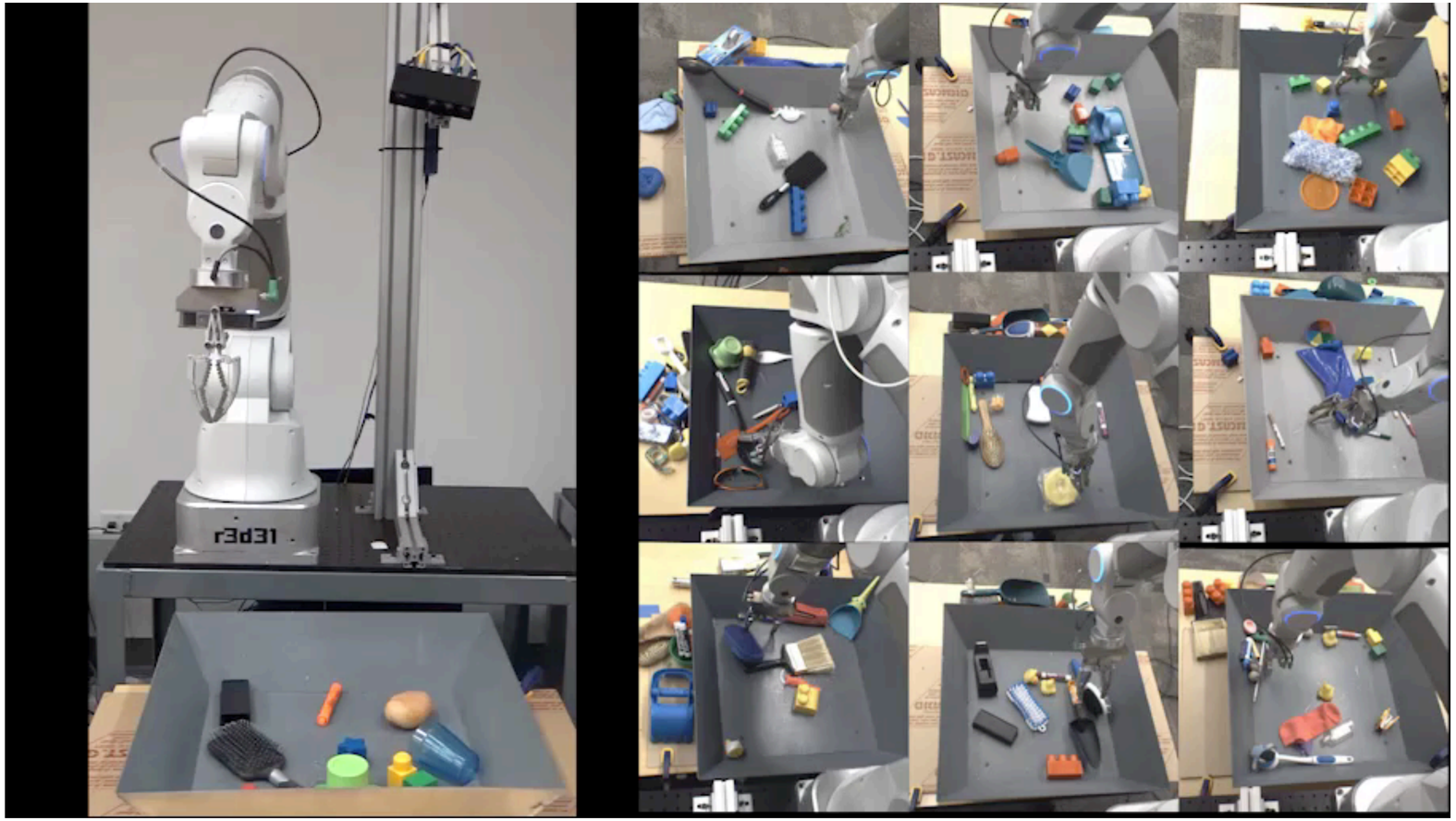
Finn et al. ICML '17
Finn*, Yu* et al. '17 (under review)

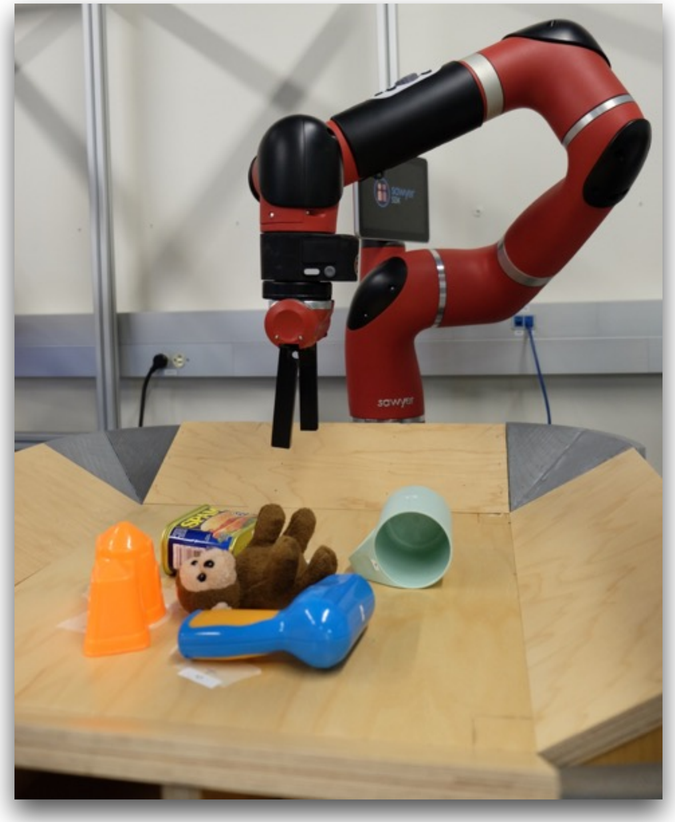
Learning about the Physical World

self-supervised setting

- use **raw sensory inputs** (i.e. vision)
- scale up data collection to **learn without supervision**
- instead of learning single-purpose policy, learn **reusable model**

Data collection - 50k sequences (1M+ frames)





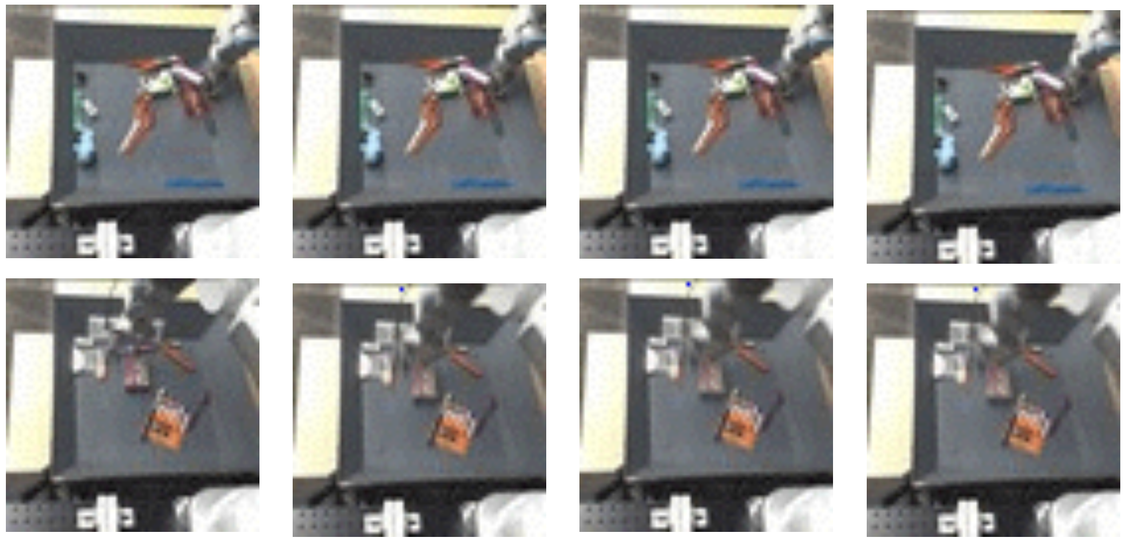
Finn, Goodfellow, Levine, NIPS '16

Ebert, Finn, Lee, Levine '17 (under sub.)

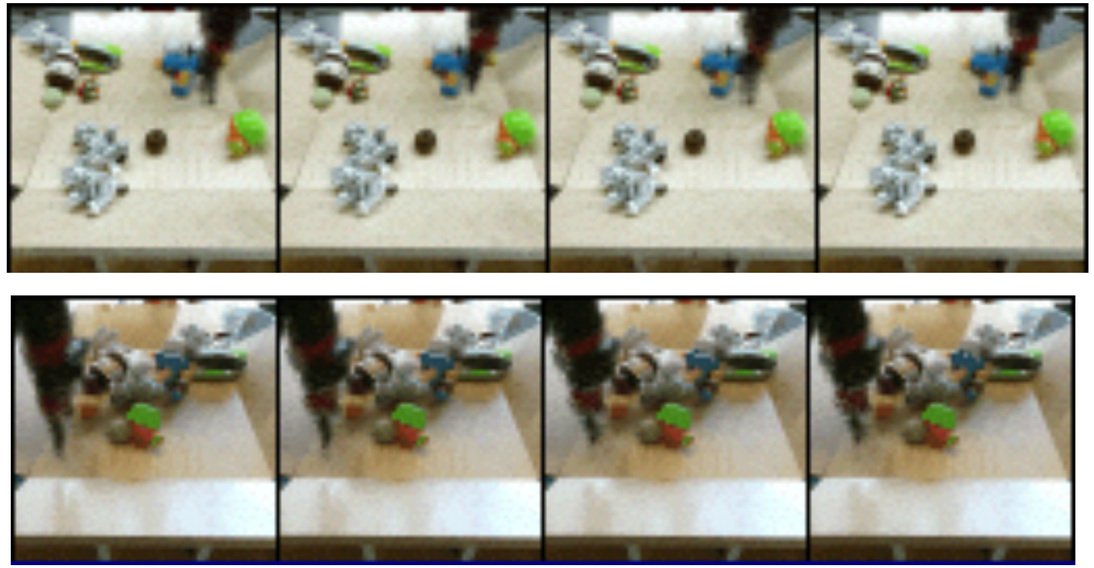
Train predictive model

$$I_t,\ a_{t:t+H} \longrightarrow I_{t:t+H}$$

<— varying actions —>



<— varying actions —>

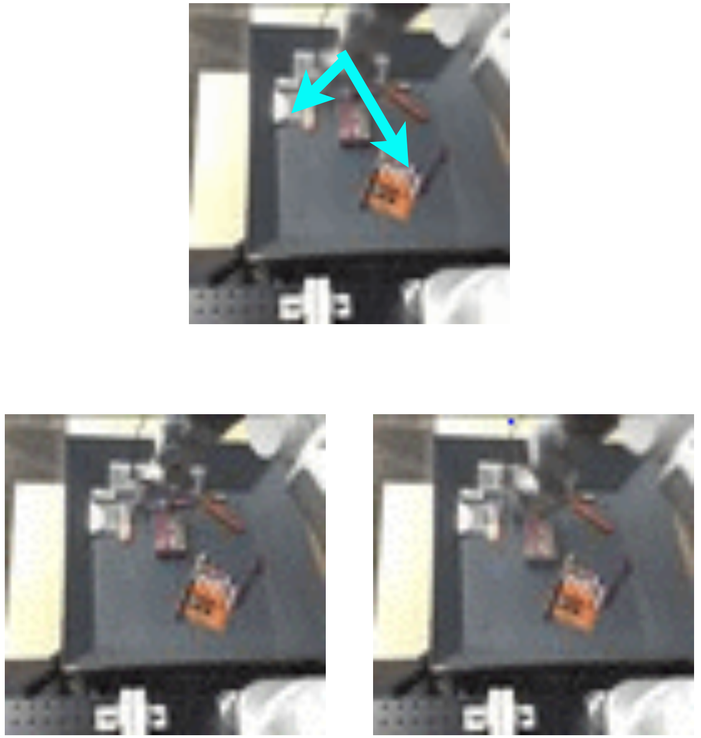


Finn, Goodfellow, Levine, NIPS '16

Ebert, Finn, Lee, Levine '17 (under sub.)

Planning with Visual Foresight

1. Consider potential action sequences
2. Predict the future for each action sequence
3. Pick best future & execute corresponding action
4. Repeat 1-3 to replan in real time



visual model-predictive control (MPC)

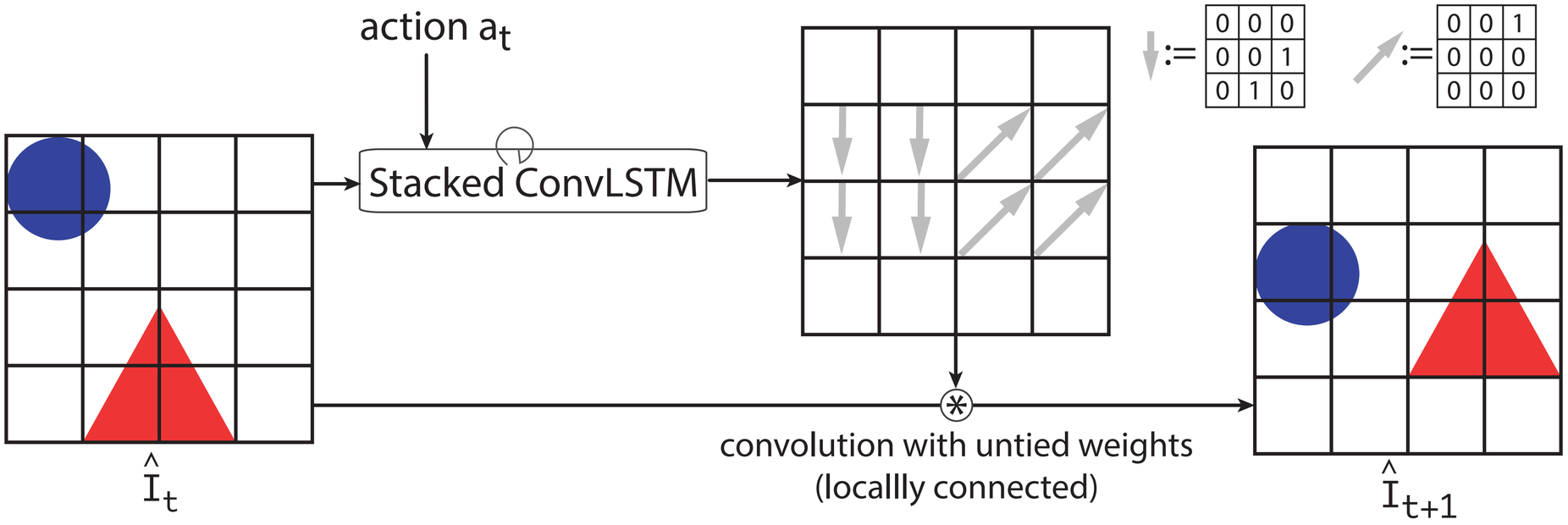
Finn & Levine ICRA '17

How to predict video?

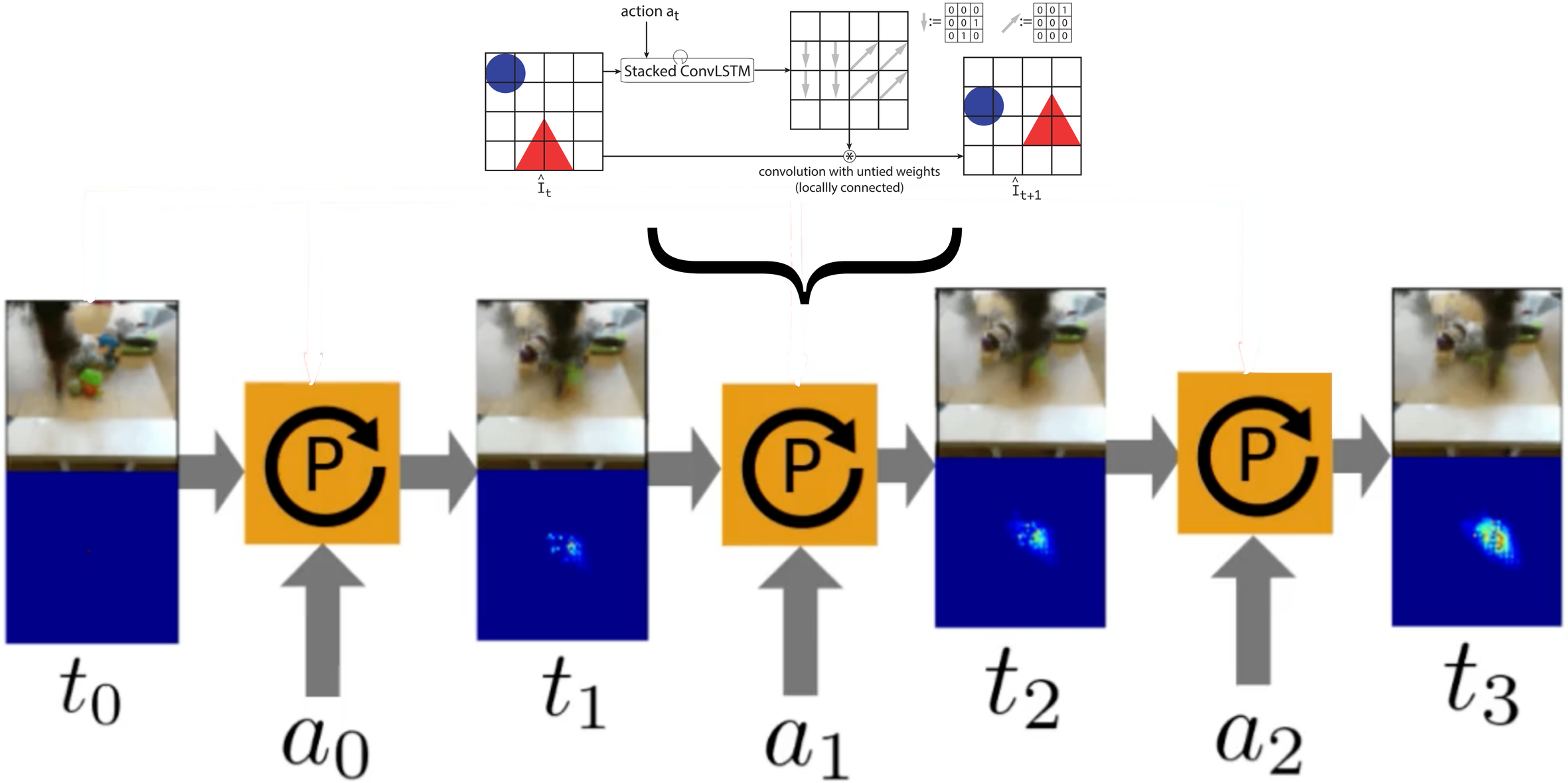
<u>Goal</u>: Design a prediction model that's good for control

- action-conditioned
- multi-frame prediction
- explicitly model motion

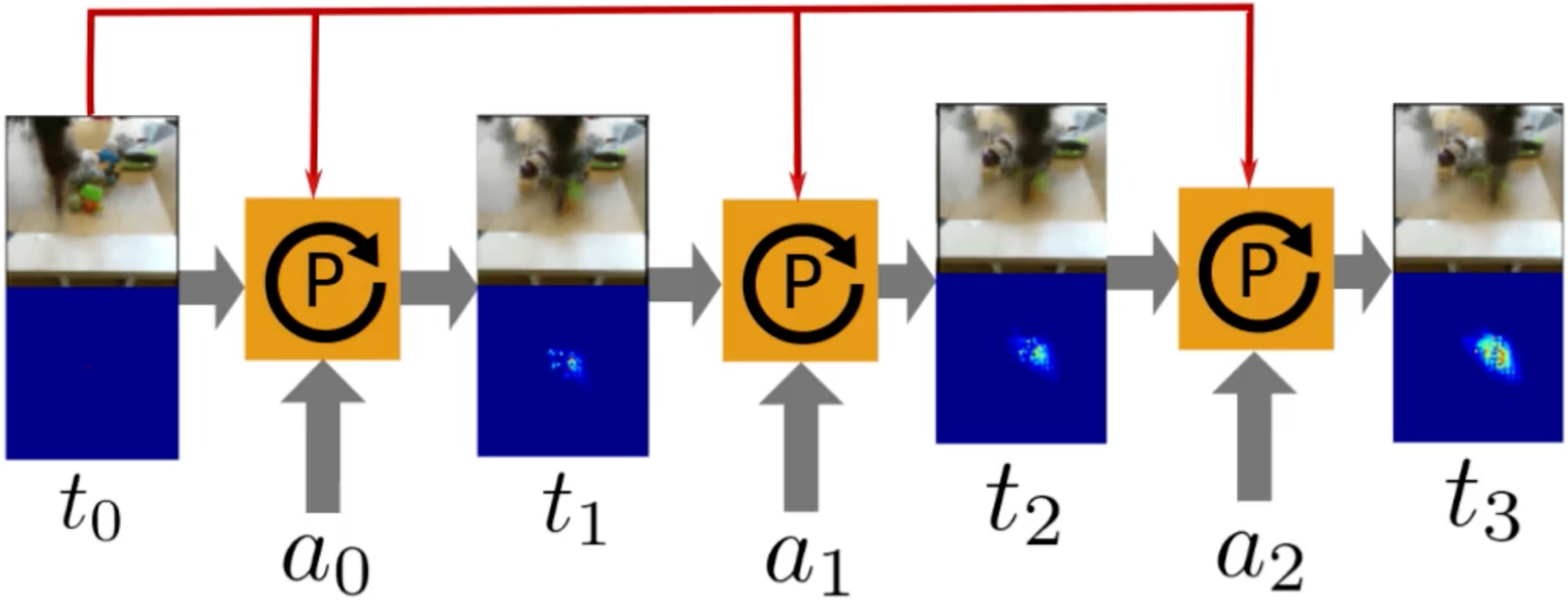
Key idea: output how pixels will move, rather than pixels themselves



Finn, Goodfellow, Levine, NIPS '16

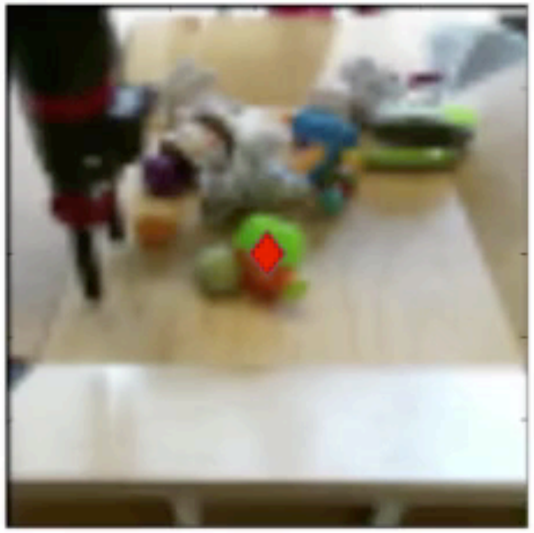

action a_t
Stacked ConvLSTM
:=
:=
convolution with untied weights
(locallly connected)
$\hat{I}_t$
$\hat{I}_{t+1}$
P
P
P
t_0
a_0
t_1
a_1
t_2
a_2
t_3

Use both first frame and most recent frame for prediction



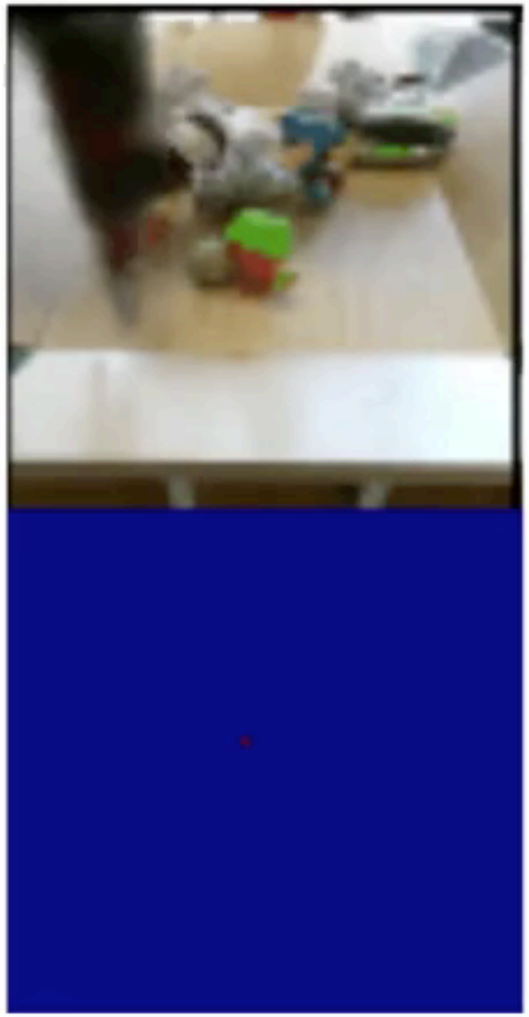
Ebert, Finn, Lee, Levine, '17 (under sub.)

Use both first frame and most recent frame for prediction

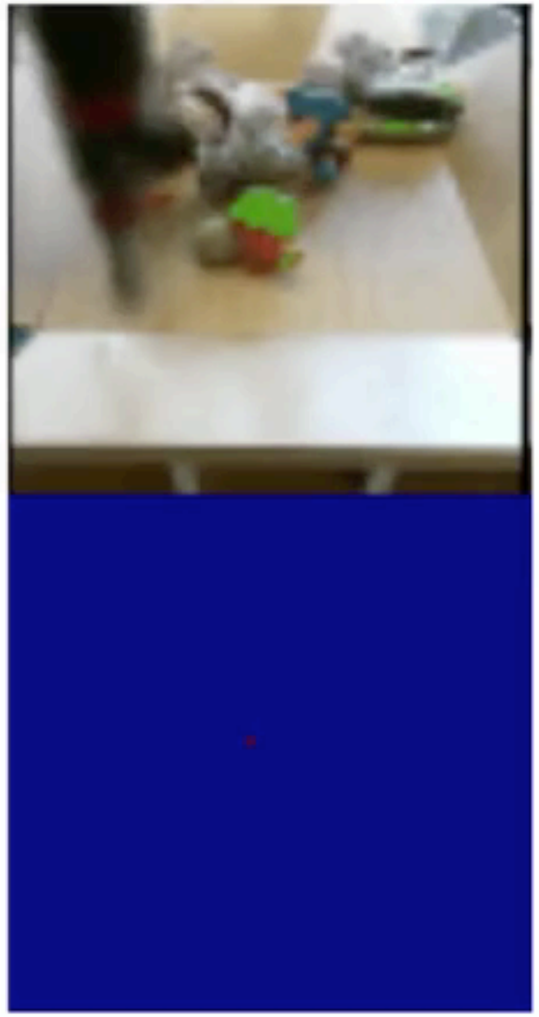


Designated Pixel

most recent frame only



both

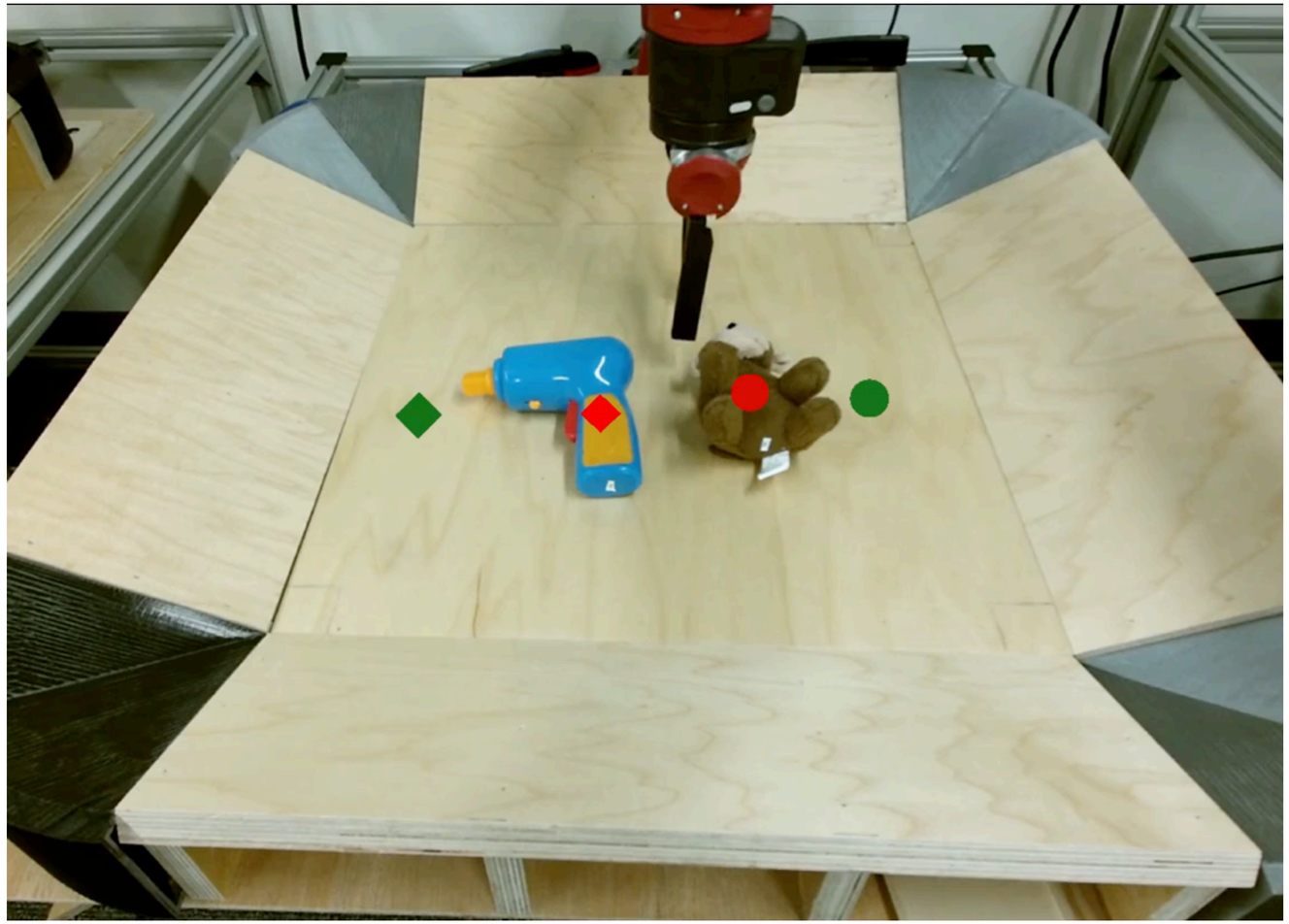


Ebert, Finn, Lee, Levine, '17

How it works

Finn & Levine ICRA '17

How does this approach do?



- evaluation on maneuvering **seen** & **novel** objects

- model trained on 8 days of **unlabeled** robot data

Completely self-supervised:
Only human involvement during training is:
programming initial motions and providing objects

Finn & Levine ICRA '17
Ebert, Finn, Lee, Levine, '17

Planning with raw sensory inputs

Pros: learn for a **variety of tasks**, entirely **self-supervised**

Cons: can't [yet] learn **complex skills**, compute intensive at test time

Next steps for this approach:

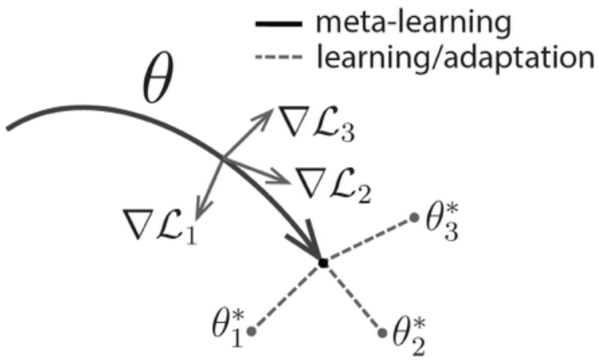
- stochastic model for **handling uncertainty** in the long term
- collect data using model to achieve **more complex goals**
- **long term planning** in more abstract spaces

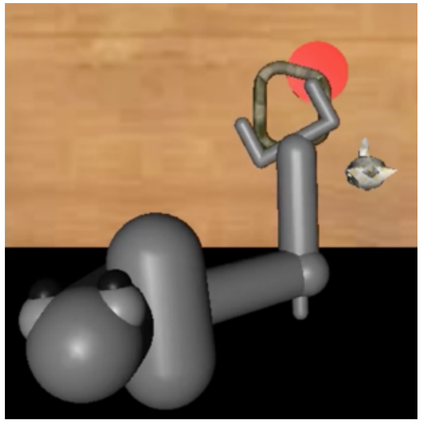
How can robots use past experience?

1. Learn about the physical world

Finn et al. NIPS '16, *Finn* & Levine ICRA '17
Ebert et al. '17 (under review)

2. Learn to learn

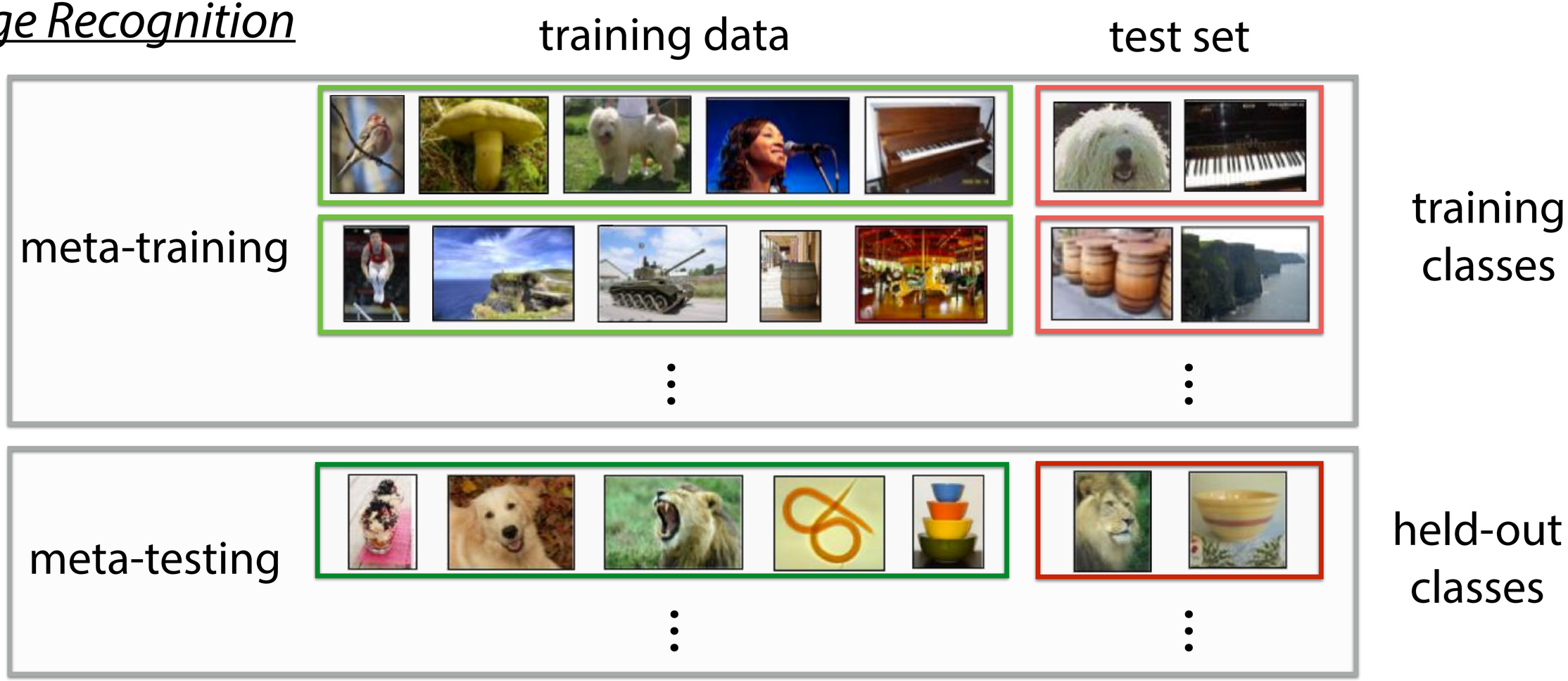




Finn et al. ICML '17
Finn*, Yu* et al. '17 (under review)

Few-shot learning: incorporate prior knowledge from other tasks for fast learning

Image Recognition

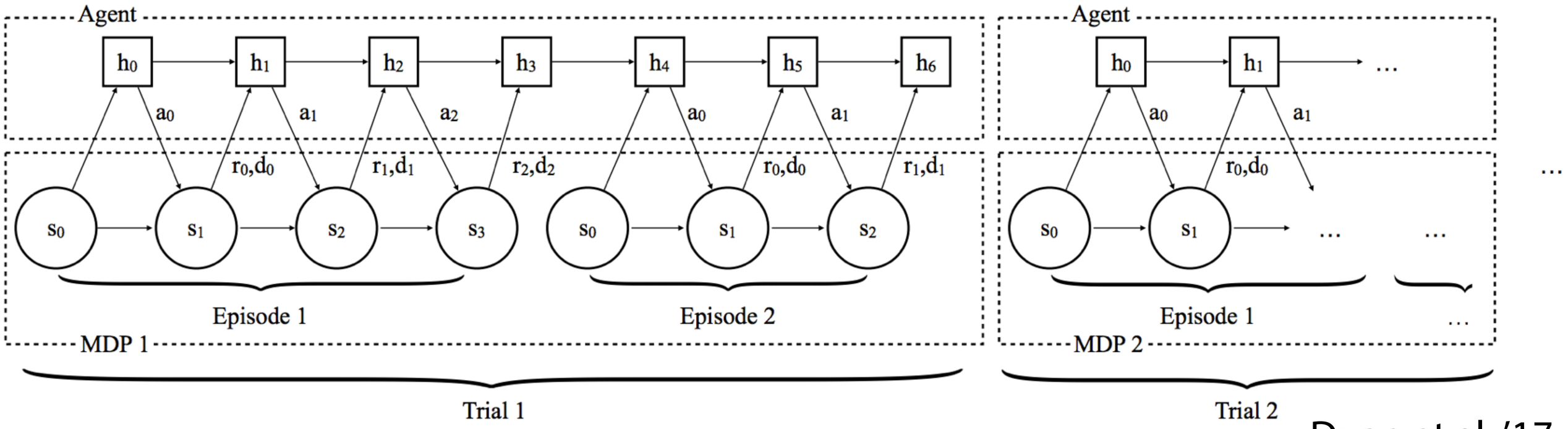


What is few-shot learning for behavior?

diagram adapted from Ravi & Larochelle '17

Few Shot Learning via Recurrence

(Santoro et al. ’16, Duan et al. ’17, Wang et al. ’17)



Duan et al. ’17

Learning Few-Shot Adaptation

Transfer learning: finetune from ImageNet-trained features (Deng et al. '09, Donahue et al. '14)

+ simple, works well, same learning rule - no ImageNet for behavior…

How can we get transferable features for behavior?

Learning Few-Shot Adaptation

Transfer learning: finetune from ImageNet-trained features (Deng et al. '09, Donahue et al. '14)

+ simple, works well, same learning rule - no ImageNet for behavior…

How can we get transferable features for behavior?

pretrained parameters

Fine-tuning: $$\theta \leftarrow \theta - \alpha \nabla_\theta \mathcal{L}_\mathcal{T}(\theta)$$

[test-time]

task

Our method: $$\min_\theta \sum_{\text{task } \mathcal{T}} \mathcal{L}_\mathcal{T}(\theta - \alpha \nabla_\theta \mathcal{L}_\mathcal{T}(\theta))$$

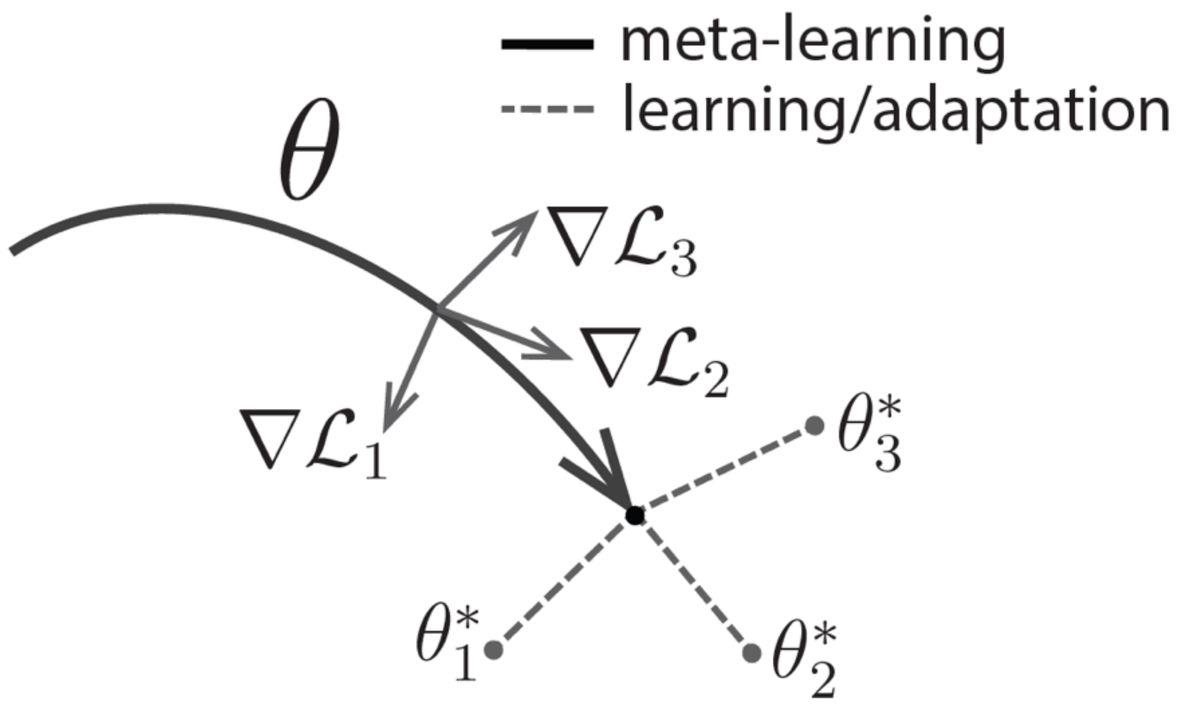
Key idea: Train over many tasks, to learn parameter vector θ that transfers

Learning Few-Shot Adaptation

$$\min_{\theta} \sum_{\text{task } \mathcal{T}} \mathcal{L}_{\mathcal{T}}(\theta - \alpha \nabla_{\theta} \mathcal{L}_{\mathcal{T}}(\theta))$$

θ parameter vector being meta-learned

θ_i^* optimal parameter vector for task i

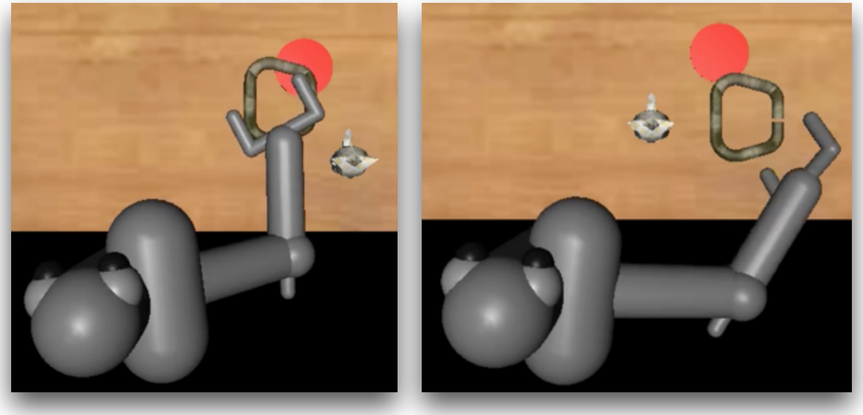


Model-**A**gnostic **M**eta-**L**earning

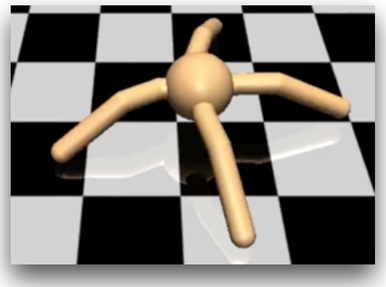
Few-Shot Learning Experiments

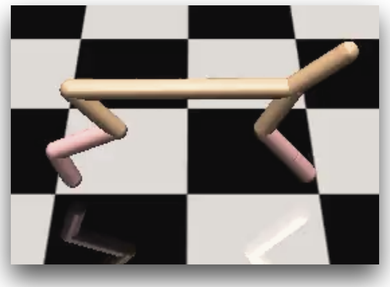
Few-Shot Classification
compare to prior methods

One-Shot Imitation

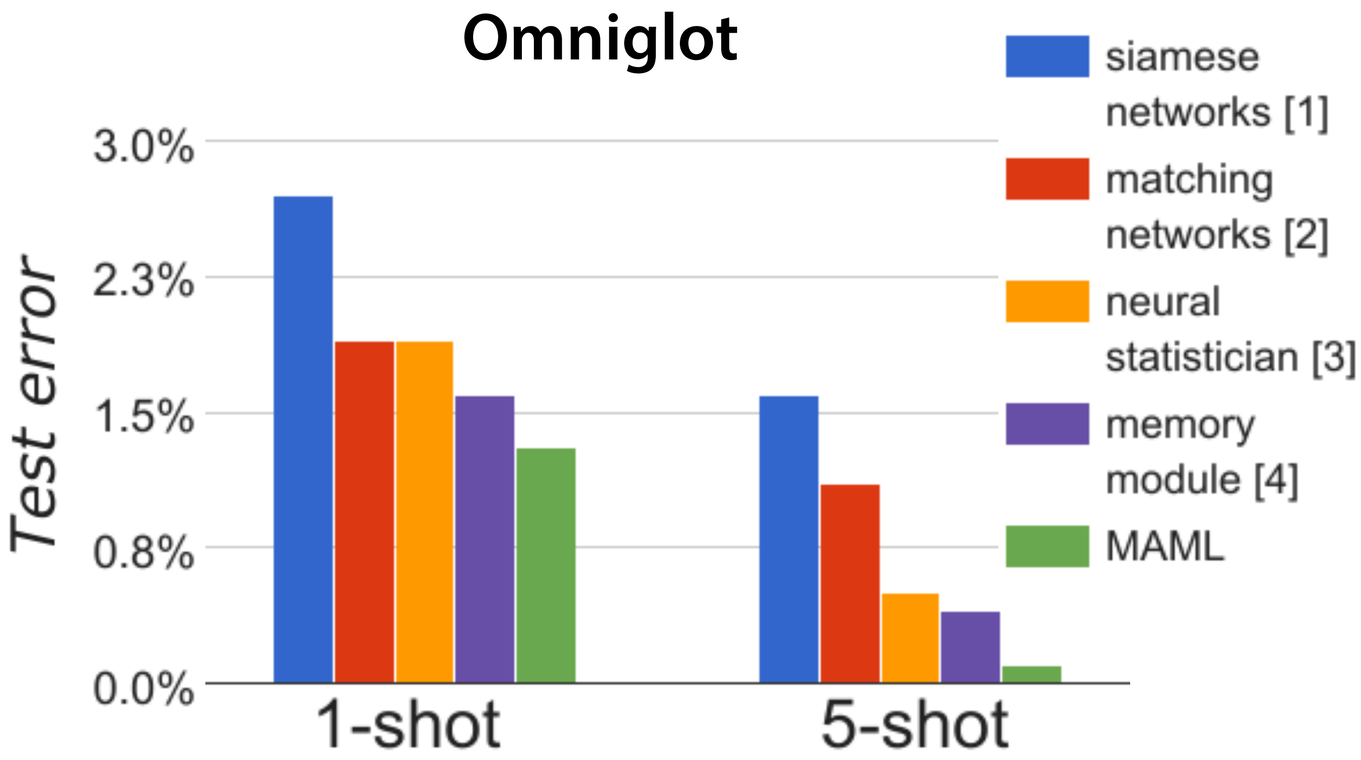


Fast Adaptation in Reinforcement Learning

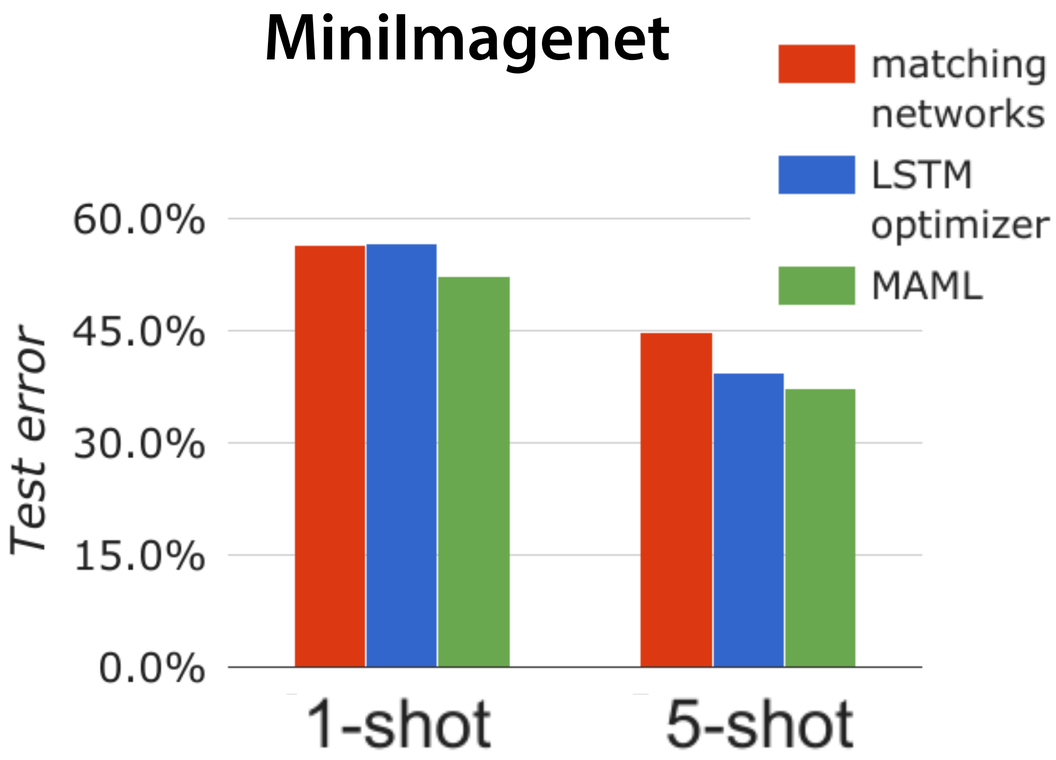




Few-Shot Image Recognition



1200 training classes, 423 test classes

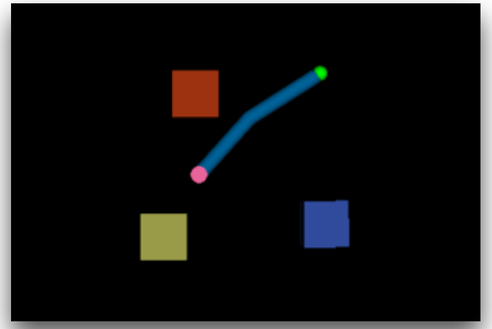


64 training classes, 24 test classes

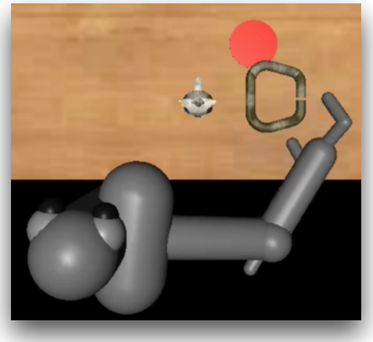
[1] Koch '15 [2] Vinyals et al. '16
[3] Edwards & Storkey '17 [4] Kaiser et al. '17

One-Shot **Visual** Imitation Learning

Vision-Based Manipulation Problems



simulated reaching



simulated pushing

one demonstration provided

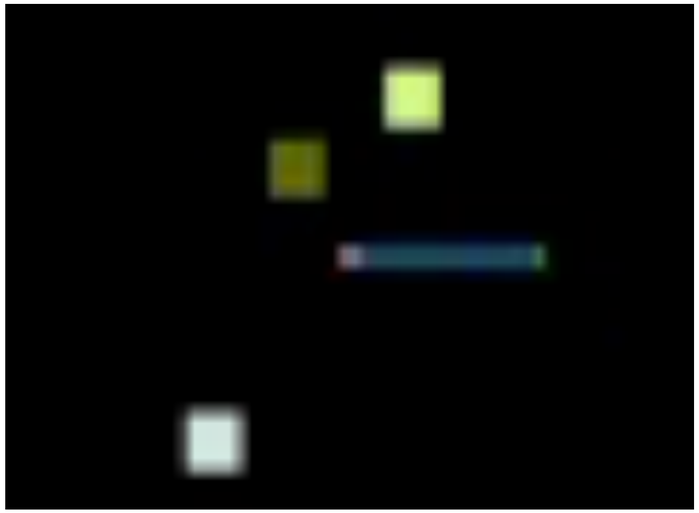
Methods:

MAML | LSTM (Duan et al. '17) | contextual

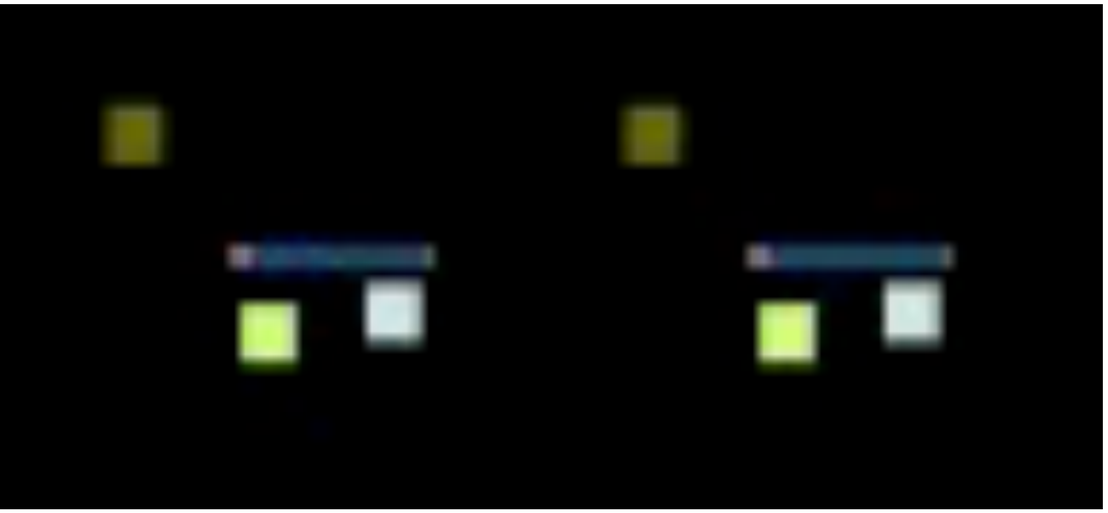
supervised learning for all objectives

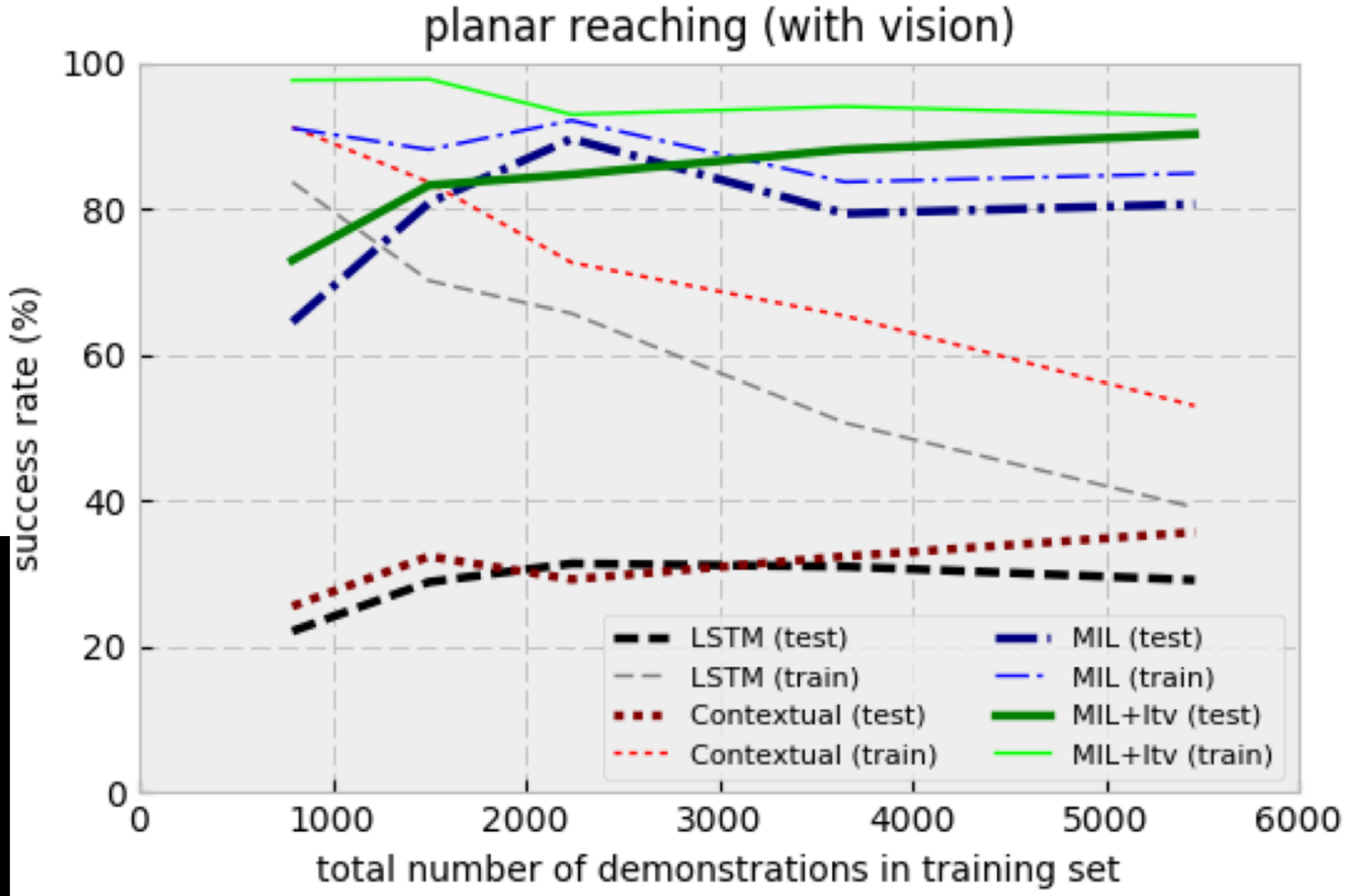
One-Shot Imitation: simulated reaching from pixels

input demonstration



MAML contextual

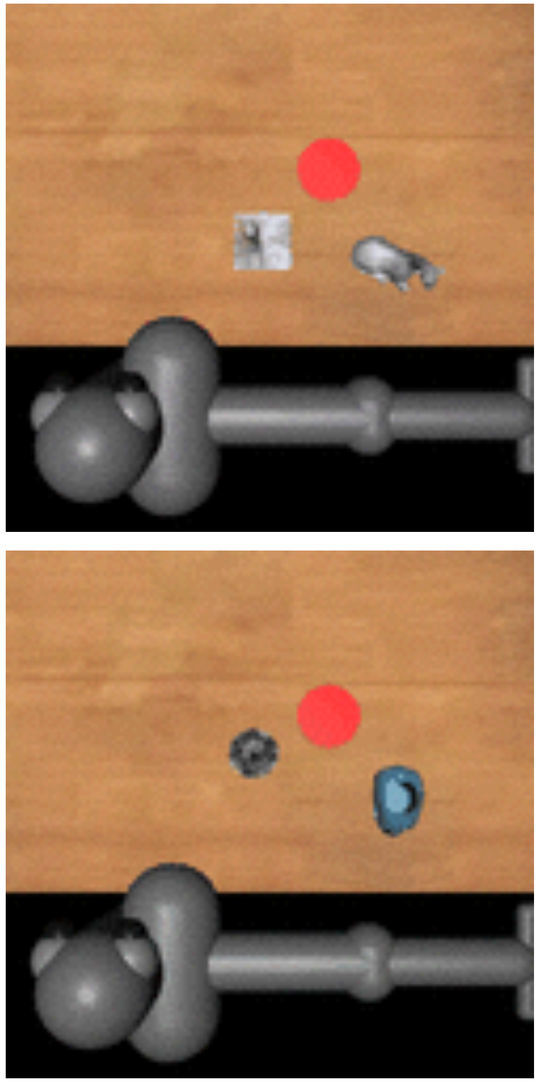




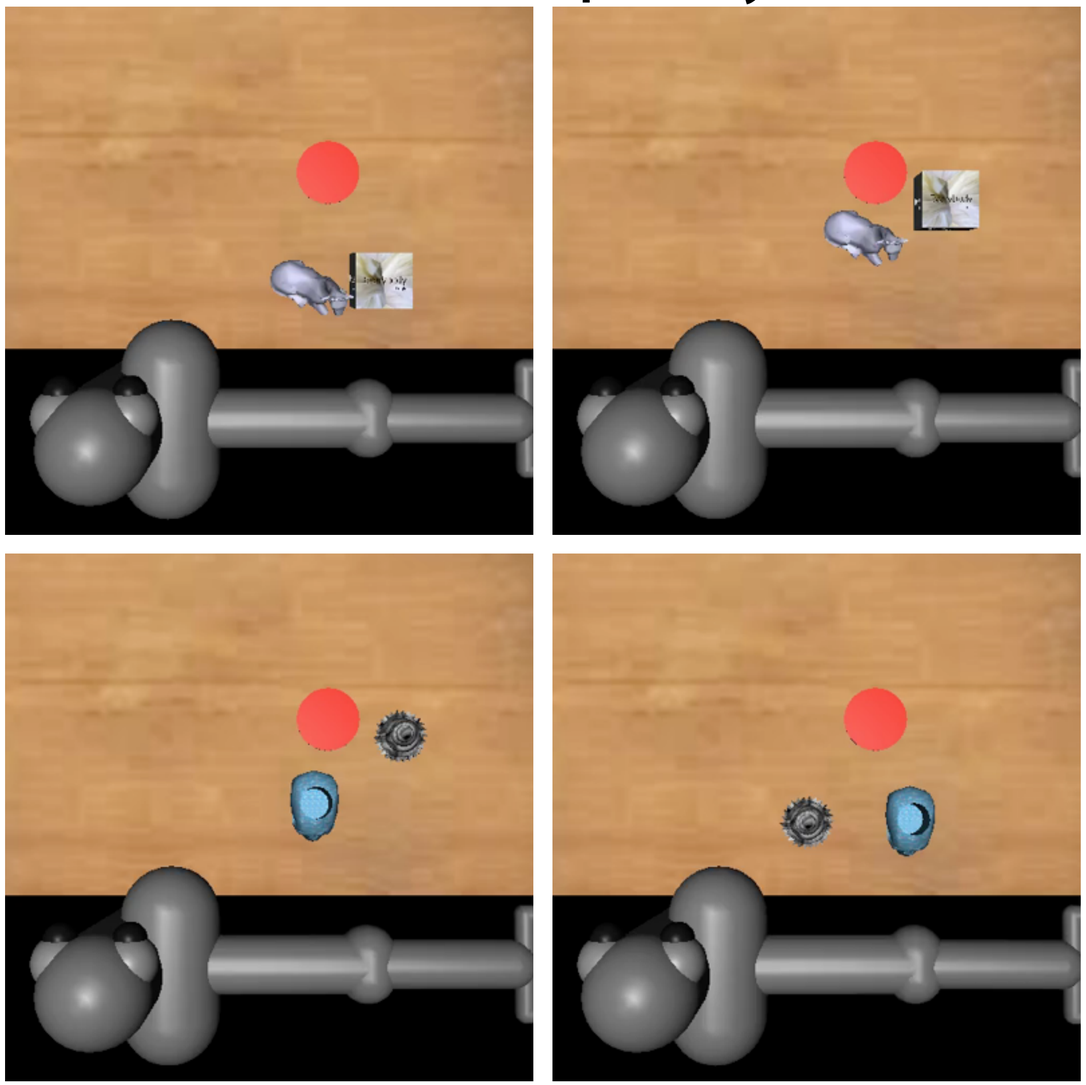
One-Shot Imitation: simulated pushing from pixels

115 random objects
with random textures,
masses, frictions, etc.

input demonstration



learned policy



Takeaway: reuse experience across objects when learning to interact with new objects

Fast Adaptation in Reinforcement Learning

Locomotion problems:

1. <u>run at goal velocity</u>
   (continuous range of tasks)
2. <u>run forward or backward</u>
   (2 tasks)

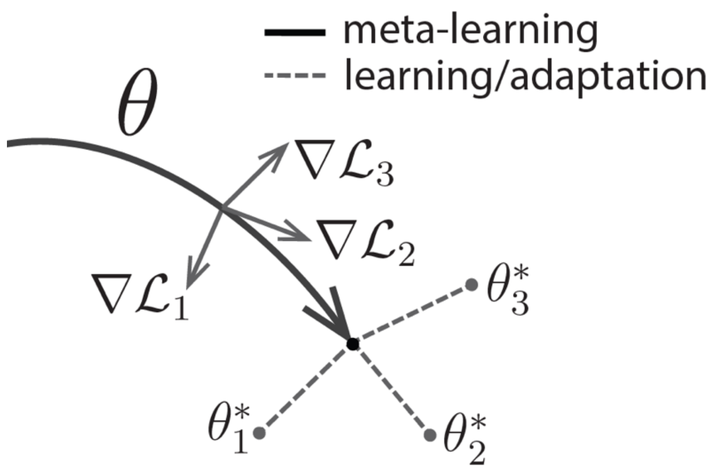
Methods:

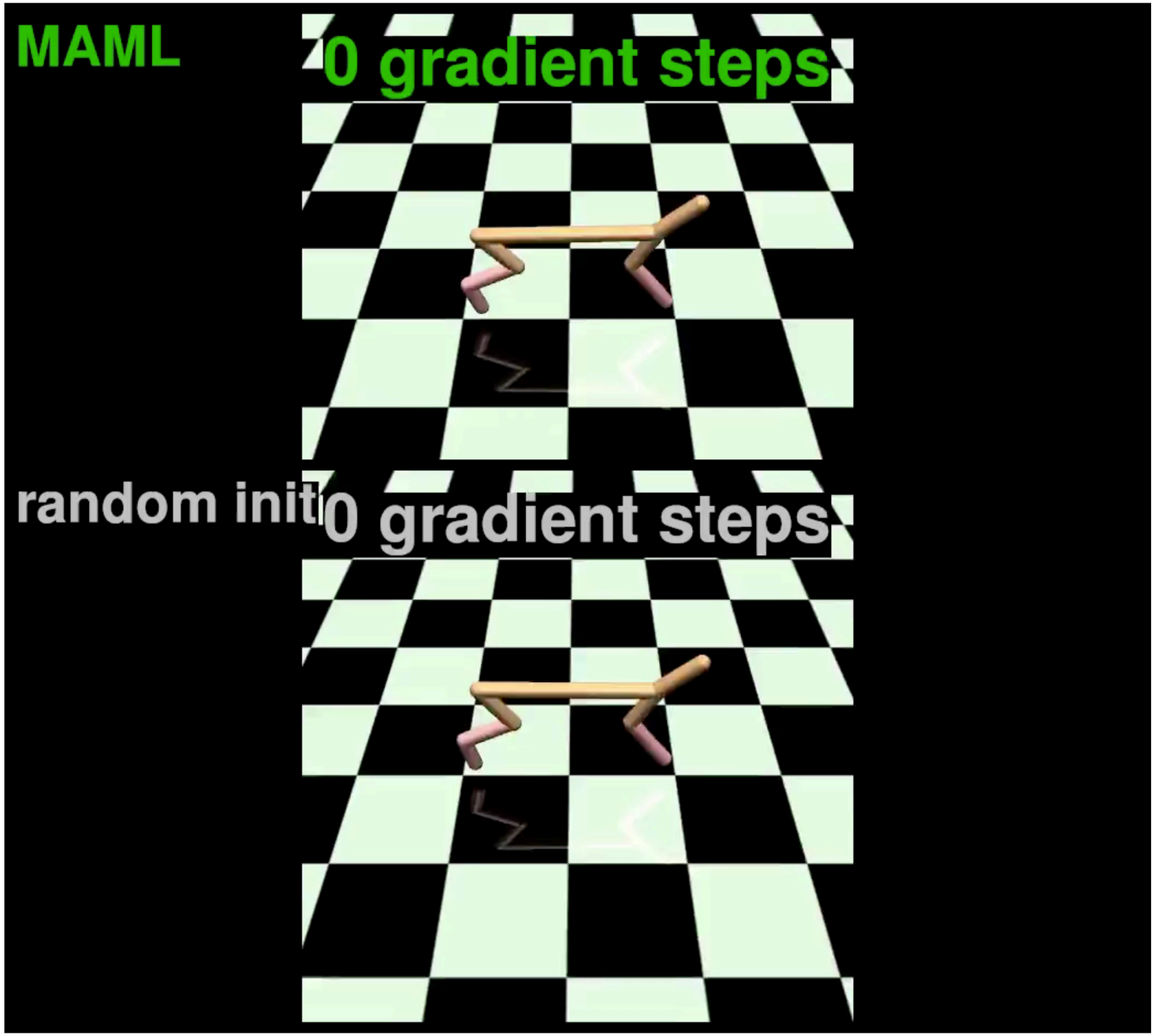
MAML | pretrain on all tasks | random init

meta-learning and adaptation with policy gradients

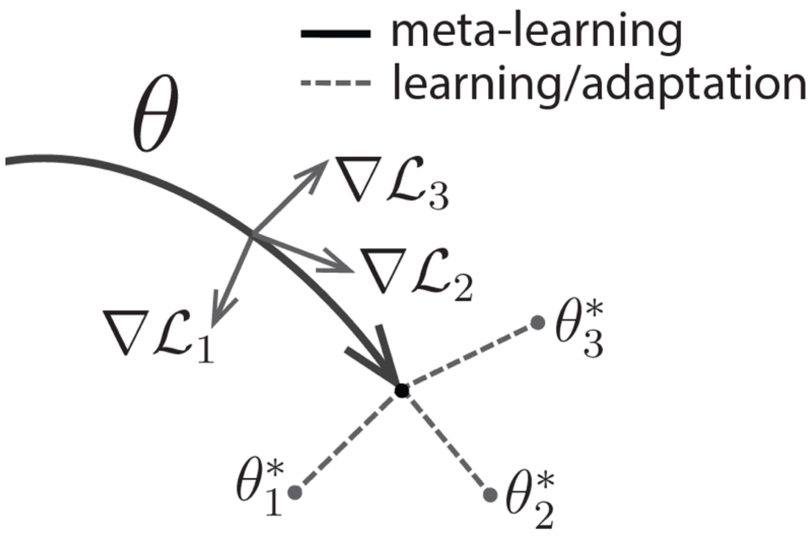
all: 20 roll-outs for learner update

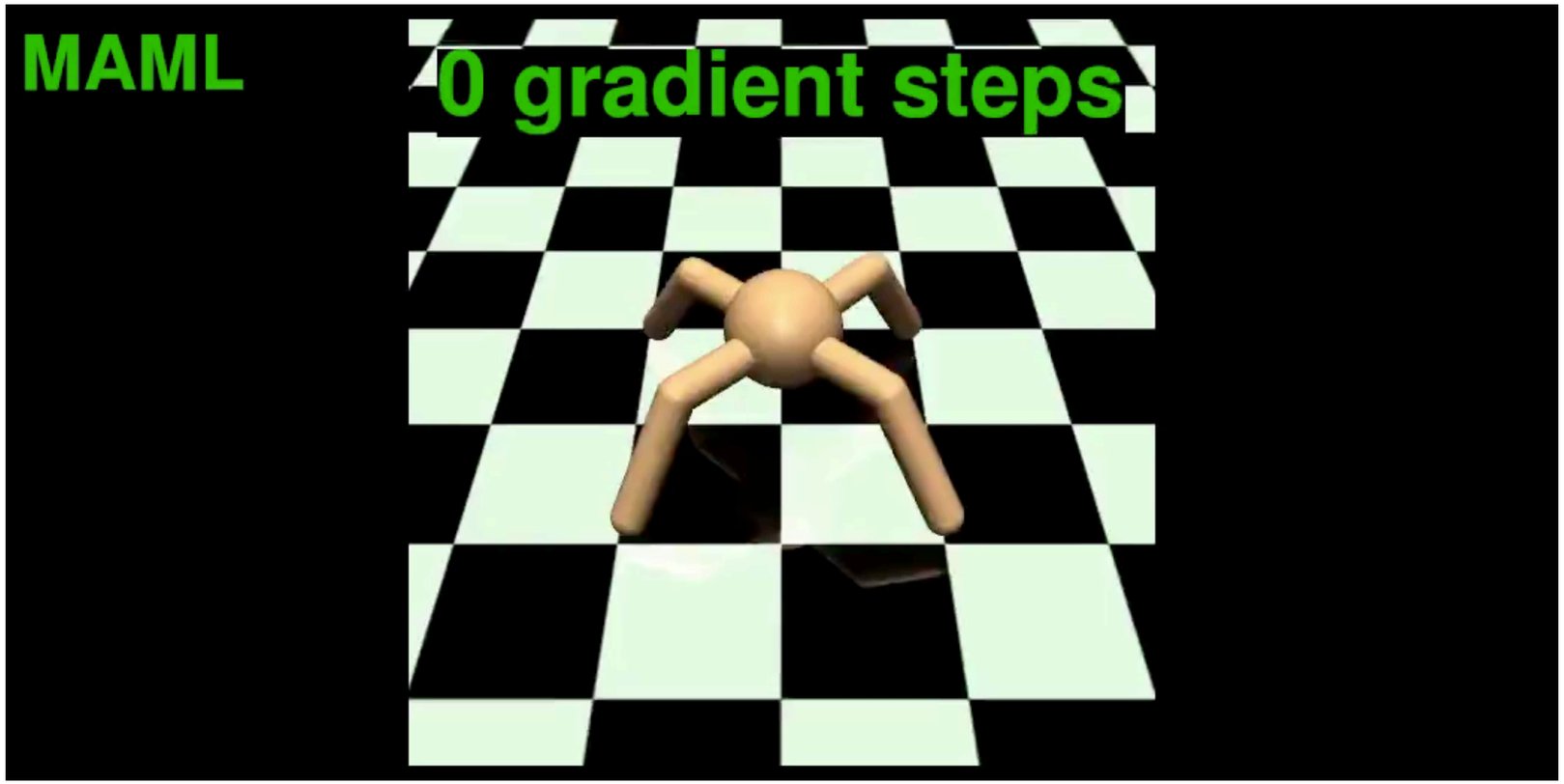
run backward or forward





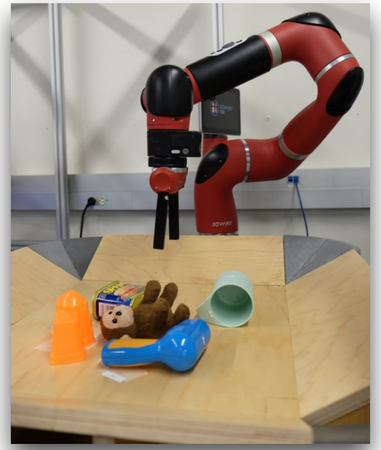
walk/run at goal velocity





Main Takeaway: Robots can reuse prior experience for faster learning

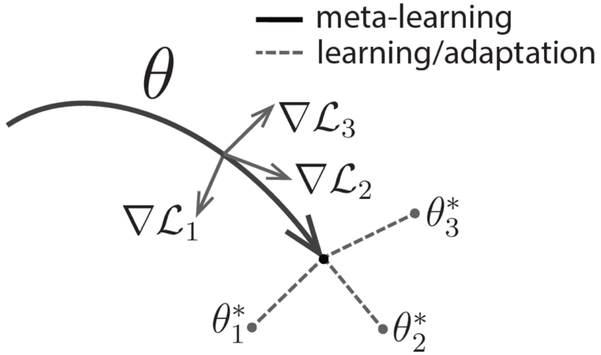
1. Learn about the physical world

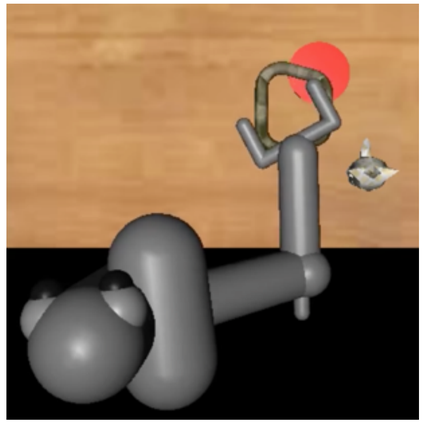


Finn et al. NIPS '16, *Finn* & Levine ICRA '17

Ebert et al. '17 (under review)

2. Learn to learn



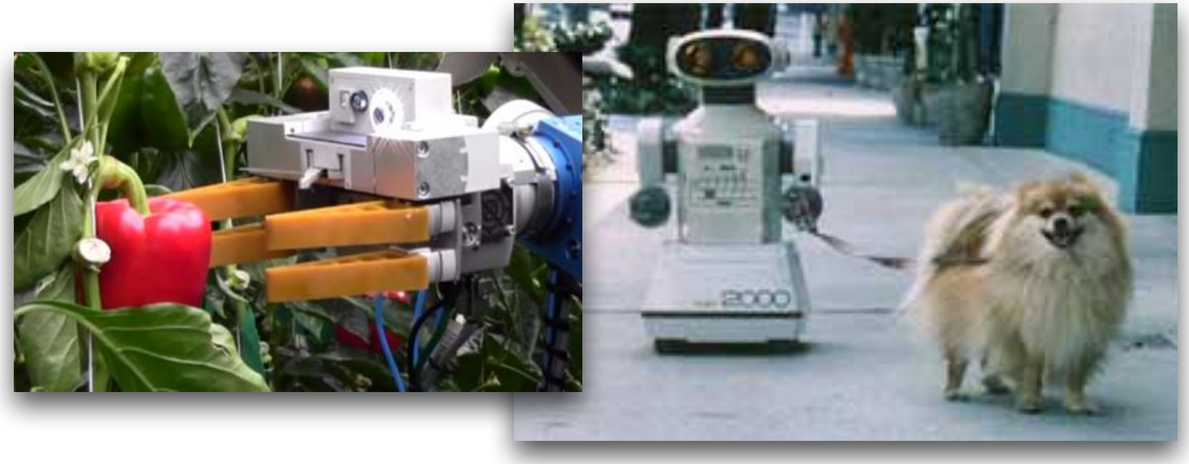


Finn et al. ICML '17

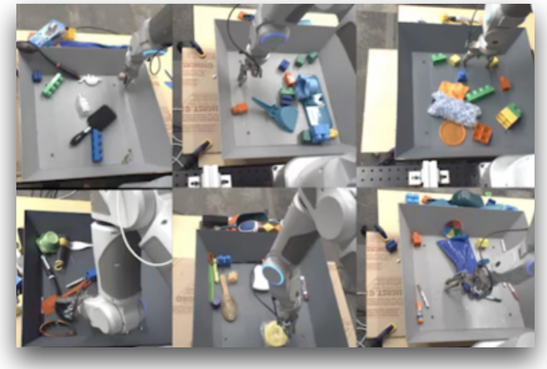
Finn*, Yu* et al. '17 (under review)

Future Directions

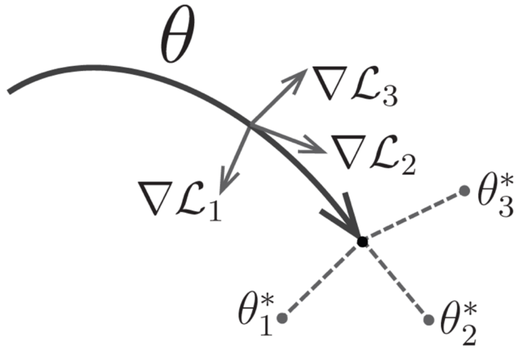
continual learning in the real world



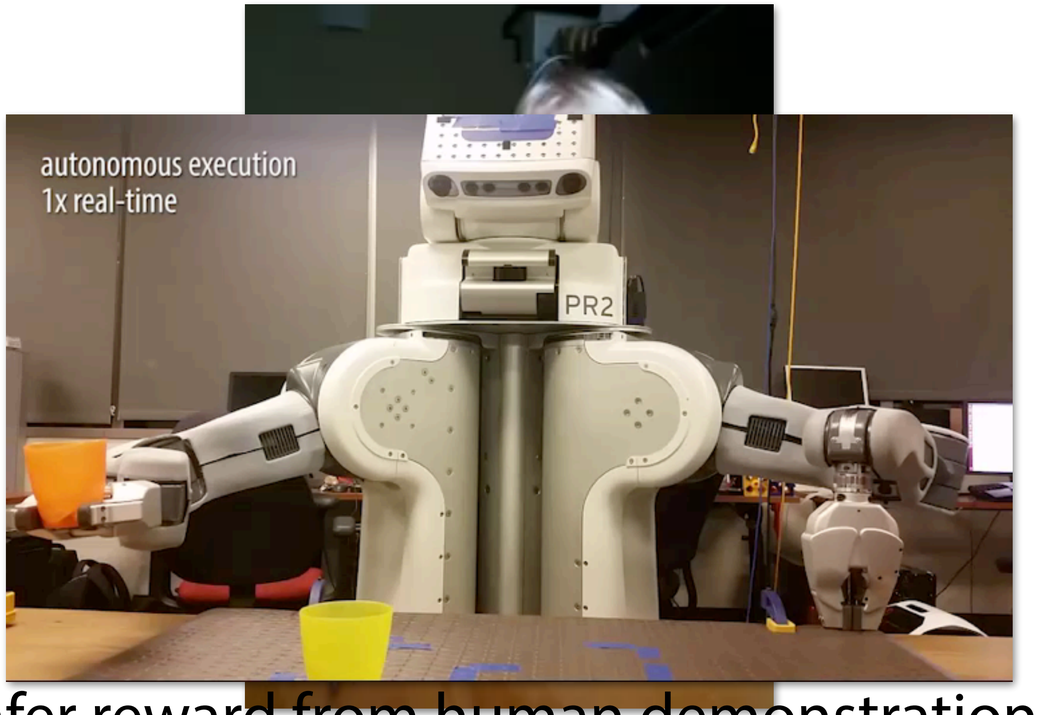
learning about the world



learning to learn



richer ways to specify goals



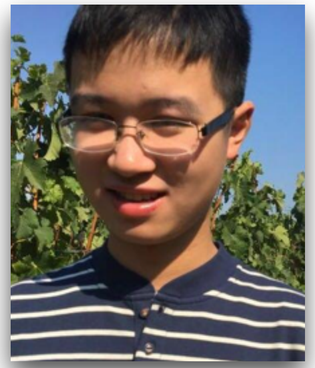
infer reward from human demonstrations
what is the reward?
Finn, Levine, Abbeel, ICML '16

Collaborators

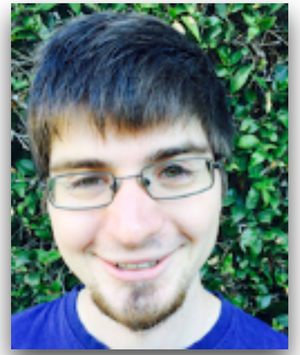
Sergey Levine

Pieter Abbeel

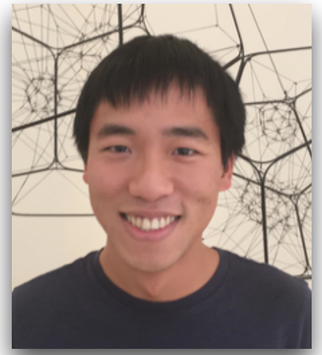
Frederik Ebert



Tianhe (Kevin) Yu



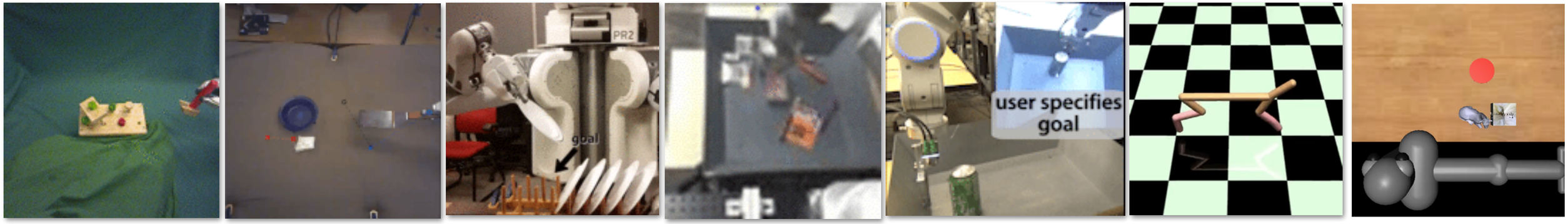
Ian Goodfellow



Alex Lee

Tianhao Zhang

Questions?



All **data** and **code** linked at: people.eecs.berkeley.edu/~cbfinn

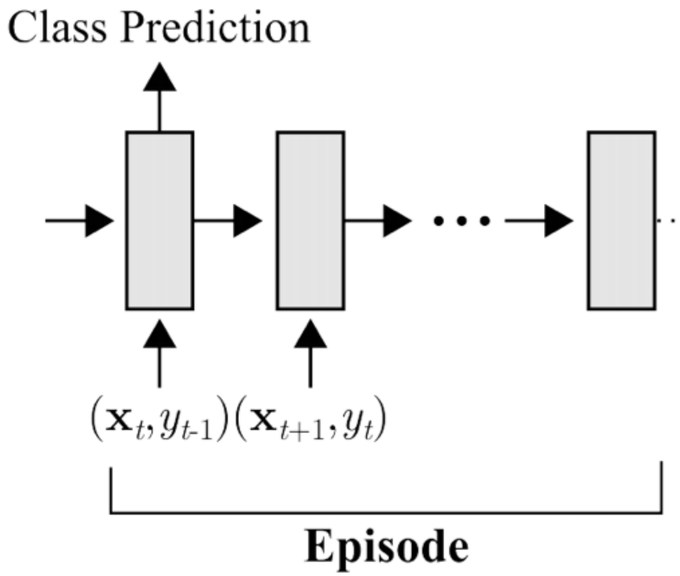
cbfinn@eecs.berkeley.edu

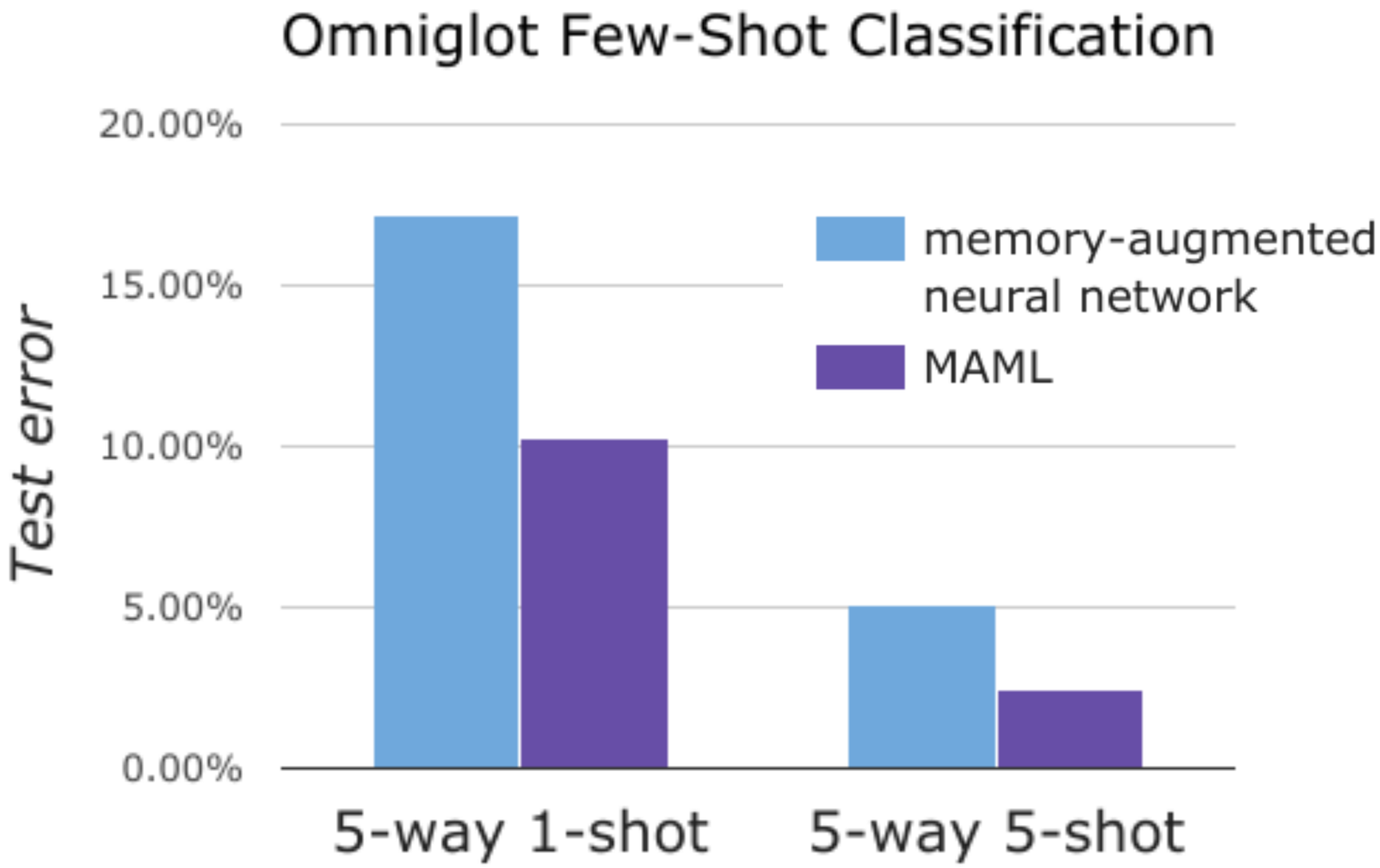
Omniglot Few-Shot Classification

Omniglot dataset

Memory-Augmented Approach

Santoro et al.'16





non-convolutional network

Foresight quantitative comparison