# Large-Scale Self-Supervised Robotic Learning

Chelsea Finn

In collaboration with Sergey Levine and Ian Goodfellow

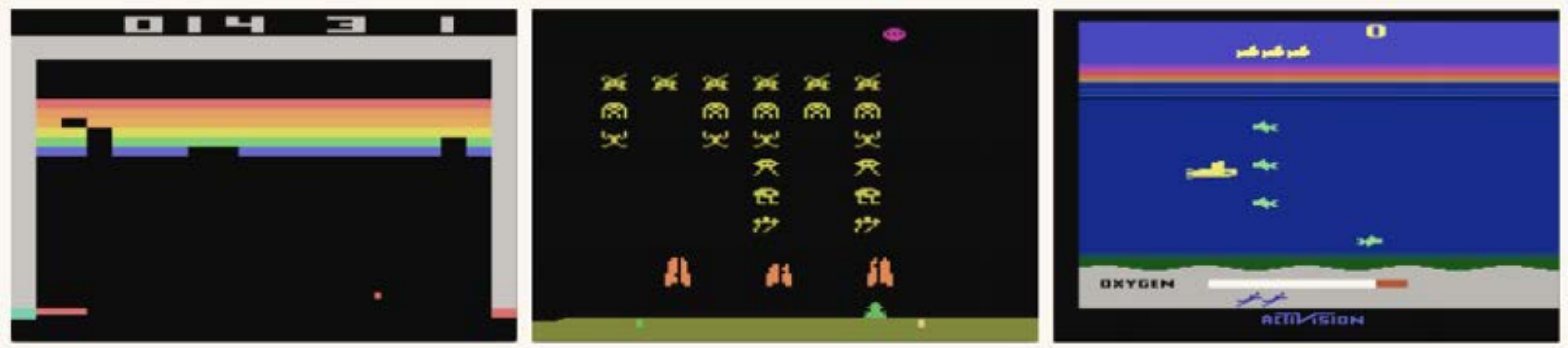
Berkeley Artificial Intelligence Research Laboratory

Google

OpenAI

# Generalization in Reinforcement Learning

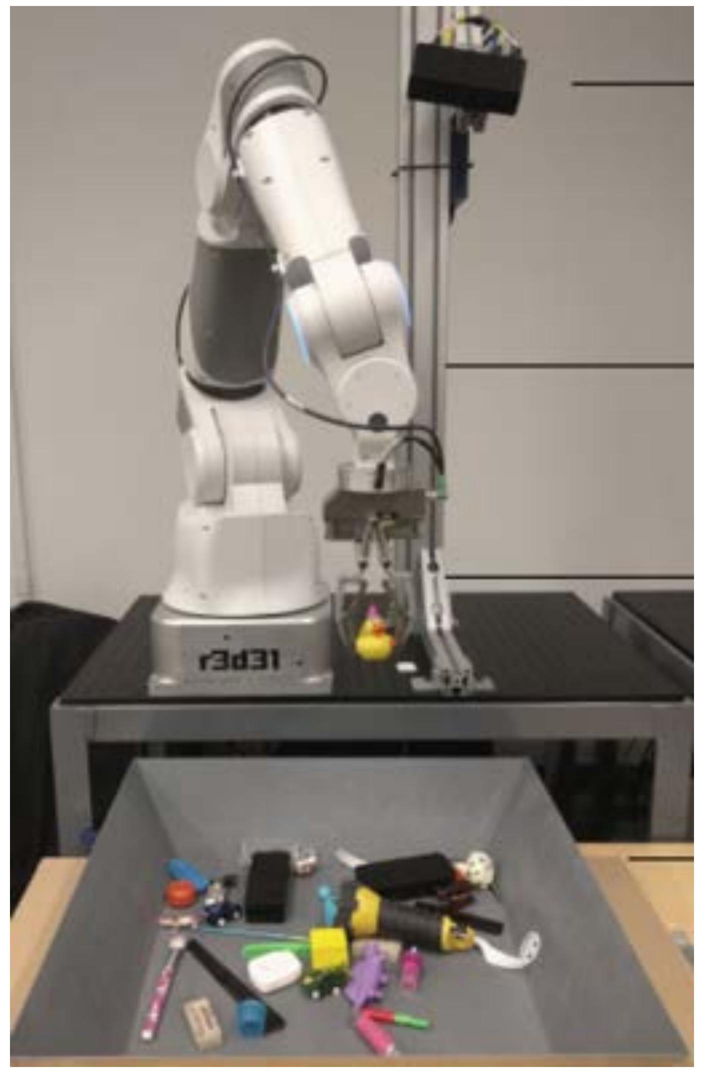
## to object instances



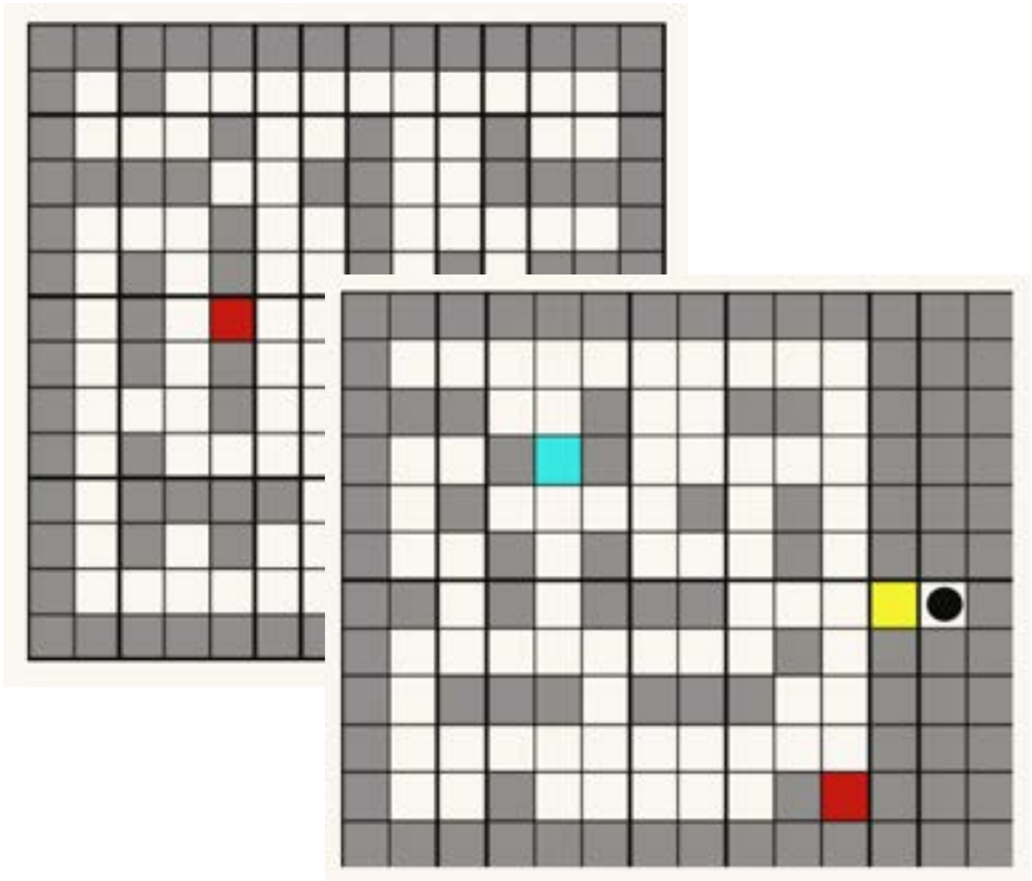
Mnih et al. '15

## to tasks and environments

Pinto & Gupta '16



Levine et al. '16



Oh et al. '16

# Generalization in Reinforcement Learning

need <u>data</u> → **scale up**


**First lesson**: human supervision doesn't scale
(providing rewards, reseting the environment, etc.)

# Generalization in Reinforcement Learning

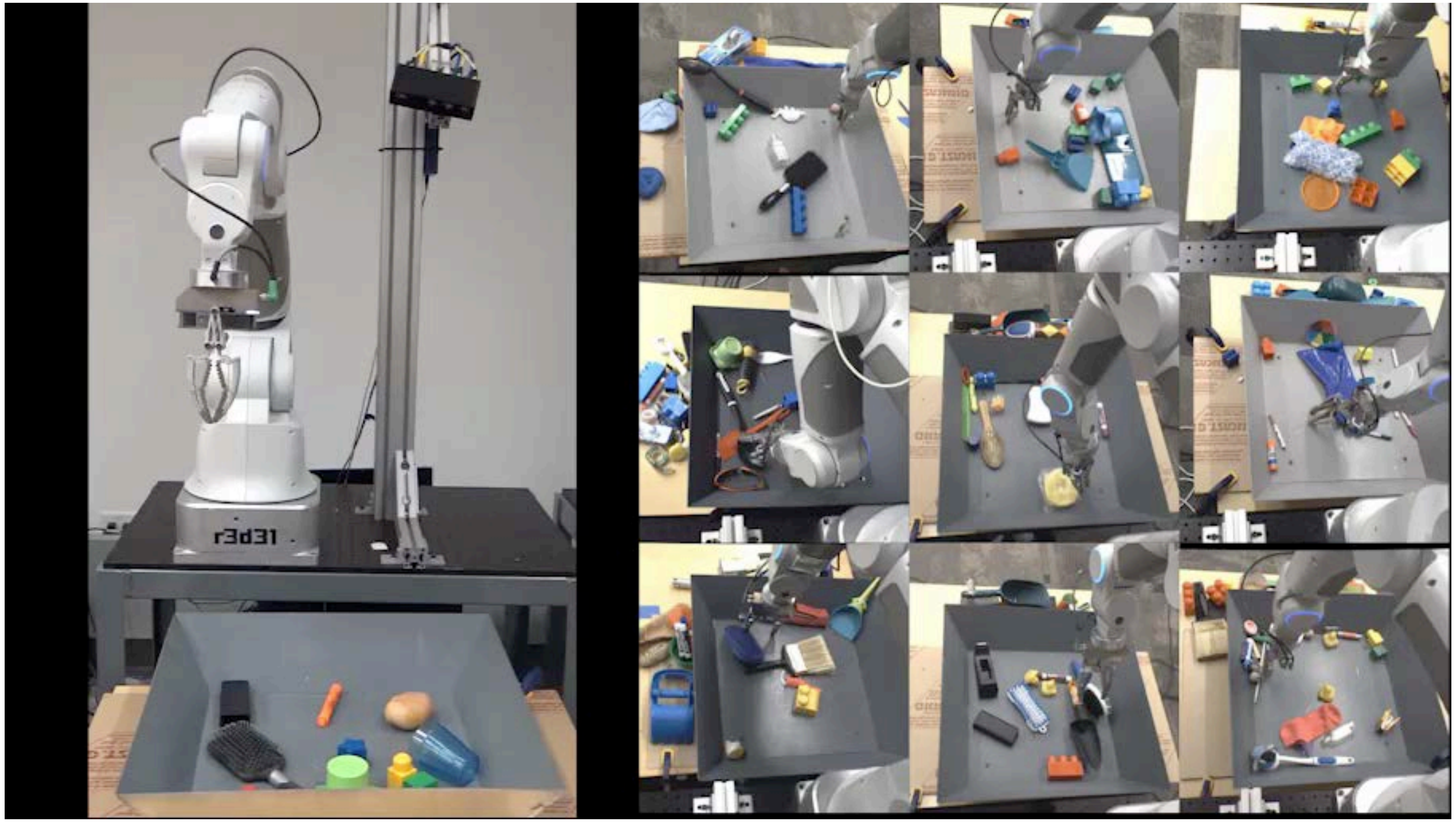
need data → **scale up**

where does the supervision come from?
→ **self-supervision**

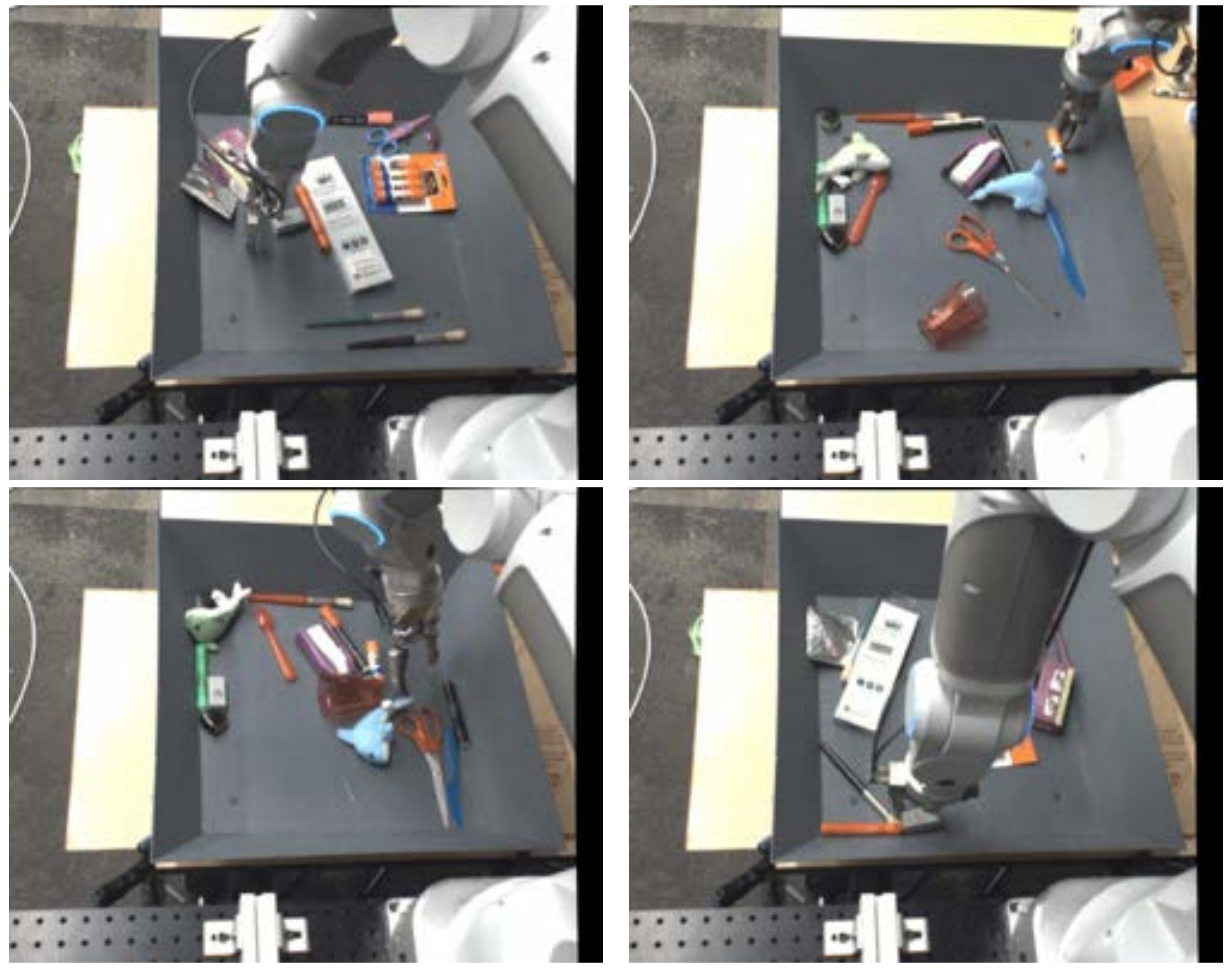
most deep RL algorithms learn a single-purpose policy
→ **learn general-purpose model**

Evaluating unsupervised methods?
lacking task-driven metrics for unsupervised learning

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



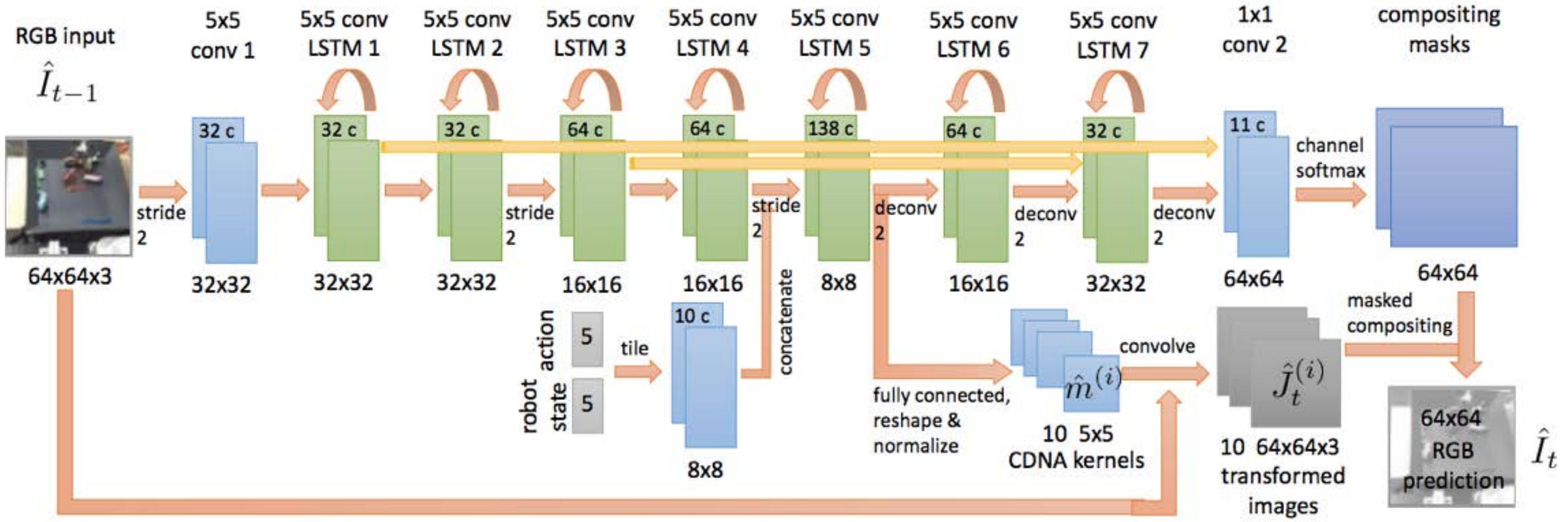

test set with
novel objects



data publicly available for download **sites.google.com/site/brainrobotdata**

# Train predictive model

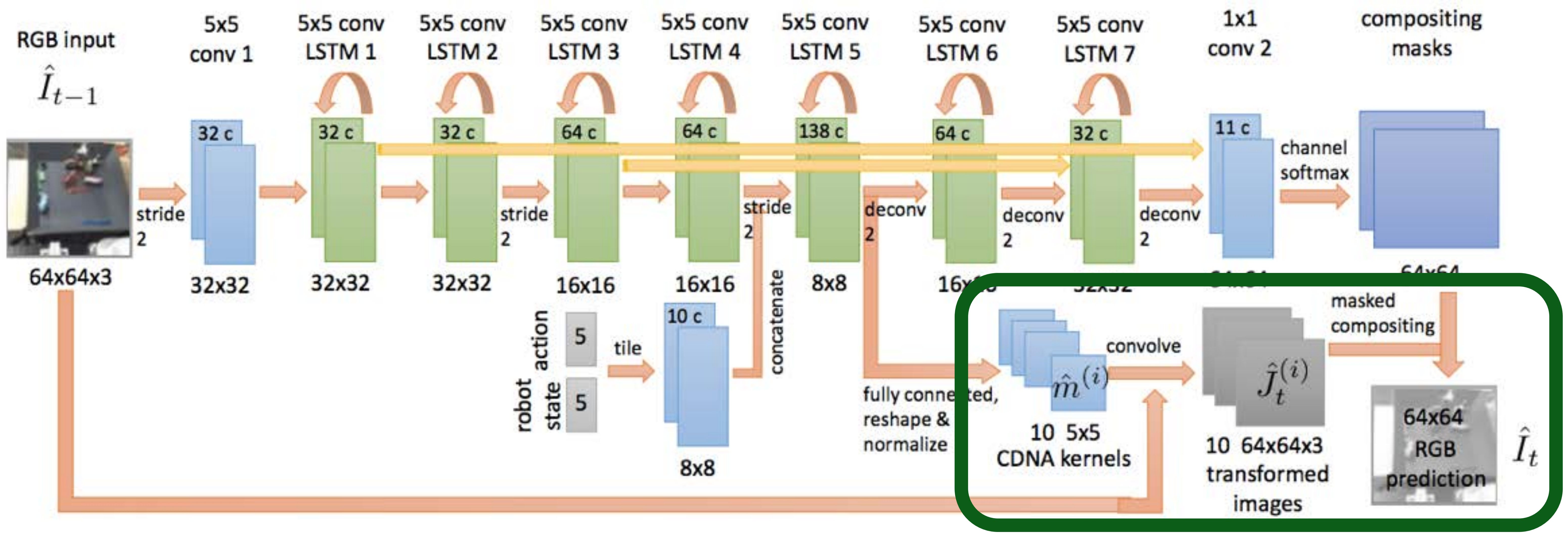
**convolutional LSTMs**



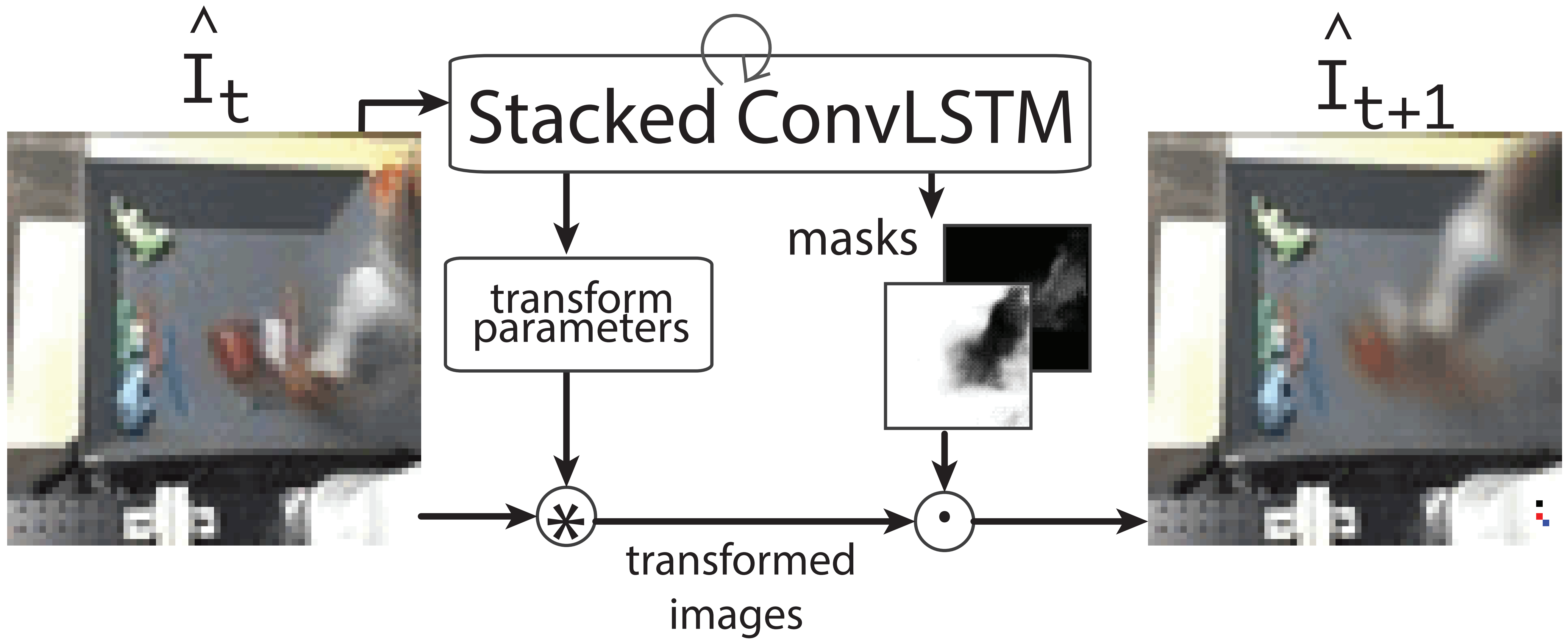

**action-conditioned** **stochastic flow prediction**

- feed back model's predictions for multi-frame prediction
- trained with $l_2$ loss

# Train predictive model

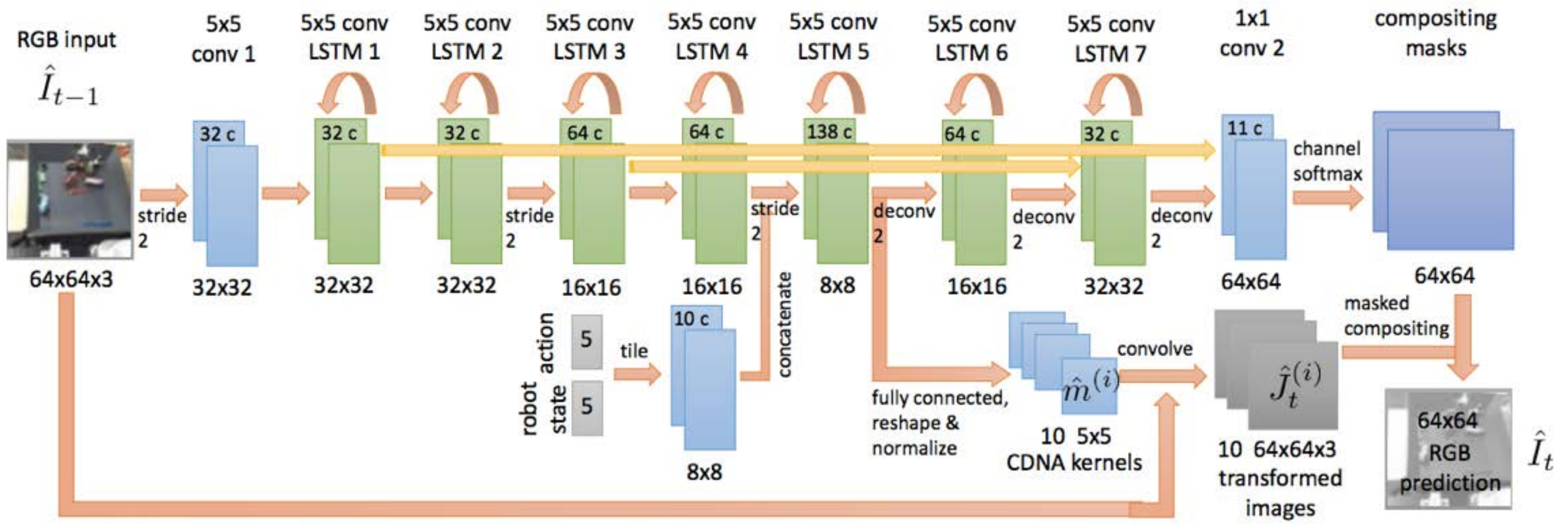



**stochastic flow prediction**



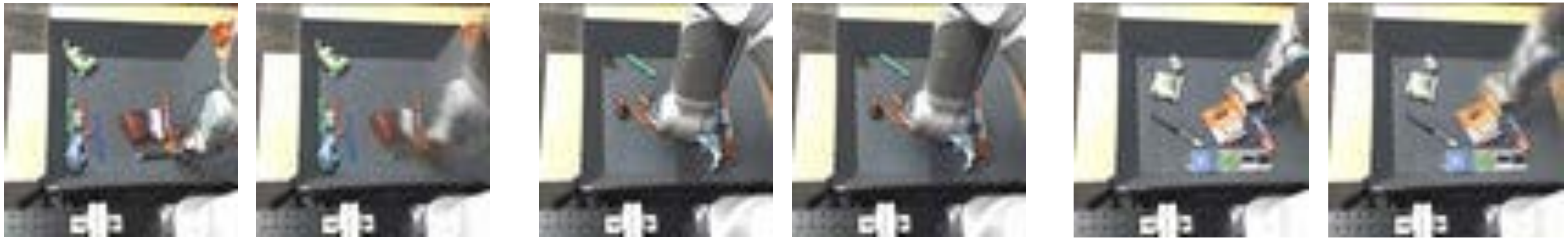
# Train predictive model

**convolutional LSTMs**




**action-conditioned** **stochastic flow prediction**

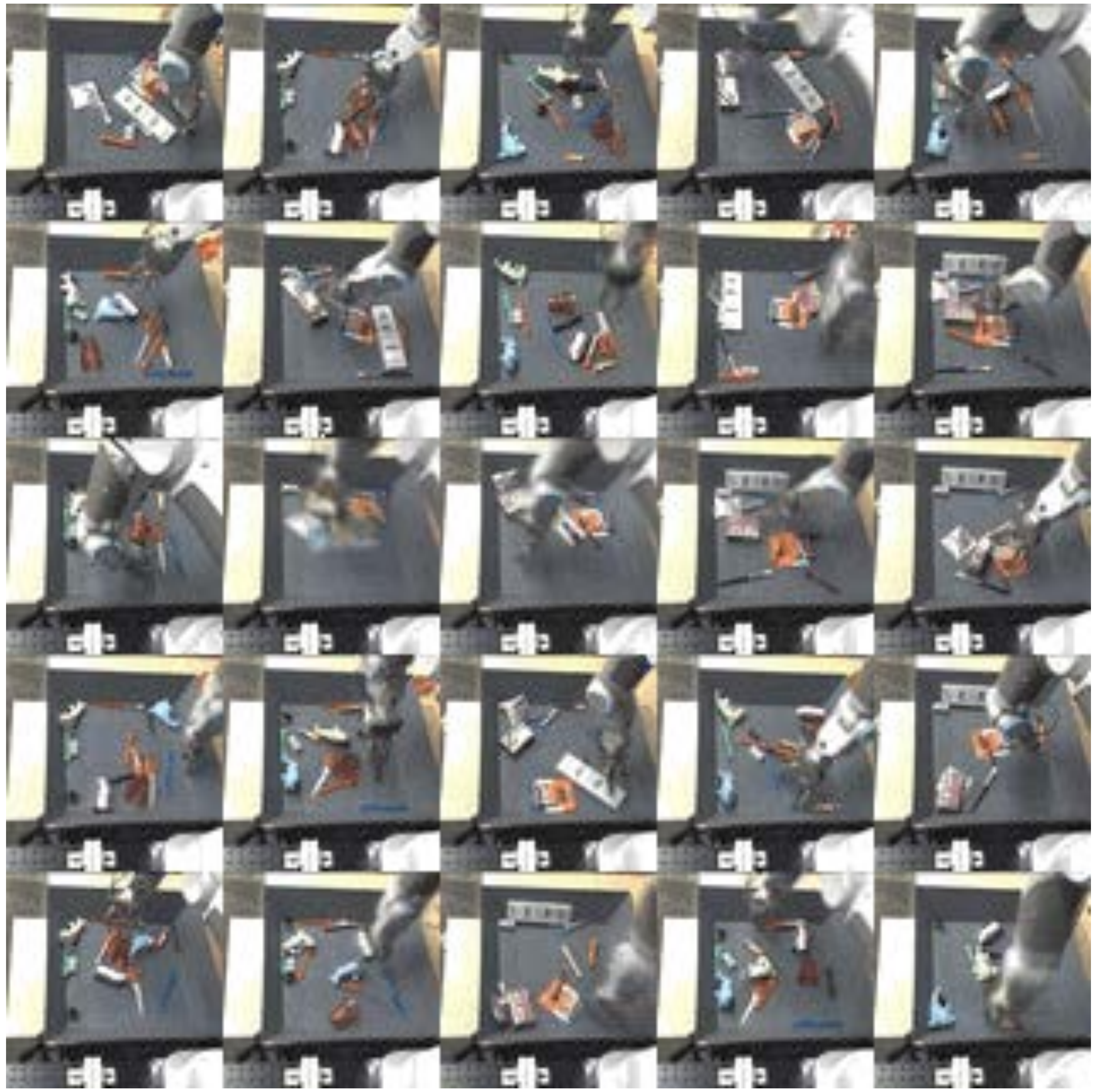
evaluate on held-out objects



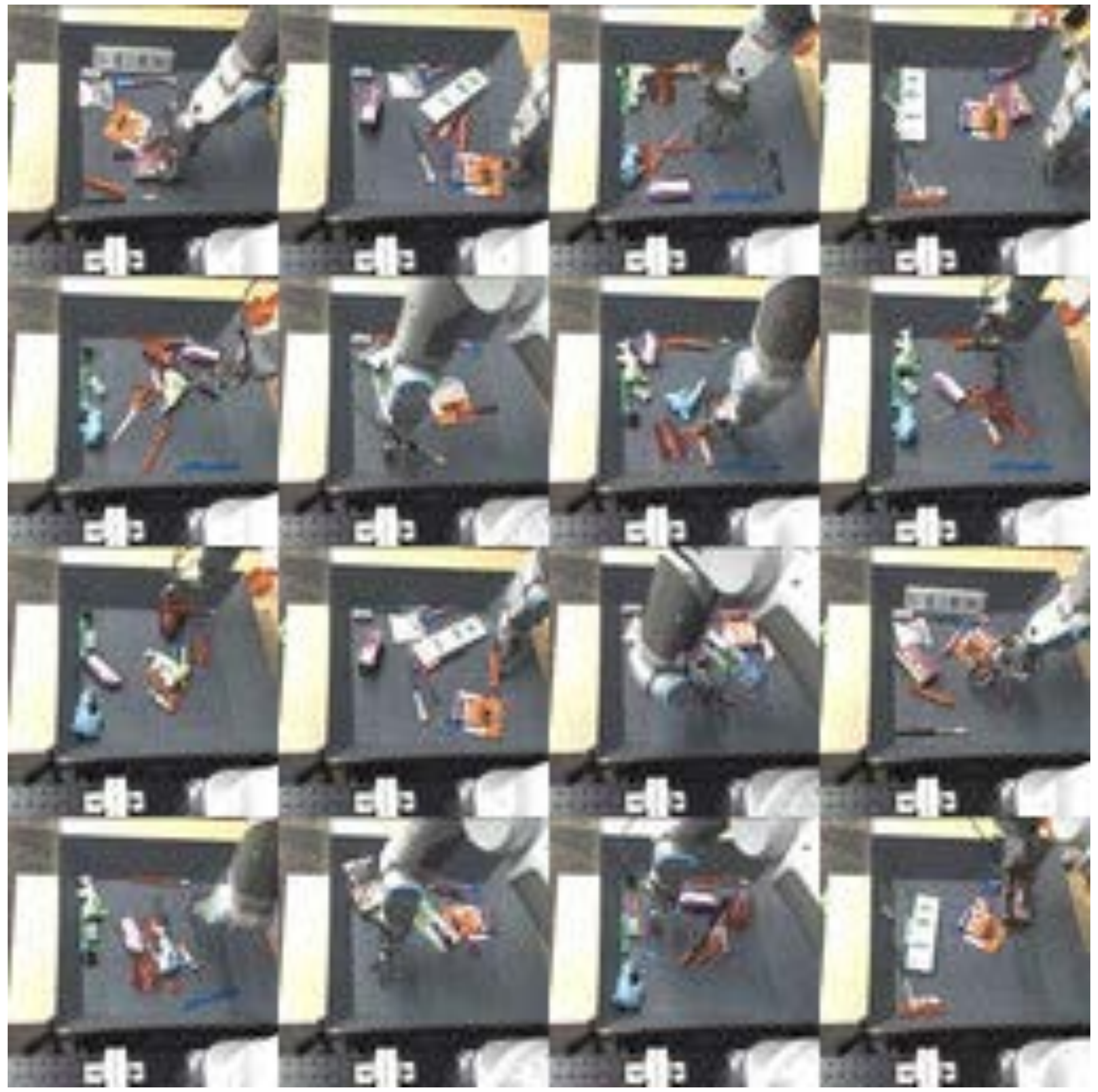
Are these predictions good?

# Train predictive model

Finn et al., '16



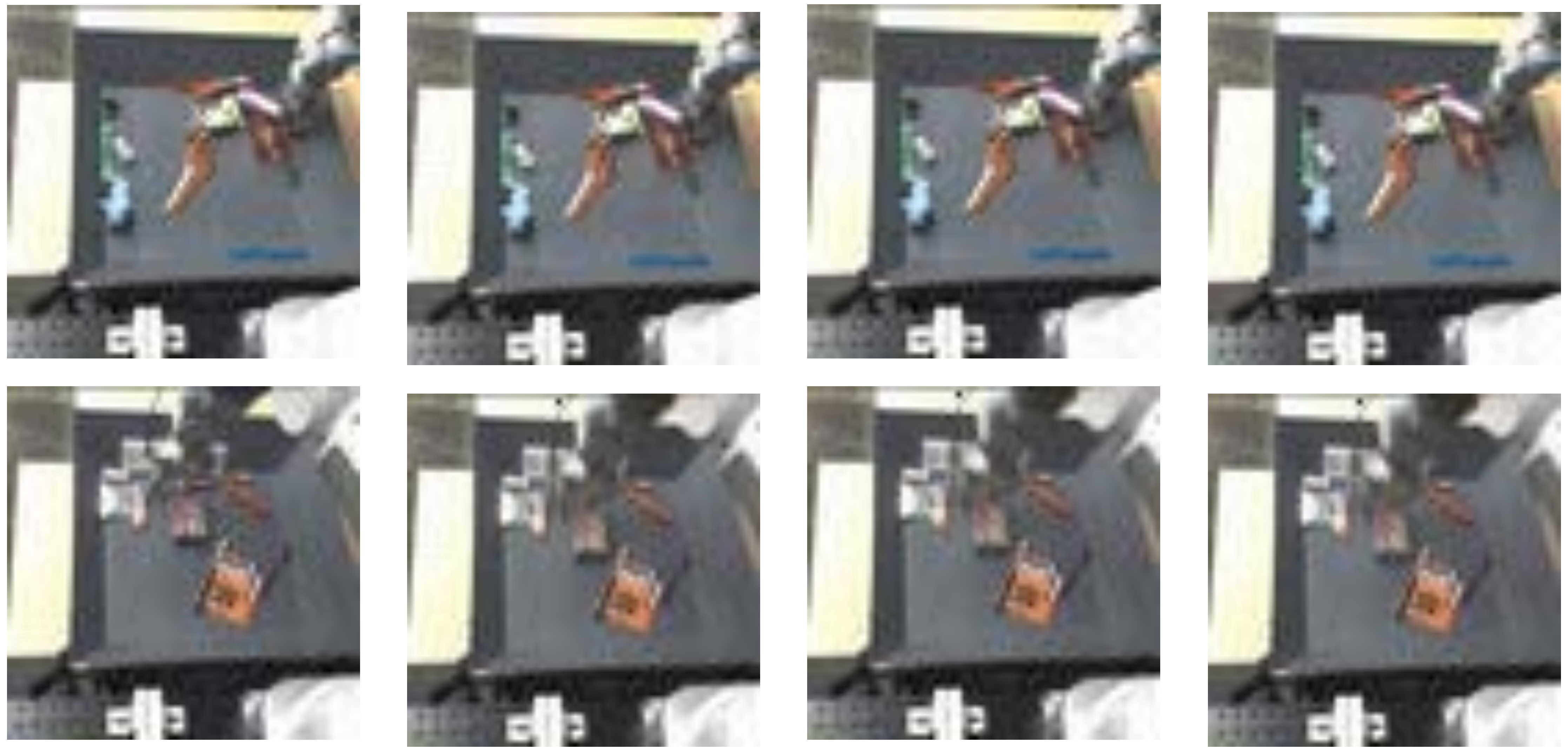
Kalchbrenner et al., '16



Are these predictions good? accurate? useful?

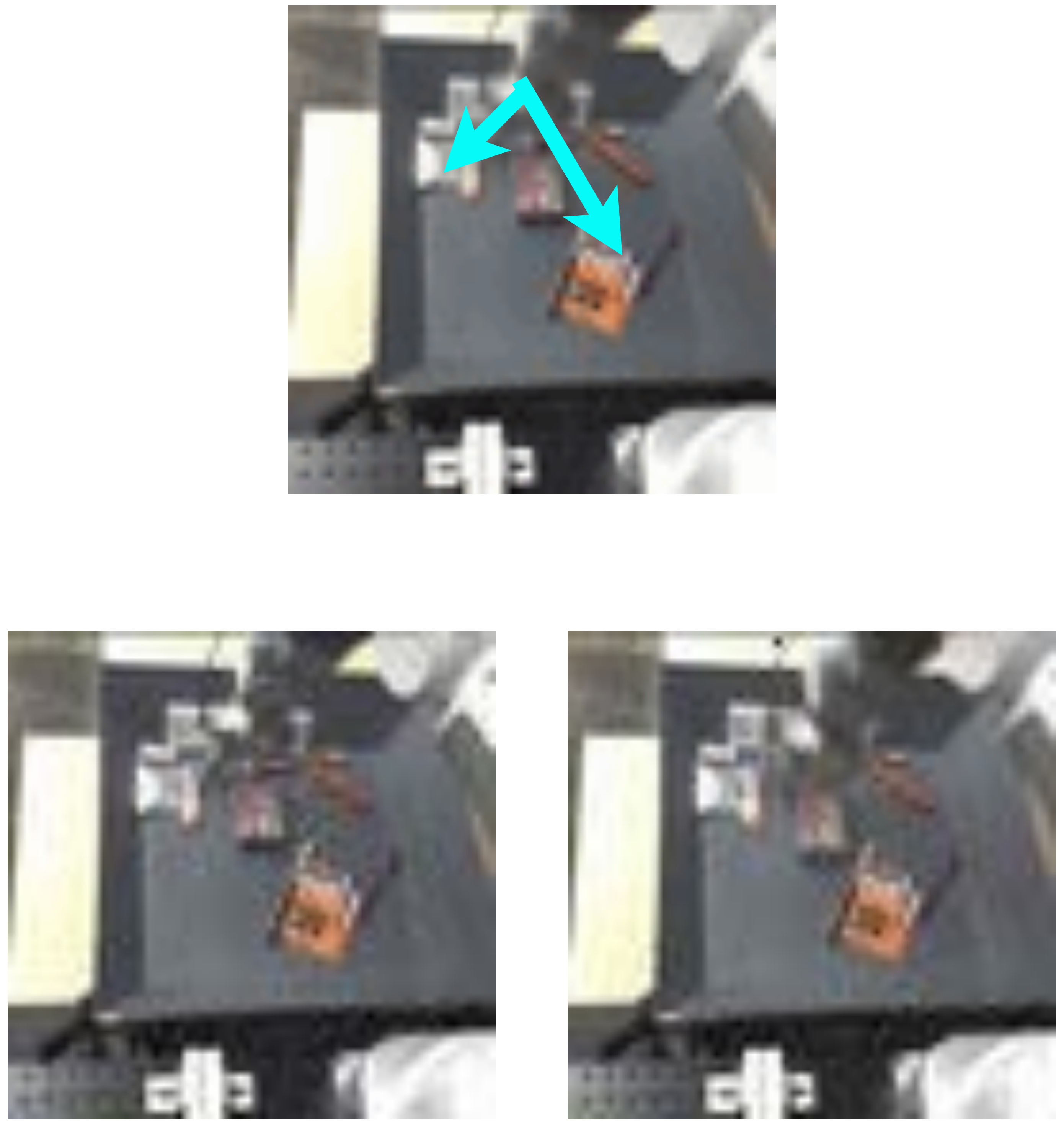
# What is prediction good for?

action magnitude: 0x 0.5x 1x 1.5x



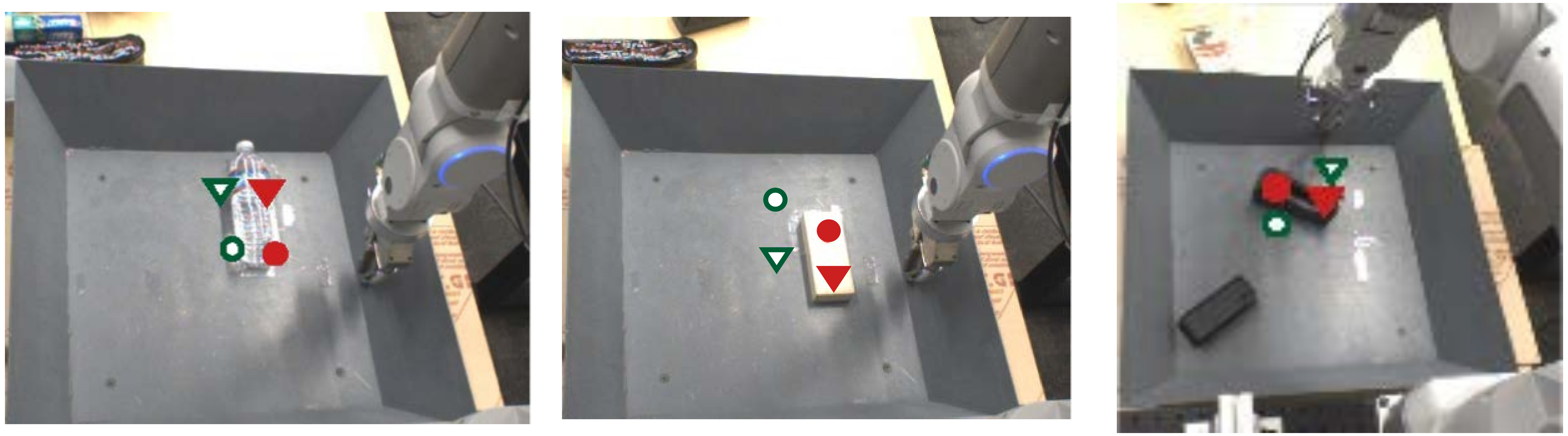
# Visual MPC: Planning with Visual Foresight

1. Sample N 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

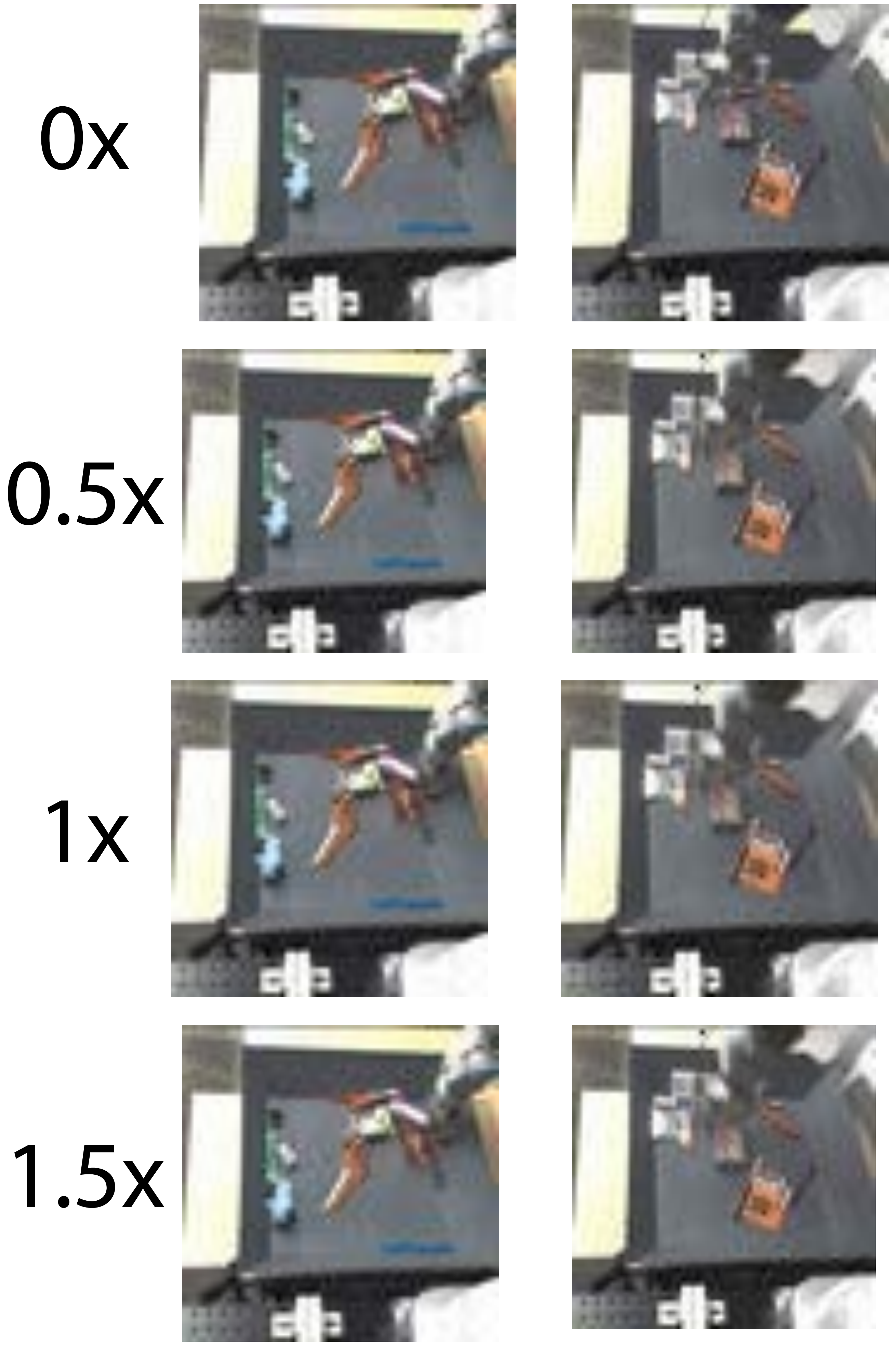


# Which future is the best one?

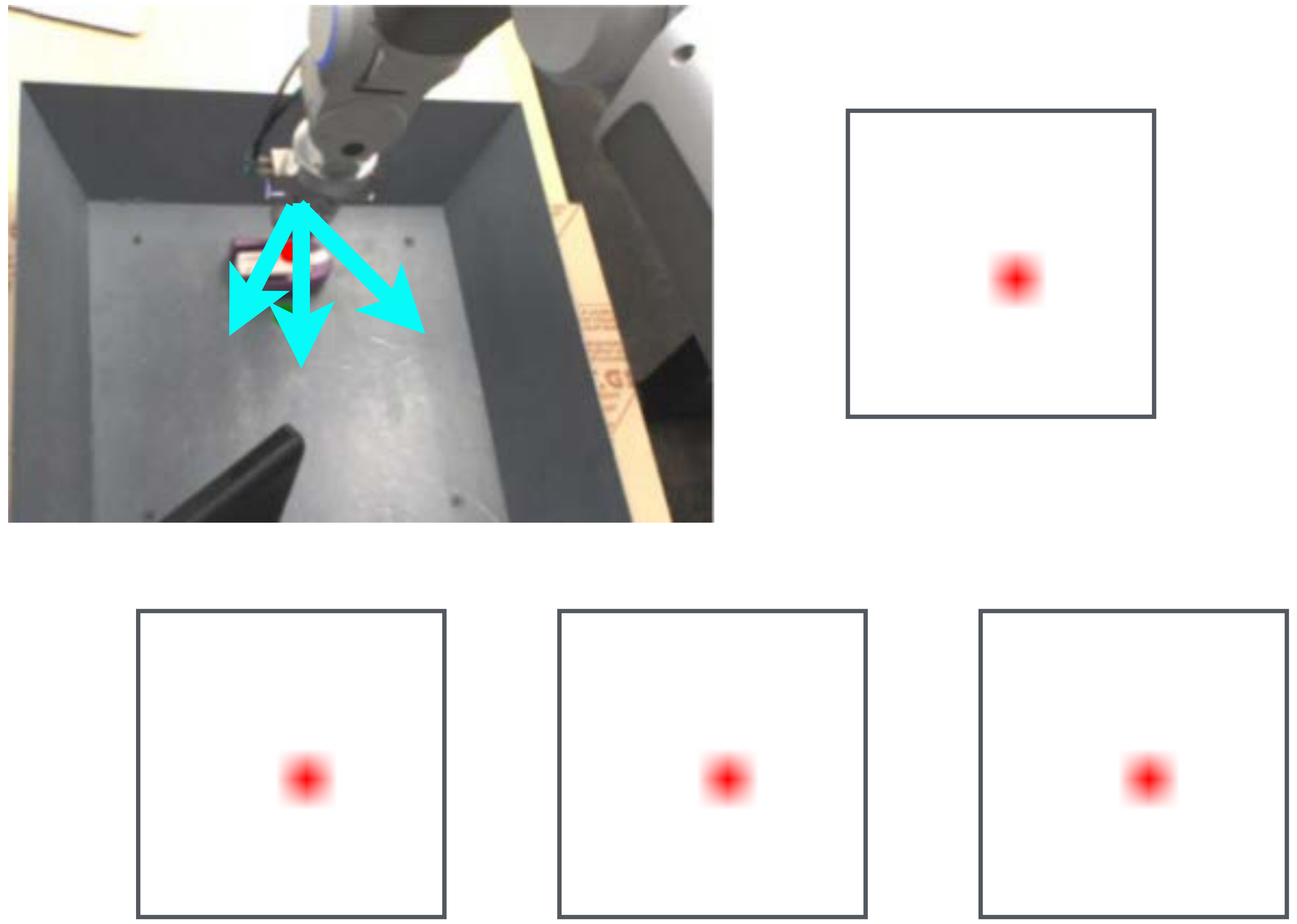
Specify goal by selecting where pixels should move.



Select future with maximal probability of pixels reaching their respective goals.




We can predict **how pixels will move** based on the robot's actions



*output is the mean of a probability distribution over pixel motion predictions*

# How it works

# Does it work?


- evaluation on short pushes of **novel objects**

- translation & rotation

**Only human involvement** during training is:
programming initial motions and providing objects to play with.

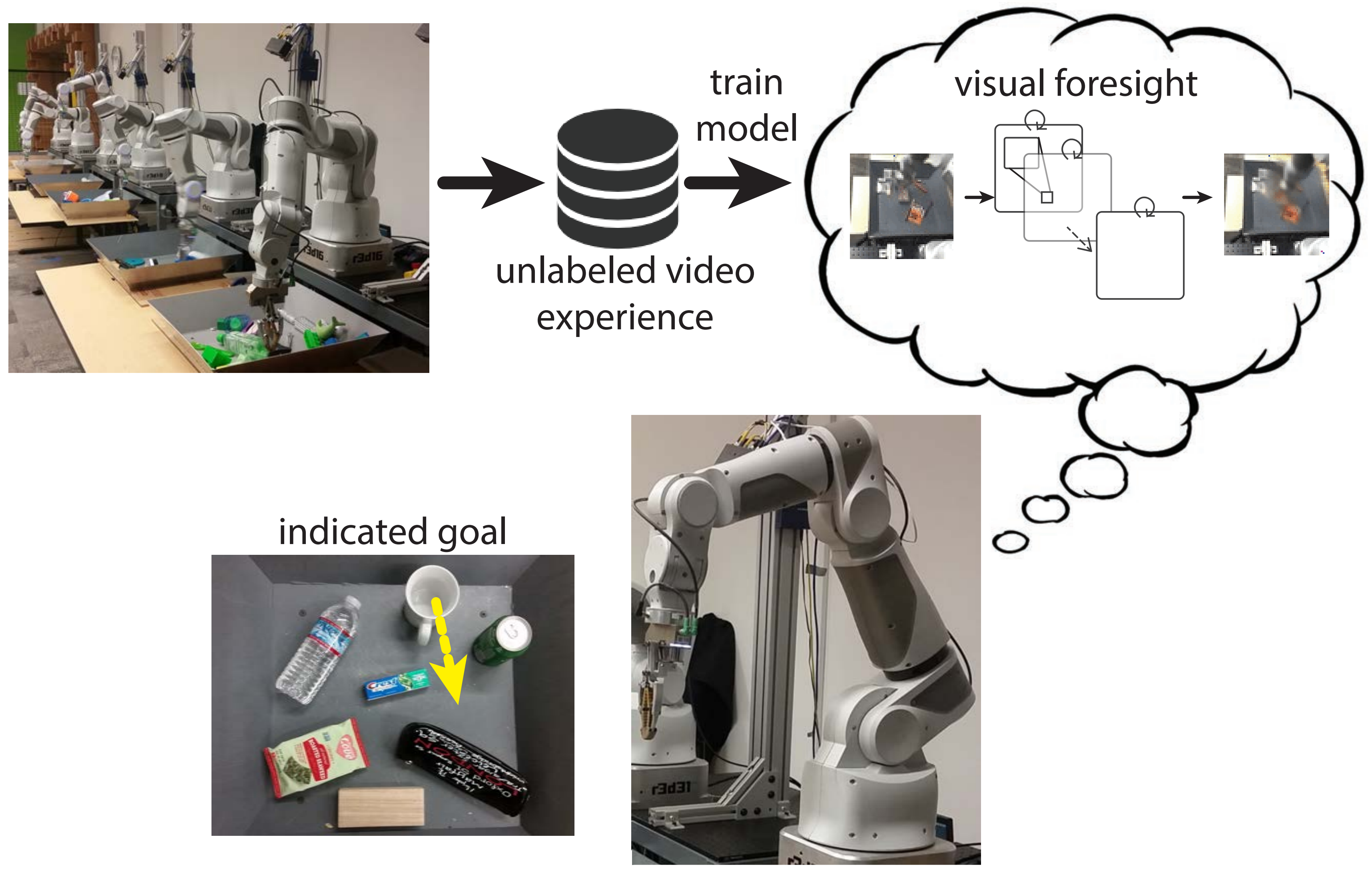
# Outperforms naive baselines

| method | mean pixel distance |
|---|---|
| initial pixel position | $5.10 \pm 2.25$ |
| 1) random actions | $4.05 \pm 1.75$ |
| 2) move end-effector to goal | $3.79 \pm 2.66$ |
| 3) move end-effector along vector (with replanning) | $3.19 \pm 1.68$ |
| visual MPC (ours) | $\mathbf{2.52 \pm 1.06}$ |

# Takeaways

**Benefits of this approach**

- learn for a **wide variety of tasks**
- **scalable** - requires minimal human involvement
- a good way to evaluate **video prediction models**

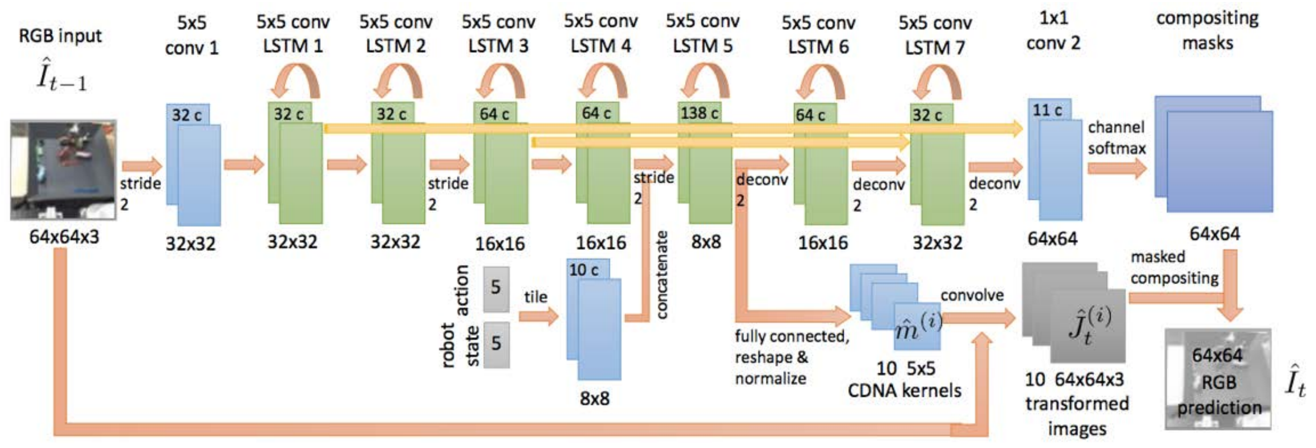



**Limitations**

- can’t [yet] learn complex skills
- **compute-intensive** at test time
- some planning methods susceptible to **adversarial examples**

# Future challenges in large-scale self-supervised learning

better **predictive models**




**long-term** planning
- hierarchy
- stochasticity

task-driven exploration, attention

learn **visual reward** functions

# Collaborators

Sergey Levine

Ian Goodfellow

# Thanks to…

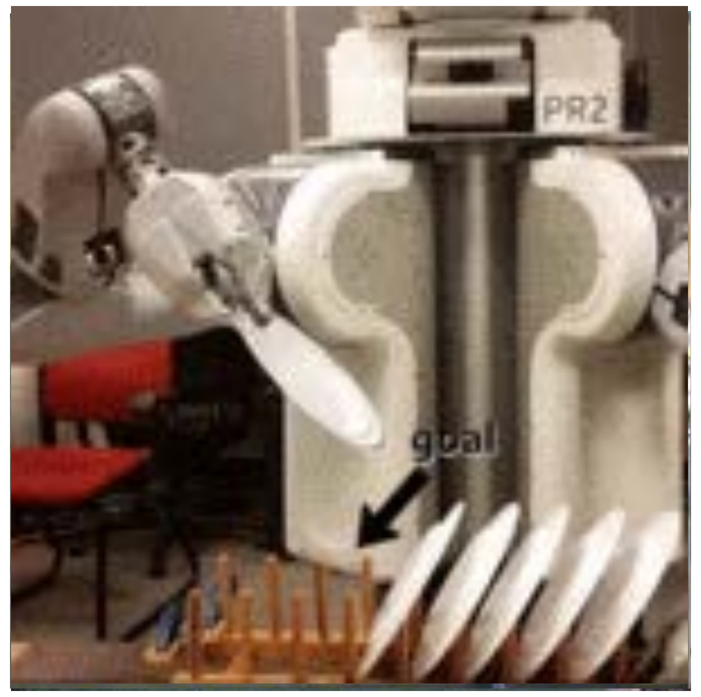
Vincent Vanhoucke
Peter Pastor
Ethan Holly
Jon Barron

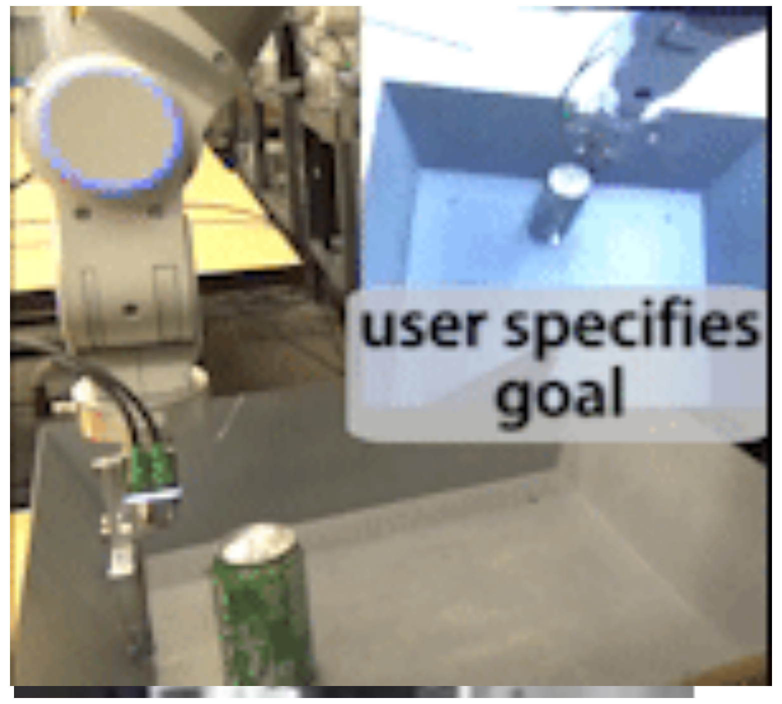
# Bibliography

Finn, C., Goodfellow, I., & Levine, S. *Unsupervised Learning for Physical Interaction through Video Prediction*. NIPS 2016

Finn, C. & Levine, S. *Deep Visual Foresight for Planning Robot Motion*. Under Review, arXiv 2016.

# Questions?




cbfinn@eecs.berkeley.edu

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

## Collaborators

Sergey Levine

Ian Goodfellow

## Thanks to…

Vincent Vanhoucke

Peter Pastor

Ethan Holly

Jon Barron

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

# Questions?

cbfinn@eecs.berkeley.edu

# Thanks!

**Takeaway**:

Acquiring a cost function is important! (and challenging)


**es of failure:**

mispredictions

more?”

ompute needed

cclusions

el tracking

seed

# This is just the beginning…

Can we design the **right** model?
- stochastic?
- longer sequences?
- hierarchical?
- deeper?

Can we handle **long-term** planning?

Collecting data with a purpose.