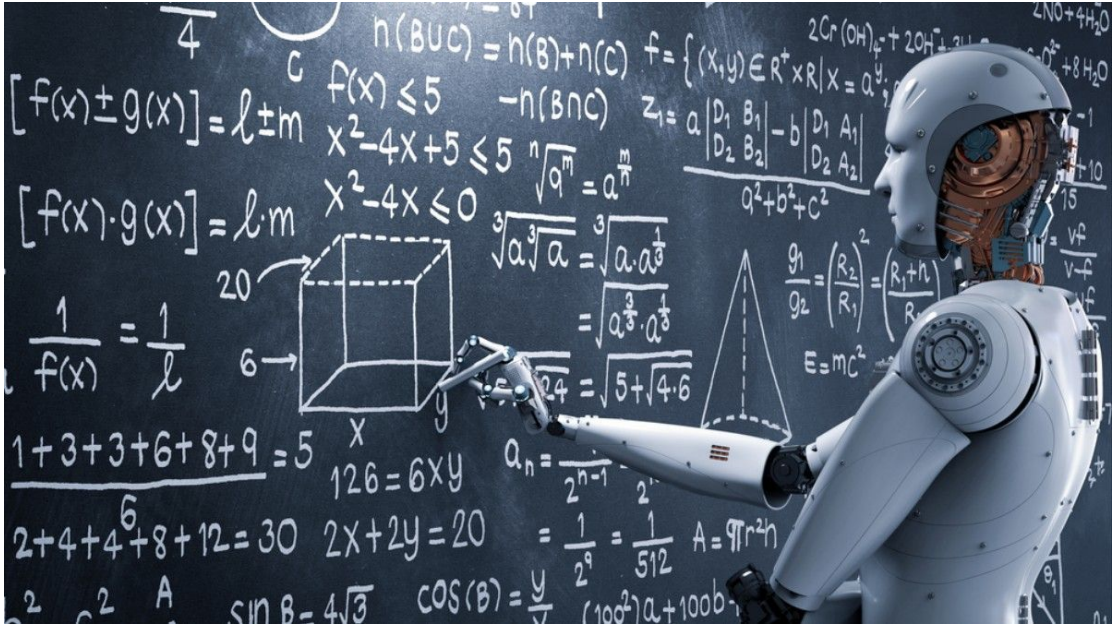
Learning to Poke by Poking: Experiential Learning of Intuitive Physics

NIPS 2016
Pulkit Agrawal, Ashvin Nair, Pieter Abbeel, Jitendra Malik, Sergey Levine
Presented by David Niu

Problem Overview

Can we learn an “intuitive" model of physics using interaction data collected directly from images of a robot's interactions?



Motivation

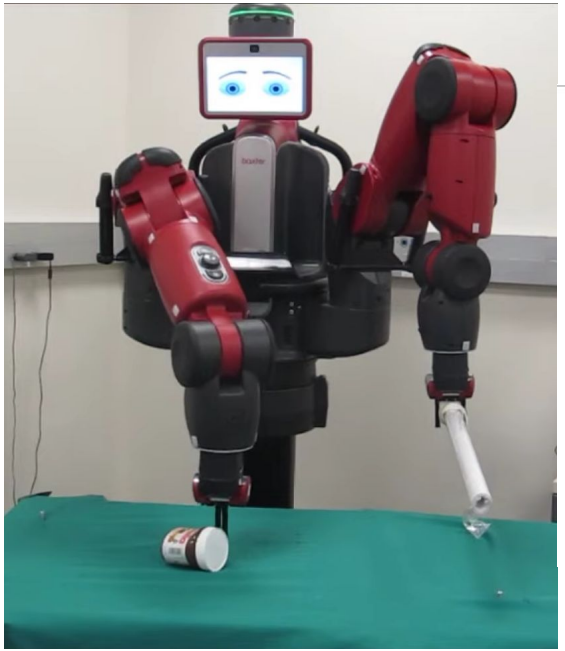
Humans are able to manipulate novel objects in coherent ways, possibly due to an internal model of physics

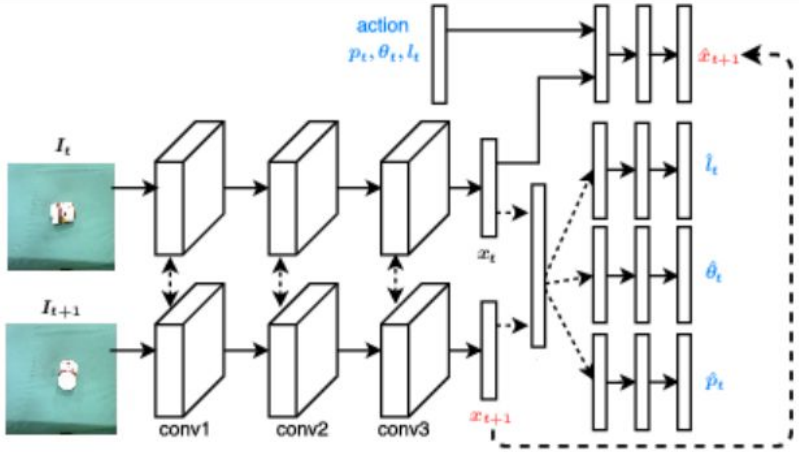
Motivation

Infants spend years "randomly" playing with toys, allowing them to develop an intrinsic sense of physics

Proposed Approach

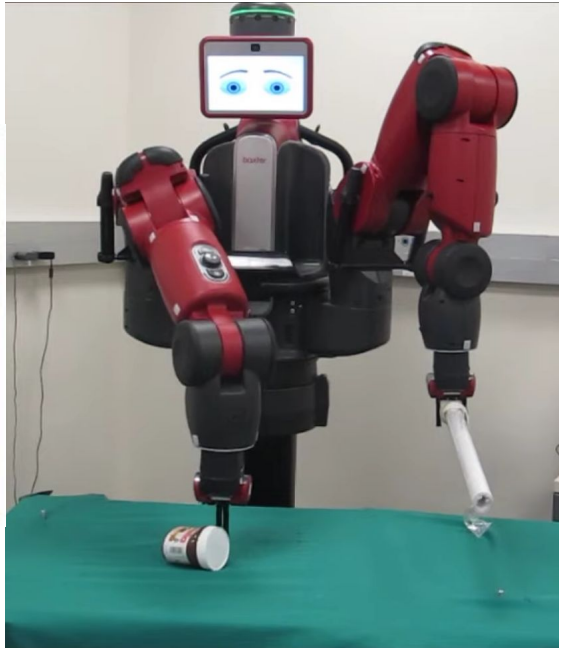
Approach has two components: physical robot and learned model





Proposed Approach

Setup consists of a robot with a Kinect camera and an arm for poking objects on a table in front of it



Related Work

Previous models for visual control have learned goal-specific policies

In addition, many prior methods use hand-designed visual features, or use Newtonian physics in combination with neural nets

Combination of using poking to learn object displacement had not been widely explored before

Method

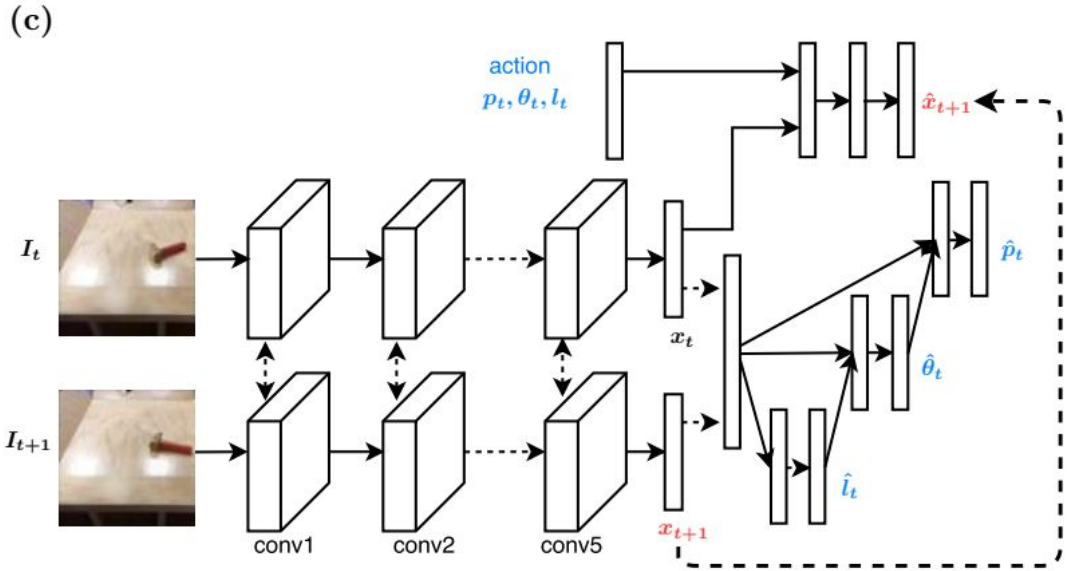
Joint model consists of forward and inverse models

$$\hat{x}_{t+1} = F(x_t, u_t; W_{fwd})$$

$$\hat{u}_t = G(x_t, x_{t+1}; W_{inv})$$

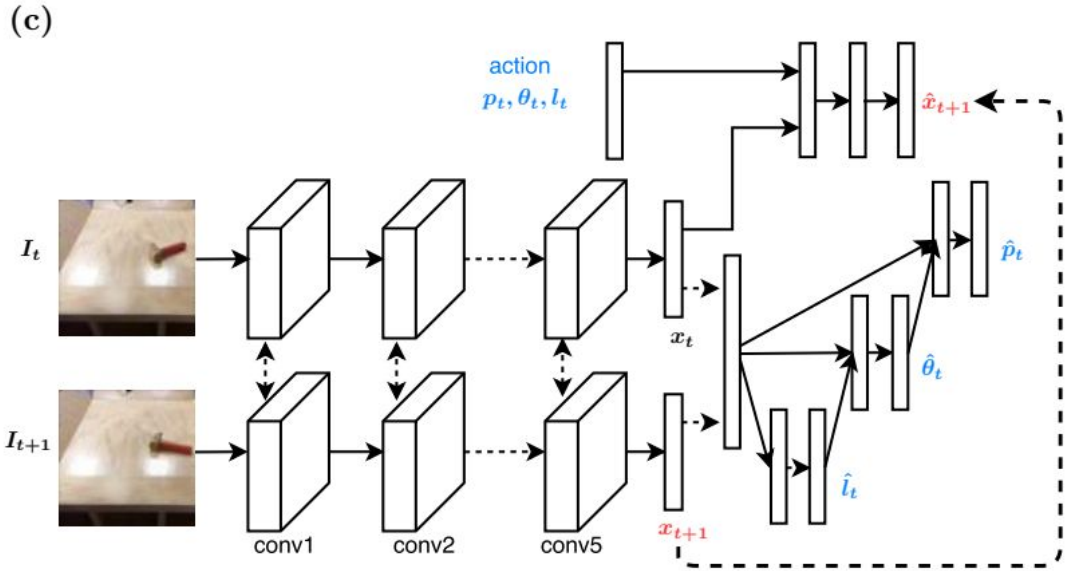
Method

Pairs of images are passed into 5 convolutional layers with the same architecture as the first 5 layers of AlexNet to create feature representations



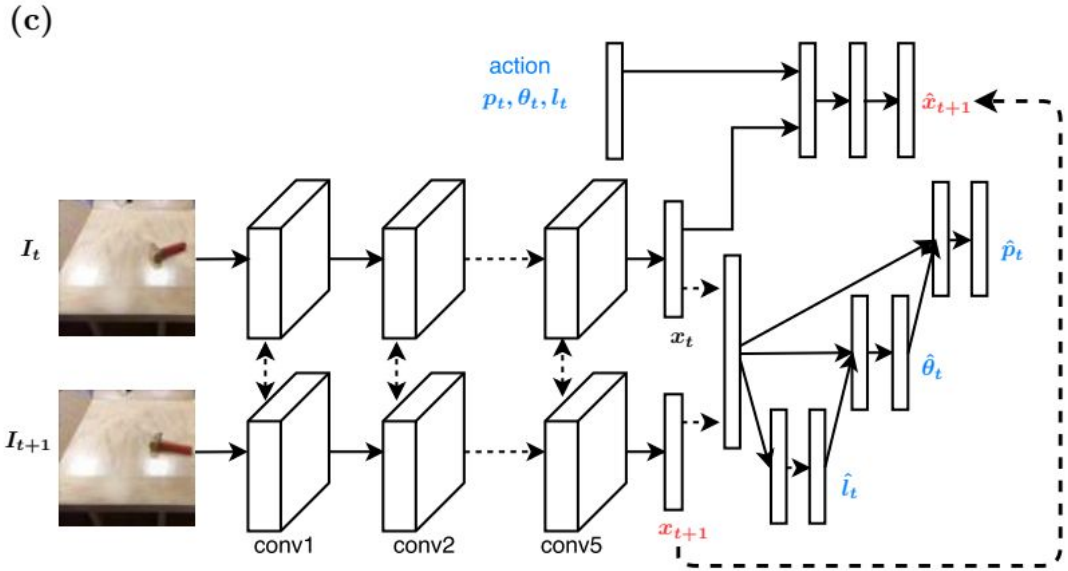
Method

Forward model takes the action and the feature representation of the current state to predict the future state



Method

Inverse model feeds the feature representations of input pair into fully connected layers to predict action



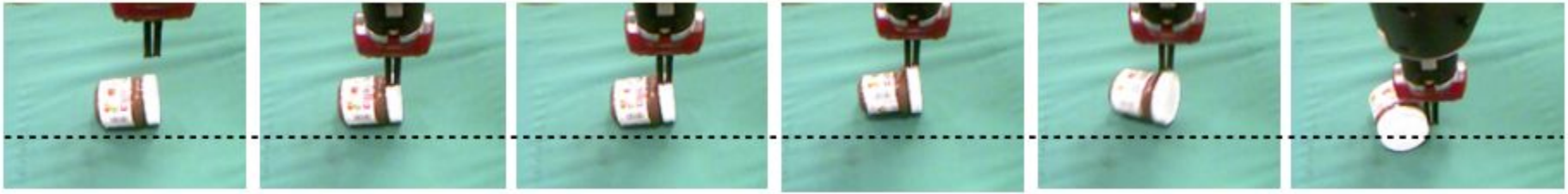
Method

The authors compare the joint model with the inverse model, with the joint model optimizing according to the below loss function

$$L_{joint} = L_{inv}(u_t, \hat{u}_t, W) + \lambda L_{fwd}(x_{t+1}, \hat{x}_{t+1}, W)$$

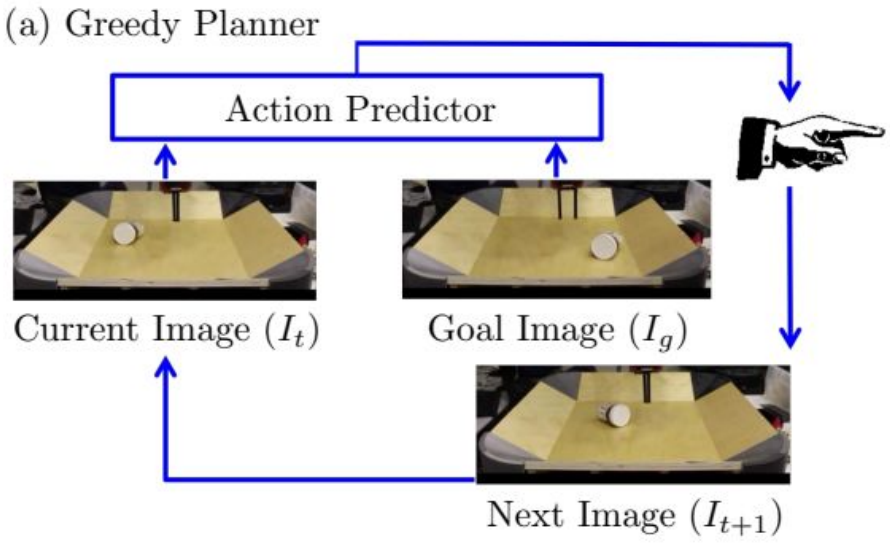
Evaluation Procedure

Task is to transform initial state to goal state by poking, where each state is fed into the model as an image



Evaluation Procedure

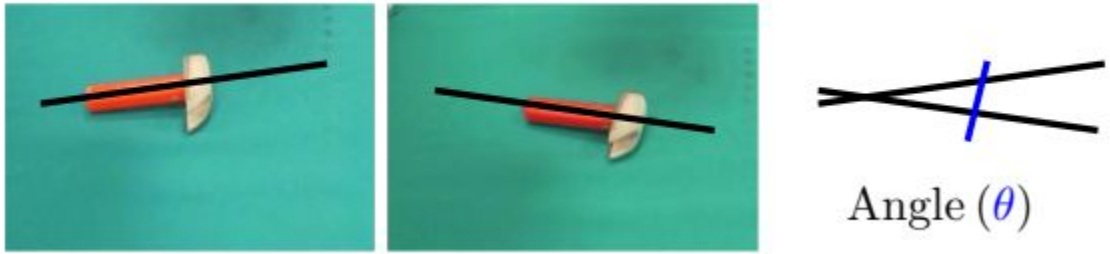
Sequence of pokes is generated using a greedy planner



Evaluation Procedure

Initial and goal images differ only by the position of a single object. Error is calculated using distances between object locations, relative to initial distance

(c) Pose Error Evaluation



Evaluation Procedure

Learned model is compared to a baseline 'blob' model that computes pokes naively using vector differences

(b) Blob Model

Results

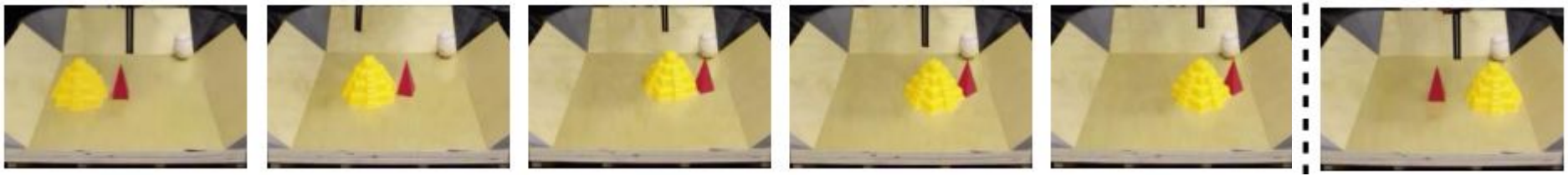
As a baseline, the robot is able to successfully displace objects in the training set

Results

In addition, the robot is able to manipulate objects with previously unseen geometries and is unaffected by the presence of distractor objects

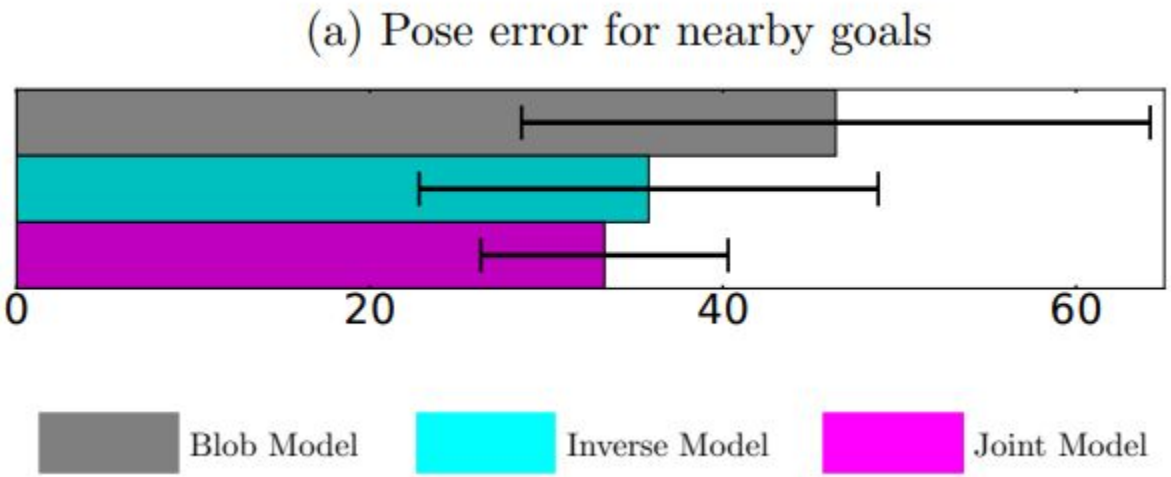
Results

However, the robot is unable to push objects around obstacles, due to the limitation of greedy planning



Results

Both joint model and inverse model outperform blob model on pose error

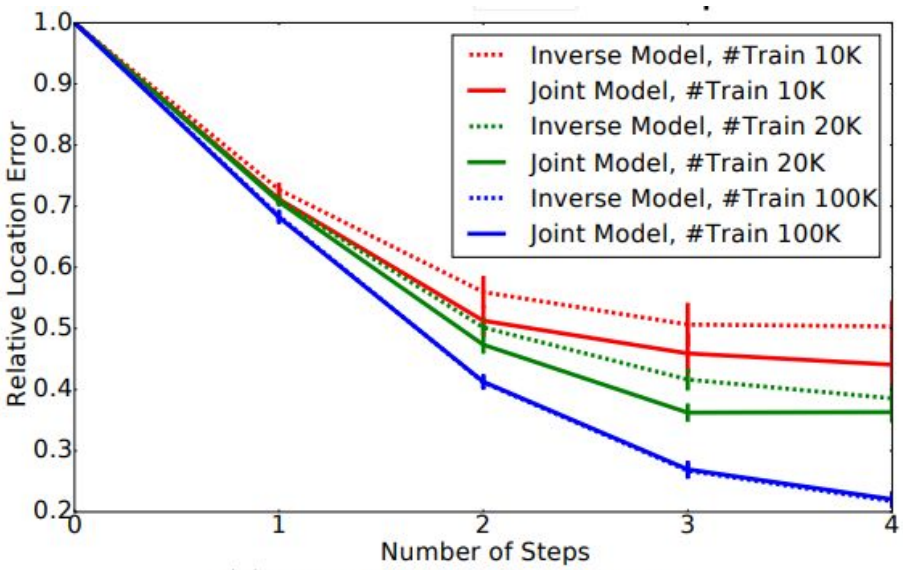


Regularization

Authors hypothesize that forward model regularizes the features space learned by the inverse model, test hypothesis on 2D simulator environment

Regularization

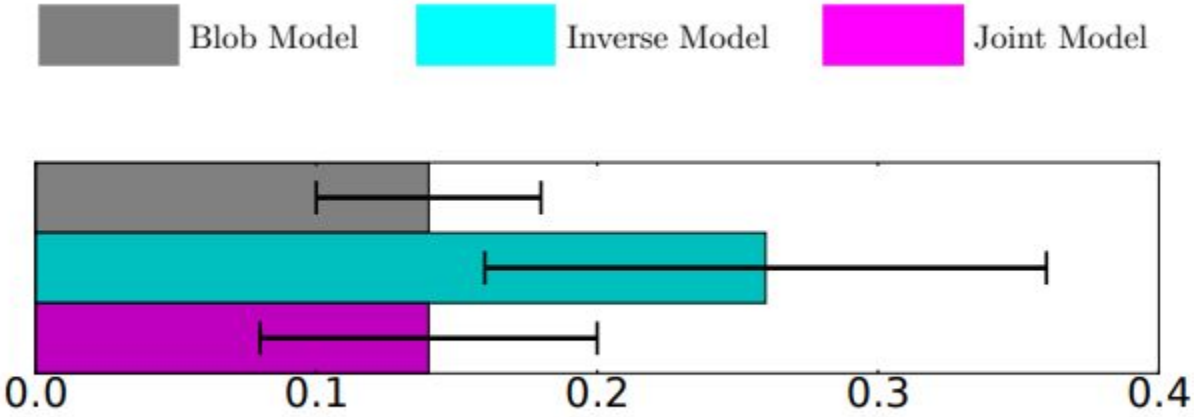
When training on less data, we see that the forward model regularizes the inverse model



(c) Simulation experiments

Regularization

These findings are replicated in real-world experiments with the robot



(b) Relative location error for far away goals

Summary

It is possible for a neural network to learn an "intuitive model" of physics without using predefined parameters for predicting object dynamics

The intuitive physics approach could be more robust than parameterized models in that it doesn't make object specific assumptions, but more research is needed

The proposed approach only operates in discrete pokes, and future research could investigate continuous time control

Questions?