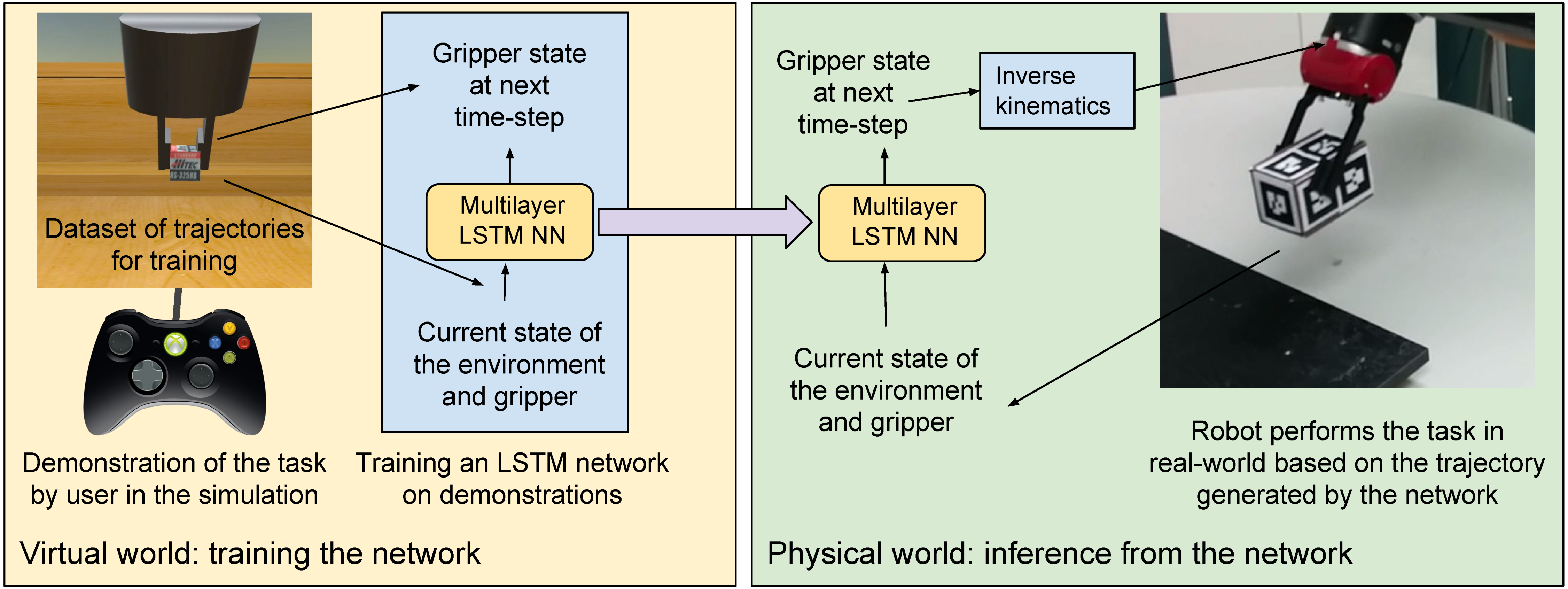
Learning from demonstration in robotics

Demonstration in simulation, deployment in real world

- Collect demonstrations in simulation of pick-and-place and push-to-position tasks
- Record object position + gripper position
- Train a controller on this data using behavioral control loss (3 layers of LSTM + MDN) - transforms object position to shelf position
- In a real world environment: collect object position using an off-the-shelf vision system (ArUco markets), collect gripper position from robot proprioception
- Use the same controller as in simulation, use an off-the-shelf inverse kinematics module to move the gripper to the new position.

Demonstration in simulation, deployment in real world



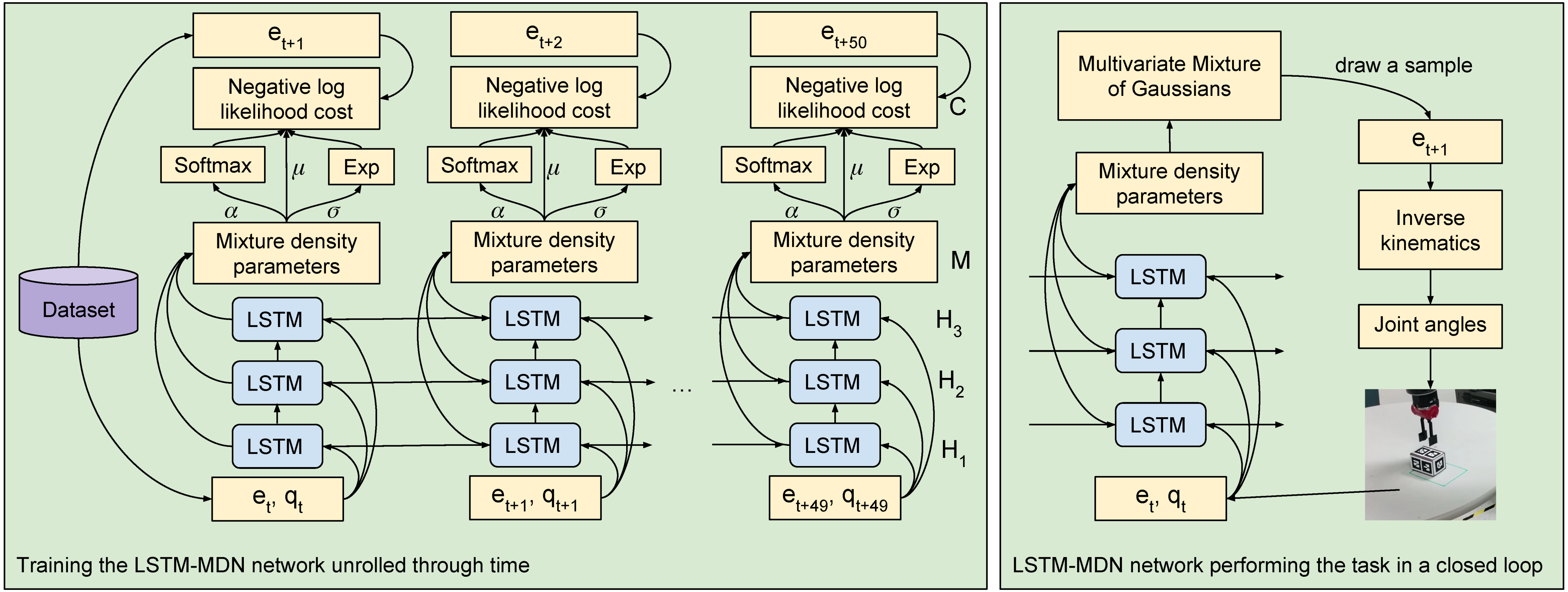
Let us see how it works

Click for video

or

Youtube

Network training and deployment



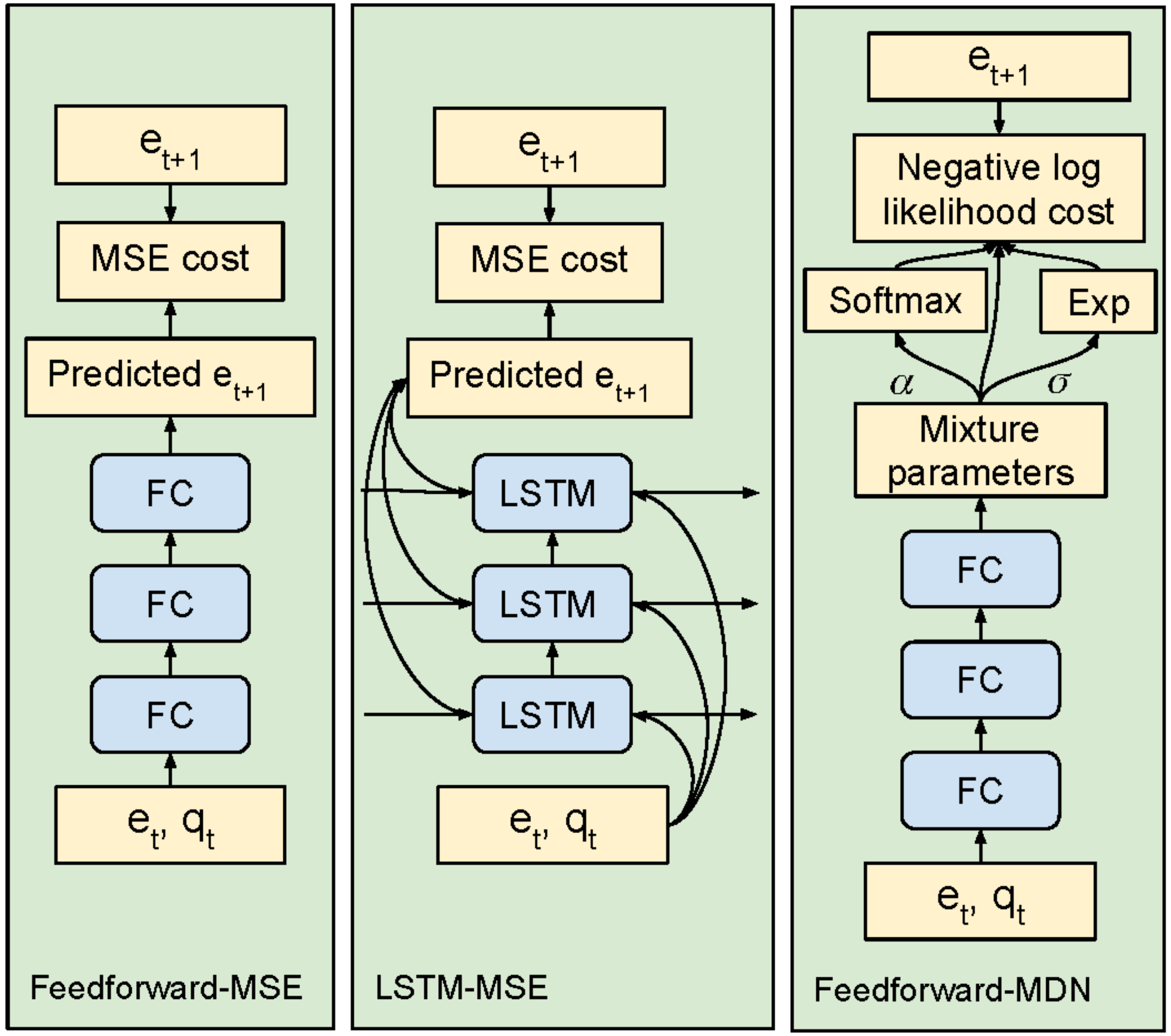
Training the LSTM-MDN network unrolled through time

LSTM-MDN network performing the task in a closed loop

Why did it work for us, but was not working previously

- This paper ended up that year in the Berkeley Deep RL class, with a commentary something like: here is an example where imitation learning is working, although they are not doing anything fancy.
- Why was it working for us, while not working for other people?
  - Mixture density network
  - LSTM
  - Errors during demonstrations

What is new in the network architecture: three networks that don't work



Ablation study - success rate

Controller	Pick and place	Push to pose
Feedfoward-MSE	0%	0%
LSTM-MSE	85%	0%
Feedforward-MDN	95%	15%
LSTM-MDN	**100%**	**95%**

The importance of making errors during demonstrations

- Our demonstrations were far from perfect, in fact there were many errors, dropped objects, overpushed objects etc.
- Turns out that this is critical to avoid the problems pointed out by Drew Bagnell
- The robot needs to learn how to correct an error (eg. go around and push from the other side)

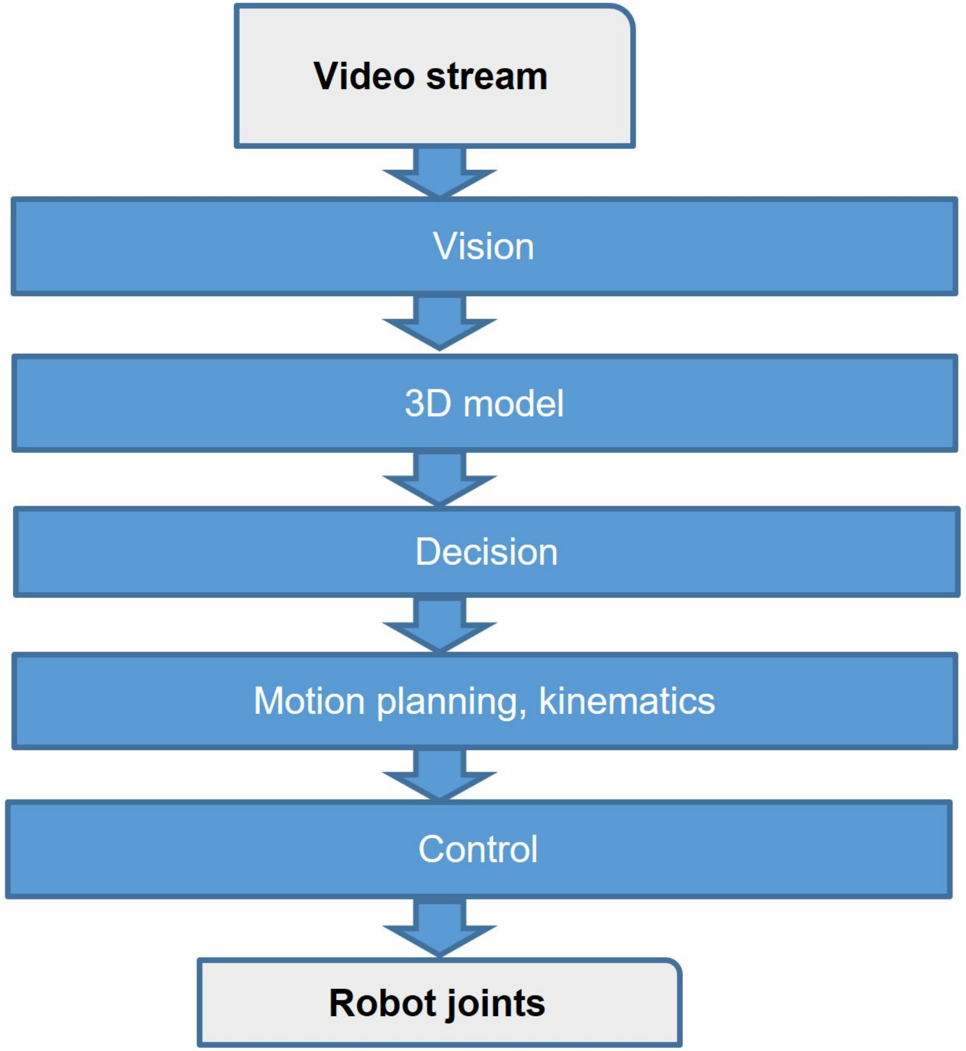
Going end-to-end all the way

- This was not quite end-to-end learning, it had off the shelf vision and inverse kinematics components
- Experiment: can we learn **everything**?

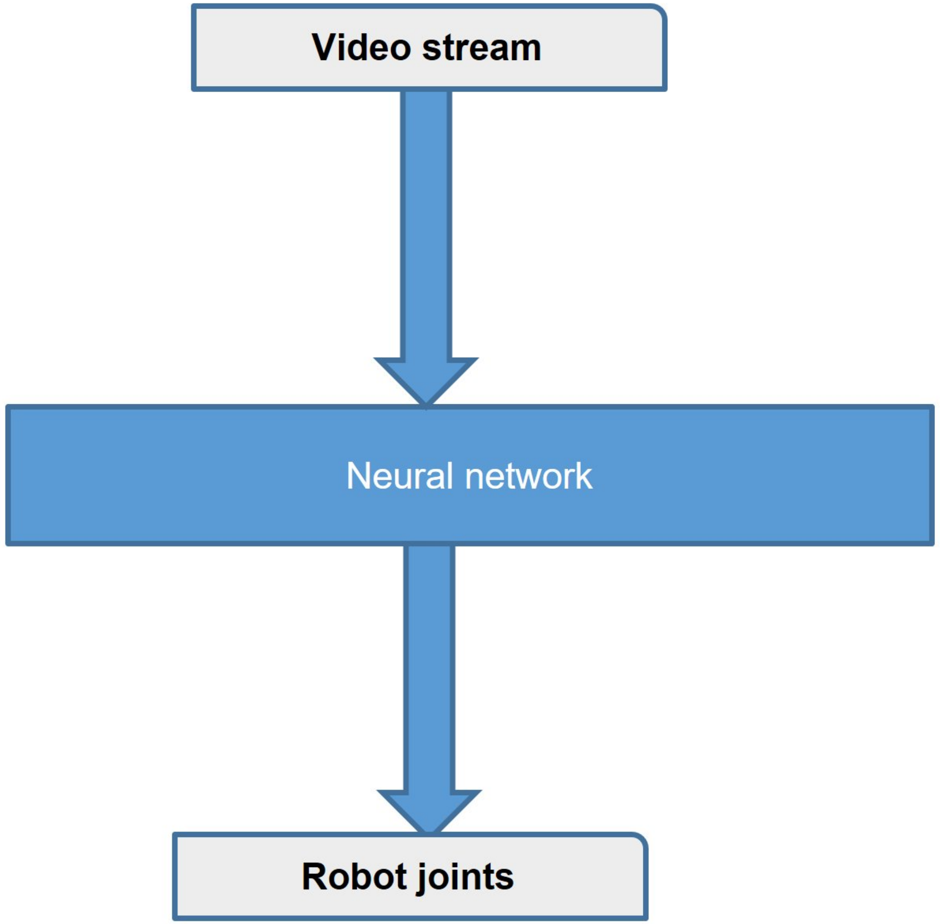
We traditionally create systems by engineering

- Decompose a problem into subproblems
- Then decompose those even further
- Until you reach either a known model, or something that you can solve

Robot control, engineered way



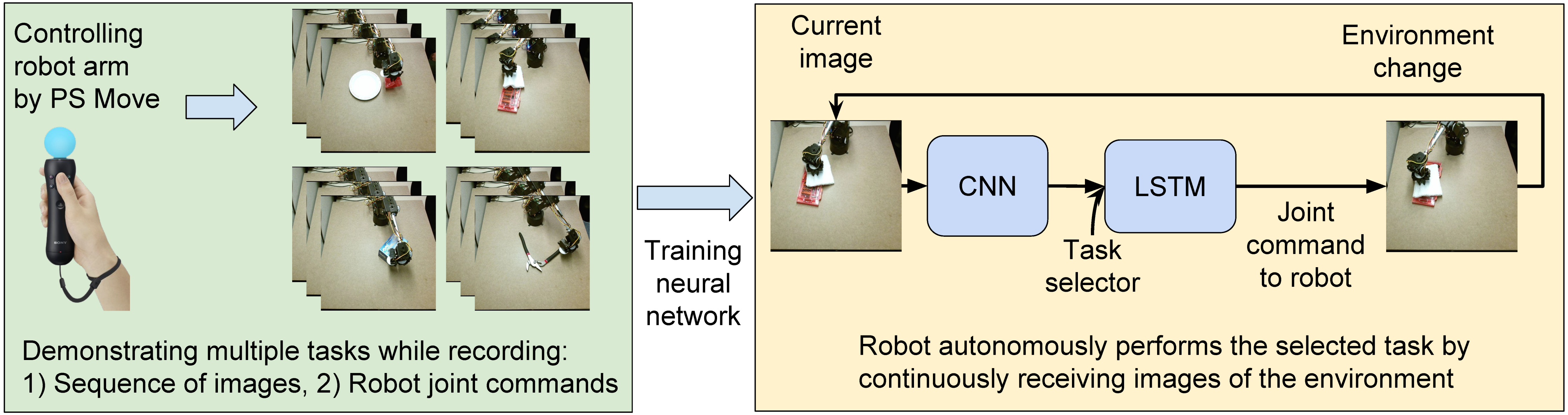
Robot control, end-to-end learned



Demonstration and deployment in real world

- Inexpensive physical robot controlled by remote control. Fixed camera captures the environment.
- Capture demonstrations, five different tasks, five different objects, but prehensile manipulation as well as pushing
- Training data: video stream from camera + controls captured before sent to robot.
- Train a network (VAE+GAN+LSTM+MDN)
- Test time: camera captures image, transfers to network, network output directly drives robot.
- No off-the-shelf components, **everything** is learned.

Demonstration and deployment in real world



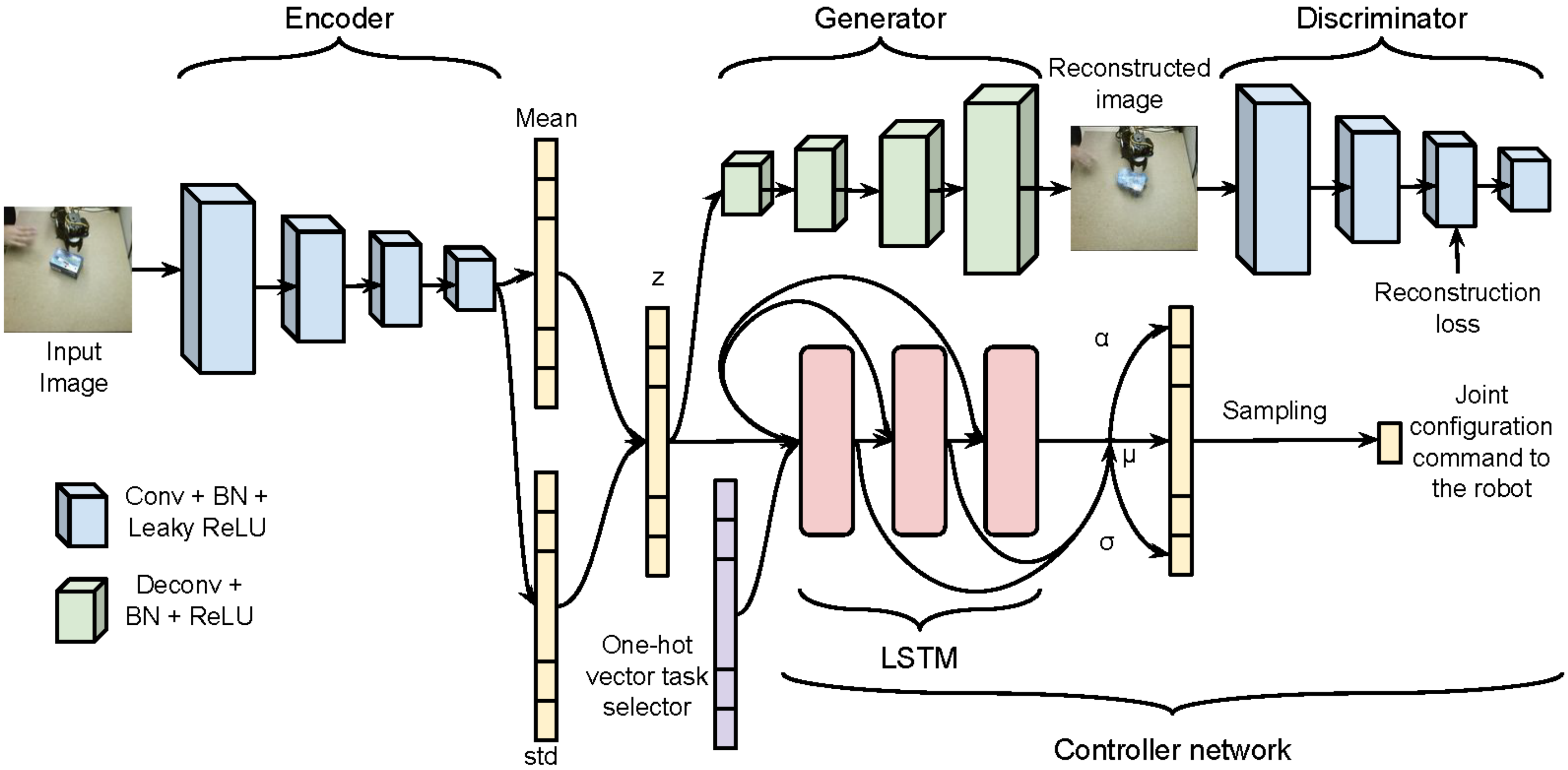
Let us see how it works

Click for video

or

Click for youtube link

Network architecture



Tasks and demonstration data

- T1: pick up a small bubble wrap and put it into a small plate
- T2: push a round plate to a specified area on the left side of the table
- T3: push a large box to a specific position and orientation close to the base of the robot arm
- T4: close a set of open pliers and orient them parallel to the borders of the table
- T5: pick up a towel and rub a small screwdriver box to clean it
- We collected 3 hours of demonstrations for each task, equivalent to 909, 495, 431, 428, 398 demonstrations for the tasks T1-5, respectively. We used 80% of this data for training and kept the remaining 20% for validation.

Ablation study

Method	T-1	T-2	T-3	T-4	T-5
Single-task (w/o autoregressive)	36%	16%	44%	16%	8%
Full network (w/o autoregressive)	16%	20%	52%	64%	20%
Full network (w/o VAE-GAN)	12%	72%	56%	48%	16%
Full network	**76%**	**80%**	**88%**	**76%**	**88%**

Performance comparison of different methods. The numbers are the percentile rate of successfully accomplishing the tasks.

Our progression on end-to-end learning for robot manipulator

- Learn manipulation only
  - Rahmatizadeh-2018-RealVirtual
  - Sim2Real2016.mp4
  - https://www.youtube.com/watch?v=9vYlIG2ozaM
- Learn manipulation + vision
  - Rahmatizadeh-2018-InexpensiveRobot
  - run:videos/MultiTask2017.mp4
  - https://www.youtube.com/watch?v=AqQFzoVsJfA
- Learn manipulation + vision + language + attention
  - Abolghasemi-2019-PayAttention
  - videos/PayAttention2018.mp4
  - https://www.youtube.com/watch?v=xdvNF_R_EkI