Vision Based Pendulum Control using Diffusion Policy

ES158 Final Project: Aneesh Muppidi

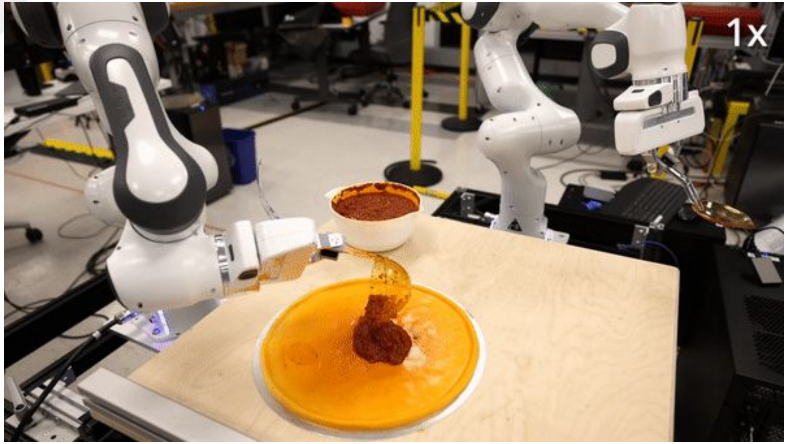
Harvard John A. Paulson School of Engineering and Applied Sciences

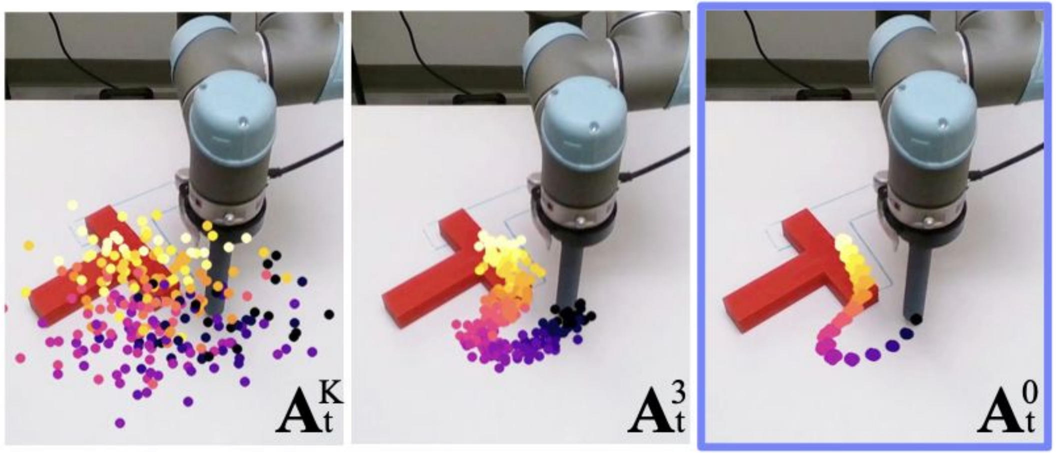
Image to control inverted pendulum swing up

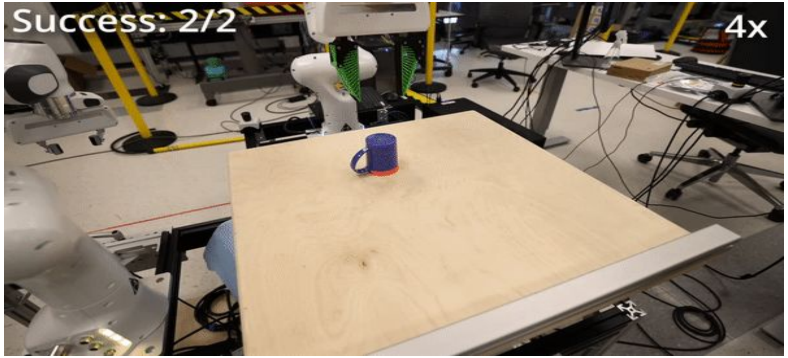
r = -(theta^2 + 0.1 * theta_dt^2 + 0.001 * torque^2)

Action Torque	1d, -2 to 2
Observation Shape	(3,) + rgb
Observation High	[1. 1. 8.]
Observation Low	[-1. -1. -8.]
Import	`gym.make("Pendulum-v1")`

Diffusion Policy in Robotics







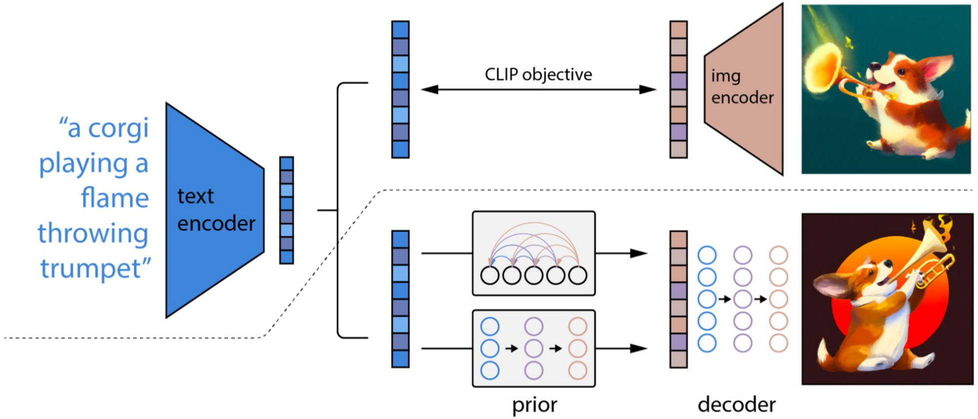
- High Dimensional Action Space
- Multiple Trajectories
- Closed Loop Action Sequences
- Very complicated tasks

Harvard John A. Paulson School of Engineering and Applied Sciences

Does a Diffusion Policy work for simple, one dimensional action space, unimodal, nonlinear control problems?

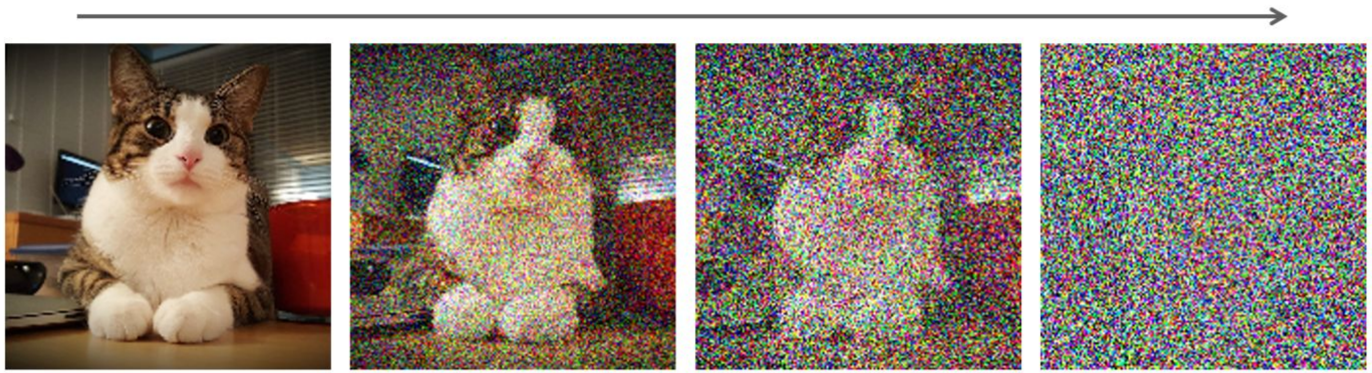
Intuition: **yes**!

What is Diffusion ??



It can be used for a lot more than just image generation

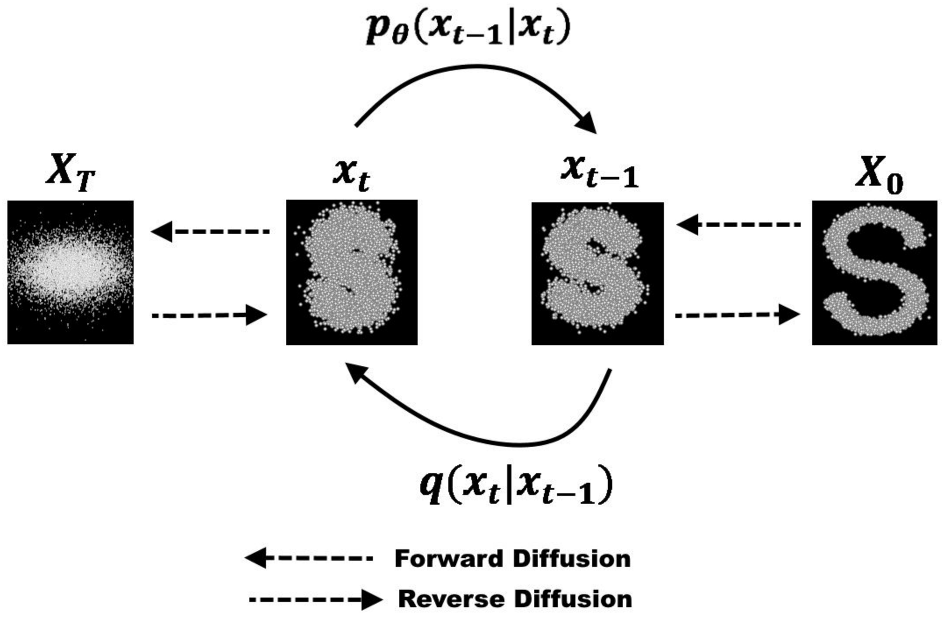
How does Diffusion Work?
Forward Pass



$$q(x_t|x_{t-1}) = \overbrace{\mathcal{N}}^{\text{normal distribution}}(\underbrace{x_t}_{\text{output}}, \overbrace{\sqrt{1-\beta_t}\,x_{t-1}}^{\text{mean}}, \underbrace{\beta_t I}_{\text{variance}})$$

Harvard John A. Paulson
School of Engineering
and Applied Sciences

Backward Pass

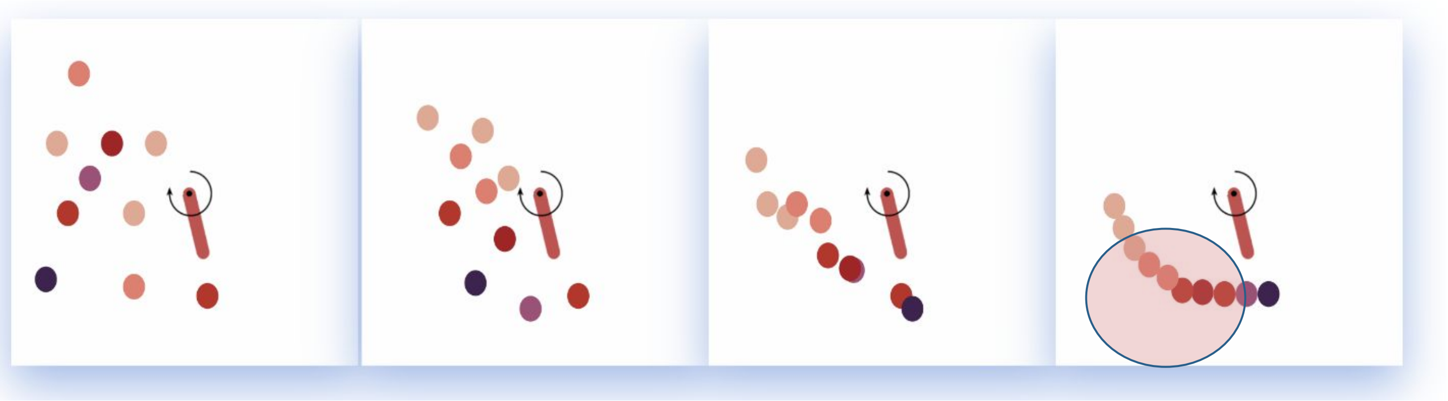


$$p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) = \mathcal{N}(\mathbf{x}_{t-1}; \mu_\theta(\mathbf{x}_t, t), \Sigma_\theta(\mathbf{x}_t, t))$$

Harvard John A. Paulson School of Engineering and Applied Sciences

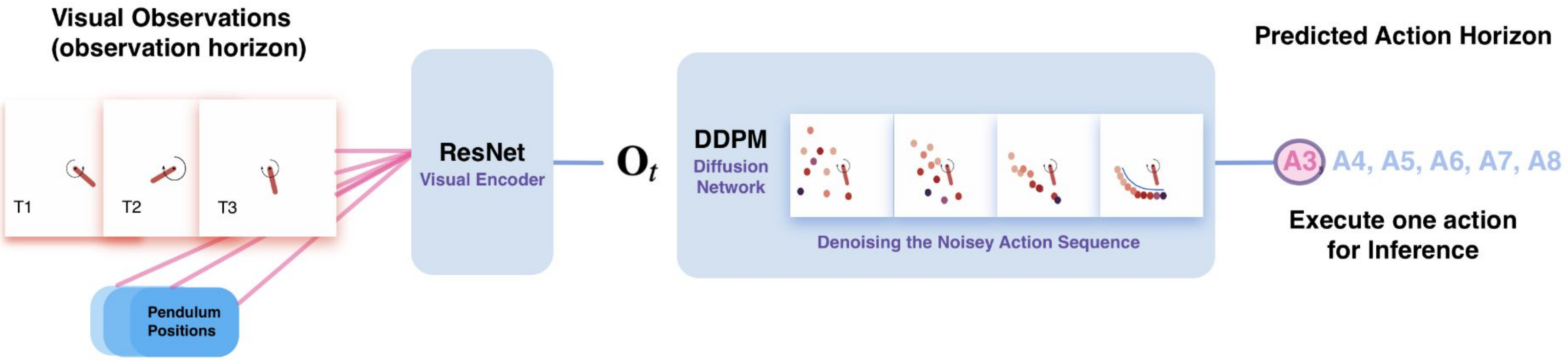
How do we adapt this for control?

We just denoise actions instead!



$$\mathbf{x}^{k-1} = \alpha(\mathbf{x}^k - \gamma\varepsilon_\theta(\mathbf{x}^k, k) + \mathcal{N}(0, \sigma^2 I))$$

The Architecture



Training - Dataset

Solved DQN

Model Training

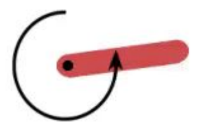
500 Episodes, 100 steps each
Observation Horizon: 8
Prediction Horizon: 16
Action Horizon: 2

FAS RC

Results (best model from 450 runs)

Model 1 (θ - not given, θ dot not given)

Model 2 (θ - given, θ dot not given)



Harvard John A. Paulson School of Engineering and Applied Sciences

Is diffusion a good policy for simple, nonlinear control problems?

Computational Expense

Fine Tuning

Does not converge easily

If I had more time, I would test...

Transformer Model (more robust for velocity)

Other control problems

Different Diffusion Models

Other ways to concatenate observations

Why not? (open question)

Stability/Target