Diffusion Policy Evaluation

Franklin Wang and Ram Goel

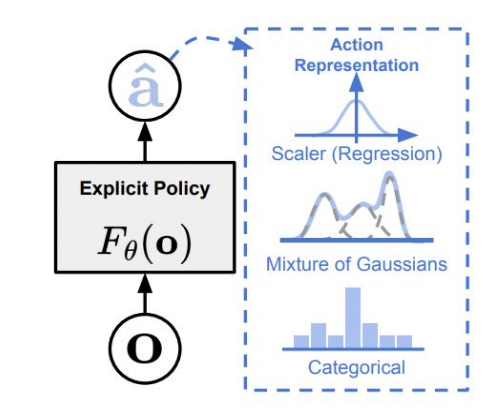
6.8200 Final Presentation

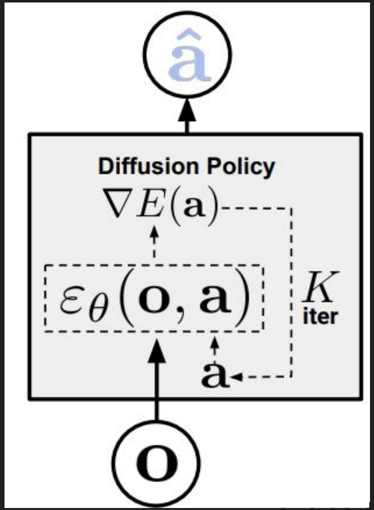
Multimodality in Policy Learning for Offline RL Problems

- Offline RL: Clone expert's policy given expert data
- Main challenge: **multimodal modeling capabilities**

- Approaches:
  - Explicit (Mixture of Gaussian, Binning)
  - Implicit/Diffusion

(From Diffusion Policy paper)





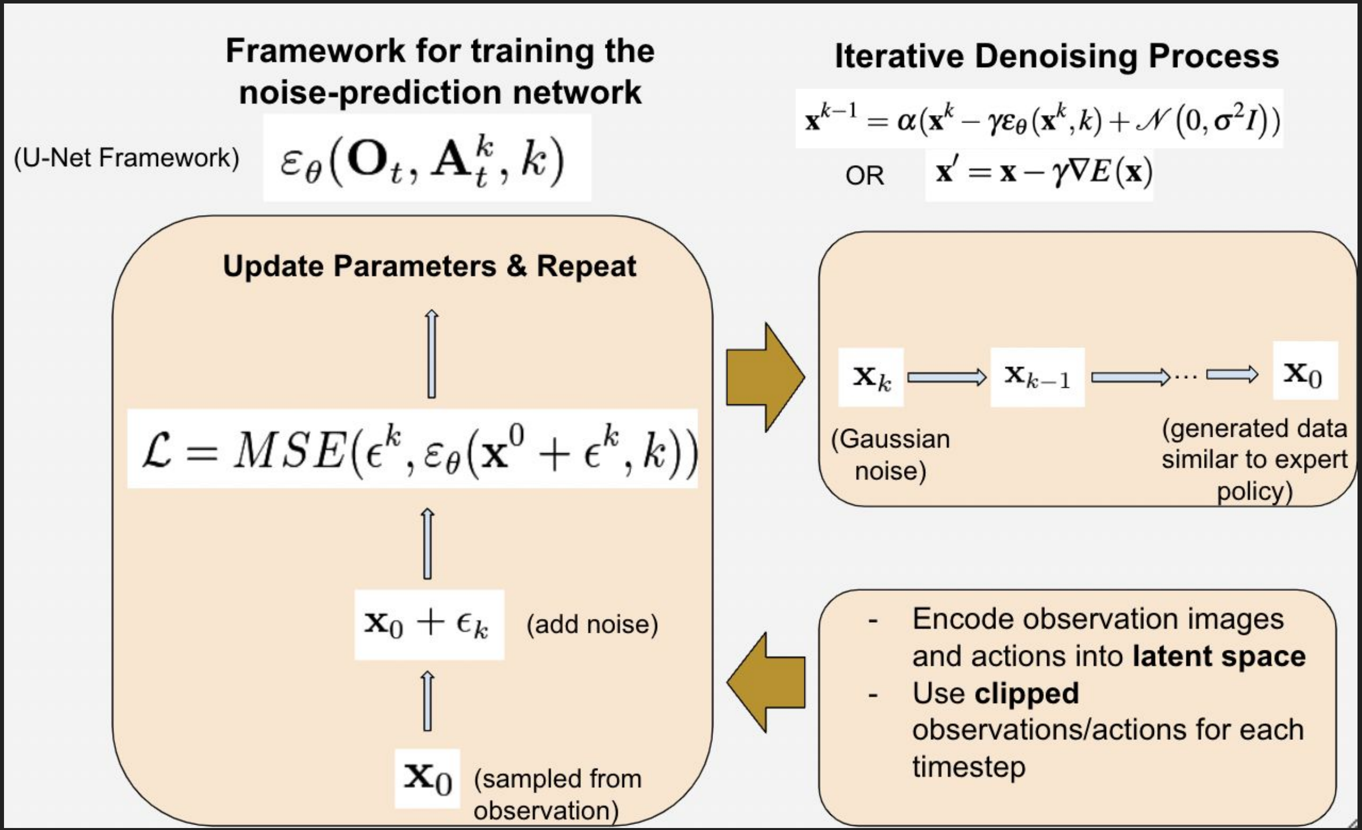
Problem Formulation: Push-T Task

- Observation space: 96 x 96 pixel image of top-down 2D view
- Action space: (x,y) coordinates for hand to move to

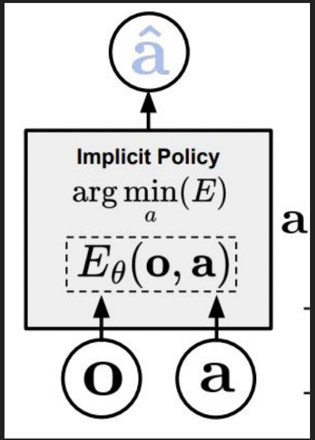
- Success Criteria: Intersection over Union (IoU) > 90%
- No inherent reward
  - Dataset of expert examples for behavior cloning

*Our trained **Diffusion Policy learning algorithm** performing Push-T*

Diffusion Policy Methodology Overview

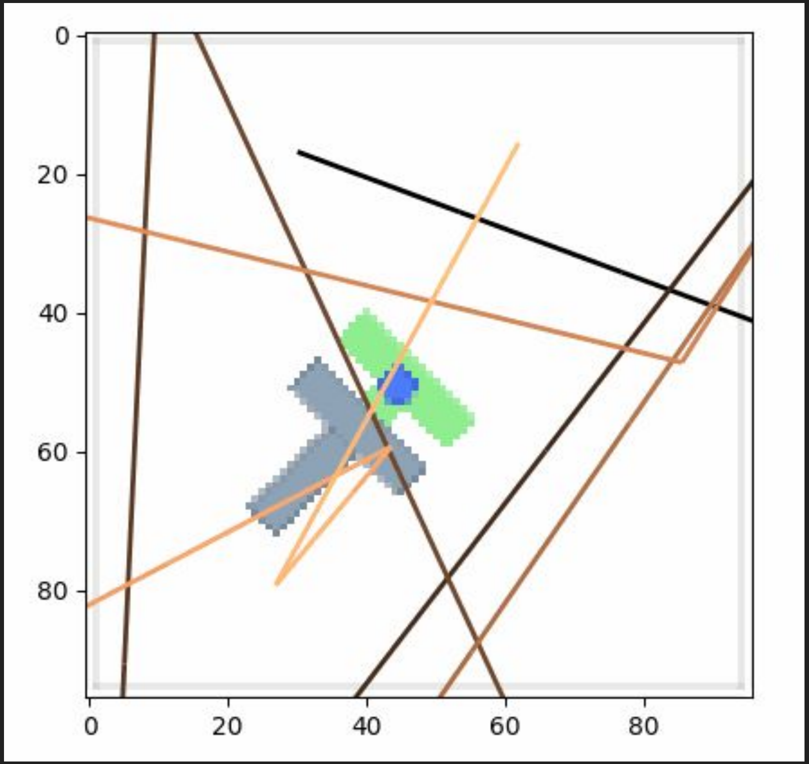


Baseline: Implicit Behavior Cloning (IBC)



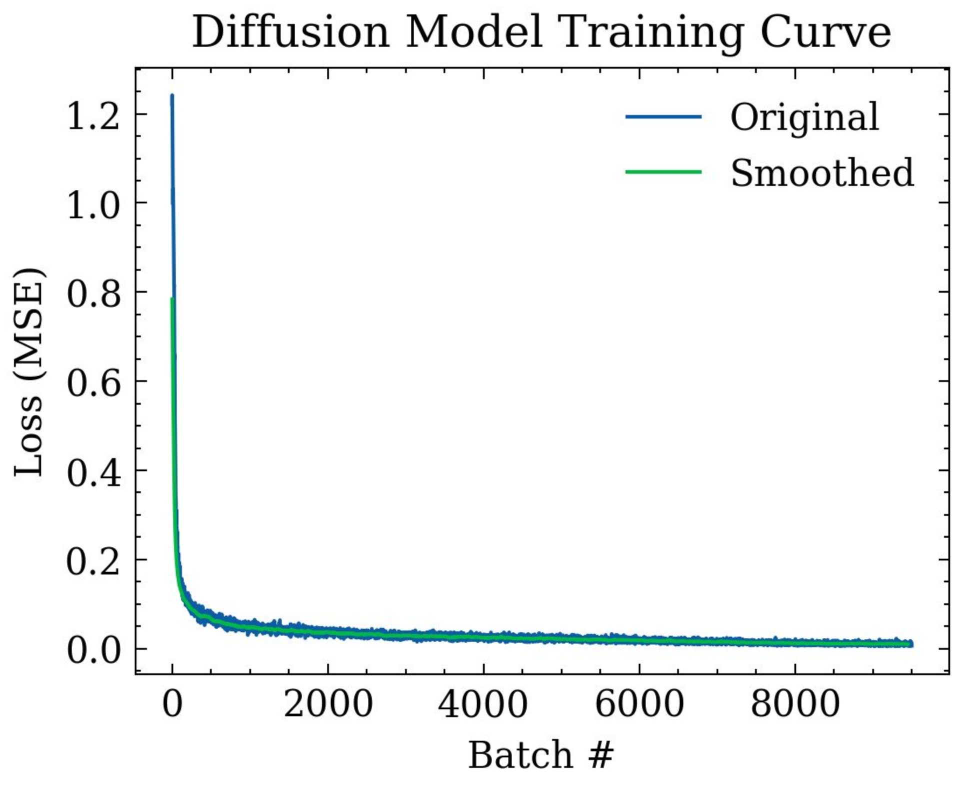
- Minimize an energy function for observation to select action

- Potential for **multimodal** modeling



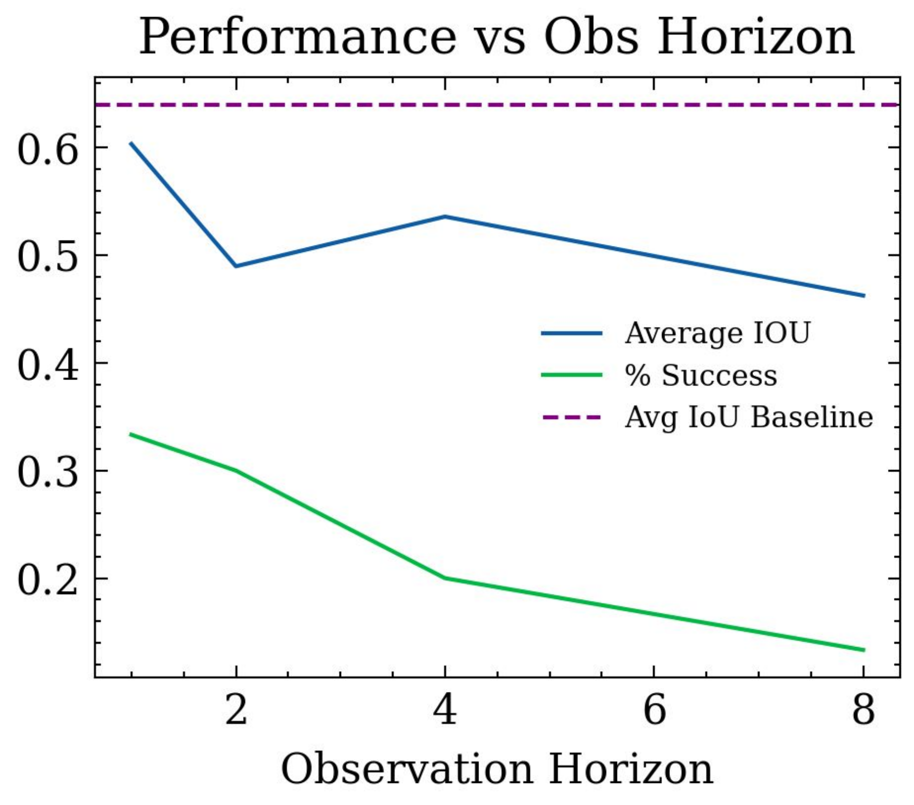
Diffusion Demonstration
of our Trained Model

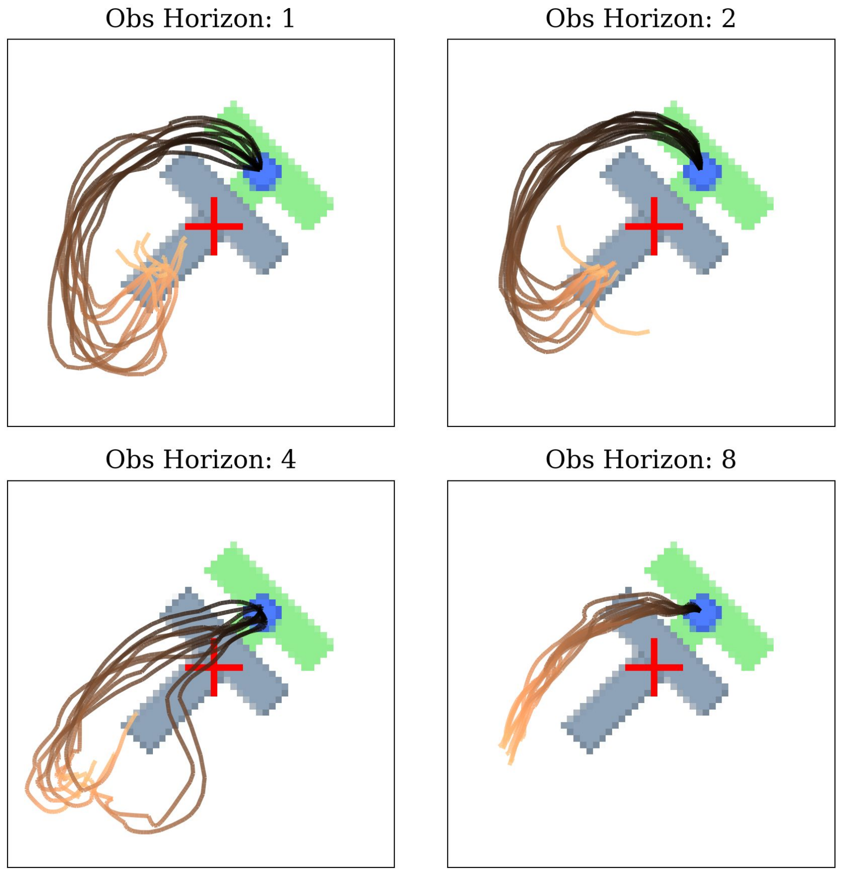
Diffusion Policy Results



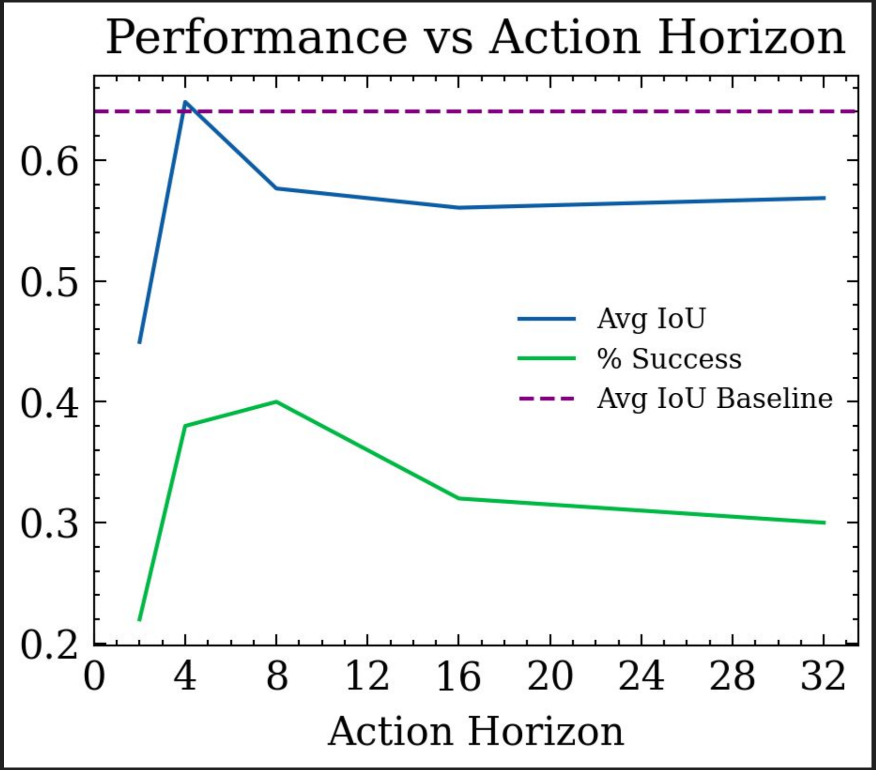
	Best Model	IBC
Average Best IoU	75%	64%
Success Rate (IoU > 0.9)	47%	42%

Observation Horizon: # of observations fed into the model per timestep

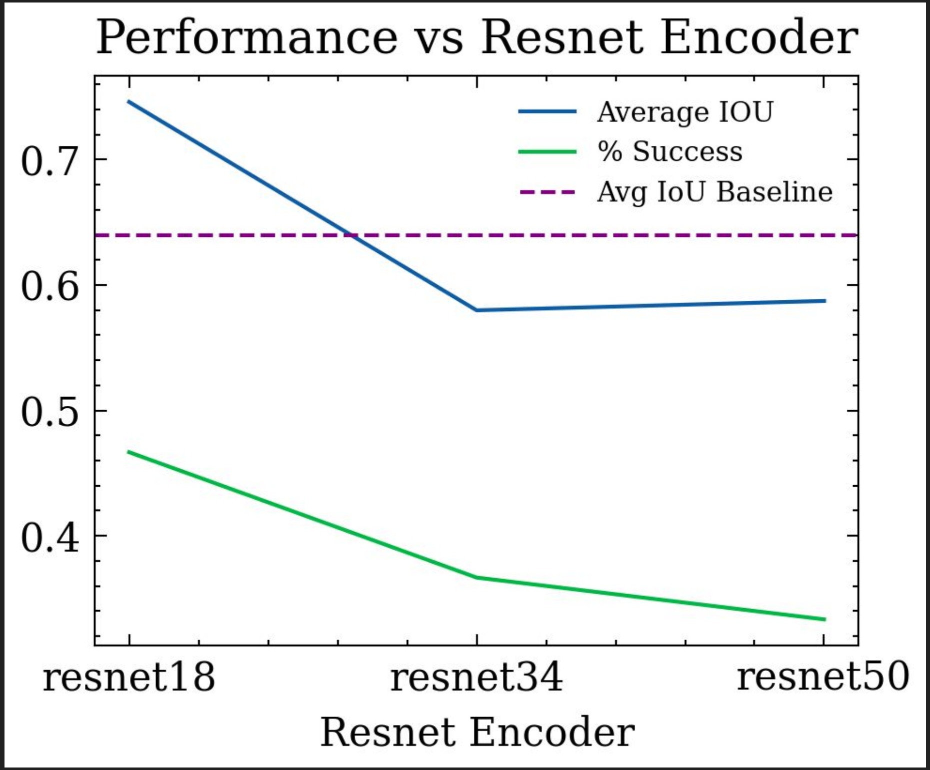




Action Horizon: # of actions we act upon at a time



Resnet Encoder is used to encode the image observations



Issues, Changes, and Learnings

Analysis:

- Main important hyperparameters
  - Action horizon
  - Observation Horizon
- Discrepancies between our findings vs original paper
  - Different “optimal” hyperparameters
  - Performance: ~90% of original
    - 75% vs 84% IoU Average

Challenges:

- Multimodality not as apparent as it should be
  - Possible overfitting to one mode due to longer training epochs
  - Size of validation/test set
- Difficulty training diffusion models