Batch-Constrained Q-Learning (BCQ)

Optimal Decision-Making: BCQ enhances decision-making in varying sensor modes, particularly for color and distance.

Targeted Learning: Ideal for datasets from specific policies, even when not all states and actions are covered.

Ease of Implementation: BCQ integrates easily into existing Q-learning frameworks with minimal modifications.

Effective Use: Best suited for discrete action spaces with limited data availability.

Implement:D4RL

Simple API Interface: D4RL offers an easy-to-use API, facilitating direct access to datasets for training agents.

Rich Baselines: D4RL provides a comprehensive set of baselines, including common offline algorithms like BCQ, BEAR, BRAC, and others.

Maybe Simulation ? mujoco210