

FCAI

fcai.fi

i. For i in number of episodes

i. Collect trajectory $\tau_i = \{\mathbf{s}_t, \mathbf{a}_t, \mathbf{s}_{t+1}, r_t\}_{t=0}^T$

iii. Add trajectory to replay buffer $\mathcal{D} \leftarrow \mathcal{D} \cup \tau_i$

iii. Perform T updates to world model

i. Sample batch from replay buffer 2

i. Update encoder, dynamics and reward

iii. Update actor and critic

DCWM



DCWM

Algorithm

- i. For i in number of episodes
 - i. Collect trajectory $\tau_i = \{\mathbf{s}_t, \mathbf{a}_t, \mathbf{s}_{t+1}, r_t\}_{t=0}^T$
 - ii. Add trajectory to replay buffer $\mathcal{D} \leftarrow \mathcal{D} \cup \tau_i$
 - iii. Perform T updates to world model
 - i. Sample batch from replay buffer \mathcal{D}
 - ii. Update encoder, dynamics and reward
 - iii. Update actor and critic

Decision-time Planning

Model Predictive Control (MPC)

