



**Previous RL Method**

- ✗ Only actor optimized, limited upper bound of capability
- ✗ Sparse 0/1 reward

**Ours (WA-RL)**

- ✓ Actor-WM co-evolve for greater capability
- ✓ Dense Recon Reward