

Learning Fused State Representations for Control from Multi-View Observations

Zeyu Wang*¹ Yao-Hui Li*¹, Xin Li†¹, Hongyu Zang², Romain Laroche³,
Riashat Islam⁴

¹Beijing Institute of Technology, Beijing, China

²Meituan, Beijing, China

³Wayve, London, UK

⁴Mila - Quebec AI Institute, Montreal, Canada

July 2025

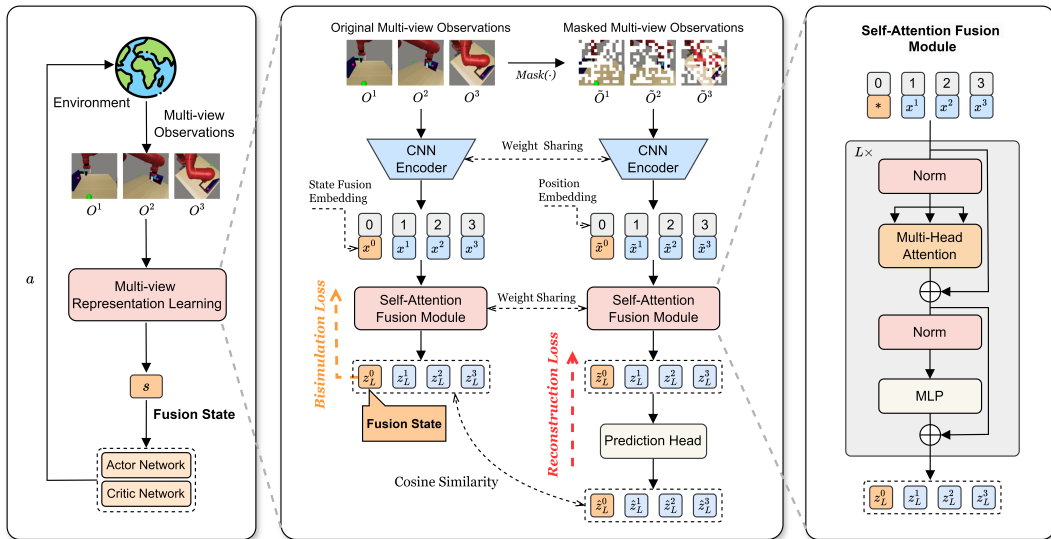
* Equal contribution

† Corresponding author

Key challenges of multi-view reinforcement learning: learning expressive and robust state representations

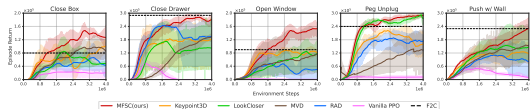
- the difficulty in learning and fusing compact representation from redundant and high-dimensional multi-view observations
- the problem of dealing with missing or interfering views in real-world scenarios

MFSC: Multi-view Fusion State for Control

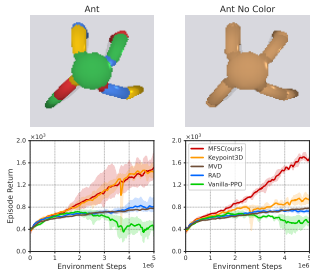


Experimental Results

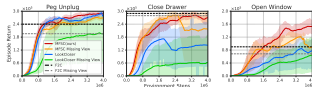
• Evaluation on Meta-World



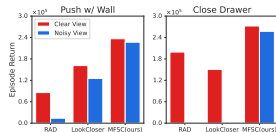
• Evaluation on PyBullet's Ant.



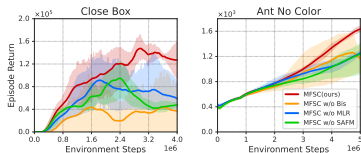
• Robustness Against Missing Views



• Robustness Against Noisy Views

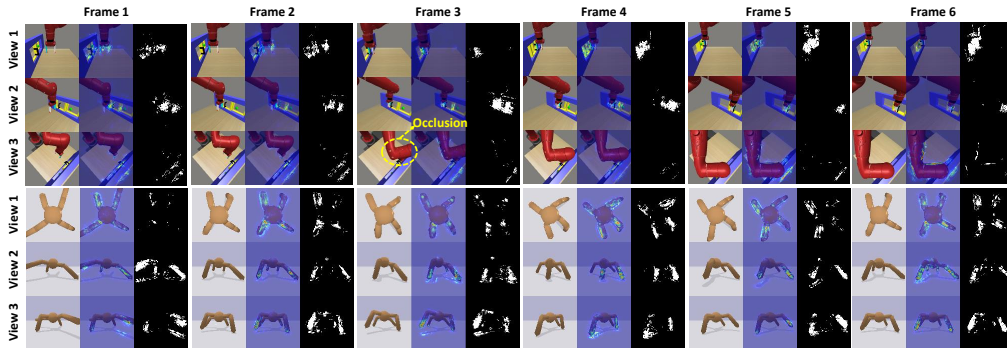


• Abalation Study



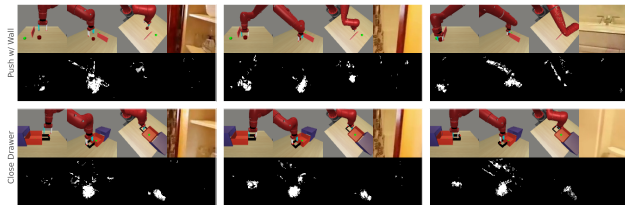
Interpretability Visualization

- Visualization of multi-view fusion for task-relevant representations.

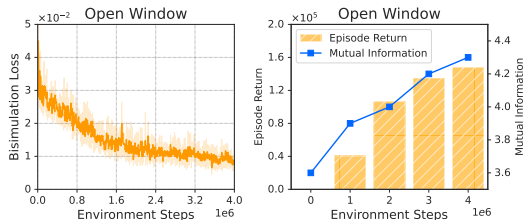


Interpretability Visualization

- Visualization of learned representation on noisy view



- Quantitative Analysis on Bisimulation Metric



Check our paper for ...

- Detailed description of our method
- Theoretical guarantees
- More empirical results



(a) paper



(b) project