Your Robots Need Memory!

MemoryVLA

One of the earliest works to explicitly model memory in VLA

See More

Homepage (Video) | Official Code (200+ Star) | Dexbotic Code (900+ Star)

Contact

shihao1895@gmail.com

u1s11024 (石昊)

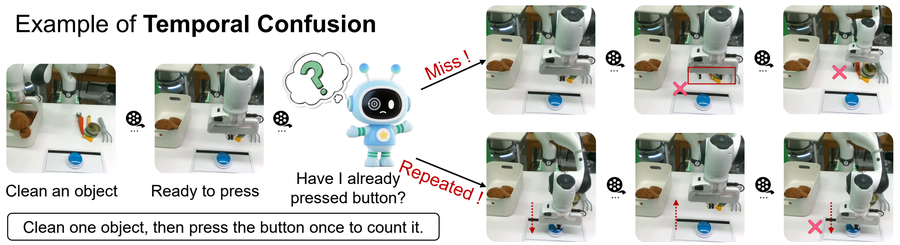
MemoryVLA: Perceptual-Cognitive Memory in Vision-Language-Action Models for Robotic Manipulation

Hao Shi, Bin Xie, Yingfei Liu, Lin Sun, Fengrong Liu, Tiancai Wang, Erjin Zhou, Haoqiang Fan, Xiangyu Zhang, Gao Huang

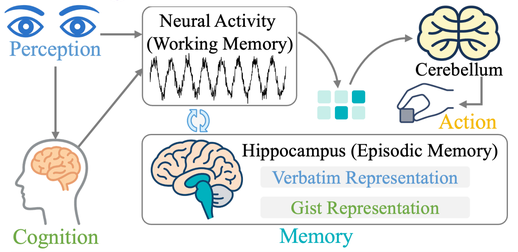
Tsinghua University | LEAP | DEXMAL | StepFun | ICLR

1 Motivation

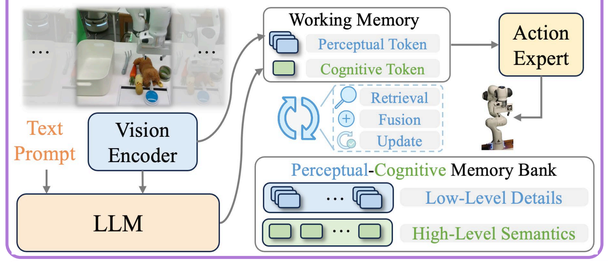
① Why Does Memory Matter in VLA?



② How Humans Do It?



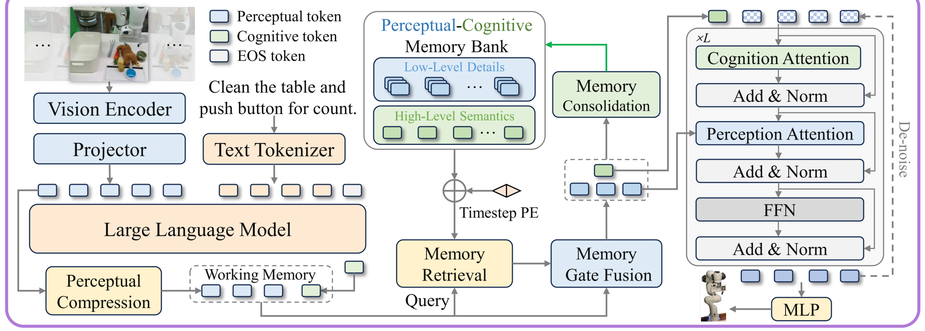
③ Human-Inspired Memory for VLA



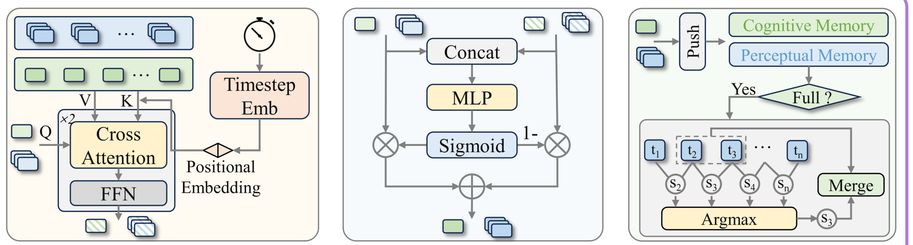
④ Insight

2 Method

Overall framework



Module Derails

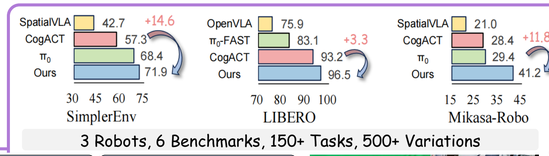


(a) Memory Retrieval — Select past info relevant to current decision

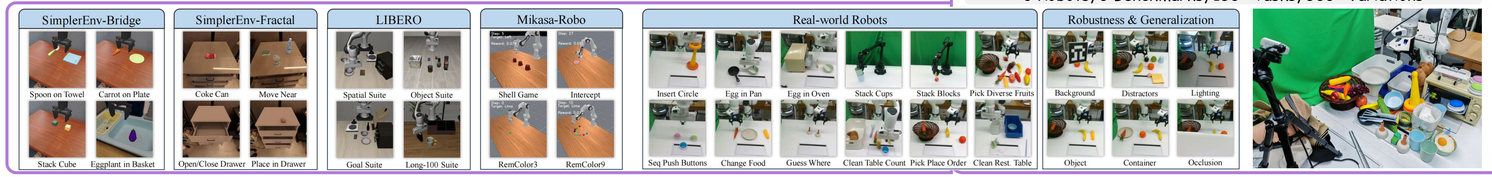
(b) Memory Gate Fusion — Adaptive fusion of past and current info

(c) Memory Consolidation — Merge nearby & similar entries for compact memory

3 Experiments



3 Robots, 6 Benchmarks, 150+ Tasks, 500+ Variations



Future Direction

① Hierarchical Memory for Efficient Retrieval
② Memory Reflection for Reasoning
③ Lifelong Memory Across Episodes, Environments, and Embodiments
④ Improved Memory Consolidation
⑤ Memory with World Models