



東南大學
SOUTHEAST UNIVERSITY

Memory-Augmented Vision-Language- Action Models in Embodied AI: Methodological Advances, Efficiency Challenges, and Future Directions

汇报人：李文卓



東南大學
SOUTHEAST UNIVERSITY

CONTENTS

01

Introduction

— Why Vision-Language-Action Models need Memory?

02

RoboMemory

— A Brain-inspired Multi-memory Agentic Framework for Interactive Environmental Learning in Physical Embodied Systems

03

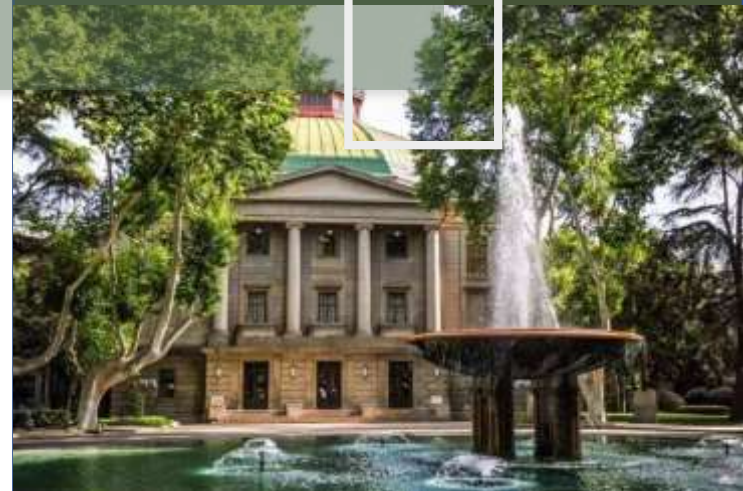
MEMER

— Scaling up memory for robot control VIA experience retrieval

04

Limitation & Future work

— What can we do next?





東南大學
SOUTHEAST UNIVERSITY

PART ONE

Introduction

— Why Vision-Language-Action Models need
Memory?

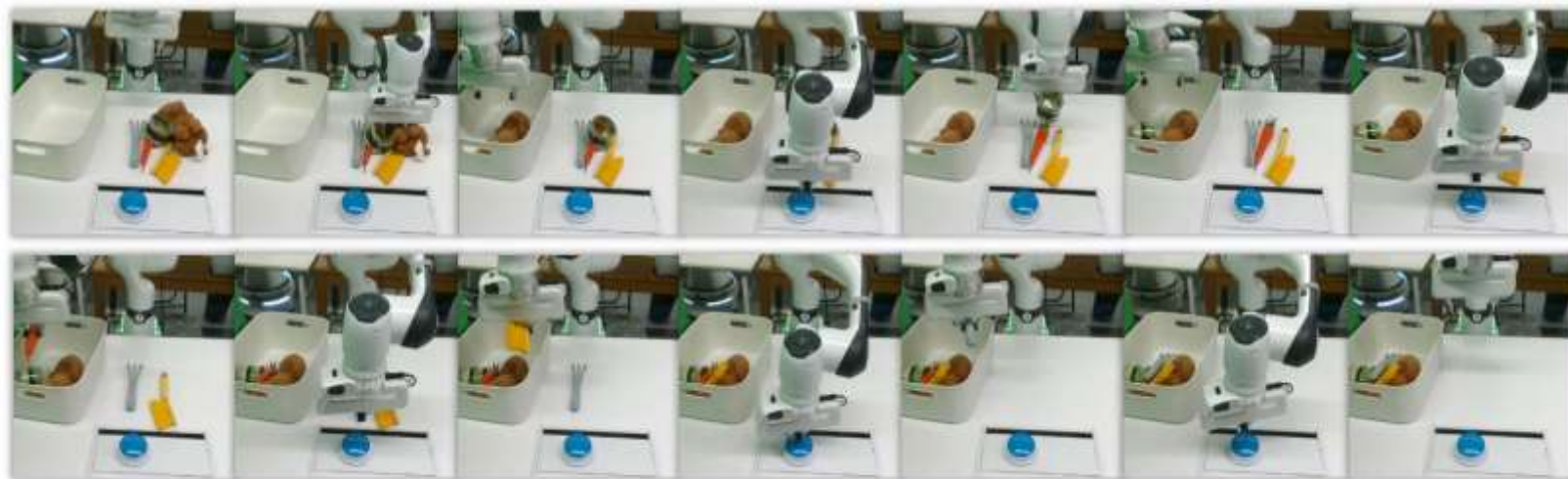
01



Introduction

Several major challenges currently facing robots in the real world:

- 无法持续学习新知识
- 复杂的记忆系统导致反应迟钝
- 难以关联不同任务的经验
- 在做决策时容易陷入死循环



(d) Clean Table & Count: clean the table one by one, and press the button for each item cleaned

These types of tasks typically involve counting repetitive actions or manipulating temporarily occluded objects.



Introduction

VLA Paradigm is still lack of:

1. 稳健的空间与长期记忆机制
2. 获取并利用具身规划经验以实现好的泛化

Assessment limitations:

1. 狭窄、非标准化的环境
2. 部分可观测性
3. 碎片化的空间推理
4. 异构记忆的低效整合



Partial observability

Temporal dependencies

Long-horizon Task



Introduction

Challenge of Incorporating historical information

利用历史上下文会带来巨大的计算开销

eg. HAMLET：在VLA输入中简单地附加额外的四帧过去观测帧，就会使前向传播速度减慢约35%，并使峰值内存消耗增加约3.6倍与单帧设置相比，膨胀的内存占用会减少可行的批处理规模，从而进一步限制了可扩展性。

False correlations caused by memory

不良的记忆系统 甚至可能损害性能

这种现象被称为“因果混淆”(Pim et al. 2019)

记忆加剧了假相关性，使机器人模仿学习变得混乱





東南大學
SOUTHEAST UNIVERSITY

PART TWO

RoboMemory

— A Brain-inspired Multi-memory Agentic
Framework for Interactive Environmental
Learning in Physical Embodied Systems

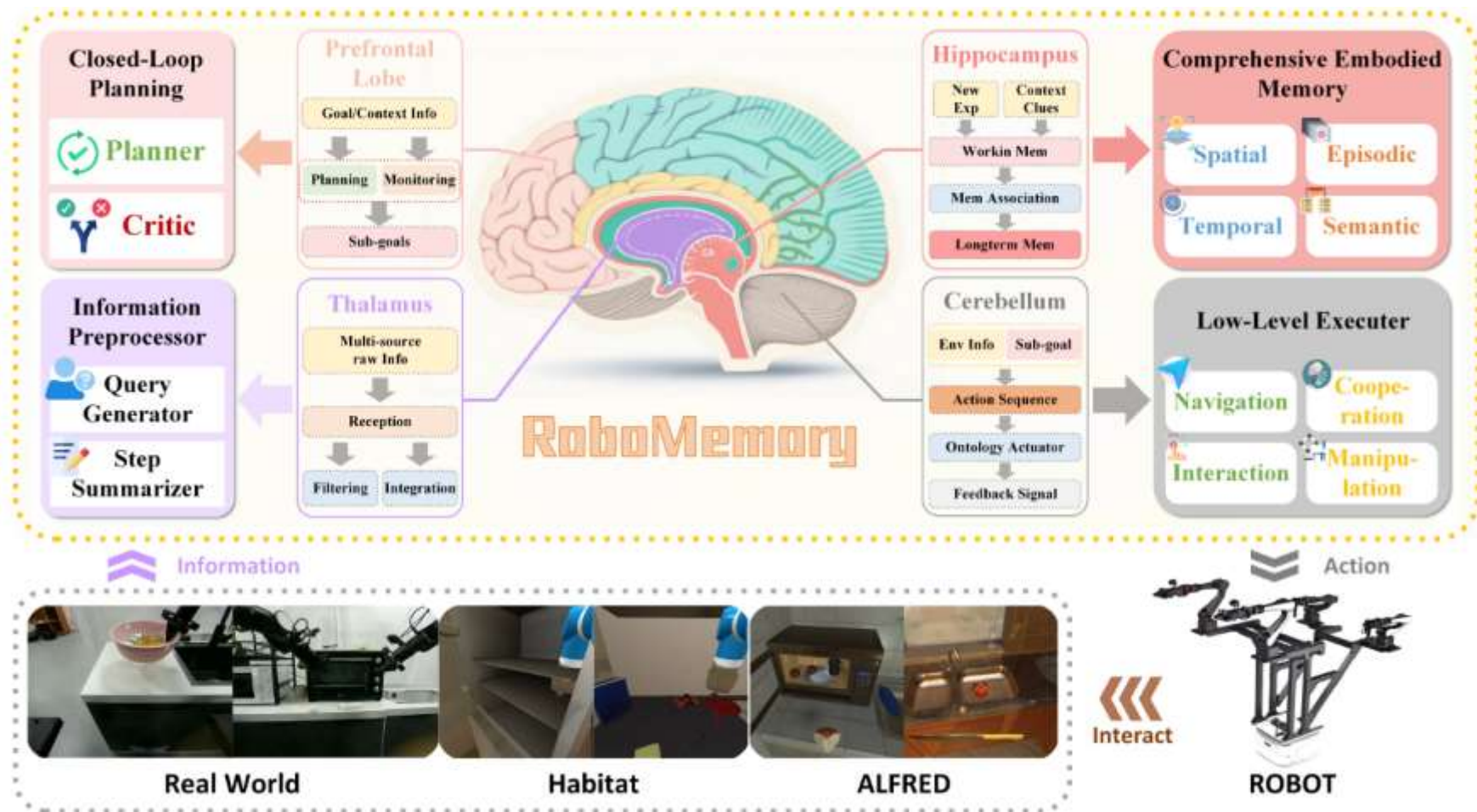
02

NeurIPS 2025



RoboMemory

1. Information preprocessor将多模态传感器输入转换为当前场景的文本摘要，作为 Comprehensive Embodied Memory的主要输入
2. Comprehensive Embodied Memory更新其内部表征，包括动作历史、物体位置与经验知识
3. 信息更新后，记忆系统检索上下文相关的条目以服务于 Closed-Loop Planning模块
4. Closed-Loop Planning模块利用该上下文记忆生成高层次、基于文本的动作指令
5. 这些指令被分派至 Low-Level Executer，由其直接控制机器人完成指令



模仿大脑结构，设计了四个协同工作的模块，特别是其创新的终身具身记忆系统，系统包含空间、时间、情景和语义四种并行工作的记忆模块



RoboMemory

人脑组件	功能	RoboMemory对应模块	在机器人中的作用
丘脑	感官信息中继	信息预处理器	接收摄像头画面，转化成模型能理解的文本信息。
海马体	长期记忆形成与整合	终身具身记忆系统	核心记忆库，存储空间、时间、情景、语义四种记忆。
前额叶皮层	规划与决策	闭环规划模块	制定行动计划，并根据环境变化进行动态调整。
小脑	运动控制与肌肉记忆	底层执行器	将高层指令（如拿起杯子）转化为具体的电机动作。



RoboMemory

信息预处理器：

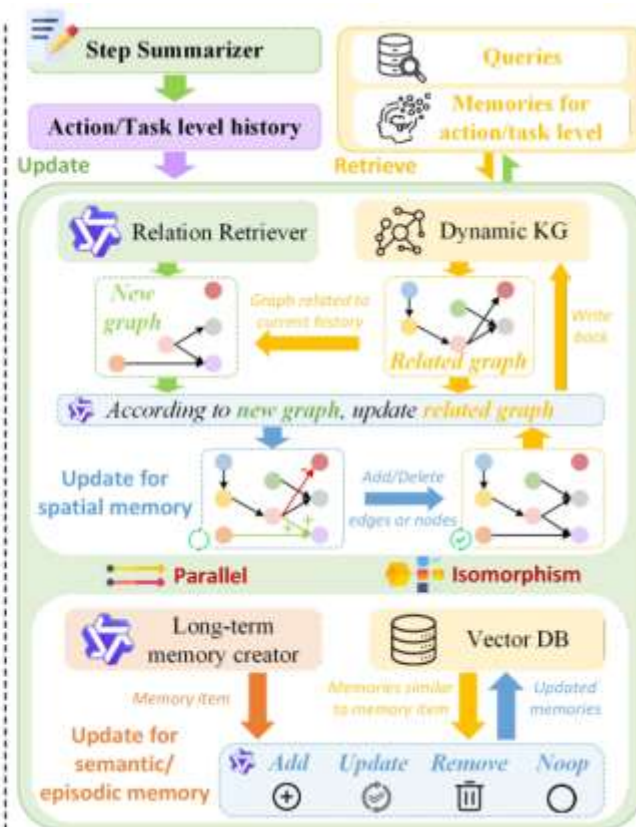
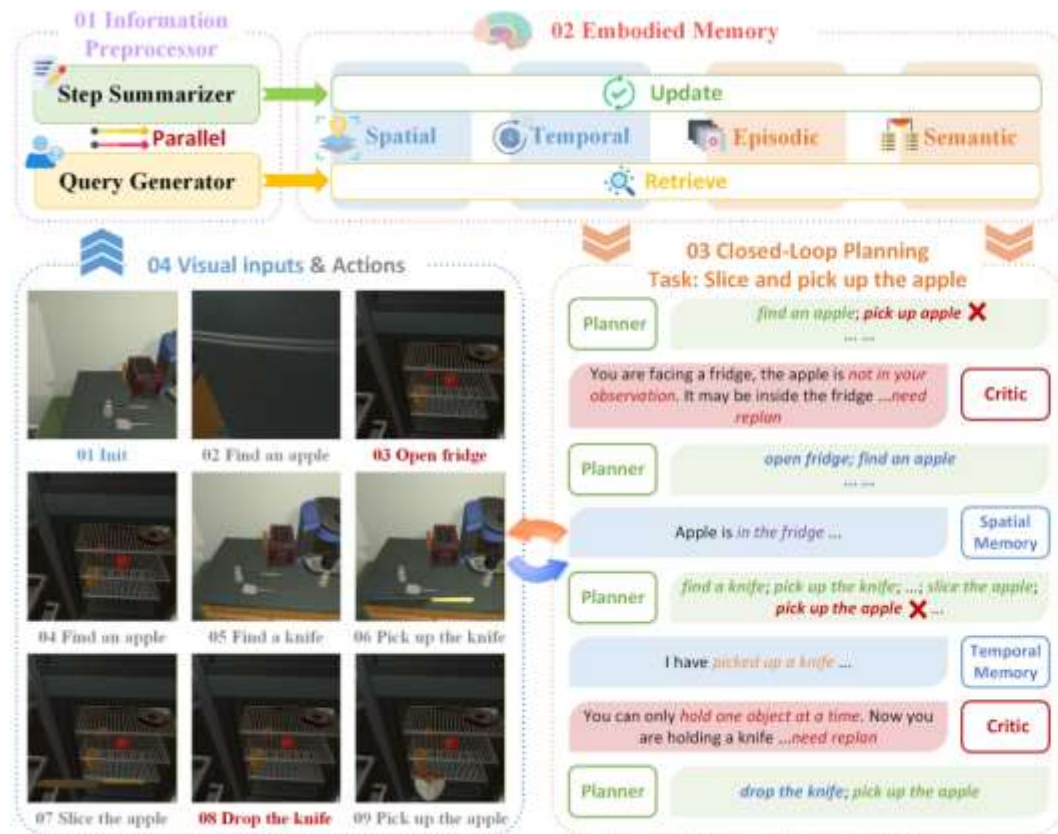
1. 将刚刚发生的事情用一句话总结下来，形成工作记忆。

eg. 我刚刚把红色的杯子放在了桌子上

2. 根据当前的所见所闻，生成一个用于提问长期记忆系统的问题。

eg. 桌子上除了红色的杯子还有什么？

或者我上次看到钥匙是在哪里？





RoboMemory

终身具身记忆系统

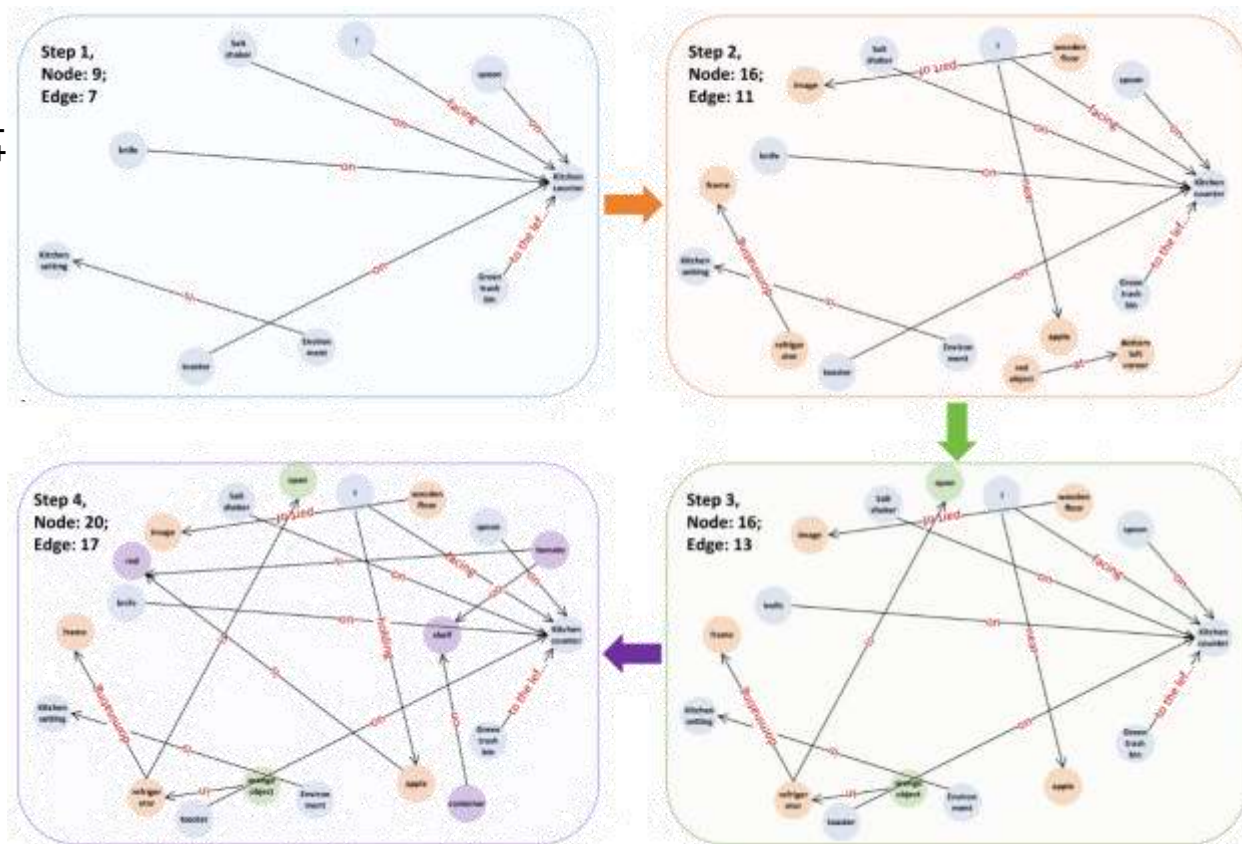
短期记忆：一个 FIFO 缓冲区

一个容量有限的便签本。比如只能记 3 件事。当第 4 件事发生时，第 1 件事就会被挤掉。但为了不完全丢失信息，在挤掉之前，一个语言模型会把这 3 件事总结成一个摘要再存入更高层的记忆中。这实现了信息的有效压缩和传递

空间记忆：一个动态更新的知识图谱。

动态更新的挑战：真实世界是动态的。机器人把苹果从桌上拿到冰箱里，知识图谱就需要实时更新。如果每次更新都要重构整个图，那效率太低了。

基于检索的增量式更新算法：当新信息（比如苹果现在在冰箱里）进来时，系统不会去动整个庞大的知识图谱。它会先检索出与苹果和冰箱相关的局部子图，然后在这个小范围内进行信息的融合、冲突的解决（比如删除苹果在桌上这个旧关系），最后再把这个更新后的小子图合并回主图





RoboMemory

情景记忆：完整地记录每一次任务的故事，包括目标、执行过程、遇到的困难和最终结果（成功或失败）。

每个任务结束后，系统会将整个过程打包成一个情景记忆实体，存储在一个使用 **RAG 技术** 的数据库里。当未来遇到相似任务时，机器人可以翻阅这本日记，寻找可借鉴的经验。

语义记忆：从多次的情景记忆中提炼出通用的知识、规律和技能。

在任务结束后，一个专门的模块会分析刚刚发生的情景，从中总结出成功的关键因素或失败的根本原因，形成抽象的语义知识，同样存入 **RAG 数据库**。

eg. 机器人可能会学到：玻璃门通常是推开的，而木门通常是拉开的，或者拿起装满水的杯子时，动作要慢一点。



$$M_t = [M_t^{(1)}, M_t^{(2)}, \dots, M_t^{(L)}]$$

$$M_t = \mathcal{U}(M_{t-1}, s_t) \quad \mathcal{U}: \text{记忆更新算法} \quad (1)$$

$$r_t = \mathcal{R}(M_t, q_t) \quad \mathcal{R}: \text{记忆检索算法} \quad (2)$$

并行更新与检索
所有记忆模块

$$r_t = [r_t^{(1)}, r_t^{(2)}, \dots, r_t^{(L)}]$$

各模块的检索结果



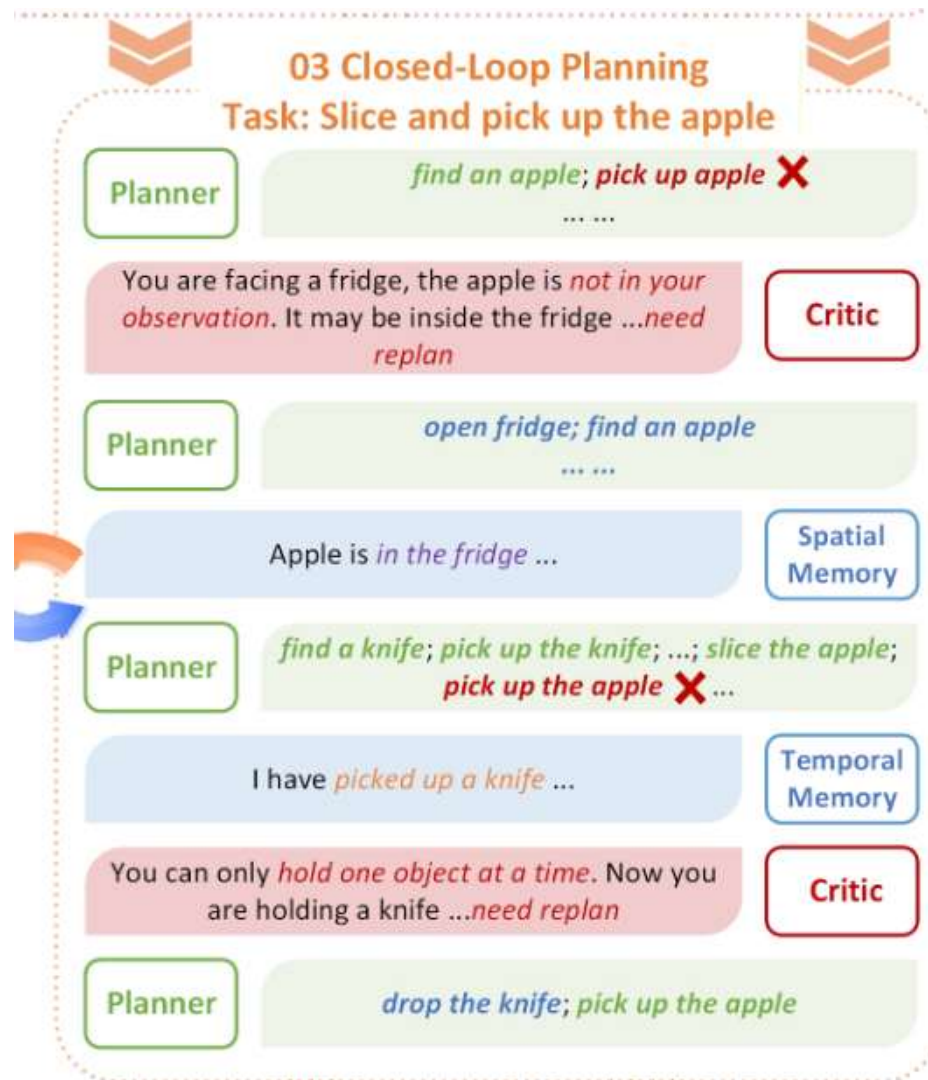
RoboMemory

闭环规划模块：Planner-Critic 机制

Planner: 根据当前任务目标和从记忆系统中检索到的信息，制定一个初步的多步骤行动计划。

eg. 1. 走向冰箱 -> 2. 打开冰箱 -> 3. 拿出牛奶。

Critic: 在执行每一步之前，都会重新审视这一步在最新的环境状态下是否还合理。如果机器人走到冰箱前，发现冰箱门已经被别人打开了，Critic 就会发现打开冰箱这一步是多余且错误的，于是它会立即要求Planner停止，并根据新情况重新规划。





RoboMemory

原始的 Planner-Critic 机制有个缺陷。如果Critic过于**挑剔**，一直对Planner的计划不满意，机器人就会陷入重新规划 ->被否决 ->再重新规划的死循环，卡住不动。

RoboMemory 做了一个简单而有效的修改：Planner的**第一步**不经过评论家审核，直接执行。这保证了机器人至少能动起来，打破僵局，避免分析瘫痪。算是一个工程实践中的技巧。

底层执行器：

负责将规划模块发出的高层抽象指令（如 `pick_up(milk)`），翻译成机器人机械臂和底盘可以执行的、精确的电机控制信号。在真实世界的实验中，研究者们使用了一个名为 $\pi 0$ 的 VLA 模型来承担这个角色，并通过 LoRA 微调技术来提升其在特定任务上的表现



实验评估指标

成功率 (Success Rate, SR): 任务完全成功的比例 :

$$SR = \mathbb{E}_{x \in \mathcal{X}} [\mathbb{1}_{SCN_x = GCN_x}] \quad (3)$$

目标条件成功率 (Goal Condition Success Rate, GC): 任务完成度的平均百分比 :

$$GC = \mathbb{E}_{x \in \mathcal{X}} \left[\frac{SCN_x}{GCN_x} \right] \quad (4)$$

在上述公式中， $\mathcal{X}$ 代表测试子集， x 表示测试任务。成功条件数（ SCN_x ）指智能体已完成的条件数量，而全局条件数（ GCN_x ）表示任务完成所需的条件总数。当 $SCN_x = GCN_x$ 时，该任务即被视为成功。



RoboMemory

终身学习能力的验证

研究团队搭建了一个真实的厨房环境，并使用一个名为 Mobile ALOHA 的双臂移动机器人平台进行实验。

他们让机器人执行 15 个不同的任务，每个任务连续执行两次，**中间不清除或重置机器人的长期记忆**。如果机器人真的从经验中学习了，那么第二次尝试的成功率应该会高于第一次。

在所有三类任务中，第二次尝试的成功率都显著高于第一次。

eg.

任务：把香蕉放进烤箱。

第一次尝试：机器人陷入了循环，反复在几个已经探索过的、没有香蕉的地点打转，最终失败。

学习过程：任务失败后，语义记忆总结出：不应该反复去已经确认过没有目标物品的地方。情景记忆则记下了这次失败的完整过程。

第二次尝试：在记忆的指导下，机器人意识到它从未去过厨房台面这个地方。它果断地导航到新地点，成功找到了香蕉，并完成了任务。

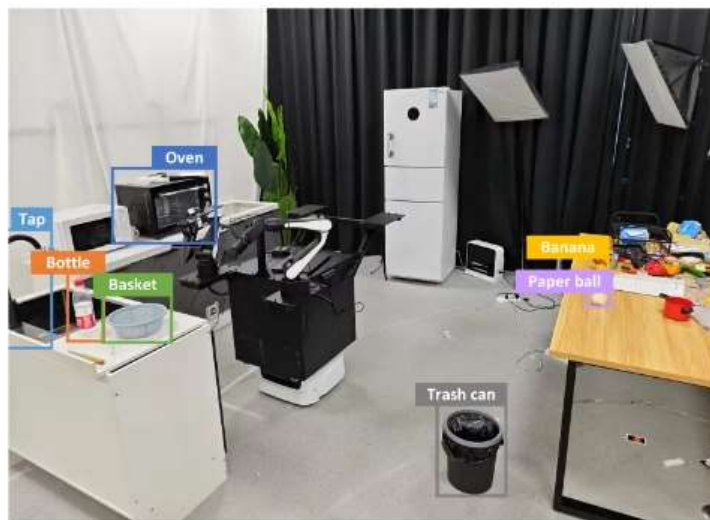


Fig. 4: Visualization of the experimental environment.

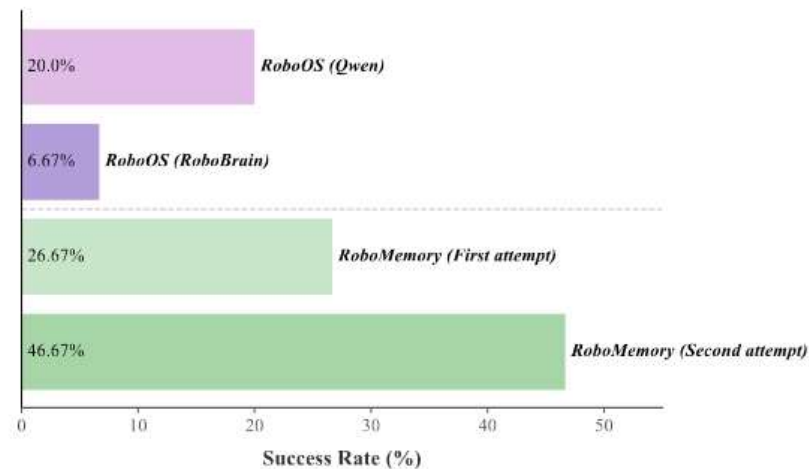


Fig. 5: Real-world experiment results. “Qwen” denotes Qwen2.5-VL-72B-Ins; “RoboBrain” denotes RoboBrain2.0-32B.



東南大學
SOUTHEAST UNIVERSITY

PART THREE

MEMER

— Scaling up memory for robot control VIA
experience retrieval

03

arXiv:2510.20328





MEMER

该论文提出了一个**分层策略与经验检索机制**

核心：“分层政策”与“动态**关键帧管理**”

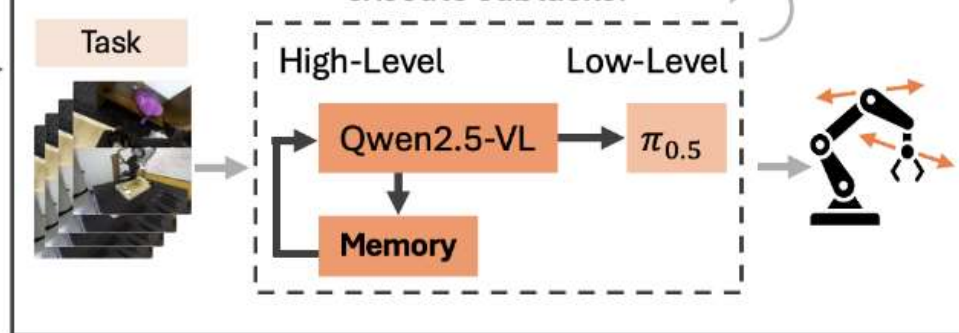
Minimal Robot Training Data



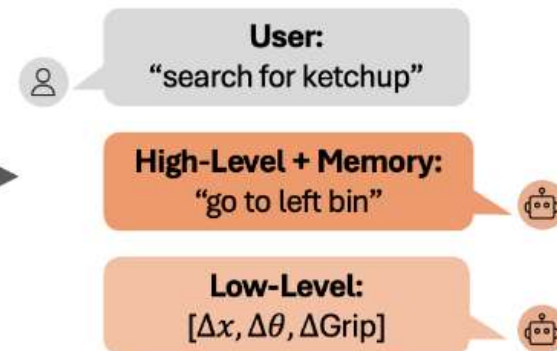
MemER

Hierarchical VLA Model

Finetune to select keyframes & execute subtasks:



Memory-Aware
Closed Loop Robot Control



MemER 将机器人政策拆分为高层与低层两个协同模块，形成“记忆决策 - 动作执行”的闭环：

高层政策（记忆决策者）基于Qwen2.5-VL-7B-Instruct 微调而成，核心职责是“管理记忆”与“分解任务”。

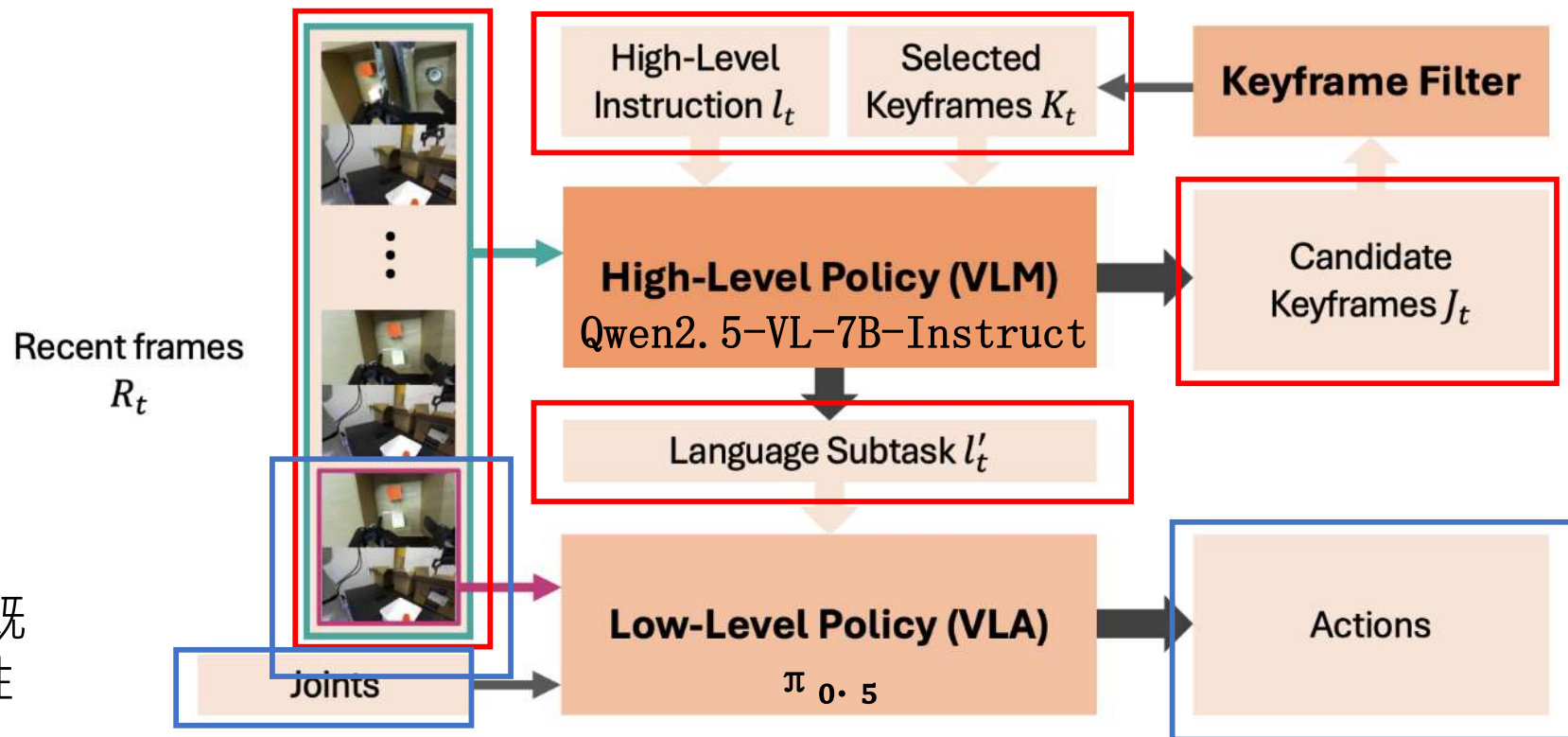
它接收三部分输入：各相机的最近 N 帧图像（实验中 $N=8$ ，即“近期上下文”）、原始任务指令（如“寻找番茄酱”）、历史筛选的关键帧集合 K （最多 8 帧，来自整个任务周期）。



高层专注于长时序记忆推理，
避免陷入高频控制的细节

低层专注于实时动作执行，无
需负担记忆管理的计算成本

两者通过“子任务指令”衔接，既
保证了记忆的有效性，又不牺牲
控制响应速度。



MemER 架构一览



关键帧管理：用聚类算法浓缩“有效记忆”

高层政策筛选关键帧的核心是

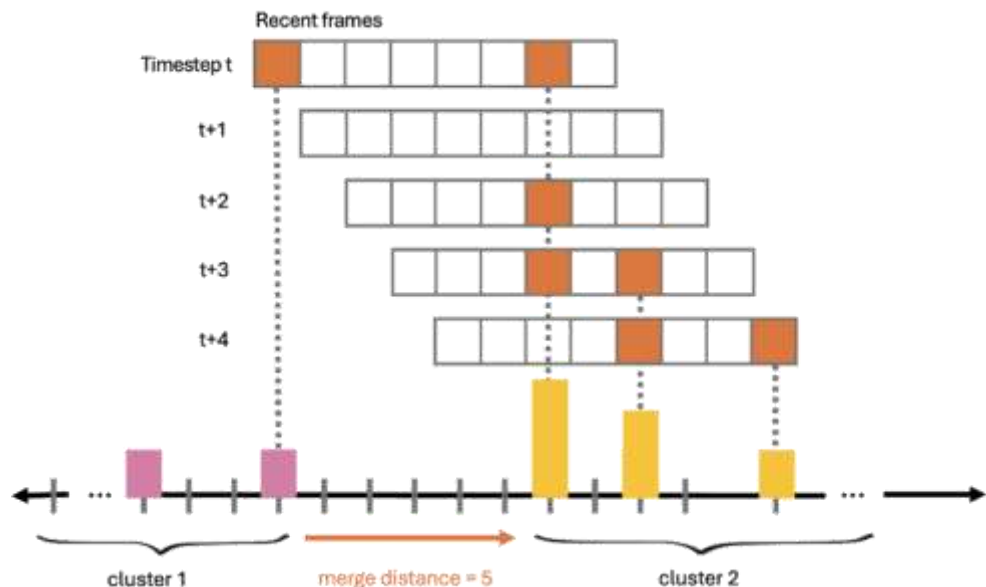
1D 单链接聚类算法，该机制确

保机器人能从数千帧的任务周

期中，动态保留最有价值的视

觉信息，避免记忆库无限膨胀。

具体流程分为三步：



1. **候选帧收集**。高层政策在每个时间步都会从近期 8 帧中提名候选关键帧 J_t ，系统收集所有时间步的候选帧，提取其时间索引并形成有序列表 G_0 。列表会保留重复索引——例如某帧被多次提名，说明其包含的信息对任务至关重要。

2. **聚类分组**。系统将时间间隔不超过 d 帧（实验中 $d=5$ ）的索引归为同一聚类。

例如索引 $\{1, 3, 3, 4, 10\}$ 会被分为两个聚类： $C_1 = \{1, 3, 3, 4\}$ （间隔均 ≤ 5 ）、 $C_2 = \{10\}$ （与前一聚类间隔 $6 > 5$ ）。这种分组方式确保同一聚类的帧代表同一关键事件（如“查看左侧箱子”的整个过程）。

3. **代表帧选择**。对每个聚类，选取**中位数索引**对应的帧作为最终关键帧。例如聚类 C_1 的中位数索引为 3，对应的帧即为该事件的“记忆锚点”。这种方式既减少了冗余（每个事件仅保留 1 帧），又能最大程度保留关键信息——中位数索引对应的帧最能代表该事件的关键信息。



通过该机制，MemER 能将整个任务周期的数千帧图像浓缩为不超过 8 帧的关键帧集合，在降低计算成本的同时，确保了记忆内容与任务强相关。

为降低实际应用门槛，MemER 在训练与部署层面做了三项关键优化：

少数据高效训练：不同于传统长时序任务需要海量标注数据，MemER 仅需两类数据就能完成微调：一是 50 条完整的长时序任务演示轨迹（每条包含图像、本体感受数据、子任务标注），二是 10-15 条干预演示（用于处理部署中的常见失败场景，如夹爪未抓稳物体）。这种少数据需求源于高层政策复用了 Qwen2.5-VL-7B-Instruct 的预训练视频理解能力，低层政策则复用了 $\pi_{0.5}$ 在 DROID 数据集上的控制能力。

模型融合策略：高层政策微调后可能会丢失预训练模型的鲁棒性（如对异常帧的容忍度）。研究团队采用“权重插值”方案：将微调后的模型权重 θ_{ft} 与预训练权重 θ_{pre} 按 $\alpha=0.8$ 的比例融合（ $\theta_{final} = 0.8 \times \theta_{pre} + 0.2 \times \theta_{ft}$ ）。实验显示，该策略能在保持任务适配性的同时，提升模型对低层政策重试、冻结等异常情况的容忍度，三项任务的平均性能提升约 5%。

异步闭环部署：为平衡推理质量与响应速度，MemER 采用异步调度。高层政策以 1Hz 频率更新子任务与关键帧，低层政策以 2Hz 频率输出动作。系统将图像以 2Hz 采样存入队列，高层政策完成当前推理后，直接从队列读取最新数据生成下一个子任务。这种方式避免了低层政策等待高层推理的延迟，确保闭环控制的稳定性。



对比实验

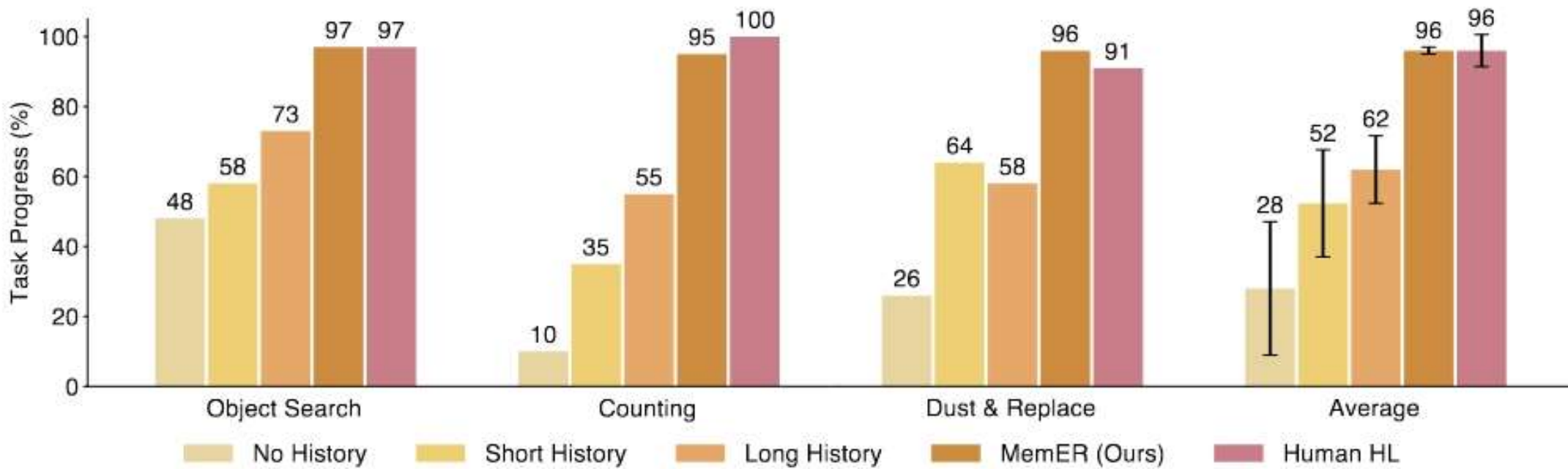
研究团队设置了四类基线方法，以验证 MemER 的优势：

无历史：高层政策仅使用当前帧，无任何历史信息，模拟现有主流机器人政策；

短历史：高层政策使用最近 8 帧上下文，代表“有限上下文”方法；

长历史：高层政策使用最近 32 帧上下文，代表“暴力扩展上下文”方法；

人类高层：由人类手动提供子任务指令，仅测试低层政策的执行能力，作为性能上限参考。





模态对比：视觉记忆优于文本记忆

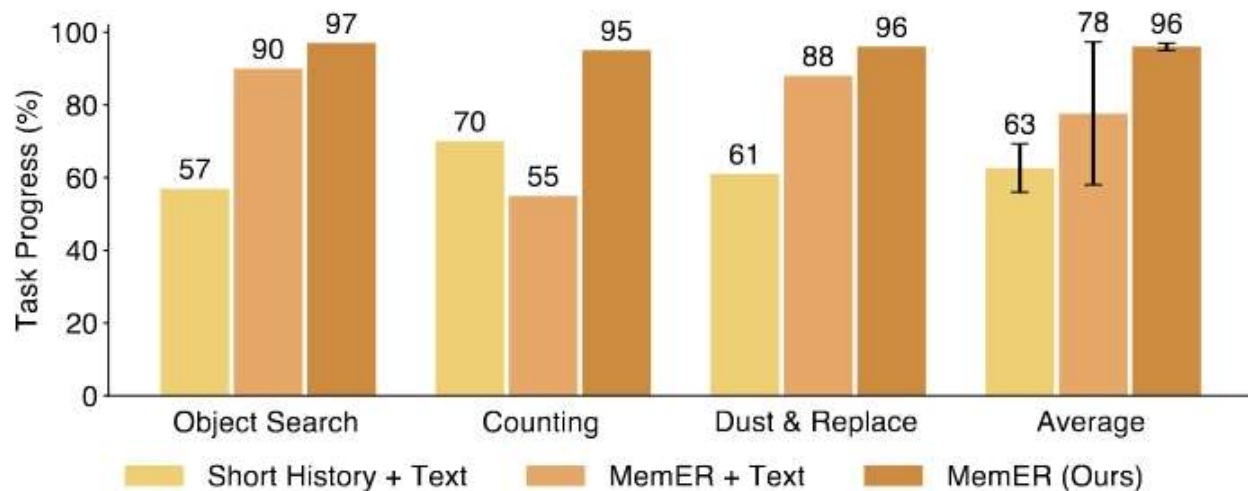
为验证“哪种模态更适合机器人记忆”，研究团队测试了两种文本记忆方案：

短历史 + 文本：高层政策使用最近 8 帧图像 + 对应的子任务文本描述；

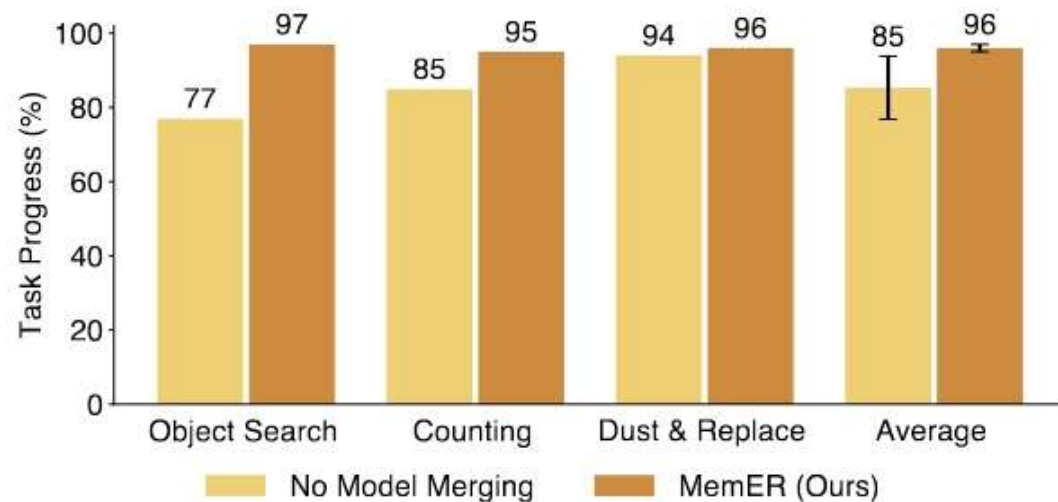
MemER + 文本：在 MemER 的关键帧基础上，添加对应的子任务文本描述。

实验结果显示，纯视觉记忆的 MemER 表现最优，文本记忆的加入反而导致性能下降：

Modality Results



Model Merging





東南大學
SOUTHEAST UNIVERSITY

PART FOUR

Limitation & Future work

— What can we do next?

04



Limitation

RoboMemory

- 规划能力瓶颈：任务失败主要源于规划错误，说明即使记忆系统完善，底层 VLM 模型的推理与规划能力仍是整体性能的天花板。
- 知行脱节问题：高层大脑规划与底层身体执行存在脱节，抽象指令在物理执行中易受环境复杂性影响。
- 泛化性与鲁棒性待提升：虽有真实环境验证，但实验场景受控，面对更复杂、开放的家庭环境时，感知、知识图谱扩展及大规模记忆管理等方面的鲁棒性和效率仍需突破。

MEMER

- 缺乏过期记忆删除机制，仅能添加关键帧，若任务延长至数小时会导致关键帧集合膨胀、计算成本上升；
- 高层 1Hz、低层 2Hz 的调度频率，难以满足装配线分拣等高速操纵任务需求；
- 仅依赖视觉信息，未整合触觉、听觉等其他感官数据，记忆维度不够全面；
- 实验仅在单机械臂上完成，尚未适配移动操纵机器人或多机器人协作场景。



Future work

RoboMemory

需进一步提升大模型的逻辑与因果推理能力

加强高层与底层的协同与交互，实现大脑 - 身体一体化进化。

MEMER

需让高层政策学会判断记忆时效性以删除无用信息

通过优化模型缓存、改进 tokenization 算法提升控制频率，将多模态数据融入关键帧筛选以丰富记忆维度

探索空间映射与情景记忆结合的方案，适配更复杂的多机器人应用场景。



東南大學
SOUTHEAST UNIVERSITY

Thanks for listening!

汇报人：李文卓