

PersonaVLM

Long-Term Personalized Multimodal LLMs

长期个性化多模态大语言模型

聂畅¹ · 付朝友¹ · 张一凡² · 杨海华² · 单彩峰¹

¹南京大学 ²字节跳动

冯朝晨组会论文汇报

目录

CONTENTS

01 背景与动机

从通用 MLLM 到个性化多模态大模型

02 方法详解

记忆架构 · 多步推理与时间感知检索 · 个性演化机制 · 训练方案

03 数据集与基准

合成数据管道 · Persona-MME (7维14任务 2000+案例)

04 实验结果

Persona-MME 主表 · 对齐评估 · 7B 以 79% 胜率超越 GPT-4o

05 消融与分析

记忆类型消融 · PEM 消融

06 总结

核心贡献 · 局限性

多模态大语言模型（MLLMs）为数百万用户提供日常助理服务。然而，它们生成符合个体偏好响应的能力仍然有限。先前的方法仅通过输入增强或输出对齐实现静态、单轮个性化，因而无法捕捉用户随时间不断演变的偏好和个性



Figure 1. Illustration of PersonaVLM's three core capabilities for long-term personalization. PersonaVLM proactively remembers user preference shifts, performs multi-turn reasoning with retrieval, and generates responses aligned with the user's personality. In contrast, existing personalization strategies, such as input augmentation and output alignment, will result in poor recommendations based on outdated memories and replies that are misaligned with the user's personality.

案例一：偏爱转变未被捕捉

用户最初表达对雪碧的偏好，但在一次多模态交互中，为了缓解焦虑转而选择可口可乐。当用户后来再次表达压力时，传统的检索增强回复仍推荐雪碧

案例二：个性不匹配

通用对齐的模型回复可能表现得过于外向和热情，这对于正在经历焦虑的内向、神经质用户而言并不合适。用户的真实个性特征通常在大量不相关的对话中微妙展现，静态对齐方法无法捕捉这种逐步揭示的过程。

从通用到个性化：MLLM 的下一个前沿

当前 MLLM 的困境

固定上下文窗口 + “一刀切” 优化范式

即使在 GPT-4o 等先进专有模型中，
生成迎合用户独特偏好的回复能力仍然有限

用户期望的转变

通用问题解决 → 个性化与共情的长期体验

核心问题

如何将通用 MLLM 转变为真正的个性化助手，针对个体用户知识和偏好定制响应

01

准确推断用户意图

02

动态对齐偏好和个性

03

持续记住多模态信息

应用价值：推荐系统 · 医疗保健 · 教育 · 日常 AI 助手

现有应对方案的主要方向

基于自适应 Adaptation-based

模型参数层面

核心思路

微调将用户特定知识编码到模型参数

代表工作

MyVLM (可学习嵌入)

Yo'LLaVA (软提示)

致命缺陷

每个新用户概念都需要微调，可扩展性差，且无法捕捉用户偏好的动态演化

基于增强 Augmentation-based

输入增强层面

核心思路

外部数据库 + 检索用户记忆

代表工作

RAP (开放词汇检测器)

A-Mem / Memory OS (通用记忆)

致命缺陷

手动预定义数据库，缺乏从动态交互中主动管理和更新知识的机制

A-Mem 等局限于纯文本 + 依赖专有 API (商业闭源模型)，缺乏真正多模态输入的适用性

基于对齐 Alignment-based

输出对齐层面

核心思路

优化目标从通用改为用户特定标准

代表工作

ALIGNXPRT

PAS (用户特定探针)

致命缺陷

预设静态用户特征，无法适应多样化的用户偏好和沟通风格

训练用户特定的探针，探针的静态特性意味着随着用户个性在长期交互中演变，对齐可能变得过时

共同的失败模式：都是“静态”设计——为单轮、快照式交互设计，无法捕捉用户随时间的变化

长期个性化的两大支柱

基于失败案例分析，确定有效长期个性化的两个基础支柱

支柱一：个性化记忆架构

主动构建和管理动态的、以用户为中心的多模态数据库

解决案例一的偏好转变问题：

能够跟踪用户从“雪碧→可口可乐”的动态变化

而非仅仅存储静态的初始偏好

支柱二：记忆利用与响应对齐

运用推理和检索，生成与用户演化特征深度对齐的回复

解决案例二的个性不匹配问题：

追踪用户在长期交互中逐步揭示的内向、神经质个性

动态调整回复风格以匹配个性特征

→ PersonaVLM：首个整合记忆、推理、对齐的统一框架

构建于个性化记忆架构之上，通过响应（Response）和更新（Update）两个协作阶段实现长期个性化。

响应阶段（生成对齐回复）+ 更新阶段（演化个性与记忆）

记忆

主动提取并总结多模态记忆

推理

多轮检索与整合相关记忆

响应对齐

推断演化个性确保输出一致

PersonaVLM 整体框架

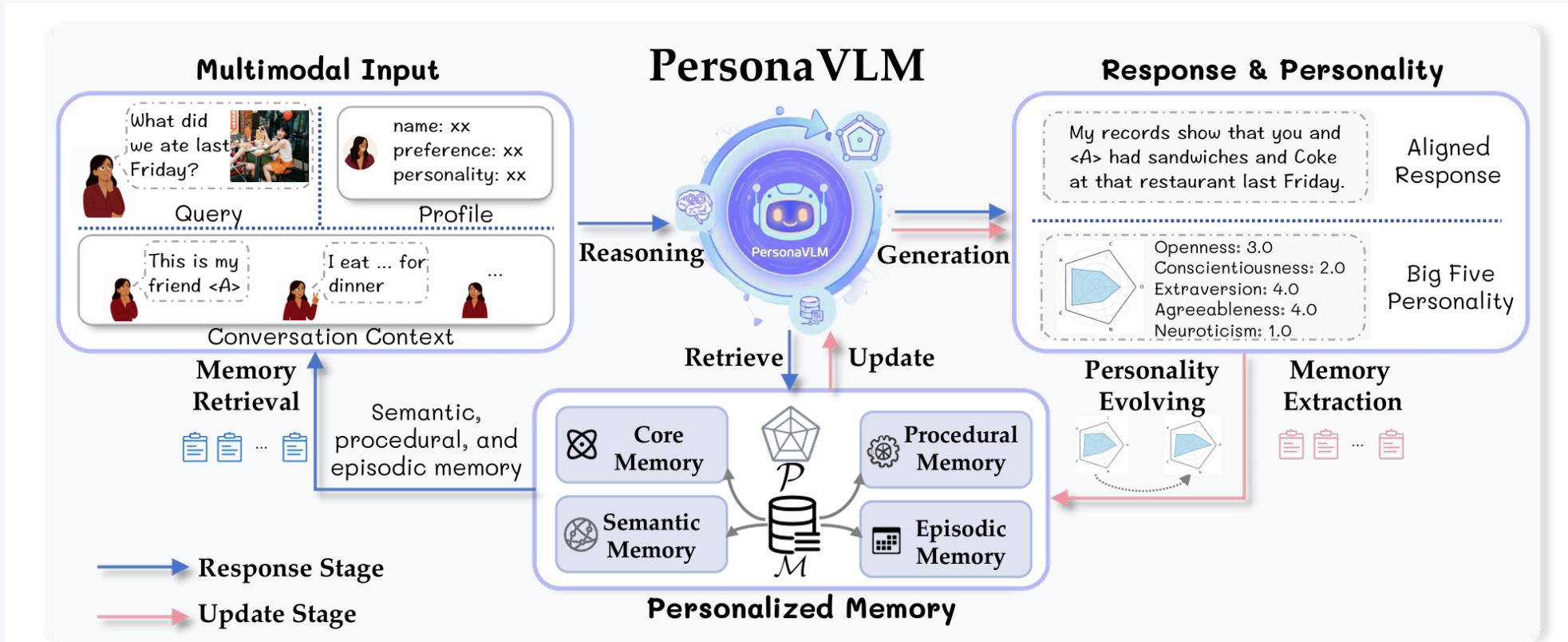


Figure 2. PersonaVLM框架概述。它利用个性化记忆架构，并通过两个协作阶段实现长期个性化。在Response Stage（蓝色箭头），它处理多模态输入，从个性化记忆中检索，并生成与个性对齐的响应。随后，在Update Stage（粉色箭头），该框架分析完成的交互以提取关键记忆，并更新用户不断发展的个性配置文件¹。

PersonaVLM 整体框架

个性化记忆架构：该架构旨在构建和维护一个全面的长期用户档案，存储两大类信息。

首先，它维护一个用户个性档案 (P)，以 Big Five 维度 (开放性、尽责性、外向性、宜人性和神经质) 的分数向量形式提供用户个性的量化表示。

其次，它具有一个多类型记忆数据库 (M)，捕获广泛的用户相关知识。该基于时间线的智能系统支持灵活的 CRUD (创建、读取、更新、删除) 操作，并分为四种不同的记忆类型。

个性化记忆架构 — Personalized Memory Architecture

个性档案 \mathcal{P}

大五人格向量 $p \in \mathbb{R}^5$

开放性·尽责性·外向性·宜人性·神经质

核心记忆

基础属性

仅保留最新版本

语义记忆

实体·偏好·视觉概念

按时间线存储

情景记忆

原子事件+时间戳

按时间线存储

程序记忆

计划·目标·习惯

仅保留最新版本

响应阶段 Response Phase

$$\mathcal{R}_m = R(Q_m, C_m, \mathcal{M}_{\{m-1\}})$$

用户查询 Q_m (文本指令+可选图像+时间戳)

对话上下文 C_m (60分钟窗口)

→ 多步推理 → 时间感知检索 (最多3次) → \mathcal{R}_m

模型自主决定：是否需要检索？检索什么？从何时检索？

更新阶段 Update Phase · 异步执行

$$(\mathcal{P}_m, \mathcal{M}_m) = U(Q_m, \mathcal{R}_m, \mathcal{M}_{\{m-1\}})$$

PEM 个性演化：EMA + 余弦衰减调度

$$p_m \leftarrow \lambda_m \cdot p_{\{m-1\}} + (1-\lambda_m) \cdot p'_m$$

选择性记忆更新：语义 (每轮) / 核心+程序 (每会话)

闲时自动执行，不阻塞用户交互

▲ 响应后异步触发更新，更新结果写回记忆供下一轮响应使用 ▲

个性化记忆架构：四种记忆类型

借鉴认知心理学分类，每种记忆有独立的存储和更新逻辑

核心记忆 Core Memory

存储用户的基本属性

Human (人员信息)：年龄、性别、偏好、兴趣 (用户名为必填)

Persona (角色设定/人格信息)：身份、角色、交互风格要求

更新策略：每次会话结束时 CRUD · 仅保留最新版本

每次会话更新

语义记忆 Semantic Memory

提取与事件无关的抽象知识

明确指令、核心事实、偏好与习惯、视觉概念

视觉概念：键值对存储，文本描述 + 图像裁剪

更新策略：纯增量 · 按时间顺序存储 · 每轮更新

每轮更新

情景记忆 Episodic Memory

将原始对话组织成带时间戳的原子事件

每段：摘要 + 关键词 + 涉及的对话轮次

消融证明：最关键的记忆组件 (-12.41% 32k)

更新策略：按主题自动构建 · 原始对话数据永不可删

自动化构建

程序记忆 Procedural Memory

记录以用户为中心的计划、目标和重复行为

长期目标：进行中的项目、计划、目标

习惯与常规：自动识别的重复行为或 workflow

更新策略：可变 · 每会话 CRUD · 仅保留最新版本

每次会话更新

记忆检索与管理

时间感知检索 + 多源并行搜索 + 自包含管理策略

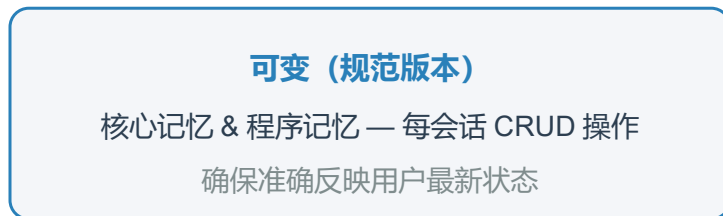
检索机制



视觉概念检索



管理策略



设计优势

不依赖 GPT-4 等商业模型

显式个性化, 以用户为中心的组织

完整多模态支持

本地运行, 消除隐私问题

克服 A-Mem/Memory OS 局限

vs 通用记忆架构 (纯文本+专有API)

响应阶段：多步推理与时间感知检索

响应阶段 Response Phase

该阶段的目的是通过执行多步推理和基于时间线的检索来生成对齐的响应。形式上，在第 m 轮对话中，该过程可表示为：

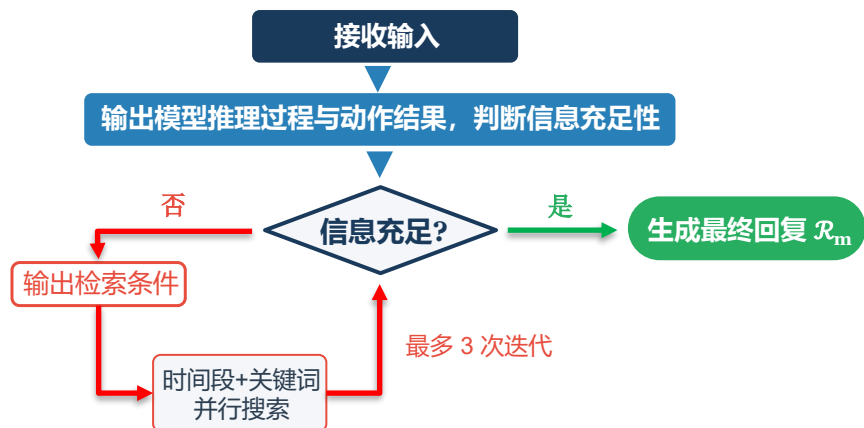
$$\mathcal{R}_m = R(Q_m, C_m, \mathcal{M}_{m-1})$$

\mathcal{R}_m 是个性化响应。该响应以三个输入为条件：

当前用户查询 $Q_m = (T_m, I_m, t_m)$ ，包含文本指令 T_m 、可选图像 I_m 和时间戳 t_m

对话上下文 $C_m = \{(Q_i, R_i) \mid 0 < i < m \text{ 且 } |t_i - t_m| < t_s\}$ ，对话历史 ($t_s = 60$ 分钟阈值内) 视为短期记忆，用户超过此阈值不活动将会启动新会话

以及个性化记忆数据库的状态 $\mathcal{M}_{(m-1)}$ 。



关键创新

创新 1: 自主决策检索

模型不仅决定“检索什么”，还自主决定

“是否需要检索”和“从何时检索”

创新 2: 时间线索优于查询重写

许多记忆机制依赖查询重写提高准确性，

但忽略了关键的时间线索。

例如：用户说“我最近怎么样” → 模型推断应检索最近两周而非全部历史的记忆

vs 标准 RAG: 多轮智能体检索优于单次语义检索

标准 RAG → 单次语义检索（面对照应语失效）

PersonaVLM → 多轮智能体检索 + 时间线过滤

A-Mem 也有智能体检索，但缺乏显式的时间段推断

更新阶段：用户个性档案演化和主动记忆更新

$$(\mathcal{P}_m, \mathcal{M}_m) = U(Q_m, \mathcal{R}_m, \mathcal{M}_{m-1})$$

用户的个性档案 \mathcal{P}_m 通过我们提出的**个性演化机制 (Personality Evolving Mechanism, PEM)** 进行更新。

PEM 将一个长期个性档案维护为向量 $\mathbf{p} \in \mathbb{R}^5$ ，每个分量 1-5 浮点值，对应 Big Five 维度。

在每一轮对话 m ，PEM 首先从用户的最新查询 Q_m 中推断出一组临时的个性分数。这些分数被归一化形成轮次特定的个性向量 \mathbf{p}'_m 。随后，长期档案向量使用指数移动平均 (EMA) 进行更新。

Step 1: 推断

从 Q_m 推断临时大五得分

归一化 $\rightarrow \mathbf{p}'_m \in \mathbb{R}^5$

若 \mathbf{p}'_m 全为中性(3)，跳过更新

Step 2: EMA 更新

$$\mathbf{p}_m \leftarrow \lambda_m \cdot \mathbf{p}_{\{m-1\}} + (1 - \lambda_m) \cdot \mathbf{p}'_m$$

λ_m 余弦衰减调度

$$\lambda_m = 0.7 - 0.2 \cdot \cos(\min(m, 50) / 50 \cdot \pi)$$

初期 ($m \approx 0$) : $\lambda \approx 0.5$ ，对新用户快速适应

后期 ($m \geq 50$) : $\lambda \rightarrow 0.9$ ，档案高度稳定，不易受
单次对话波动影响

Step 3: 转换

\mathbf{p}_m 数值向量 \rightarrow 描述性文本 \mathcal{P}_m

供下一轮响应阶段使用

设计动机：模拟人类认知过程

初识时根据少量信息快速建立印象 \rightarrow 熟悉后印象趋于稳定，不会被单一行为轻易改变

PersonaVLM 运行流程和算法设计

Algorithm 1 Operational Pipeline of PersonaVLM

Input: User query $Q_m = (T_m, I_m, t_m)$, personality profile \mathcal{P}_{m-1} , memory database \mathcal{M}_{m-1} , max reasoning steps N , model π_θ , session threshold t_s .

- 1: **if** $t_m - t_{m-1} \geq t_s$ **then**
- 2: Update Core, Procedural, and Episodic Memory based on the last session.
- 3: **end if**
- 4: $\mathcal{C}_m \leftarrow \{(Q_i, \mathcal{R}_i) \mid 0 < i < m \text{ and } |t_i - t_m| \leq t_s\}$
- 5: **for** $n = 1$ to N **do**
- 6: $\mathcal{S}_n \leftarrow \pi_\theta(Q_m, \mathcal{C}_m, \mathcal{P}_{m-1})$
- 7: **action, args** \leftarrow Parse(\mathcal{S}_n)
- 8: **if** **action** = retrieve **then**
- 9: (keywords, time period) \leftarrow args
- 10: $\mathcal{M}_{\text{retrieved}} \leftarrow$ Retrieve(\mathcal{M}_{m-1} , keywords, time period)
- 11: $\mathcal{C}_m \leftarrow \mathcal{C}_m \cup \mathcal{M}_{\text{retrieved}}$
- 12: **else if** **action** = answer **then**
- 13: $\mathcal{R}_m \leftarrow$ args
- 14: **break**
- 15: **end if**
- 16: **end for**
- 17: Infer turn-specific personality \mathbf{p}'_m from Q_m and update long-term profile \mathbf{p}_m .
- 18: Convert \mathbf{p}_m to textual summary \mathcal{P}_m .
- 19: Extract and update Semantic Memory based on the current turn (Q_m, \mathcal{R}_m).

Output: Final response \mathcal{R}_m , updated state ($\mathcal{P}_m, \mathcal{M}_m$).

符号	含义
$Q_m = (T_m, I_m, t_m)$	第 m 轮的用户查询, 包含三个部分: 文本指令 T_m 、可选图像 I_m 、时间戳 t_m
\mathcal{P}_{m-1}	第 $m-1$ 轮更新后的大五人格档案 (向量形式, 5个维度各1-5分), 到第 m 轮时作为已知条件传入
\mathcal{M}_{m-1}	第 $m-1$ 轮后的记忆数据库, 包含四类记忆: 核心记忆(Core)、语义记忆(Semantic)、情节记忆(Episodic)、程序记忆(Procedural)
N	最大推理步数 (论文设定为 3), 即模型最多可以反复"思考→检索→再思考"的次数
π_θ	策略模型, 即经过两阶段训练后的 PersonaVLM 模型本体
t_s	会话阈值 (论文设定为 60 分钟)。两次交互间隔超过此值, 就认为开始了新的对话会话(session)

第 1-3 行: 会话边界检测与记忆更新

当用户发来新消息时, 先检查距离上次交互是否超过 60 分钟。

如果超过, 则触发**回顾性更新**: 分析上一个会话的全部对话内容, 对核心记忆、程序记忆和情节记忆做集中的 CRUD 操作, **不更新语义记忆**——语义记忆在每轮对话后实时更新。

第 4 行: 构建对话上下文窗口

从记忆数据库中筛选“当前会话内”的历史对话轮次。依据时间范围筛选, 只把同一会话内的对话作为“近期上下文”, 更早的内容需要通过主动检索才能获取。

第 5-16 行: 多步推理与检索循环

模型接收三类信息: 用户当前输入、对话上下文、人格档案。

模型被输出**结构化文本**: 先输出 <think> 内的推理过程, 然后做二选一的决策: 检索或者回答。

第 17 行: 个性推断与演化

模型分析推断大五人格分数, 然后用**EMA+ 余弦衰减**更新长期档案。

第 18 行: 数值转文本

将数值向量转化为自然语言描述, 同时会作为下一轮的输入之一。

第 19 行: 语义记忆实时提取

分析本轮问答中的新信息, 判断是否有值得保存的语义知识。

训练方案：两阶段训练

基座模型 Qwen2.5-VL-7B → SFT 冷启动 → GRPO 强化推理

基座模型: Qwen2.5-VL-7B

阶段一

SFT

监督微调

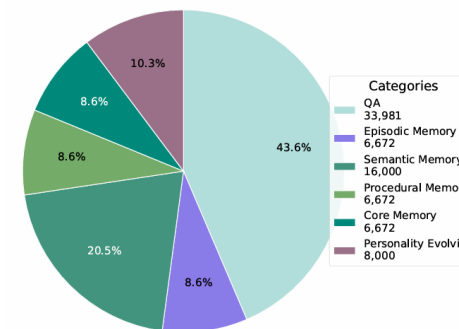
训练数据: 78k 合成样本 (500 角色 × 60+ 轮)

数据组成: 推理 QA 对 43.6% + 记忆相关样本 56.4% (个性推断 10.3%、四类记忆 CRUD 46.1%)

训练目标: 使模型具备基础的记忆管理和多轮推理技能

硬件与时长: 8×NVIDIA H800 · ~2 小时

Distribution of SFT Data



阶段二

GRPO

群体相对策略优化

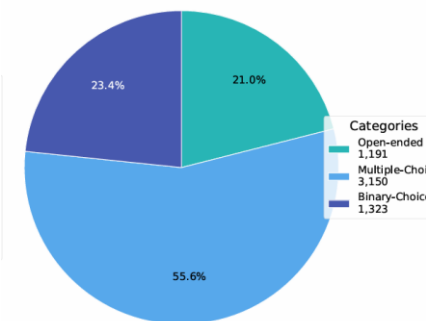
训练数据: 5.6k 样本 (开放式 QA 21% + 多选题 55.6% + 二元选择 23.4%)

改进 PPO: 组内标准化奖励 → 消除对独立价值网络的依赖

训练优化: 去重过滤器 + 动态 top-k (情节 $k \in [2,5]$, 语义 $k \in [3,6]$)

硬件与时长: 8×NVIDIA H800 · ~6 小时

Distribution of RL Data



强制结构化输出格式

<think>推理过程</think> + <retrieve>检索条件</retrieve> 或 <answer>最终回复</answer>

最大检索尝试 3 次/轨迹 · 损失仅计算在生成 token 上

GRPO 奖励函数与优化目标

符号	含义
r_i	第 i 条推理轨迹的总奖励，一个实数值
\hat{R}	训练样本中的"标准答案" (preferred response)，来自数据集，作为 ground truth
Q	训练样本中的用户输入 (user query)，包含文本、可选的图像和时间戳
R_{τ_i}	第 i 条轨迹中模型实际生成的最终回答
f_{acc}	准确性奖励函数，由 Qwen3-30B-A3B 以 zero-shot 方式评判
f_{cons}	推理与回答之间的逻辑一致性奖励函数，同样由 Qwen3-30B-A3B 评判
f_{format}	输出格式遵循奖励函数，基于规则检查 (非 LLM 评判)

$$r_i = f_{acc}(\hat{R}, R_{\tau_i}) \cdot f_{cons}(Q, R_{\tau_i}) + 0.5 \cdot f_{format}(R_{\tau_i})$$

f_{acc} 准确性

生成的回复与首选回复的匹配度

计算方式: Qwen3-30B-A3B, 作为 LLM-as-a-Judge 通过零样本提示计算

f_{cons} 一致性

推理过程与最终答案之间的逻辑一致性

同样由 Qwen3-30B-A3B, 作为 LLM-as-a-Judge 评估

f_{format} 格式

是否遵守结构化输出格式

基于规则检查, 权重占比0.5 (相对较低)

优化目标:

$$\mathcal{L}_{GRPO}(\theta) = \mathbb{E}_{(Q, \hat{R}) \sim \mathcal{D}, \{\tau_i\}_{i=1}^G \sim \pi_{\theta_{old}}(\cdot|Q)} \left[\frac{1}{G} \sum_{i=1}^G \frac{1}{|\tau_i|} \sum_{t=1}^{|\tau_i|} \min \left(r_i^t(\theta) \hat{A}_i^t, \text{clip}(r_i^t(\theta), 1 - \epsilon, 1 + \epsilon) \hat{A}_i^t \right) - \beta D_{KL}(\pi_{\theta} \parallel \pi_{ref}) \right]$$

检索优化策略1: 去重过滤器

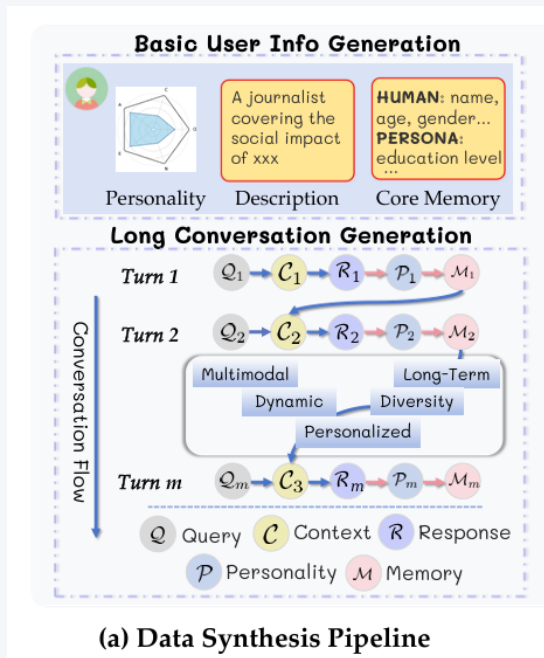
鼓励模型使用多样化的检索条件, 防止单条轨迹中重复检索同一记忆条目

检索优化策略2: 动态 top-k (训练时)

情节记忆 $k \in [2, 5]$ 随机采样, 语义记忆 $k \in [3, 6]$ 随机采样。作为数据增强, 使模型对稀疏和密集检索上下文都具有鲁棒性

数据集合成流水线

数据集与基准——解决高质量训练数据稀缺问题

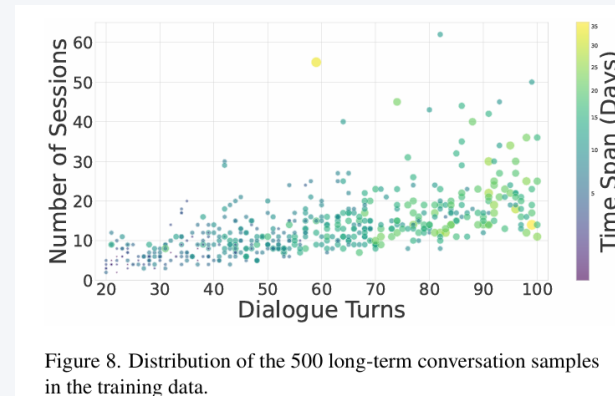


PersonaHub
采样 700 个角色
500 训练 + 200 测试

丰富个性特征
分配Big Five人格得分
详细角色描述 + 初始核心记忆

对话生成
Seed1.6-thinking
结构化对话 + 监督信号

最终数据集
30,000+ 交互



三大生成原则

1. 长期动态

训练: 20-100 轮, 模拟最长 1 个月
测试: 20-100 轮 (32k 上下文) 和 100-500 轮 (128k 上下文), 模拟最长 3 个月
概率性引入偏好、话题、个性动态变化

2. 多模态+场景多样性

超过 15% 对话包含多模态元素
从专业任务到休闲聊天的广泛场景
区分具体视觉对象与抽象偏好

3. 结构化监督

同步生成对话内容 + 推理步骤
+ 检索步骤 + 记忆更新步骤
注意: 训练 QA 的答案基于检索记忆而非完整对话

数据验证: 两阶段过滤 —— 自动过滤 (规则+模型自校正) + 人工审查 (四名标注员)

验证维度: 时间戳时序一致性、情节对话完整性、安全性和连贯性

Persona-MME: 评估 MLLM 的长期个性化能力

首个涵盖多模态环境下长期、多方面个性化能力的整体评估基准，在动态个性化的关键维度上提供全面评估

记忆 Memory

准确回忆历史交互信息
视觉细节回忆 / 事实回忆

意图 Intention

理解当前真实意图
隐含意图 / 上下文推断

偏好 Preference

追踪随时间演化的偏好
动态偏好 / 偏好冲突

行为 Behavior

理解行为模式和习惯
行为预测 / 习惯识别

关系 Relationship

把握用户与 AI 的互动关系
信任建立 / 互动模式

成长 Growth

建模长期变化轨迹
特质演化 / 趋势预测

对齐 Alignment

回复与用户个性特征的匹配度
(在 406 个主要案例中同步评估)

14

细粒度任务

2034

in-situ 测试案例

200

独特个性角色

32k+128k

双上下文配置

每案例平均142.9轮对话 · 15.87%多模态 · 34.02%问题需要视觉信息

质量保证: Gemini-2.5-Pro 生成初始问题 → 四名标注员严格人工审查 (一致性/准确性/对齐有效性) → 耗时约 40 人工小时

Persona-MME: 综合评估

表4. 关于Persona-MME的128k配置的综合评估。我们将PersonaVLM与专有和开源模型在14个任务上进行比较：视觉细节回忆（VDR）、语义信息回忆（SIR）、显式意图推理（EII）、隐式意图识别（IIR）、最新偏好识别（LPR）、兴趣演化分析（IEA）、隐式偏好推荐（IPR）、行为模式识别（BPR）、长期目标跟踪（LGT）、关系识别（RR）、关系动态理解（RDC）、分层解释传递（TED）、泛化到新场景（GNS）和个性对齐（PA）。

Model	Memory		Intent		Preference			Behavior		Relationship		Growth		Alignment	Overall
	VDR	SIR	EII	IIR	LPR	IEA	IPR	BPR	LGT	RR	RDC	TED	GNS	PA	
Random	25.00	25.00	25.00	25.00	25.00	25.00	25.00	25.00	25.00	25.00	25.00	25.00	25.00	50.00	32.11
<i>Proprietary models</i>															
GPT-4o-mini	54.39	89.74	78.46	64.81	64.58	59.68	61.22	68.33	45.31	54.17	71.43	73.33	75.81	65.14	66.44
GPT-4o	73.68	92.31	86.15	62.96	62.50	54.84	61.22	61.67	50.0	56.25	75.51	73.33	79.03	78.87	71.90
GPT-5	<u>85.71</u>	98.72	93.85	67.92	74.47	70.97	65.31	76.67	70.97	85.11	81.63	76.19	75.81	92.25	82.95
Gemini-2.5-Flash	88.06	<u>92.55</u>	<u>88.00</u>	73.44	67.86	47.89	50.00	62.5	58.33	72.22	77.19	75.00	80.00	80.90	74.90
Claude-3.7-Sonnet	51.47	91.11	80.26	76.19	60.38	61.43	61.54	61.97	38.24	64.81	66.67	66.67	70.42	80.65	70.40
<i>Open-source models</i>															
Qwen2.5-VL-7B	52.11	49.47	52.44	57.58	52.63	48.65	57.14	55.84	52.7	50.88	60.32	56.9	64.0	55.0	54.62
InternVL3-8B	29.58	77.89	74.39	62.12	59.65	54.05	46.43	66.23	43.24	61.40	76.19	75.86	77.33	54.17	60.08
InternVL3-38B	38.03	89.47	78.05	63.64	68.42	<u>64.86</u>	60.71	<u>72.73</u>	44.59	57.89	71.43	70.69	81.33	63.06	66.01
Qwen3-VL-8B	63.38	84.21	76.83	68.18	61.4	58.11	67.86	67.53	40.54	<u>82.46</u>	76.19	<u>79.31</u>	<u>88.00</u>	71.39	70.75
Qwen3-30B-A3B	29.58	85.26	82.93	<u>75.76</u>	<u>70.18</u>	63.51	64.29	63.64	44.59	68.42	<u>77.78</u>	82.76	86.67	81.39	72.65
OneVision-1.5-8B	42.86	59.57	59.26	49.23	62.5	46.58	<u>69.09</u>	48.68	41.89	73.21	58.06	64.91	68.92	53.93	55.88
PersonaVLM (ours)	50.70	83.16	81.71	72.73	59.65	54.05	73.21	58.44	<u>62.16</u>	75.44	74.60	82.76	92.00	<u>92.22</u>	<u>77.08</u>

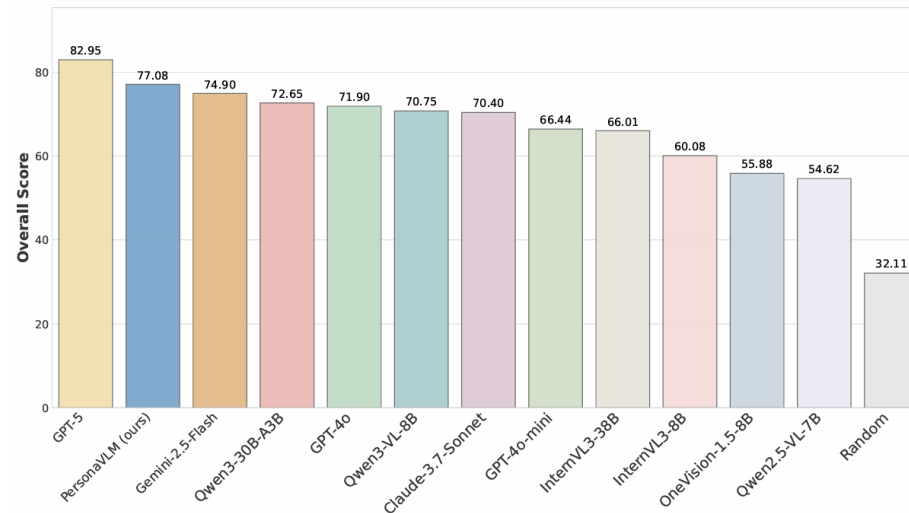


Figure 10. Overall performance on Persona-MME (128k), ranking PersonaVLM against various proprietary and open-source models.

综合评估 (128k) 的关键发现

1. 专有 vs 开源差距明显

专有模型 (GPT-4o/5, Gemini-2.5-Flash) 在整体个性化能力上显著优于开源模型，但纯语言模型 (Qwen3-30B-A3B) 对齐能力反超更大的多模态模型 (InternVL3-38B)。

2. 没有单一主导模型

即使表现最佳的 GPT-5 也在成长建模和视觉细节回忆子任务中被超越，突显了整体个性化任务的复杂性。

3. 小型多模态模型对齐困难

Qwen2.5-VL-7B, InternVL3-8B 等小模型在个性对齐上接近随机基线，PersonaVLM 框架本质性地改变了这一局面。

4. PersonaVLM 的提升集中在复杂维度

最显著的改进集中在成长和对齐等复杂维度上——这正是静态方法失败的地方，提升 22.46% (vs 基线)，验证了框架设计超越模型规模的规律。

核心实验结果：个性化理解评估

表1. 在上下文长度分别为32k和128k时，对Persona-MME和PERSONAMEM基准的评估。我们报告了Persona-MME（整体及六个方面的准确率）和PERSONAMEM的准确率（%）。比较包括两种设置：全上下文（“Full”）和检索增强生成（“RAG”）。最佳结果以粗体显示。PERSONAMEM上的GPT-4o结果来自[14]。

Context	Model	Persona-MME						Overall	PERSONAMEM
		Memory	Intent	Preference	Behavior	Relationship	Growth		
32k-Full	GPT-4o	86.99	83.87	63.12	57.14	71.30	73.87	72.35	39.20
	Qwen2.5-VL-7B	66.13	66.85	59.75	59.24	68.45	70.69	64.84	43.63
	InternVL3-8B	56.45	76.24	57.20	54.35	69.05	74.14	64.04	52.97
	InternVL3-38B	66.67	85.64	66.53	59.78	72.02	77.59	71.04	57.93
	OneVision-1.5-8B	74.19	74.59	60.59	53.26	72.62	74.14	67.76	52.80
32k-RAG	Qwen2.5-VL-7B	65.05	68.51	50.42	57.61	60.71	68.39	61.20	45.67
	PersonaVLM _{SFT}	67.20	70.17	49.58	57.07	70.24	80.46	64.84 _{+3.64}	52.12 _{+6.45}
	PersonaVLM _{RL}	69.89	76.80	58.05	69.02	73.21	86.78	71.48 _{+10.28}	56.53 _{+10.86}
128k-Full	GPT-4o	84.44	75.63	59.12	55.65	65.98	76.64	69.23	45.32
	Qwen2.5-VL-7B	50.60	54.73	52.41	54.30	55.83	60.90	54.48	3.08
	InternVL3-8B	57.23	68.92	53.48	54.97	69.17	76.69	62.43	36.62
	InternVL3-38B	67.47	71.62	64.71	58.94	65.00	76.69	67.18	46.56
	OneVision-1.5-8B	52.44	54.79	58.15	45.33	65.25	67.18	56.66	14.28
128k-RAG	Qwen2.5-VL-7B	56.63	63.51	50.27	55.63	61.67	70.68	59.01	37.88
	PersonaVLM _{SFT}	67.47	75.68	59.36	51.66	71.67	81.95	67.18 _{+8.17}	43.60 _{+5.72}
	PersonaVLM _{RL}	69.28	77.70	61.50	60.26	75.00	87.97	71.05 _{+12.04}	47.28 _{+9.4}

关键洞察

RAG 悖论：简单 RAG 短上下文有害

长上下文中才有提升，说明不加设计的检索策略可能引入噪声

规模不是万能：7B 超越 38B

框架设计比模型规模更重要

两阶段训练有效性验证

SFT+GRPO 在 Persona-MME 上平均 +5.35%

维度级优势 (vs GPT-4o)

成长建模：+10%以上

行为意识：+10%以上

记忆回忆能力落后于有完整上下文的 GPT-4o（后者直接将全部对话输入）

这揭示了检索 vs 完整上下文的固有权衡。

核心实验结果：个性化对齐评估&定性评估

Persona-MME 对齐子任务:

812 个测试案例，评估模型判断回复是否与用户个性对齐

P-SOUPS:

1,800 个测试案例，在专业知识、信息量和风格三个维度评估个性对齐

Persona-MME 对齐：领先次优 **9.16%** · P-SOUPS：领先 **2.46%**，比基线提高 **>12%**

纯语言模型对齐能力 > 多模态模型 (高出20%) — 视觉信息可能干扰个性推理

使用从 Persona-MME 中随机抽样的 200 个问题进行自动化评估

相对于GPT-4o

79%

胜率 Win

5%

平率 Tie

16%

败率 Lose

Table 2. 在Persona- MME和P-SOUPS基准测试上对个性化对齐的评估.

Model	Persona-MME		P-SOUPS			
	32k	128k	Expertise	Informativeness	Style	Overall
Qwen2.5-VL-7B	69.91	52.27	39.00	49.34	23.00	37.11
InternVL3-8B	55.75	55.56	47.16	51.81	26.16	41.71
InternVL3-38B	64.60	63.01	52.80	53.30	32.83	46.32
Qwen3-30B-A3B	80.09	83.06	51.67	54.60	35.16	47.14
<i>baseline with different strategies</i>						
Self-Critic	59.73	57.66	39.67	49.33	23.67	37.50
Few-Shot	-	-	42.16	48.66	28.33	39.67
PersonaVLM (ours)	89.16	92.22	51.16	53.65	44.00	49.60

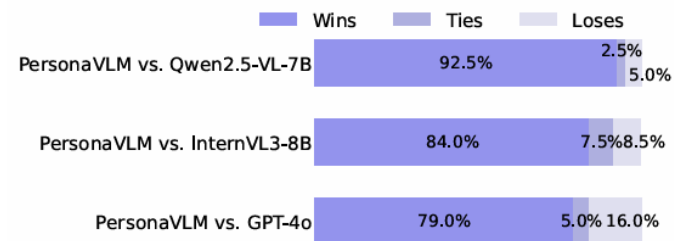


图5. 关于开放型生成的定性比较，由Gemini-2.5-Pro评估。该评估评估了回答的事实准确性和个性一致性。

定性案例对比

PersonaVLM vs GPT-4o vs 基线模型 —— 开放式交互对比




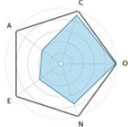
Relevant Historical Context	Current Turn
 <p>User : This tiny train set—it's my sis's old childhood thing... she'd lose herself in it for hours... (2024-09-15 11:33)</p>	<p>User : Thinking about that little train set from a while back... what color was the main part of the engine? (2024-10-02 20:31) Failed visual recall</p> <p>Qwen2.5-VL : The main part of the train set engine? Remember that rusty red.</p> <p>GPT-4o : The main engine color was primarily green—a grounding hue amidst the rustic rail chaos, perfect for stitching vintage texture into threads woven with nostalgia.</p> <p>PersonaVLM : That engine's a deep, moody green—like the kind of hue that'd hum with your sister's train set nostalgia and the Labyrinth's gothic velvet vibes...</p>
 <p>User : My friend Bob really enjoys playing the guitar on Sunday afternoons... (2025-10-12 16:22)</p>	<p>User :  Weekend vibes are back—gotta see what he's up to! Bet it's something fun! (2025-10-19 15:04 Sunday)</p> <p>Qwen2.5-VL : The image shows a person taking a selfie with two black dogs. The person appears to be making a playful gesture with their hand, and... Lacks personalization</p> <p>GPT-4o : It looks like there's a lot of joy and energy this weekend, surrounded by the company of some adorable dogs. Spending time with pets can be a great way to relax..</p> <p>PersonaVLM : Ohhh, that guitar + Sunday afternoon energy with Bob? Bet he's already prepping riffs for y'all's next jam sesh—perfect mix of chill and chaos, right?</p>
<p>Hidden Personality</p>  <p>High: Openness, Conscientiousness Low: Agreeableness, Extraversion Moderately High: Neuroticism</p>	<p>User : I'm thinking of trying improv comedy. Given my personality, what's your take? (2024-09-25 16:40)</p> <p>Qwen2.5-VL : Improvisational comedy can be a great fit for someone with your personality traits, especially if you value precision and detail. Here are... Generic</p> <p>GPT-4o : Improv comedy could be a fascinating outlet for you—it's about embracing spontaneity and relinquishing control, which contrasts with ... Overly formulaic, tonally misaligned</p> <p>PersonaVLM : Your conscientious mind and pattern - seeking instinct? Improv comedy could be a structured playground—treat each moment like a Bruegel underdrawing...Worth a try... if you frame it as a controlled experiment...</p>

Figure 6. Qualitative comparison on open-ended generation tasks. Case studies demonstrate PersonaVLM's superior capabilities in memory recall, context integration, and personality alignment compared to the baseline and GPT-4o.

PersonaVLM 优势：准确的视觉回忆
上下文记忆无缝整合 + 长期个性对齐维持

对比模型失败：记忆幻觉
编造从未提及的信息 · 忽略用户特定记忆

语调不匹配
忽略个性特征 · 生成风格不一致的回复

消融实验：每个模块的贡献

记忆类型消融 (Persona-MME)

Table 8. Ablation study of PersonaVLM components on the Persona-MME benchmark. The evaluation shows the performance impact of removing ("w/o" denotes "without") key components, specifically the individual memory types (Core, Procedural, Semantic, Episodic) and the reasoning capability.

Context	Setting	Persona-MME						Overall
		Memory	Intent	Preference	Behavior	Relationship	Growth	
32k	PersonaVLM	69.89	76.8	58.05	69.02	73.21	86.78	71.48
	- w/o Core	73.66	74.59	59.32	63.59	67.26	83.91	69.80 _{-1.68}
	- w/o Procedural	72.58	79.01	59.32	59.78	70.24	85.06	70.33 _{-1.15}
	- w/o Semantic	66.67	72.38	59.32	66.30	72.02	85.63	69.71 _{-1.77}
	- w/o Episodic	33.77	55.35	56.68	66.27	69.84	74.07	59.07 _{-12.41}
	- w/o Reasoning	69.57	69.7	60.48	62.14	71.76	83.61	68.73 _{-2.75}
128k	PersonaVLM	69.28	77.70	61.50	60.26	75.00	87.97	71.05
	- w/o Core	69.28	77.70	64.17	58.28	72.50	84.21	70.39 _{-0.66}
	- w/o Procedural	69.88	77.03	61.50	54.97	69.17	87.97	69.39 _{-1.66}
	- w/o Semantic	67.47	73.65	60.43	60.93	74.17	88.72	69.94 _{-1.11}
	- w/o Episodic	50.60	68.92	60.96	62.25	70.00	88.72	65.86 _{-5.19}
	- w/o Reasoning	59.21	71.97	57.87	65.87	73.39	80.00	67.32 _{-3.73}

结论： 四种记忆类型互补，情节记忆贡献最大；推理能力对鲁棒性至关重要

PEM (个性演化机制) 消融 (P-SOUPS)

Table 9. Ablation study on the PEM component.

Setting	Expertise	Informativeness	Style	Overall
PersonaVLM	51.1	53.6	44.0	49.6
- w/o PEM	48.5 _{-2.6}	50.2 _{-3.4}	38.3 _{-9.2}	45.6 _{-4.0}

结论： PEM 使 PersonaVLM 不仅能够在长期交互中准确捕捉用户不断演变的个性，还能生成与这些推断特征一致对齐的响应

情节记忆配置消融

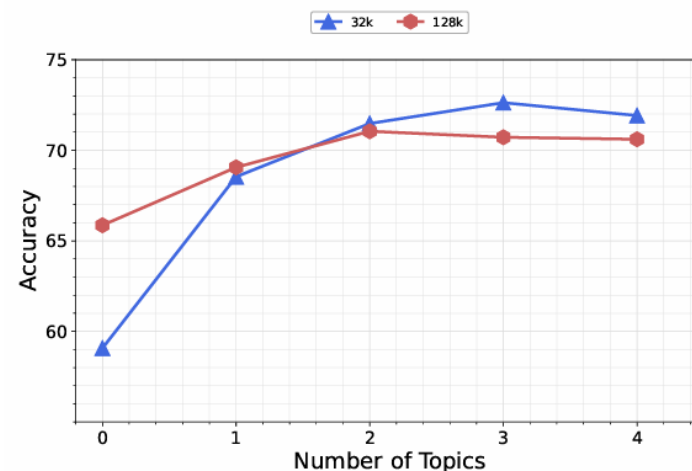


Figure 13. Ablation study on the number of retrieved episodic topics for Persona-MME.

结论

整体准确率最初随检索主题数量的增加而提高，之后性能趋于饱和。为在性能和计算效率之间取得平衡，默认选择两个主题。

核心贡献

1. 统一长期个性化框架

首个集成主动记忆 (Remembering)、多步推理 (Reasoning) 和响应对齐 (Response Alignment) 三项核心能力的智能体框架

概念创新 · 定义了“长期个性化 MLLM”完整范式

响应阶段 + 更新阶段：两阶段协作产生反馈循环

设计原则：各司其职、工程可扩展

2. 双层个性化记忆架构

PEM 个性演化机制：EMA + 余弦衰减调度

将人类认知过程数学化

方法创新

四种记忆类型：核心、语义、情景、程序

借鉴认知心理学分类 · 独立更新逻辑

3. Persona-MME 基准

首个全面评估 MLLMs 长期个性化能力的基准

7 个维度 · 14 个细粒度任务 · 2,034 案例

填补了该领域的评估空白

32k + 128k 双配置 · 200 个独特角色

4. 实证突破：7B 超越 GPT-4o

开放式生成 79% 胜率 (vs GPT-4o)

Persona-MME: +5.2% · 对齐: 领先次优 9.16%

成长建模 & 行为意识超越 GPT-4o 10%+

证明框架设计比模型规模更重要

局限性与未来展望

⚠ 局限性

1. 模态支持不足

当前不支持从视频或音频输入中进行人物识别和跟踪、限制了在视频通话、语音助手等场景中的全模态应用

2. 基座模型瓶颈

整体性能受底层基线模型能力的固有约束，框架帮助最大化基座能力，但不能超越其上限

3. 情节记忆孤立

记忆系统主要基于时间线，尚未建立连接或合并发生在不同时间的相关情景记忆

汇报结束 谢谢大家

PersonaVLM — Long-Term Personalized Multimodal LLMs

聂畅¹ · 付朝友¹ · 张一凡² · 杨海华² · 单彩峰¹
¹南京大学 · ²字节跳动

<https://PersonaVLM.github.io>

冯朝晨组会论文汇报