



湖南师范大学

大模型与Agentic AI赋能的6G通信

6G Communications Empowered by Large AI Models and Agentic AI

江沸菠

湖南师范大学 信息科学与工程学院

邮箱: jiangfb@hunnu.edu.cn



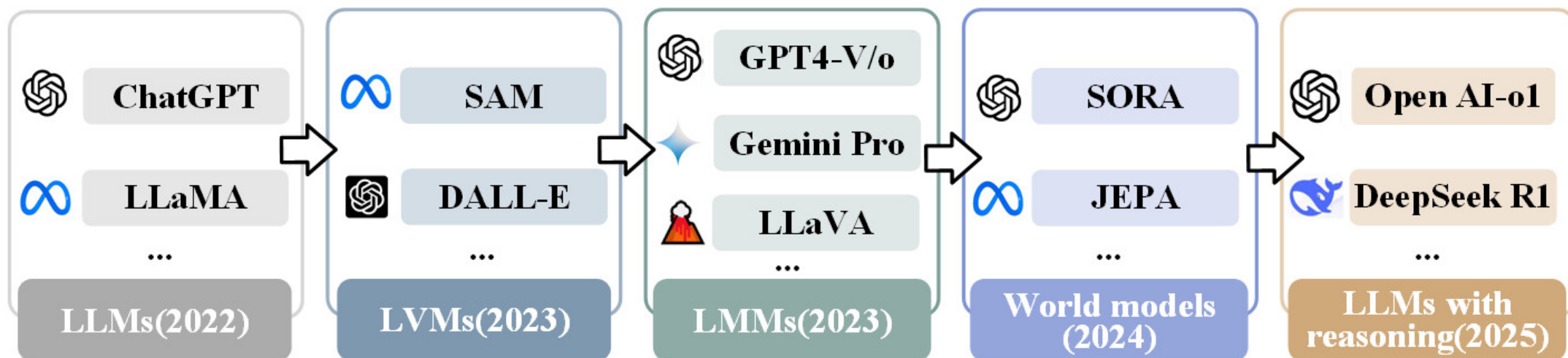
报告目录

- 1 大模型与Agentic AI基础
- 2 多模态通信基础大模型CommGPT
- 3 通信Agentic AI系统CommLLM
- 4 大模型赋能语义通信
- 5 挑战与机遇



1.1 大模型概述

- 大模型 (Large AI Model, LAM) 指包含**数千亿(或更多)**参数的神经网络模型
- 大模型学习海量的数据和任务, 具有显著的**鲁棒性和泛化性**
- 大模型支持**零/少样本学习**和**推理**, 具有惊人的理解力和创造力
- 大模型的出现和发展是人类迈向**通用人工智能**的重要阶段



Feibo Jiang, et al. "A comprehensive survey of large AI models for future communications: Foundations, applications and challenges." *arXiv preprint arXiv:2505.03556* (2025).



1.2 大模型的思维系统

快系统思维系统

反应迅速、无意识且自动化，易受经验偏差影响，在复杂问题上**容易出错**（例如GPT-4就是模拟人类的快思维系统）

慢思维系统

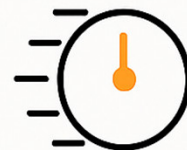
依赖有意识的**多步推理**，虽速度较慢，却能提供**更高的可靠性**，适用于下棋、复杂数学运算等需要深度思考的场景（例如DeepSeek-R1就是模拟人类的慢思维系统）

总结

快系统处理简单、常见任务，**效率高**；慢系统处理复杂、需精确推理的任务，确保结果的**准确性**和**深度**，两者协作使得大模型能够同时具备**高效性**与**精确性**

快系统

快速



无意识



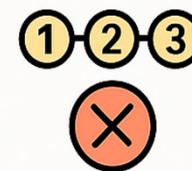
自动化



日常决策



易出错



慢系统

缓慢



有意识



需付出努力



复杂决策

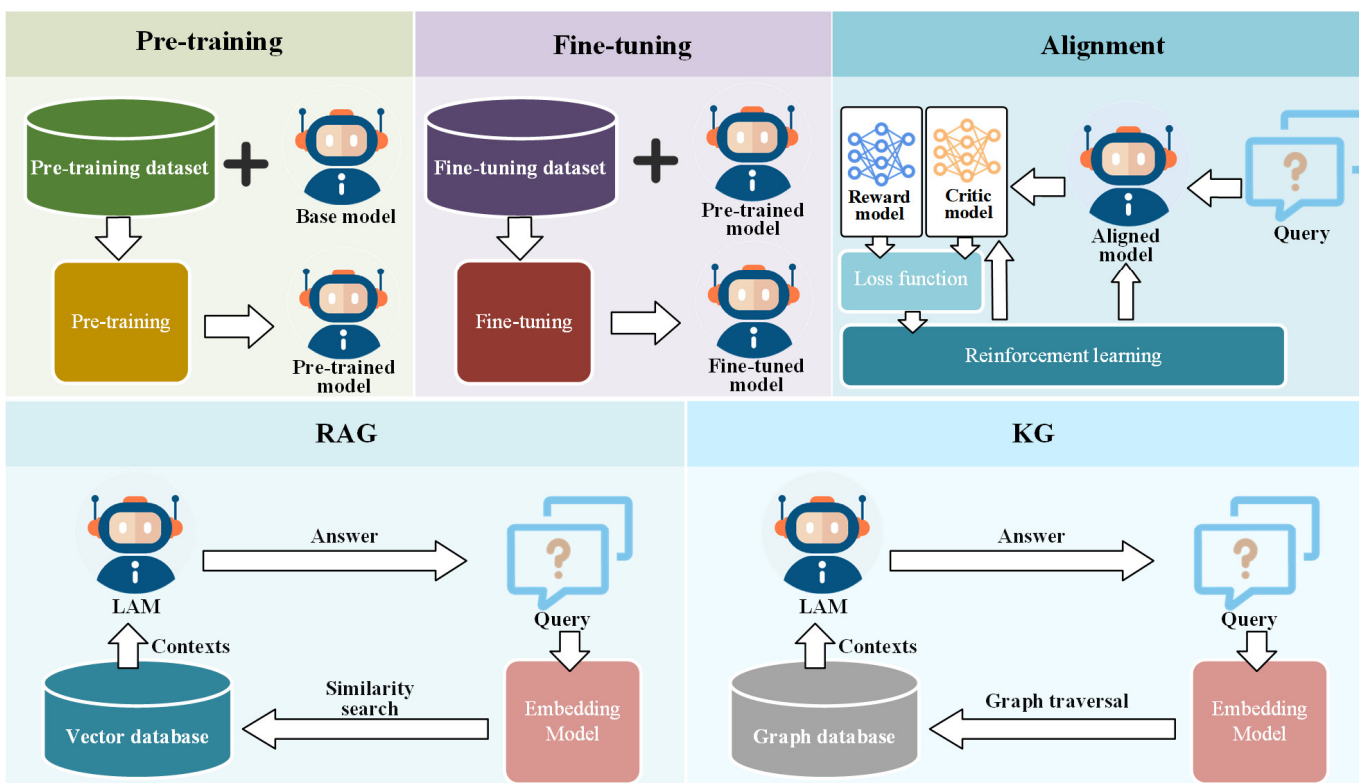


高可靠性



1.3 从大模型到Agentic AI

大模型在通信中的发展经历了以下阶段：



大模型 + 通信方向的提示工程

上下文学习，零/少样本学习，思维链

通信领域大模型

继续预训练，指令微调，偏好对齐

基于通信知识增强的大模型

检索增强生成，知识图谱

面向通信的Agentic AI

自主决策、主动学习和适应环境

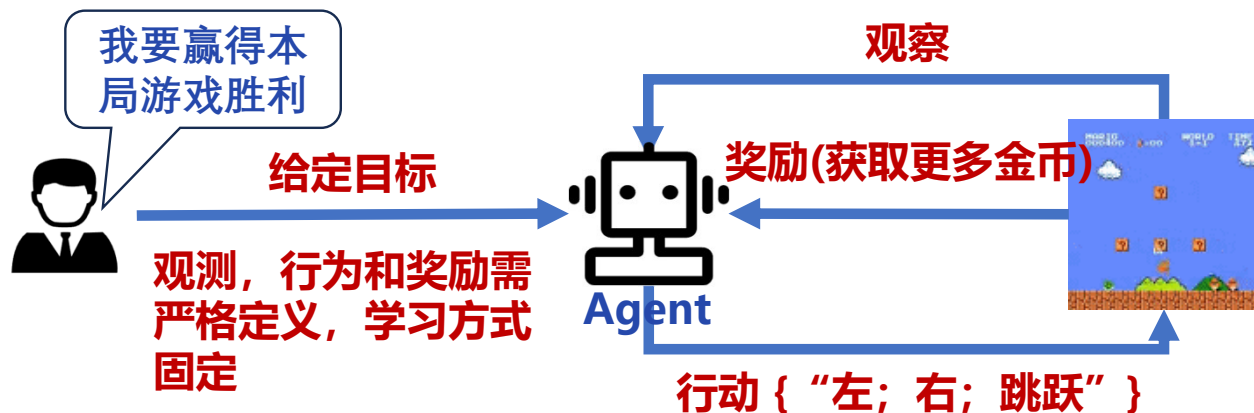


1.4 Agentic AI概述

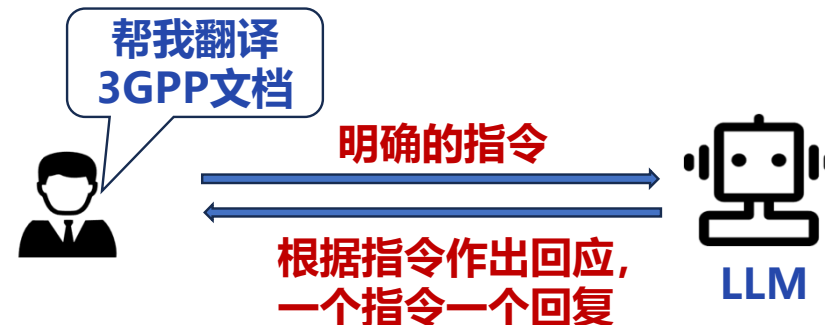
定义

Agentic AI 是一种由**多个智能体**协同构成的系统，具备持久记忆、自主计划与目标驱动能力，能够在**复杂、动态环境**中持续学习和适应，标志着从被动响应任务向**主动理解、规划和执行**任务的智能跃升。

AI Agent



大模型



Agentic AI



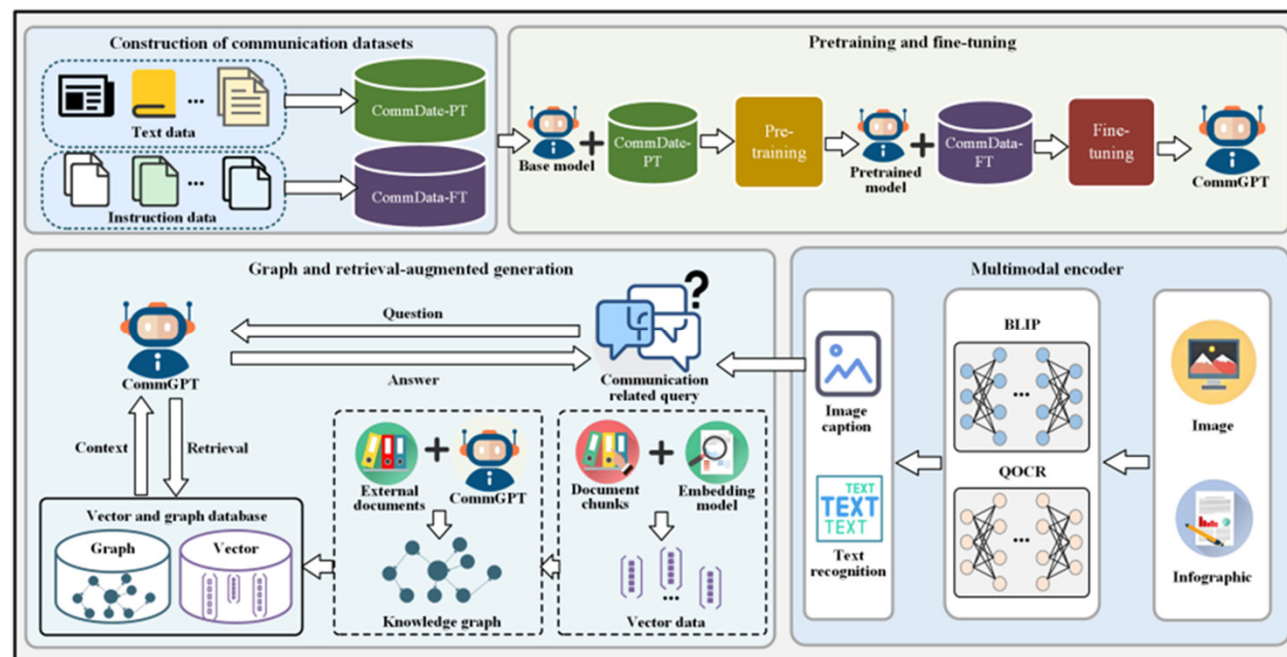


将LLM应用于通信领域面临三大挑战：1) **通信领域数据不足**；2) **输入模态受限**；3) **知识检索困难**

面临挑战

主要创新

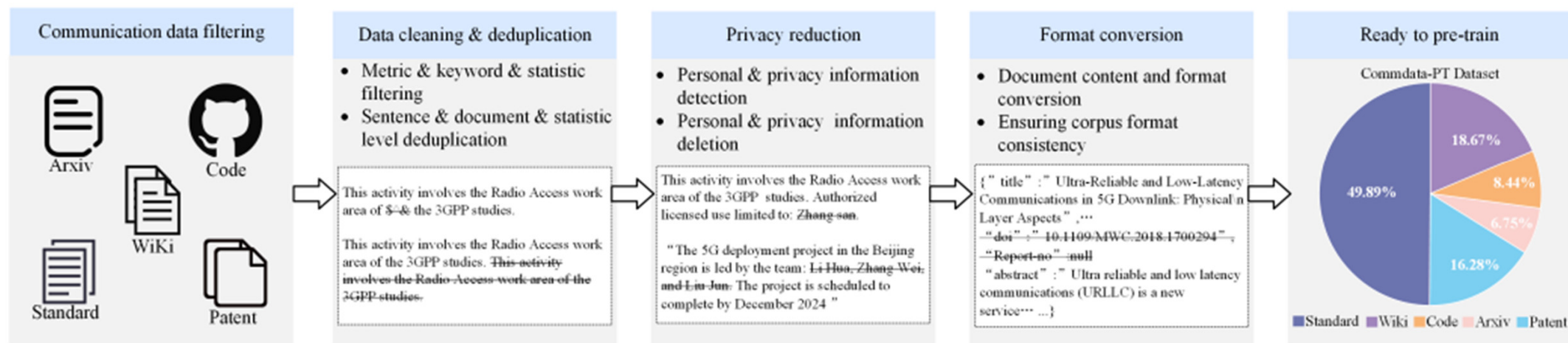
- **高质量通信数据集**：构建了面向通信领域的高质量专用数据集CommData，包含预训练数据集与微调数据集
- **多模态数据理解**：构建了多模态编码器，结合如BLIP和QOCR等多模态嵌入模型，实现对图像、表格等多种模态数据的解析
- **图与检索增强生成**：融合KG与RAG机制，构建了图知识与向量知识的高效耦合流程





数据预处理

- **CommData-PT**: 整合 3GPP/IEEE 标准、专利、ArXiv论文、开源代码与WiKi百科数据, 去噪去重后构建预训练语料
- **CommData-FT**: 基于预训练语料, 利用 LLaMA 3-8B-Instruct 自动生成高质量问答指令对, 用于基础模型的指令微调
- **数据处理流程**: 1)通信数据筛选 (ArXiv、WiKi、专利、代码等); 2)数据清洗与去重; 3)隐私信息删除; 4)格式转换 (统一结构)

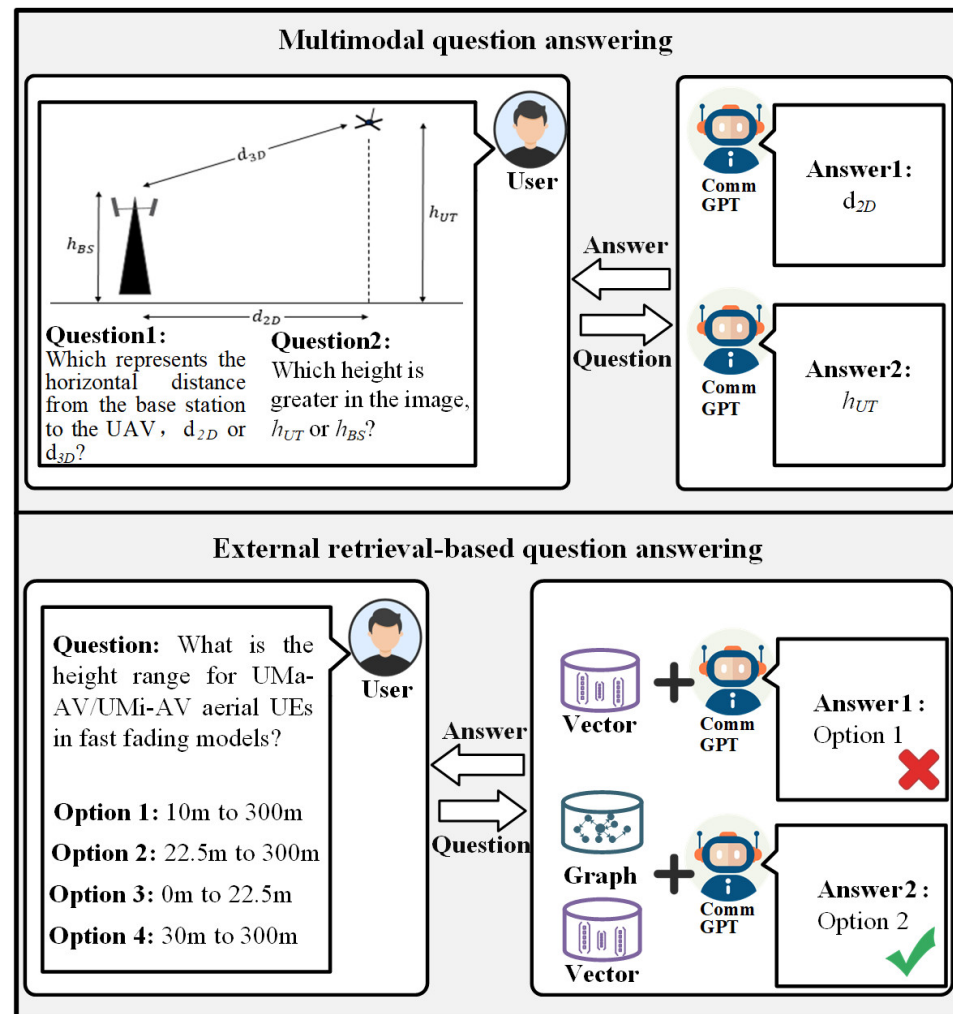
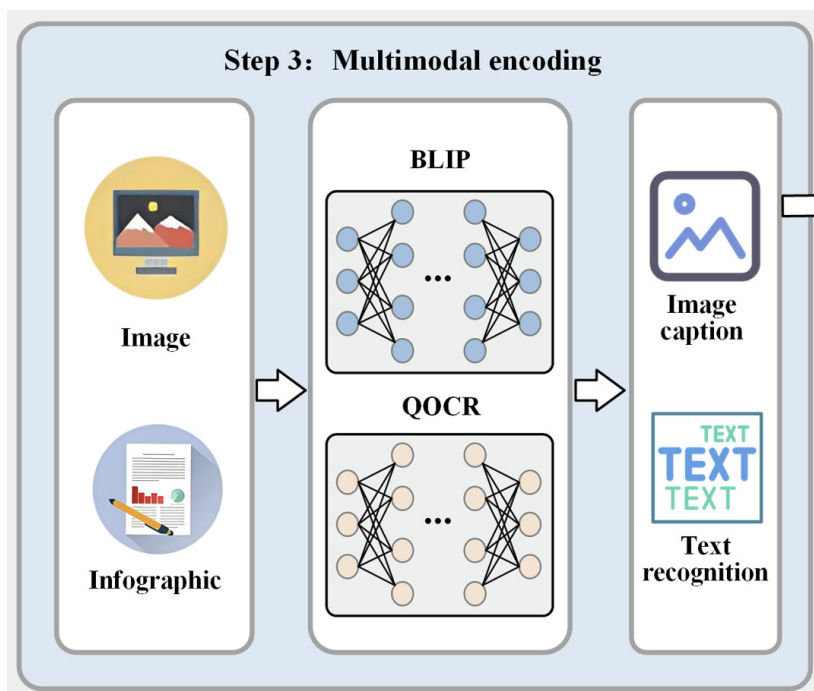




多模态输入

为通信文档中的多模态数据提供全面支持

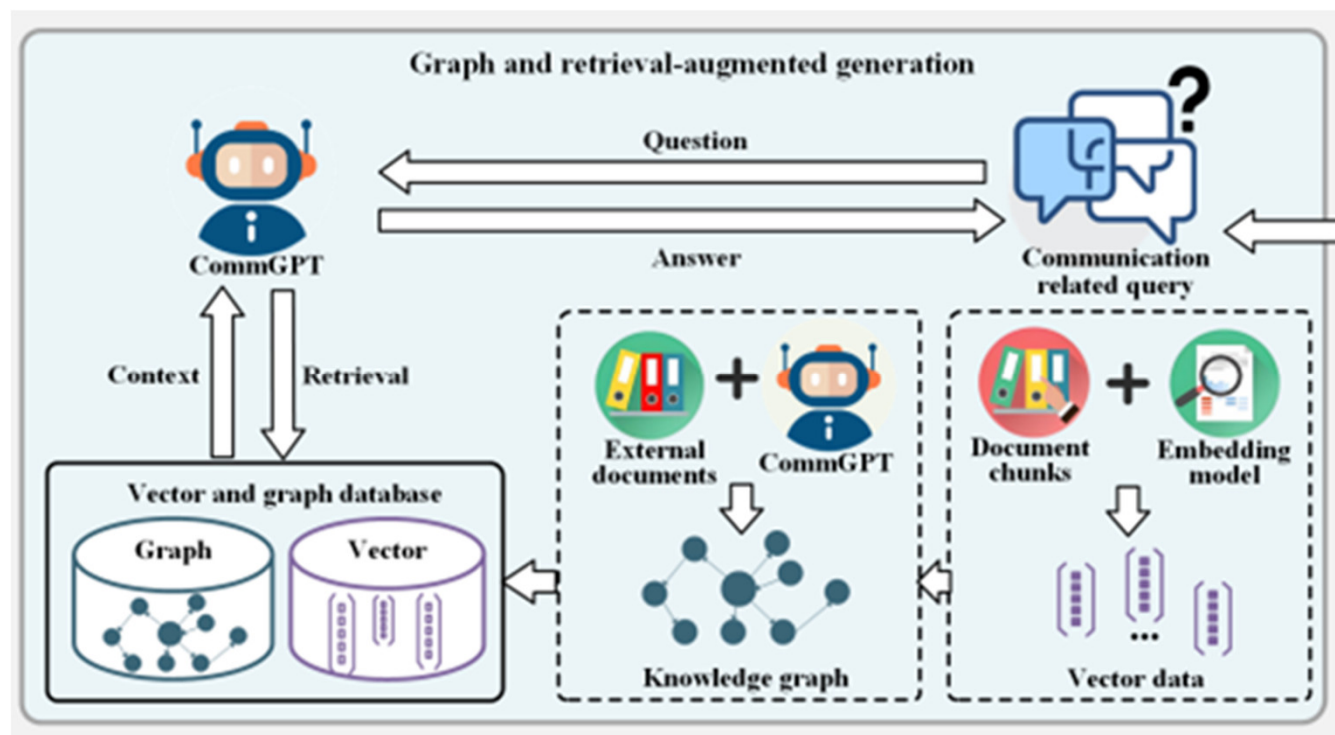
- **BLIP 编码器**: 从图像中提取高级语义特征
- **QOCR 编码器**: 提取信息图表中的低级文本信息

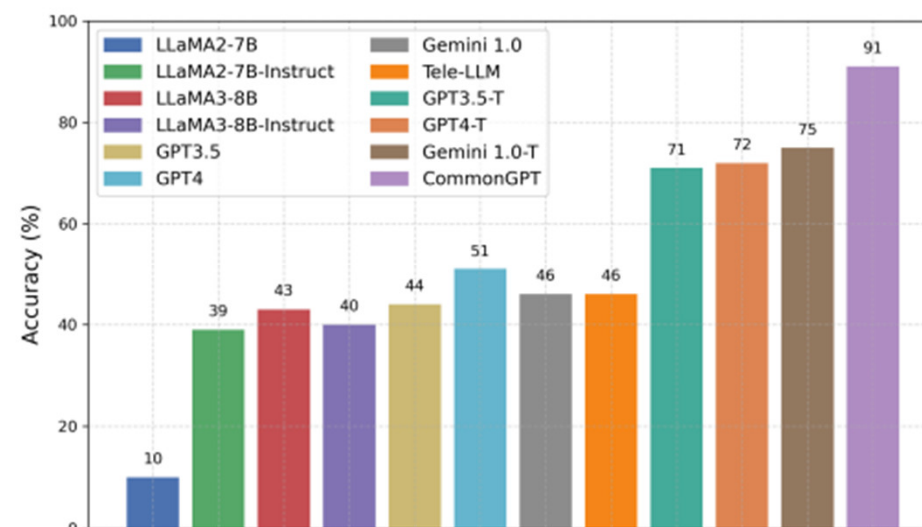
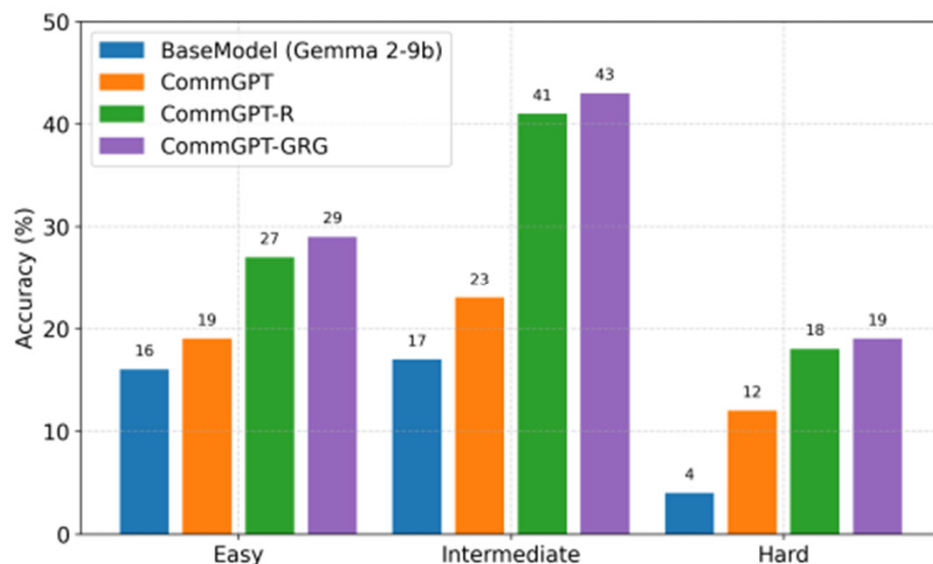




图与向量检索增强

- **向量数据库构建**: 通过文档切块 和 向量嵌入模型 构建向量数据库
- **图谱构建**: 结合外部文档与 CommGPT构建结构化通信知识图谱
- **图谱检索**: 例如知识图谱检索问答中的关键实体和相关实体子图(全局知识)
- **向量检索**: 根据实体子图检索向量数据库获得上下文知识(局部知识)
- **上下文反馈**: CommGPT根据用户问题和检索上下文内容生成精准回答





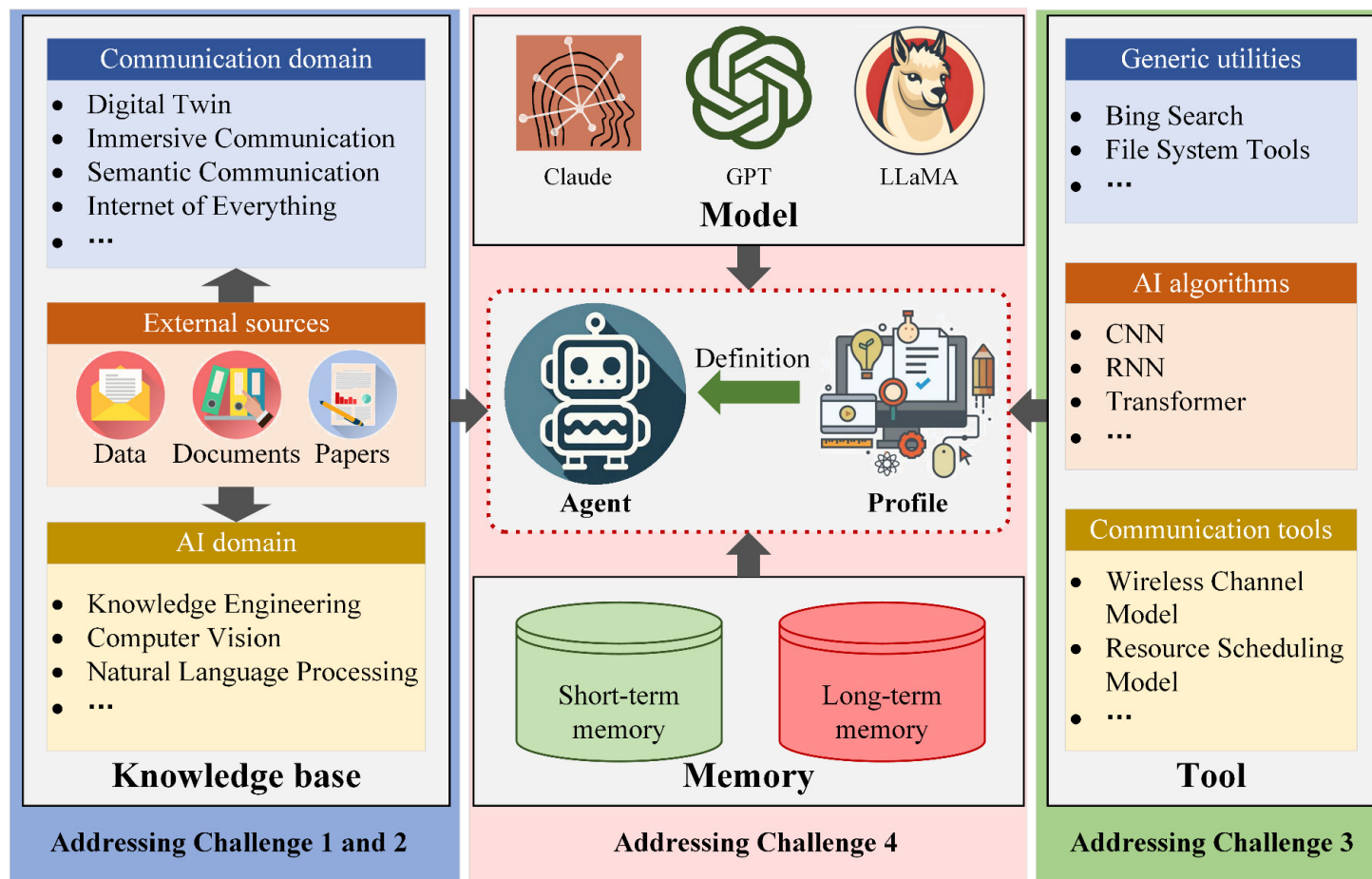
CommGPT 凭借图检索与向量检索的协同，在3GPP_TR数据集上取得最高准确率，进一步验证了其在通信领域的领先性。



3 通信Agentic AI系统CommLLM

Agentic AI 系统的组成

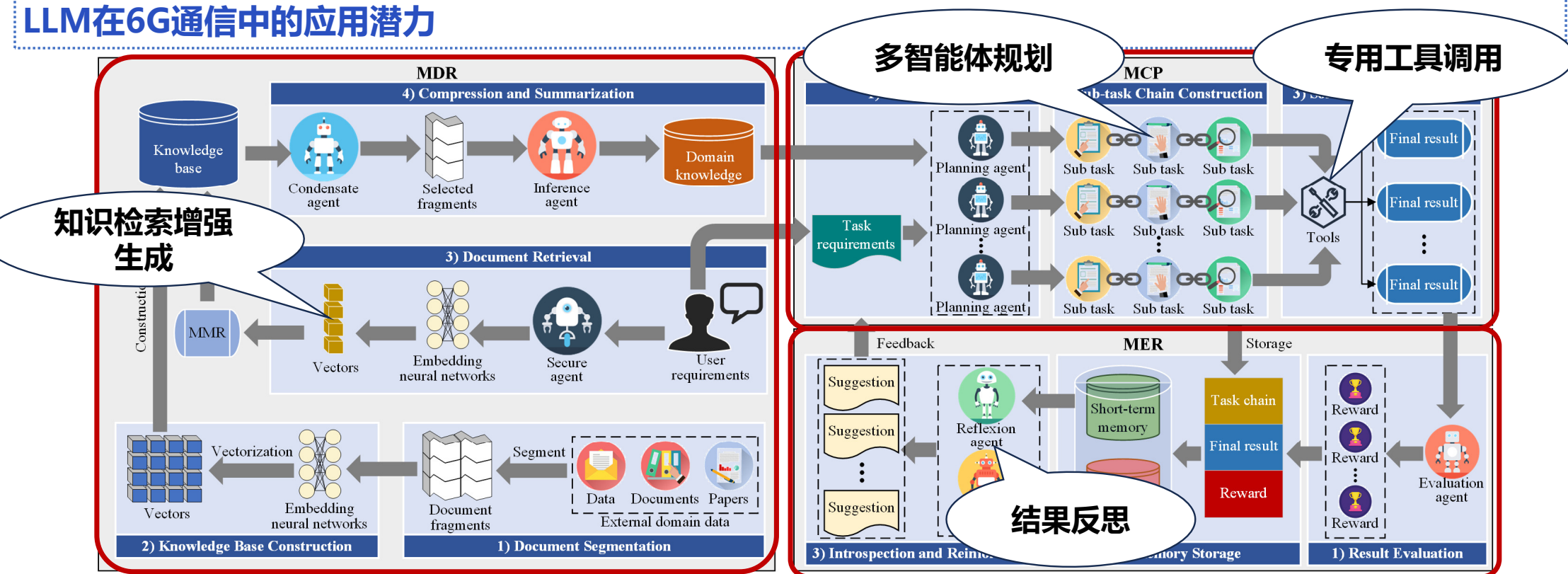
- **知识库**: 汇集了来自通信领域与AI领域的数据、文献与资料
- **模型与智能体**: 使用大模型结合用户画像定义智能体行为
- **记忆体**: 对智能体的对话和思维过程进行存储, 用于补充上下文或进行反思
- **工具**: 提供通用工具、AI及通信工具, 支持智能体执行复杂专业任务





LLM在6G中的挑战与机遇

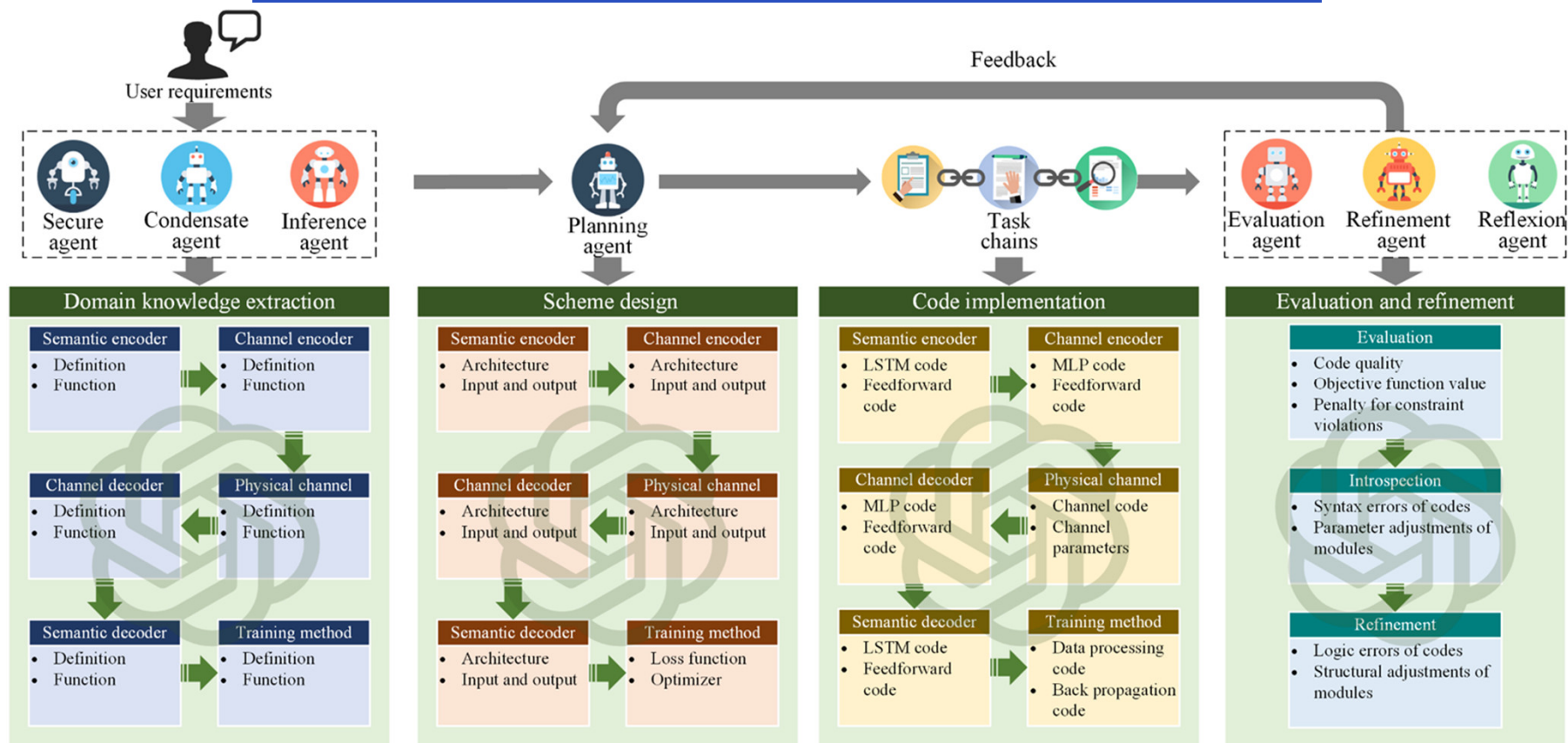
在6G中直接应用原生的LLM会遇到各种挑战，例如缺乏专业的通信数据和知识，模型逻辑推理、评估和优化能力有限。将LLM与智能系统中的检索、规划、工具、记忆、评估和反思能力相结合，可以大大增强LLM在6G通信中的应用潜力

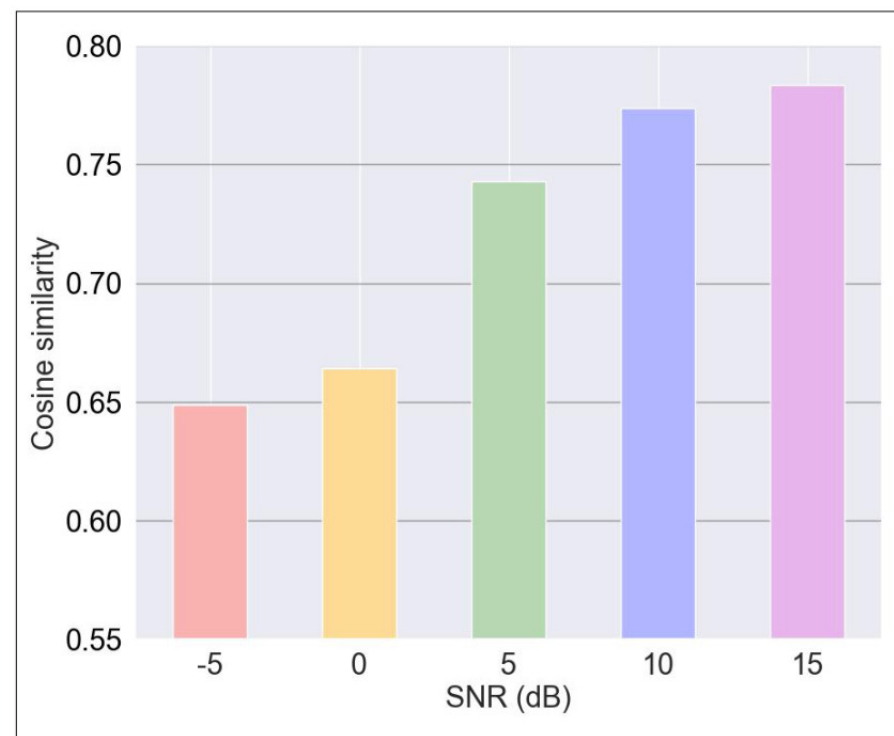
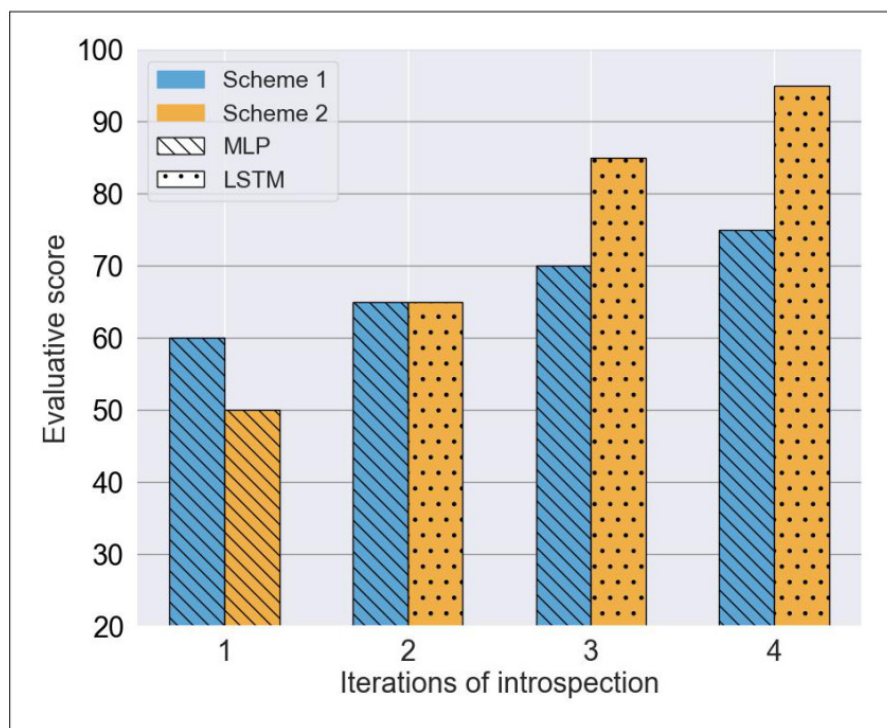




3 通信Agentic AI系统CommLLM

基于CommLLM设计语义通信系统的案例



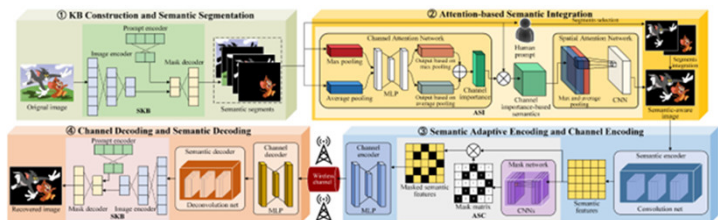


从实验结果来看，CommLLM能够通过反思自动选择更好的语义编码器架构，从而提升系统的性能

4.1 大模型赋能的语义通信

大模型赋能图像语义通信

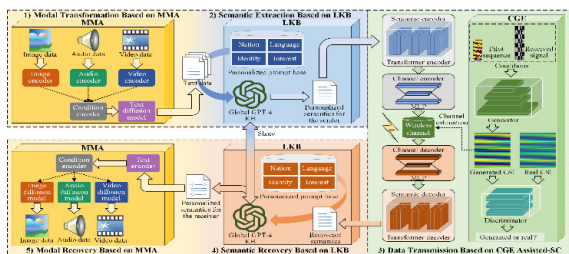
- 引入SAM大模型进行语义分割，实现**像素级语义压缩**
- 引入遮罩网络，实现自适应**特征级语义压缩**
- 与传统语义通信系统相比，**传输数据量减少60%**



Feibo Jiang, Yubo Peng, Li Dong, Kezhi Wang, Kun Yang, Cunhua Pan, Xiaohu You, "Large AI Model-Based Semantic Communications," in *IEEE Wireless Communications*, vol. 31, no. 3, pp. 68-75, June 2024, doi: 10.1109/MWC.001.2300346.

大模型赋能多模态语义通信

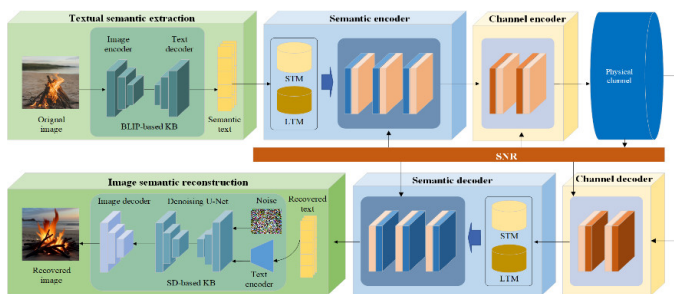
- 利用多模态大模型（如CoDi）进行**多模态语义对齐**
- 利用大语言模型（如GPT）进行**个性化语义提取**
- 相比单模态语义通信系统，**数据压缩率提高了约76%**



Feibo Jiang, Li Dong, Yubo Peng, Kezhi Wang, Kun Yang, Cunhua Pan, and Xiaohu You, "Large AI Model Empowered Multimodal Semantic Communications," in *IEEE Communications Magazine*, vol. 63, no. 1, pp. 76-82, January 2025, doi: 10.1109/MCOM.001.2300575.

大模型赋能动态环境下的语义通信

- 利用**视觉语言模型BLIP**和**Stable Diffusion**进行**模态转换**
- 通过**SNR注意力**和**长短期记忆**实现鲁棒的语义通信
- 在 CIFAR、BIRDS 等多个常用数据集上均能保持**图像语义一致性**



F. Jiang, C. Tang, L. Dong, K. Wang, K. Yang and C. Pan, "Visual Language Model-Based Cross-Modal Semantic Communication Systems," in *IEEE Transactions on Wireless Communications*, vol. 24, no. 5, pp. 3937-3948, May 2025



4.3 大语言模型驱动的人脸语义通信

面临的挑战

- **带宽利用率低**：像素级视频编码效率低下，导致带宽资源浪费
- **语义模糊**：信息表达不清晰，易引发语义误解
- **用户体验下降**：有损压缩与传输误差导致视频质量劣化，影响观看体验

主要创新

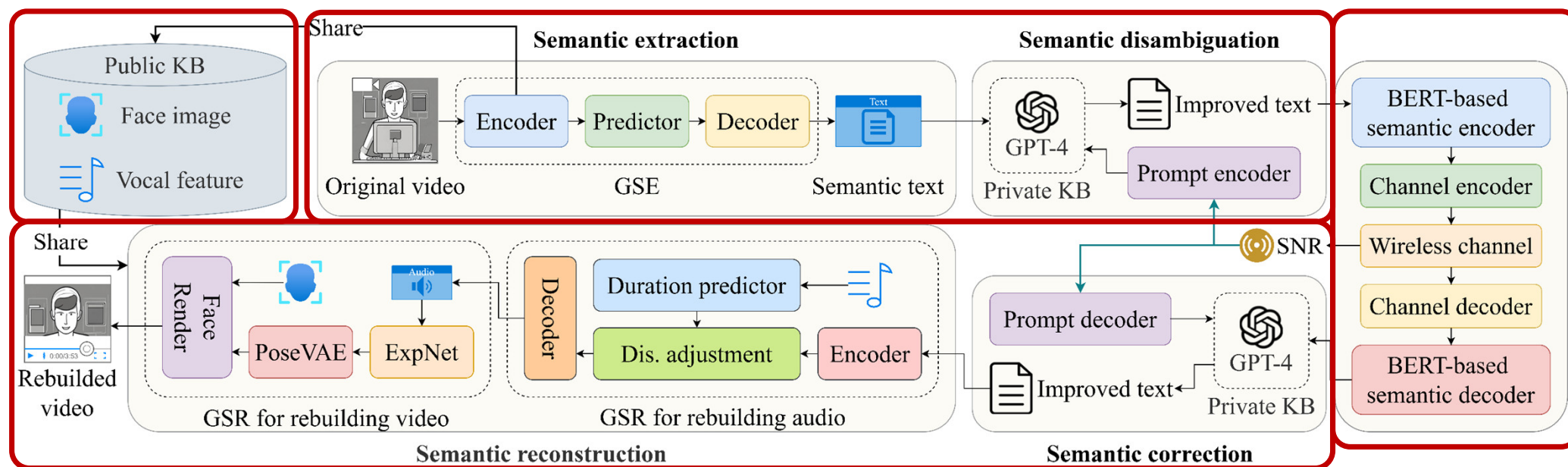
- **生成式语义提取器**：利用生成式人工智能模型将说话人脸视频转换为高密度文本信息
- **基于大语言模型的知识库**：在通信双方对语义进行消歧与纠错，确保语义传输的准确性与连贯性
- **生成式语义重建器**：在接收端采用生成式人工智能模型重建高质量的人脸说话视频，保障用户体验



4.3 大语言模型驱动的人脸语义通信

系统组件

- **生成式语义提取器**: 将人脸说话视频转换为高语义密度的文本表示
- **私有知识库**: 发送端通过大语言模型对文本语义进行消歧和纠错, 提升语义准确性
- **公共知识库**: 存储用户相关的个性化数据, 如人脸图像和语音特征, 用于视频重建
- **语义编码器与信道编码器**: 对语义数据进行压缩和编码, 保障传输效率
- **生成式语义重建**: 接收端重建音频, 并生成高质量的人脸说话视频, 提升用户体验



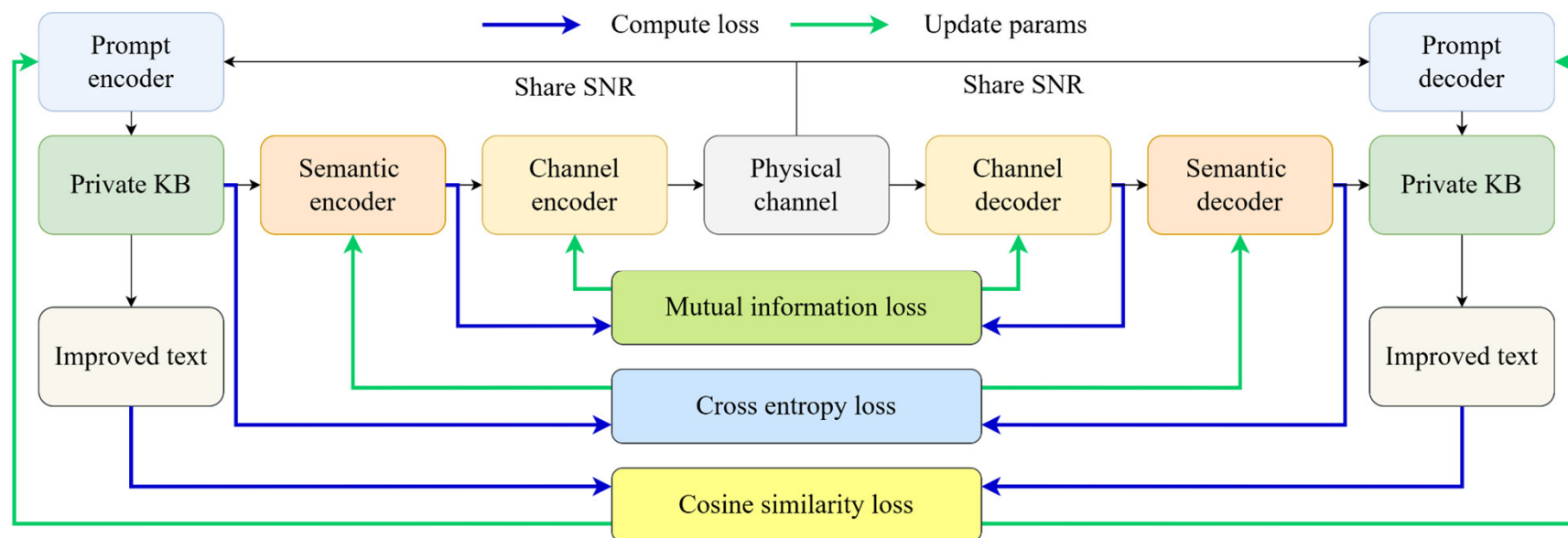


4.3 大语言模型驱动的人脸语义通信

结合知识库的语义信道联合编码

大语言模型体积过大，无法参与信道联合编码→ 采用提示词（Prompt）控制大语言模型输出

- **联合编码模块：**将提示词编码、语义编码和信道编码整合为统一框架
- **损失函数：**信道编码器与解码器：互信息（Mutual Information）；语义编码器与解码器：交叉熵（Cross Entropy）；提示词编码器与解码器：余弦相似度（Cosine Similarity）





提示编码器示例

编码器结构：模糊推理系统接收实时的信噪比数值作为输入，信噪比被映射至三个模糊类别之一：高（High）、中（Medium）或低（Low）

提示编码规则：高信噪比对应生成全文长度且语义详尽的文本；中等信噪比对应生成适度压缩的文本；低信噪比对应生成高度压缩但语义完整的文本

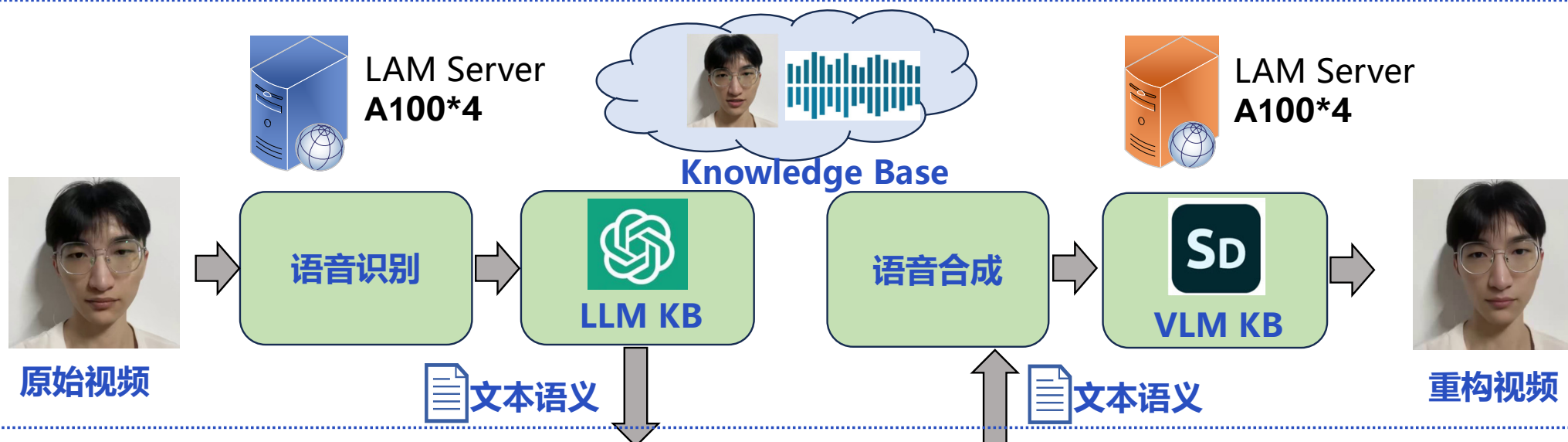
Input text: A young child, with a beaming smile, eagerly slides down the slide.	
Fuzzy rules	Output text after knowledge base encoding
If SNR is low, then length = 70%-80%	A young child, smiling, slides down the slide.
If SNR is middle, then length = 80%-90%	A young child, with a smile, eagerly slides down the slide.
If SNR is high, then length = 100%	A young child, with a beaming smile, eagerly slides down the slide.

The length in fuzzy rules means the output length of LLM when proceeding knowledge base encoding.



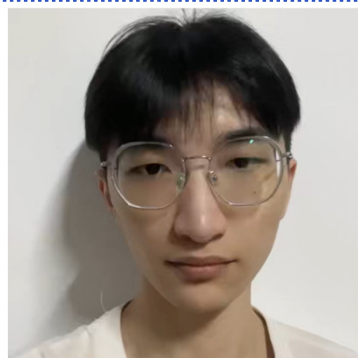
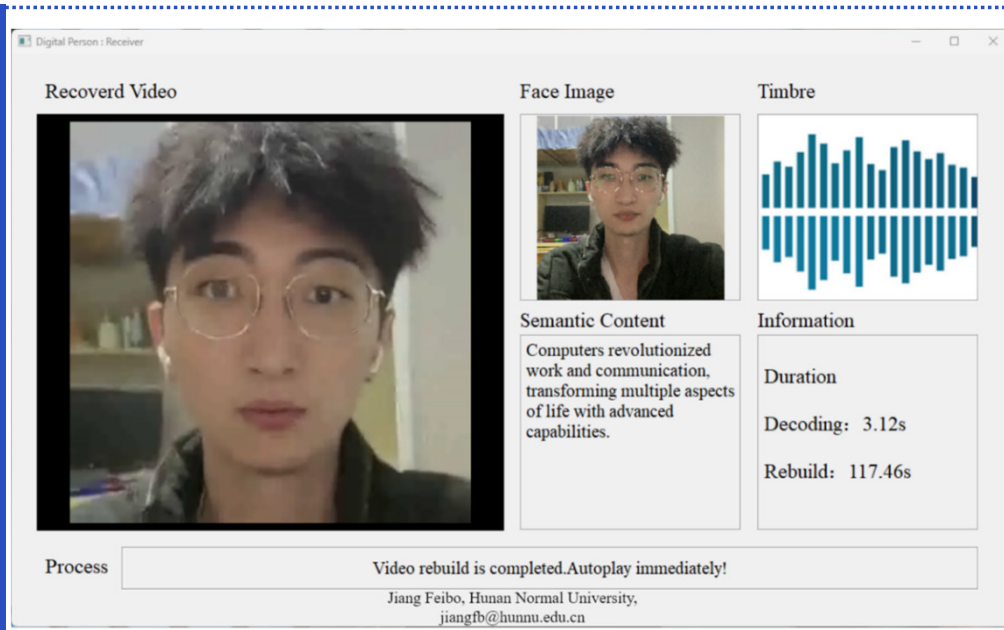
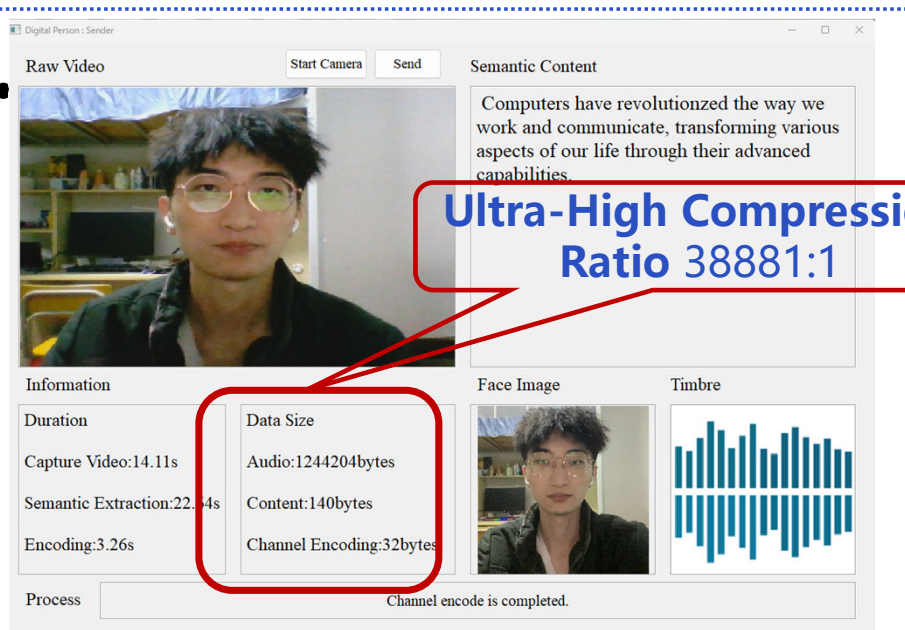
4.4 语义通信系统原型机

语义生成与重构

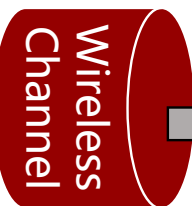
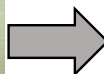
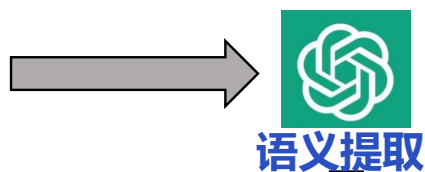




4.4 语义通信系统原型机



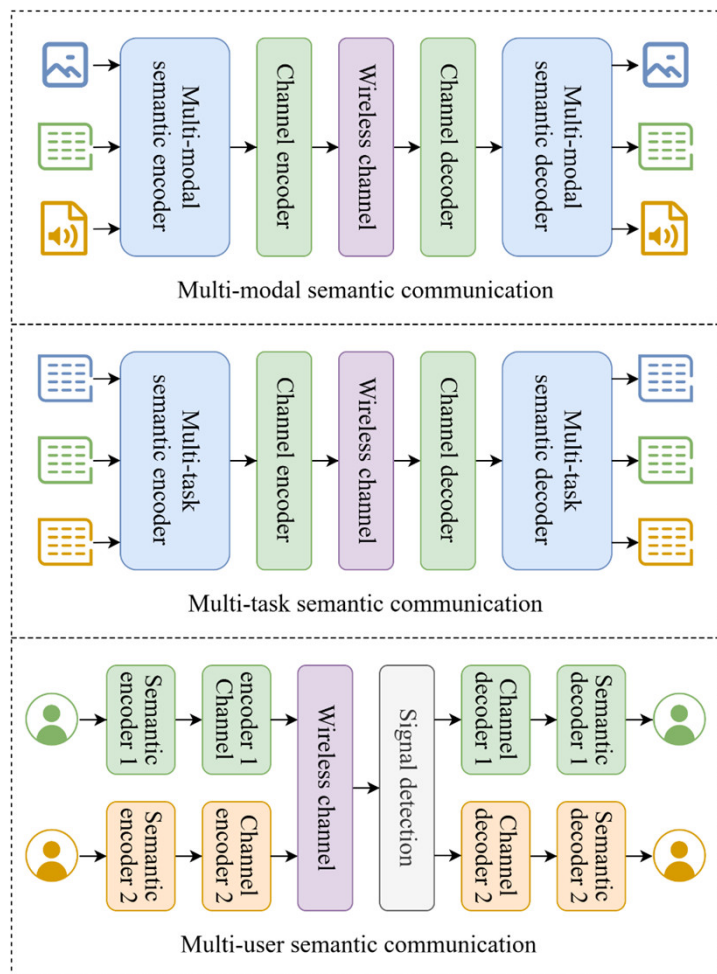
原始视频



重构视频



湖南师范大学 4.5 面向多模态、多任务与多用户的语义通信



面临的挑战

- **跨模态对齐困难:** 各模态间语义不一致
- **多任务遗忘:** 任务相互干扰, 导致知识丢失
- **多用户冗余传输:** 共享语义被重复发送

主要创新

- **多模态投影对齐:** 有效对齐多模态特征
- **多任务指令微调:** 促进泛化能力, 减少遗忘
- **多用户语义共享:** 向所有用户广播共享语义

Jiang F, Tu S, Dong L, et al. M4SC: An MLLM-based Multi-modal, Multi-task and Multi-user Semantic Communication System, in *IEEE Wireless Communications*, Accepted.



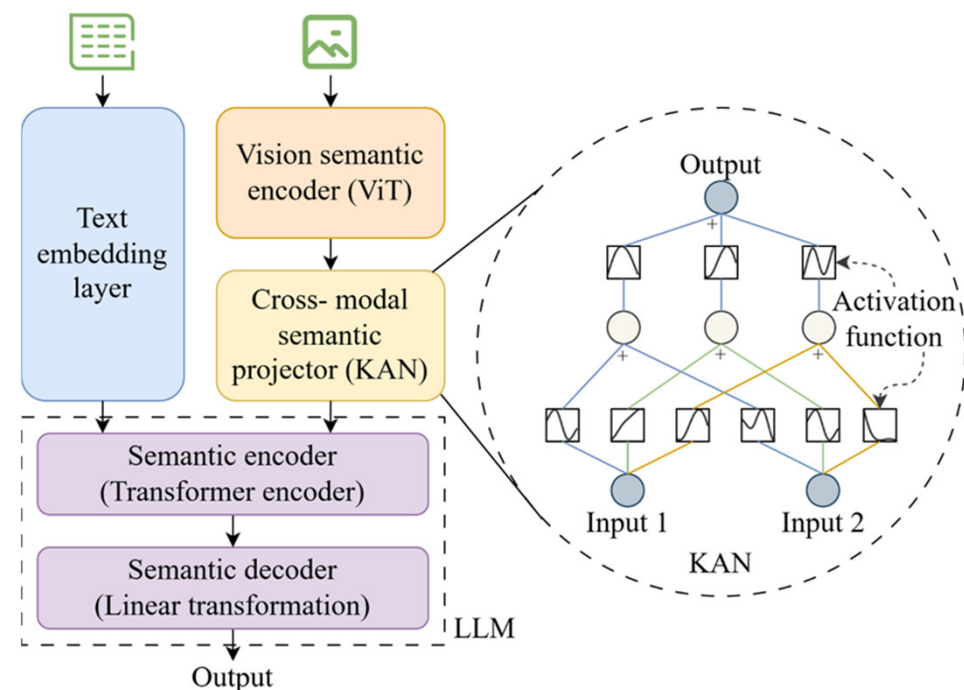
湖南师范大学 4.5 面向多模态、多任务与多用户的语义通信

MLLM的优势

- **高精度多模态对齐**: 通过跨模态投影, 将图像与文本对齐至共享语义空间
- **鲁棒的多任务泛化**: 结合上下文学习和思维链, 支持零样本/少样本学习
- **精准的语义识别**: 精准识别识别共享与差异化语义, 降低冗余传输

定制的MLLM结构

- **视觉语义编码器**: 采用 Siglip-ViT 从图像中提取语义特征
- **跨模态语义投影器**: 利用KAN对视觉与文本语义进行对齐
- **基于LLM的语义编码器**: 通过Transformer解码器提取图文特征
- **基于LLM的语义解码器**: 使用Linear+Softmax层解码生成输出 (Gemma2 2B)

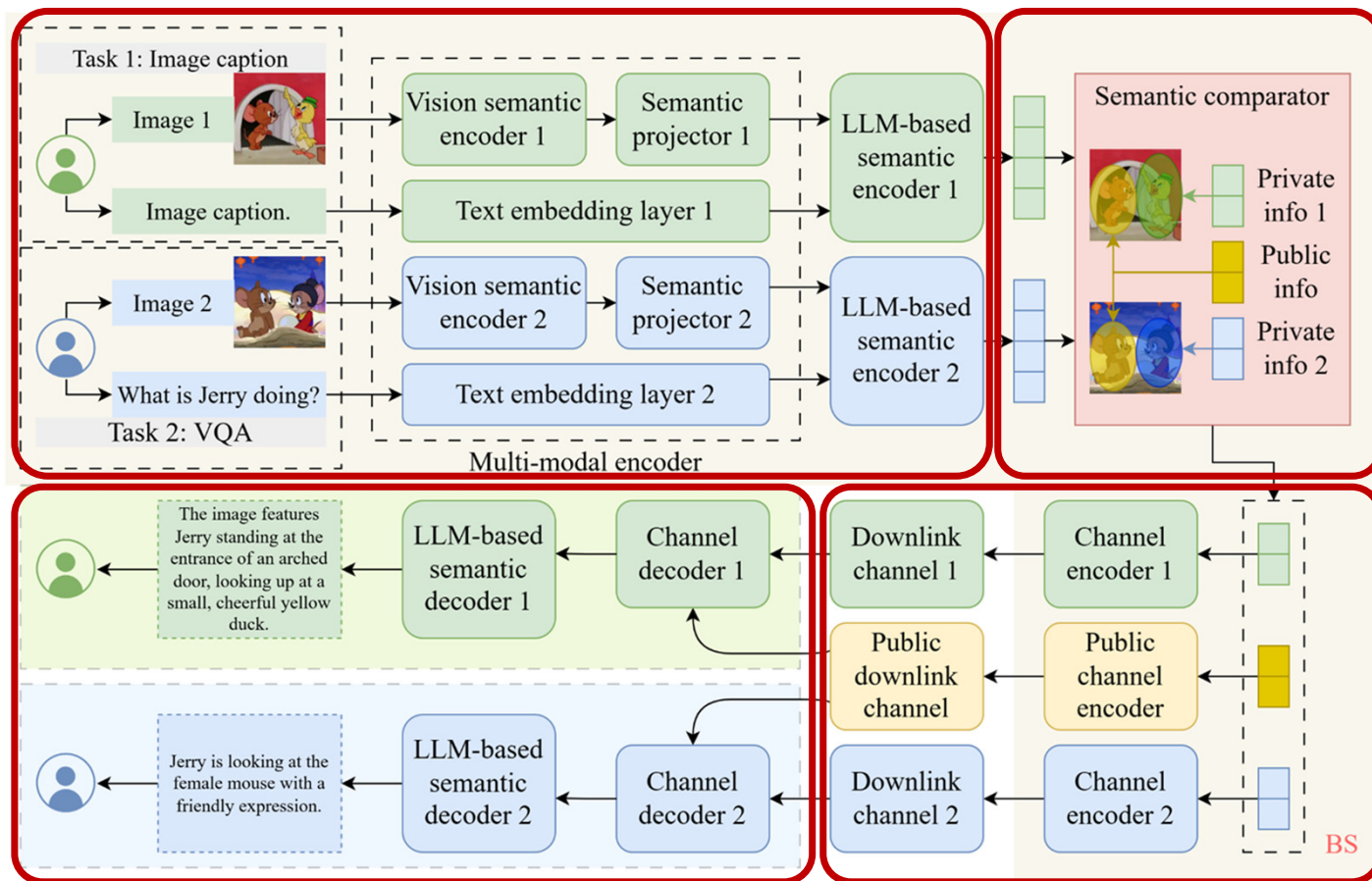


Jiang F, Tu S, Dong L, et al. M4SC: An MLLM-based Multi-modal, Multi-task and Multi-user Semantic Communication System, in *IEEE Wireless Communications*, Accepted.



系统工作流程

- 场景描述:**
基站向两个用户发送数据:
用户 1 → 图像描述
用户 2 → 视觉问答 (VQA)
- 基站编码过程:**
图像 + 文本 → 多模态编码器 → LLM编码器
- 语义对比与分发:**
公共语义: “Jerry” → 通过公共信道发送
私有语义: “黄色小鸭” (用户1)、
“雌性老鼠” (用户2) → 通过各自的独立信道发送
- 用户端解码:**
重新组合并解码 → 各自独立完成任
务



Jiang F, Tu S, Dong L, et al. M4SC: An MLLM-based Multi-modal, Multi-task and Multi-user Semantic Communication System, in *IEEE Wireless Communications*, Accepted.



湖南师范大学 4.5 面向多模态、多任务与多用户的语义通信

三阶段训练流程

阶段一：多模态对齐训练

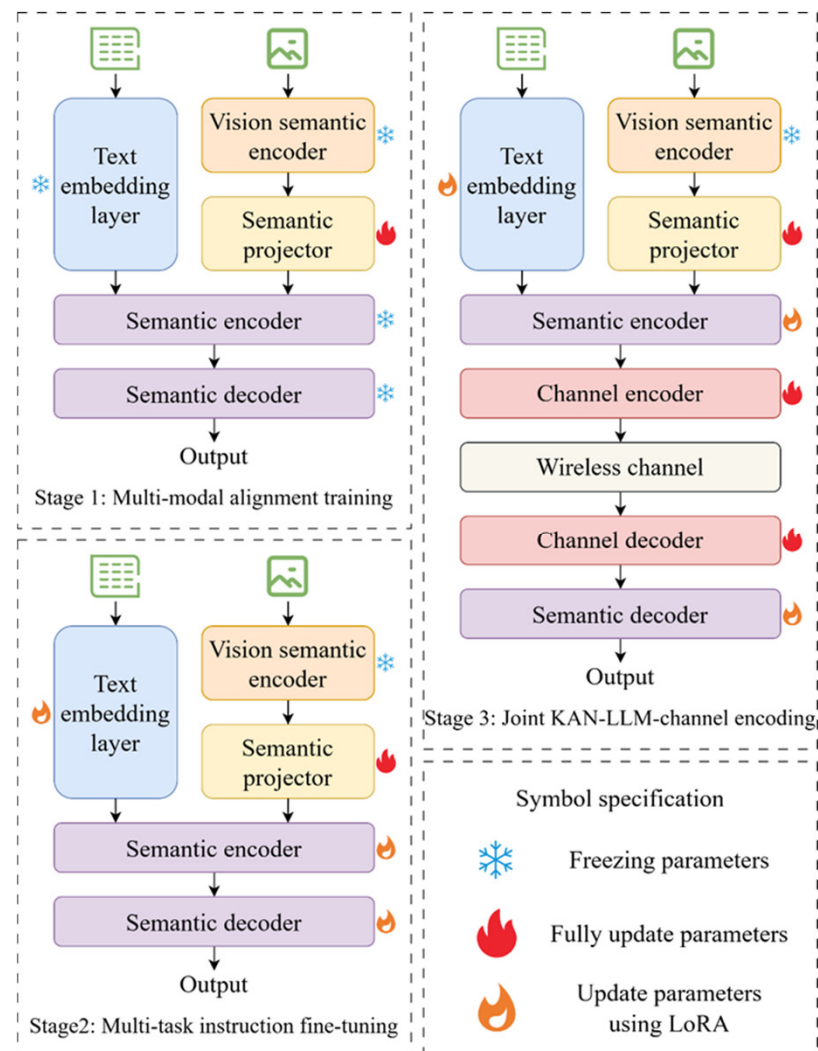
- 训练基于 KAN 的投影器
- 将图像特征映射为兼容 LLM 的输入 Token
- LLM 在此阶段保持冻结状态

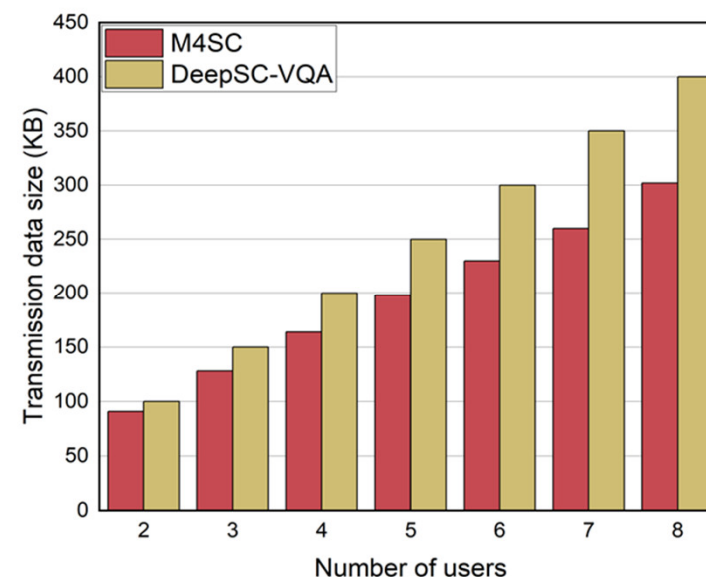
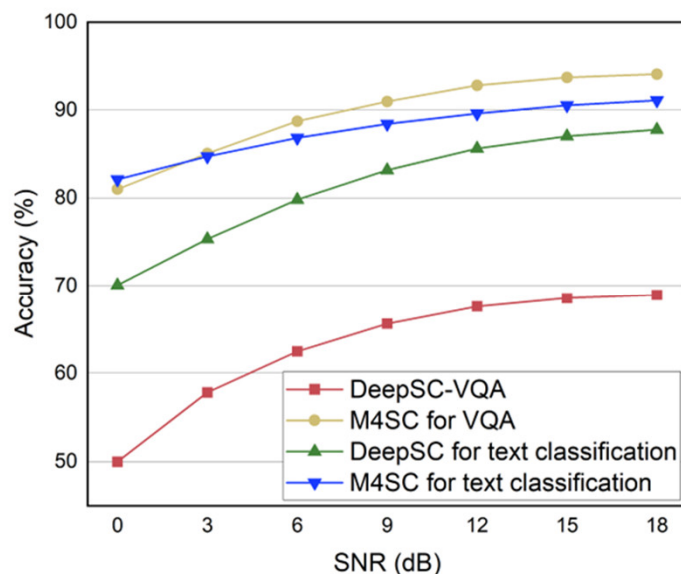
阶段二：多任务指令微调

- 联合训练 KAN 与 LLM
- 提升多模态理解能力与指令跟随能力
- 采用 LoRA 实现高效的 LLM 微调

阶段三：KAN-LLM-信道联合训

- 集成 KAN、LLM 及信道编码器/解码器
- 对 KAN 与信道模块进行完整更新
- 在 LLM 上继续使用 LoRA，实现低成本的联合编码训练





多模态与多任务条件下的准确率在视觉问答任务 (CLEVR) 和文本分类任务 (IMDB) 中, M4SC 的表现优于 DeepSC 与 DeepSC-VQA。多用户带宽效率随着用户数量增加, M4SC 显著减少了传输数据的大小



5 挑战与机遇

大模型的挑战与研究方向

- **通信数据学习的滞后性**：采用自主与鲁棒持续学习，提升在动态通信环境中的自适应与更新能力
- **通信推理能力不足**：结合过程奖励、长链推理和强化学习，增强模型推理能力
- **模型可解释性弱**：通过因果建模、特征归因和可视化提升模型透明度与可信度
- **实际部署难度高**：通过模型压缩与知识蒸馏，实现轻量化部署与资源高效利用

Agentic AI 的挑战与研究方向

- **通信知识的缺乏**：通过Agentic RAG机制实现任务感知的知识检索与动态更新
- **Agentic AI的扩展性限制**：采用分布式与分层架构，提升系统并行性与多Agent协同能力
- **控制机制的复杂性**：构建标准化的智能体协议体系，提升通信中多智能体间的任务协调性与控制效率
- **评估Agentic AI的难度大**：构建过程导向与主观评估框架，支持通信中智能体行为的细粒度分析与泛化能力验证



湖南师范大学

请各位专家批评指正!



通信大模型公众号

江沸菠

湖南师范大学 信息科学与工程学院

<https://github.com/jiangfeibo/ComLAM> 通信大模型代码仓库

<https://github.com/jiangfeibo/ComAgent> 通信Agentic AI代码仓库

邮箱: jiangfb@hunnu.edu.cn