Huilin Song (宋惠霖)

songhlin5@mail2.sysu.edu.cn
(+86) 159-0393-0001
School of Computer Science and Engineering, Sun Yat-sen University
songhl.github.io

Education

Sun Yat-sen University, Guangdong, China — Sep. 2022 – present

BSc in Information and Computing Science | GPA: 3.9/4.0 (88.6/100)

English Proficiency

CET-6: 590 (listening 205, reading 215, writing 170)
IELTS: I expect to have obtained a qualified IELTS score by July.

Research Experience

Sun Yat-sen University(SYSU), China — 2024.8 – 2024.12

Responsible for coding and main sections of paper writing

Brief introduction: This research focuses on improving the understanding capabilities of multimodal video-language large models (VLLMs) using adaptive sparse memory for long videos.

- Under the guidance of Jisheng Dang (PhD student at SYSU), improved the memory bank compression strategy of MA-LLM and incorporated a cross-attention module to enhance the interaction between vision and language.
- Experimental results demonstrated significant improvements over the baseline in various video-question answering scenarios.
- The paper has been accepted by ICME 2025.

Sun Yat-sen University, China — 2024.12 – 2025.3

Responsible for discussion and determination of core innovations, coding, and main sections of paper writing

Brief introduction: This research focuses on addressing single-frame bias in long videos and aims to reduce hallucinations in VLLMs for long video temporal understanding tasks while improving model performance and computational efficiency.

- Identified and validated the existence of single-frame static bias in long video understanding tasks by proposing the DTP module.
- Modified the MA-LLM Memory Bank module to enable the discrimination of static bias and dynamic selection based on it, and incorporated a CLIP-based contrastive learning module to enhance visual-text alignment.
- Experimental results demonstrated improvements over the baseline in various video question answering and video captioning tasks, with significantly reduced com-

putation time.
- The paper has been submitted to ACMMM 2025.

Sun Yat-sen University, National University of Singapore (NUS) 2025.3 – 2025.5

Responsible for conception, algorithm design, coding, and manuscript writing

Brief introduction: In this work, we introduce MUPA, a collaborative multi-agent framework that seamlessly integrates video grounding, question answering, self-reflection, and evidence aggregation to enhance the trustworthiness and reliability of multimodal LLMs for Grounded VideoQA.

- Designed three complementary reasoning paths and a reflection agent (PoE + MoE fusion) that jointly boost grounding fidelity *without* sacrificing answer accuracy.
- Achieved new state-of-the-art on NExT-GQA and DeVE-QA. Remarkably, the 2B variant already surpasses all existing 7B baselines.
- Supervised by Junbin Xiao (NUS) and Jisheng Dang (SYSU). The paper has been submitted to NeurIPS 2025.

Publications and Manuscripts

- Bimei Wang*, **Huilin Song***, Jisheng Dang, et al. AS Memory: Adaptive Sparse Memory Meeting Video Language Models," accepted by *ICME 2025*. (* *indicates equal contribution*)
- Jisheng Dang*, **Huilin Song***, Bimei Wang, et al. "Mitigating Temporal Hallucinations in Long Video Understanding Models through Dynamic Temporal Probing," submitted to *ACMMM 2025* (under review).
- Jisheng Dang*, **Huilin Song***, Junbin Xiao, et al. "MUPA: Towards Multi-Path Agentic Reasoning for Grounded Video Question Answering," submitted to *NeurIPS 2025* (under review).

Personal Skills

- **Solid Mathematical Foundation**
  Completed courses including Mathematical Analysis, Linear Algebra, and Probability Theory, achieving scores above 90 in most courses.
- **Strong Computer Science Background**
  Proficient in C, C++, Python, and Linux; completed foundational courses such as Data Structures, Computer Organization, Computer Networks, and Operating Systems. Experienced with mainstream deep learning frameworks (e.g., PyTorch), skilled in model design, optimization, and debugging.
- **Research Competence and Teamwork**
  Capable of independent literature review, experimental design, and fluent in reading and writing academic papers in English. Demonstrated strong teamwork and communication skills through collaborative research experiences, effectively cooperating with supervisors, PhD students, and peers.