Shangyu Xing

E-mail: xsy@smail.nju.edu.cn
Affiliation: School of Artificial Intelligence, Nanjing University
[Profile Page, Google Scholar, GitHub]

EDUCATION

Nanjing University **Nanjing, China**
Master of Computer Science *Sep 2023 - Jun 2026 (expected)*
Natural Language Processing Group (advised by Prof. Xinyu Dai)
Average Score: 88 / 100

Nanjing University **Nanjing, China**
Bachelor of Computer Science *Sep 2019 - Jun 2023*
Average Score: 91 / 100 (Top 10%)

RESEARCH INTERESTS

Multimodality: Multimodal Alignment / Multimodal Large Language Models
NLP: Natural Language Generation / Large Language Models

PUBLICATIONS

[1] **Shangyu Xing**, Fei Zhao, Zhen Wu, Tuo An, Weihao Chen, Chunhui Li, Jianbing Zhang, Xinyu Dai. *EFUF: Efficient Fine-grained Unlearning Framework for Mitigating Hallucinations in Multimodal Large Language Models*. **EMNLP' 2024**.

[2] **Shangyu Xing***, Fei Zhao*, Zhen Wu, Chunhui Li, Jianbing Zhang, and Xinyu Dai. *DRIN: Dynamic Relation Interactive Network for Multimodal Entity Linking.* **ACMMM' 2023**.

[3] Yiyang Zhou*, Zhaoyang Wang*, Tianle Wang*, **Shangyu Xing***, Peng Xia, Bo Li, Kaiyuan Zheng, Zijian Zhang, Zhaorun Chen, Wenhao Zheng, Xuchao Zhang, Chetan Bansal, Weitong Zhang, Ying Wei, Mohit Bansal, Huaxiu Yao. *AnyPrefer: An Automatic Framework for Preference Data Synthesis.* **ICLR' 2025**.

[4] **Shangyu Xing**, Changhao Xiang, Yuteng Han, Yifan Yue, Zhen Wu, Xinyu Liu, Zhangtai Wu, Fei Zhao, Xinyu Dai. *GePBench: Evaluating Fundamental Geometric Perception for Multimodal Large Language Models.* **Submitted to ACL' 2025**, under review**.**

[5] Fei Zhao, Taotian Pang, Chunhui Li, Zhen Wu, Junjie Guo, **Shangyu Xing**, Xinyu Dai. *AlignGPT: Multi-modal Large Language Models with Adaptive Alignment Capability*. Priprint. *Arxiv: 2405.14129*.

[6] Fei Zhao, Chunhui Li, Zhen Wu, **Shangyu Xing**, Xinyu Dai. *Learning from Different text-image Pairs: A Relation-enhanced Graph Convolutional Network for Multimodal NER*. **ACMMM' 2022**.

RESEARCH EXPERIENCES

Enhancing Multimodal Alignment in Multimodal LLMs *Sep 2023 - Present*

- **Pretraining - AlignGPT:** Create distinct alignment vectors for differently aligned text-image pairs during pretraining, and allocate them to various subtasks in finetuning and inference.
- **Hallucination Mitigation - EFUF:** Leverage external expert knowledge to reinforce the alignment between language and vision, thereby reducing multimodal hallucinations with no manually annotated data and minimal computational resources.

- **Benchmarking - GePBench:** Construct a large-scale dataset for evaluating fundamental geometric perception capabilities in MLLMs. We find that even the advanced models struggle, and models trained on our dataset exhibit stronger performance on real-world tasks.

Enhancing Multimodal Alignment in Information Extraction *Sep 2022 - Sep 2023*

- **Multimodal Named Entity Recognition:** Utilize Graph Neural Networks to capture external matching relationships across different text-image pairs.
- **Multimodal Entity Linking:** explicitly model four types of alignment between multimodal mentions and entities and uses a dynamic Graph Convolutional Network to automatically select appropriate alignment relations for different input samples.

Enhancing Multimodal Alignment in Cross-Domain Applications *May 2024 - Present*

- **Fine-Grained Paleontological Fossils Captioning** (research project led by advisor): train expert visual modal to recognize core visual features, and then feed the numerical information into a pretrained open-source MLLM to perform finetuning and inference.

INTERNSHIP

Huawei Technologies Co., Ltd. **Nanjing, China**

Software R&D Engineer *Jul 2022 - Sep 2022*

- Developing a deep-learning based voice cloning module in Text-to-Speech system
  - Integrated the open-source SOTA model Tacotron2 with a proprietary model optimized for handling Chinese spoken language pitch and rhythm

INFLY Tech (Shanghai) Co., Ltd. **Shanghai, China**

Software R&D Engineer *Jul 2023 - Sep 2023*

- Exploring preference alignment algorithms for training Large Language Models
  - Implemented preference alignment algorithms RLHF/PPO and its variations DPO, RRHF
  - Trained a BLOOM model with billion-level parameters using the Deepspeed and Megatron-LM frameworks, experimenting different algorithms

UNC Chapel Hill **North Carolina, USA**

Research Intern led by Huaxiu Yao (Online) *Jun 2024 - Present*

- Participating in multiple project-based collaboration
  - Implemented data synthesis algorithm with feedback on target vision-language model (AnyPrefer)
  - Developing an explainable reward model on robotics (working paper)

Nanjing University **Nanjing, China**

Teaching Assistant: *Compilers - Principles, Techniques, and Tools* *Sep 2024 - Jan 2025*

- Assisted in lectures, graded assignments and final exams

SKILLS

Programming: Python, C/C++, Assembly, Java

Frameworks: PyTorch, Hugging Face, Deepspeed, Megatron-LM

Tools: Linux, Docker, Git, LaTeX

Languages: Chinese (Native), English (TOEFL 105)

AWARDS

National Scholarship (2020)

Tencent Scholarship (2021, 2024)

Outstanding Student Model of Nanjing University (TOP 1%) (2022)

Outstanding Graduate of Nanjing University (TOP 10%) (2023)

First-class Academic Scholarship for Master's Students of Nanjing University (2024)