Huanyu WANG

Email: whyisverysmart@sjtu.edu.cn Tel: (86) 176-1218-9497 GitHub: https://github.com/whyisverysmart

Education

Shanghai Jiao Tong University (SJTU), Shanghai, China   Sept. 2022-Jun. 2026 (Expected)

Bachelor of Computer Science, Double Major in Mathematics and Applied Mathematics

GPA (overall): **3.81/4.0**

Course: Operating System (95/100), Programming Languages and Compilers (95/100), Data Structure (95/100), Principles and Methods of Program Design (95/100), Design and Analysis of Algorithms (95/100), Program Design Practice (94.5/100), Computer Network (93/100).

Publication

- Fourier-LLaVA: Fast Vision Language Models with FFT Operator
  Huanyu Wang, Jushi Kai, Zhouhan Lin
- Towards Temporal Accuracy: An Evaluation Benchmark for Understanding Time in Videos
  Chengyang Hu, Xinyu Zhou, **Huanyu Wang**, Danyu Shen, Lizhuang Ma

Research Experiences

SJTU, LUMIA Lab (Language Understanding and Machine Intelligence Algorithms)   Oct. 2024-Mar. 2025

Research Assistant, Advisor: Prof. **Zhouhan Lin**

- Addressed the challenge of computational redundancy in multi-modal large language models (LLMs), significantly improving the efficiency and accuracy of vision-language data processing.
- Developed an innovative approach by integrating the **2D Fourier Transform** with the **LLaVA** architecture, achieving a remarkable **88.9%** reduction in hidden-state redundancy.
- Utilized the LCS-558K dataset for further pretraining, optimizing feature alignment for better model performance.
- Conducted extensive evaluations across multiple benchmarks, delivering state-of-the-art results with a **2.3x** speedup in inference time.

SJTU, DMCV Lab (Digital Media & Computer Vision Laboratory)   Sept. 2024-Jan. 2025

Research Assistant, Advisor: Prof. **Lizhuang Ma**

- Designed a large-scale long-video dataset to evaluate the temporal recognition capabilities of existing models.
- Developed a high-efficiency multi-processing Python pipeline for batch downloading over **30000** videos.
- Implemented advanced speech-to-text conversion to generate precise subtitles with timestamps, utilizing GPT models for data cleaning and translation tasks.
- Trained a custom model on the dataset, achieving a **35%** improvement in temporal accuracy over existing models.

Work Experiences

Gauss Lab, Huawei   Jul. 2024-Sept. 2024

Software Development Intern, Advisor: Jinming Liao

- Collaborated with the Standby-Read team on the development of the **Gauss database kernel**.
- Designed and implemented novel C++ methods to address issues related to the database kernel state being unreadable, enhancing debugging and development efficiency by **60%**.
- Elevated functions to system-level and tested their effectiveness utilizing custom SQL scripts on a database cluster.

Skills

Computer: Python, PyTorch, Transformers, TensorFlow, Hugging Face, C++, LaTeX.

Model: LLaVA, CLIP, GPT, BERT, BLIP, LLaMA, VILA, Qwen, mPlugOwl3.

Language: English (TOEFL: **107**)