WANQI ZHONG

2021113666@stu.hit.edu.cn

7175380@gmail.com

EDUCATION

Harbin Institute of Technology *Aug 2023 - Present*
Undergraduate, Computer Science and Technology
Harbin Institute of Technology (Shenzhen) *Aug 2022 - Aug 2023*
Undergraduate Exchange Student, Computer Science and Technology
Harbin Institute of Technology *Aug 2021 - Aug 2022*
Undergraduate, Computer Science and Technology
Rank: 1/125
GPA: 96.367/100

HONOR

National Scholarship (Comprehensive ranking 1/435) *Oct 2022*
Innovation Achievement Award (18 out of all undergraduates from three campuses of HIT) *May 2024*
First Class Academic Scholarship（top 3% awarded by academic achievement) *May 2024*
Outstanding Youth League Member（3 out of all undergraduates in the same year) *Aug 2022*
Studying Stars *Jun 2022*
First Class Academic Scholarship（top 3% awarded by academic achievement) *March 2022*

EXPERIENCE

Research on Unified Multimodal Large Model based on MoE (Uni-MoE) *Aug 2023 - Apr 2024*
"Tencent-HIT" Joint Lab Advisor: Prof. Baotian Hu *Harbin Institute of Technology (Shenzhen)*

- **Introduction**: Developed a multimodal LLM based on MoE, aiming to handle various modalities such as audio, image, video, and text within a unified framework.
- **Involvement**:
  - Developed code for multimodal MoE architecture, designed to integrate information from various modalities like video through the collaboration of two experts.
  - Implemented Feature Modulation and Aux Loss code, guiding Router to distribute specific experts.
  - Led testing of Deepspeed MoE, enabling expert parallelism for large-scale multimodal data training across multiple nodes.
- **Outcome**:
  - **Paper under review: "Scaling Unified Multimodal LLMs with Mixture of Experts" (TPAMI).**
  - **Main contributor to the engineering code**, open-sourced on GitHub with over 700 stars.

Research on Robot Navigation in Complex HCI Scenarios based on LLM *Jan 2024 - Present*
Institute of Artificial Intelligence and Robotics (AIR) *Tsinghua University*

- **Introduction**: Inspired by the fast-slow dual system mechanism of the human brain, attempted to integrate complex reasoning and real-time navigation in human-machine interaction scenarios.
- **Involvement**:
  - Researched and discussed using LLM in the slow system for long-term planning; implemented short-term motion planning in the fast system and utilized cost maps as an intermediate fusion representation.
  - Independently developed a customized robot deployment framework using ROS2-Nav2.
- **Outcome**: This is an ongoing research project, and the results will be published in relevant papers upon completion.

Research on Integration of Large Model-Assisted Visual Capabilities *Jul 2023 - Present*

"Tencent-HIT" Joint Lab Advisor: Prof. Baotian Hu *Harbin Institute of Technology (Shenzhen)*

- **Introduction**: Utilized expert models like RAM to generate fine-grained descriptions and correct VLM generation results, training model's capabilities in visual storage, response, and correction through iterative dialogue and visual storage unit design.
- **Involvement**:
  - Designed and implemented the learning process of the model Learning, Storage, Review, and Reflection, and integrated capabilities.
  - Used GPT3.5 API to obtain expert model results, correcting and explaining original model illusions.
  - Independently completed code writing and data collection, allowing automated data and model iteration.
- **Outcome**: This is a research project nearing completion, and the results will be published in relevant papers upon completion.

Research on Performance of Visual Language Models in Knowledge-Intensive Tasks *Jul 2023 - Nov 2023*

"Tencent-HIT" Joint Lab Advisor: Prof. Baotian Hu *Harbin Institute of Technology (Shenzhen)*

- **Introduction**: Evaluated visual language models in basic common sense, fine-grained world knowledge, comprehensive knowledge tasks, analyzing limitations in visual retrieval capability and long-tail problems in knowledge storage.
- **Outcome**: Paper under review: "A Comprehensive Evaluation of GPT-4V on Knowledge-Intensive Visual Question Answering" (ACM MM).

Practice on Automated Gripper Robot for Designated Visual Features Objects *Nov 2022 - Aug 2023*

Experiment and Innovation Education Center (Robomaster) *Harbin Institute of Technology (Shenzhen)*

- **Introduction**: Used traditional vision and deep learning algorithms to accurately identify specific visual features in highly skewed angles and complex environments, performing PnP calculations to automatically grip objects with specific visual labels.
- **Involvement**:
  - Led algorithm code design and implementation, based on Yolo algorithm for coarse recognition, combining visual schemes like affine transformation to recognize labels.
  - Constructed the engineering code framework and led the coordination control of software and hardware such as communication with electronic controls.

PUBLICATIONS

Y. Li, S. Jiang, B. Hu, L. Wang, **W. Zhong**, W. Luo, L. Ma, M. Zhang. "Uni-MoE: Scaling Unified Multimodal LLMs with Mixture of Experts." *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, under review.

Y. Li*, L. Wang*, B. Hu, X. Chen, **W. Zhong**, C. Lyu, W. Wang, M. Zhang. "A Comprehensive Evaluation of GPT-4V on Knowledge-Intensive Visual Question Answering." *ACM International Conference on Multimedia (ACM MM)*, under review.

COURSES AND SKILLS

Mathematics Courses: Linear Algebra (100), Mathematical Logic (100), Calculus (99) and Discrete Mathematics (99) , with a weighted average score of 96.9 in mathematics courses.

Computer Science Courses: Computer Systems (98), Object-Oriented Programming (98), Algorithm Design and Analysis (97), Computer Networks (97), Operating Systems (96) and Computer Architecture (96).

Programming: C/C++, Python, Java

Tools: Pytorch, ROS2, Linux, Transformers, Deepspeed, Accelerate, Git, LaTeX, OpenCV

COMPETITION AWARDS

Robomaster University Championship	**Team National First Prize**	*Aug 2023*
Robomaster University Championship	**National Second Prize**	*Aug 2023*
Mathematical Contest in Modeling (MCM)	**Meritorious Winner**	*Feb 2023*
iCAN International Contest of Innovation	**National Second Prize**	*Dec 2022*
China Undergraduate Mathematical Contest in Modeling	**Provincial First Prize**	*Sep 2023*

OTHER EXPERIENCE

National Innovation and Entrepreneurship Project (First Prize) & Project Leader *Oct 2021 - May 2023*

- **Introduction**: Developed an automated biological perception and tracking system using 3D structured light sensors and deep learning algorithms.
- **Involvement**:
  - **Proficient in using Jetson series development boards (Nano and NX series)**, leading the full-process code development. Used image recognition algorithms on the TensorRT framework, structured light sensors and servos to track and capture videos of moving animals.
  - Worked on the underlying operating system, modifying the Linux device tree to support high performance and low power consumption, enabling collaborative operation and wake-up of the development board, achieving long-term operation of edge computing devices.
- **Outcome**:
  - Won several national-level awards in IoT competitions and innovation and entrepreneurship contests, selected for the Internet+ cultivation project and annual conference reserve project pool.