

Yuansheng Ni

Email: yuansheng.ni@uwaterloo.ca

Personal Website: <https://yuanshengni.github.io>

EDUCATION

University of Waterloo

MS in Computer Science

• Advisor: Wenhui Chen

Waterloo, CA

Sept 2024-Present

Zhejiang University

B.E. in Computer Science and Technology

University of Waterloo

Mathematics, Non Degree, Exchange

Zhejiang, China

Sept 2018-Jun 2024

Waterloo, CA

Sept 2023-Dec 2023

RESEARCH EXPERIENCE

TIGER Lab, University of Waterloo

Research Assistant; Advisor: Wenhui Chen

Waterloo, CA

Sept 2023-Present

ZJU NLP Group, Zhejiang University

Research Assistant; Advisor: Ninyu Zhang

Zhejiang, China

Jul 2023-June 2024

PUBLICATIONS

(* indicates equal contribution.)

1. A Massive Multi-discipline Multimodal Understanding and Reasoning Benchmark for Expert AGI

Xiang Yue*, Yuansheng Ni*, Kai Zhang*, Tianyu Zheng*, Ruoyi Liu, Ge Zhang, Samuel Stevens, Dongfu Jiang, Weiming Ren, Yuxuan Sun, Cong Wei, Botao Yu, Ruibin Yuan, Renliang Sun, Ming Yin, Boyuan Zheng, Zhenzhu Yang, Yibo Liu, Wenhao Huang, Huan Sun, Yu Su, and Wenhui Chen

In CVPR(**Best Paper Finalist**), 2024.

2. EasyEdit: An Easy-to-use Knowledge Editing Framework for Large Language Models

Peng Wang, Ningyu Zhang, Bozhong Tian, Zekun Xi, Yunzhi Yao, Ziwen Xu, Mengru Wang, Shengyu Mao, Xiaohan Wang, Siyuan Cheng, Kangwei Liu, Yuansheng Ni, Guozhou Zheng, Huajun Chen

In ACL SDT 2024.

PREPRINTS

1. MMMU-Pro: A More Robust Multi-discipline Multimodal Understanding Benchmark

Xiang Yue*, Tianyu Zheng*, Yuansheng Ni*, Yubo Wang, Kai Zhang, Shengbang Tong, Yuxuan Sun, Ming Yin, Botao Yu, Ge Zhang, Huan Sun, Yu Su, Wenhui Chen, Graham Neubig

arXiv:2409.02813, 2024.

2. VideoScore: Building Automatic Metrics to Simulate Fine-grained Human Feedback for Video Generation

Xuan He, Dongfu Jiang, Ge Zhang, Max Ku, Achint Soni, Sherman Siu, Haonan Chen, Abhranil Chandra, Ziyang Jiang, Aaran Arulraj, Kai Wang, Quy Duc Do, Yuansheng Ni, Bohan Lyu, Yaswanth Narsupalli, Rongqi Fan, Zhiheng Lyu, Yuchen Lin, Wenhui Chen

arXiv:2406.15252, 2024.

3. II-Bench: An Image Implication Understanding Benchmark for Multimodal Large Language Models

Ziqiang Liu, Feiteng Fang, Xi Feng, Xinrun Du, Chenhao Zhang, Zekun Wang, Yuelin Bai, Qixuan Zhao, Liyang Fan, Chengguang Gan, Hongquan Lin, Jiaming Li, Yuansheng Ni, Haihong Wu, Yaswanth Narsupalli, Zhigang Zheng, Chengming Li, Xiping Hu, Ruifeng Xu, Xiaojun Chen, Min Yang, Jiaheng Liu, Ruibo Liu, Wenhao Huang, Ge Zhang, Shiwen Ni

arXiv:2406.05862, 2024.

4. **GenAI Arena: An Open Evaluation Platform for Generative Models**
Dongfu Jiang, Max Ku, Tianle Li, Yuansheng Ni, Shizhuo Sun, Rongqi Fan, Wenhui Chen
arXiv:2406.04485, 2024.

5. **MMLU-Pro: A More Robust and Challenging Multi-Task Language Understanding Benchmark**
Yubo Wang, Xueguang Ma, Ge Zhang, Yuansheng Ni, Abhranil Chandra, Shiguang Guo, Weiming Ren, Aaran Arulraj, Xuan He, Ziyang Jiang, Tianle Li, Max Ku, Kai Wang, Alex Zhuang, Rongqi Fan, Xiang Yue, Wenhui Chen.
arXiv:2406.01574, 2024.

6. **A Comprehensive Study of Knowledge Editing for Large Language Models**
Ningyu Zhang, Yunzhi Yao, Bozhong Tian, Peng Wang, Shumin Deng, Mengru Wang, Zekun Xi, Shengyu Mao, Jintian Zhang, Yuansheng Ni, Siyuan Cheng, Ziwen Xu, Xin Xu, Jia-Chen Gu, Yong Jiang, Pengjun Xie, Fei Huang, Lei Liang, Zhiqiang Zhang, Xiaowei Zhu, Jun Zhou, Huajun Chen
arXiv:2401.01286, 2024.

SELECTED HONORS & AWARDS

CVPR Best Paper Finalist	2024
Meritorious Winner, MCM/ICM	Consortium for Mathematics and Its Application, 2022
Second Prize, Mathematical Modeling Competition	School of Mathematical Science, Zhejinag University, 2022
Third Prize, CUMCM	Industrial and Applied Mathematics of P.R.China, 2018
Second Prize, Collegiate Mathematics Competition	Department of Education of Zhejiang, 2018
Second Prize, Mathematical Modeling Competition	School of Mathematical Science, Zhejinag University, 2018

SKILLS

Programming Languages: Python; Java; C/C++; Matlab; \LaTeX
Deep Learning Frameworks: PyTorch; TensorFlow