Yiyang Du

Carnegie Mellon University, Pittsburgh, PA, US
yiyangd@andrew.cmu.edu ◇ GitHub : ADu2021 ◇ adu2021.github.io

EDUCATION

Carnegie Mellon University, Pittsburgh, PA, US 2025.8 - current
Ph.D student, Language and Technologies Institute, School of Computer Science

Tsinghua University, Beijing, China 2021.9 - 2025.6
B.Eng in Computer Science and Technology (**Graduate with Honors**)
GPA: 3.98/4 (Rank 1/173)

Honors and Scholarships:

- Outstanding Graduate
- CCF Elite Collegiate Award
- Champion of China International College Students' Internet+ Innovation and Entrepreneurship Competition, Beijing Division
- First Prize, CSP-S (Equivalent to the original National Olympics in Informatics in Provinces, NOIP)

Achieved **A+** grades in the following courses:

- Linear Algebra
- Discrete Mathematics
- Algorithm Design and Complexity Analysis
- Programing and Training
- Introduction to Computer Systems
- Cybersecurity Fundamentals

PUBLICATIONS AND RECENT MANUSCRIPTS

Chi Chen*, **Yiyang Du***, Zheng Fang, Ziyue Wang, Fuwen Luo, Peng Li, Ming Yan, Ji Zhang, Fei Huang, Maosong Sun, and Yang Liu. **Model Composition for Multimodal Large Language Models**. *In Proceedings of Annual Meeting of the Association for Computational Linguistics (ACL) 2024.*

Yiyang Du, Xiaochen Wang, Chi Chen, Jiabo Ye, Yiru Wang, Peng Li, Ming Yan, Ji Zhang, Fei Huang, Zhifang Sui, Maosong Sun, and Yang Liu. **AdaMMS: Model Merging for Heterogeneous Multimodal Large Language Models with Unsupervised Coefficient Optimization**. *In Proceedings of Computer Vision and Pattern Recognition Conference (CVPR) 2025.*

Yiyang Du, Yanzhe Zhang, William Held, and Diyi Yang. **Empowering LLM with Streaming Speech Generation Ability via Text-to-Unit Conversion**. *In preparation.*

RESEARCH EXPERIENCES

Foundation Models for Robotics 2025.8-current
Advised by Prof. Chenyan Xiong *LTI, Carnegie Mellon University*

· Develop foundation models for robotics that benefit from large language and vision models.
· Ground physical knowledge and environmental feedback into VLMs via pre-training and mid-training.
· Bridge the gap between the benchmarks of VLMs on its spatial or embodied abilities and the actual real-world problem solving abilities of VLAs.

Model Merging for Multimodal LLMs 2023.10-2025.3
Advised by Prof. Yang Liu *AIR, Tsinghua University*

· Proposed a model composition framework to fuse the modalities of multimodal large language models (MLLMs) without the burden of additional training.
· Designed parameter decoupling strategy and adaptive parameter adjustment algorithm to improve model merging performance in MLLMs by mitigating the parameter interference problem.
· Demonstrated the effectiveness of the framework by merging text, audio, vision, video, and 3D point cloud modalities into one model, while preserving the capabilities of each modality.
· Advanced the model merging strategy from identical architectures to heterogeneous architectures.

Streaming Audio LLMs 2024.7-2025.1
Supervised by Prof. Diyi Yang *NLP Group, Stanford University*

· Proposed a two-stage training paradigm to empower an arbitrary LLM with the ability to generate speech in a real-time streaming fashion through lightweight text-to-unit conversion.

· Leveraged a training-based forced alignment technique to align text tokens and unit tokens for data preparation, and studied the influence of contextual information in encoder-decoder transformer architecture on the performance of text-to-unit conversion.

· Compared with cascading a TTS model, our approach reduces latency and provides richer contextual information, leading to improved performance and real-time application.

Large-Scale Pretraining of LLMs 2023.2 - 2023.7

Supervised by Prof. Maosong Sun *DeepLang AI, Beijing, China*

· Investigated the influence of pretraining data on model performance to determine the optimal corpus composition regarding source, quality, and knowledge distribution.

· Studied the deployment of large-scale training on multi-node clusters with Megatron-LM training framework.

Semantic Retrieval 2022.11 - 2023.3

Supervised by Prof. Maosong Sun *THUNLP Lab, Tsinghua University*

· Applied language model-based semantic retrieval strategy on English and Chinese dictionary corpus.

· Proposed WantQuotes, an online reverse dictionary system that helps users retrieve sentences or phrases based on their meaning.

ACHIEVEMENTS

Outstanding Graduate (2%), *Tsinghua University* *2025.6*

Elite Collegiate Award (100 students nationwide), *China Computer Federation* *2024.10*

National Scholarship (2%), *Ministry of Education* *2023.11*

Comprehensive Outstanding Scholarship (Highest Award), *Tsinghua University* *2023.11*

Champion of Beijing Division, *China International College Students' Internet+ Innovation and Entrepreneurship Competition* *2023.8*

National Scholarship (2%), *Ministry of Education* *2022.11*

Comprehensive Outstanding Scholarship (Highest Award), *Tsinghua University* *2022.11*

First Prize, 2019 CSP-S Non-Professional Software Capability Certification (Equivalent to the original National Olympics in Informatics in Provinces, NOIP), *China Computer Federation* *2019.11*

SERVICES

Teaching Assistant, Qinghai University 2022 Fall

· Provided teaching support and prepared assignments for the course *The Foundation of Programming (Advanced Level)*.

Volunteer for Beijing 2022 Winter Olympics 2022.2

· Provided service to the audience at Wukesong Cadillac Arena.

Volunteer Reviewer

· ACL 2025

PROFESSIONAL SKILLS

Skilled in using C++ and Python for research tasks

Deep learning libraries & frameworks: Megatron-LM, DeepSpeed, etc.

High performance computing with CUDA

TOEFL iBT: 105 out of 120 (Speaking 23)