

# Dongfu Jiang

Updated February 13, 2025

✉ Email: [dongfu.jiang@uwaterloo.ca](mailto:dongfu.jiang@uwaterloo.ca)

🏠 Home Page: [jdf-prog.github.io](https://jdf-prog.github.io)

🐙 GitHub: [jdf-prog](https://github.com/jdf-prog)

🌐 LinkedIn: [dongfu-jiang](https://www.linkedin.com/in/dongfu-jiang)

## Research interests

Natural Language Processing; Large Language Models; Evaluation;  
Multimodality Reasoning;

## Education

### University of Waterloo

PhD in Computer Science

Advisor: Prof. Wenhua Chen.

Waterloo, Ontario

Sep 2023 – Present

### Zhejiang University

BE in Computer Science

Advisor: Prof. Zhou Zhao. GPA: 3.97/4.00.

Hangzhou, Zhejiang

Sep 2019 – June 2023

## Publications

### ACECODER: Acing Coder RL via Automated Test-Case Synthesis

Huaye Zeng\*, [Dongfu Jiang\\*](#), Haozhe Wang, Ping Nie, Xiaotong Chen, and Wenhua Chen

*arxiv preprint, 2025.*

### MEGA-Bench: Scaling Multimodal Evaluation to over 500 Real-World Tasks

Jiacheng Chene\*, Tianhao Liang\*, Sherman Siue\*, Zhengqing Wang, Kai Wang, Yubo Wang, Yuansheng Ni, Wang Zhu, Ziyang Jiang, Bohan Lyu, [Dongfu Jiang](#), Xuan He, Yuan Liu, Hexiang Hu, Xiang Yue, Wenhua Chen

*ICLR, 2025.*

### 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, Abhramil Chandra, and Ziyang Jiang, Aaran Arulraj, Kai Wang, Quy Duc Do, Yuansheng Ni, Bohan Lyu, Yaswanth Narsupalli, Rongqi Fan, Zhiheng Lyu, Yuchen Lin, Wenhua Chen

*EMNLP, 2024.*

### WildVision: Evaluating Vision-Language Models in the Wild with Human Preferences

Yujie Lu, [Dongfu Jiang](#), Wenhua Chen, William Yang Wang, Yejin Choi, and Bill Yuchen Lin

*NeurIPS 2024 Dataset and Benchmark Track, 2024.*

### **GenAI Arena: An Open Evaluation Platform for Generative Models**

Dongfu Jiang\*, Max W.F. Ku\*, Tianle Li\*, Yuansheng Ni, Shizhuo Sun, Rongqi Fan, and Wenhui Chen

*NeurIPS 2024 Dataset and Benchmark Track, 2024.*

### **MANTIS: Interleaved Multi-Image Instruction Tuning**

Dongfu Jiang, Xuan He, Huaye Zeng, and Cong Wei, Max Ku, Qian Liu, Wenhui Chen

*Transactions on Machine Learning Research, 2024.*

### **VIEScore: Towards Explainable Metrics for Conditional Image Synthesis Evaluation**

Max Ku, Dongfu Jiang, Cong Wei, Xiang Yue, Wenhui Chen

*ACL, 2024.*

### **MMMU: A Massive Multi-discipline Multimodal Understanding and Reasoning Benchmark for Expert AGI**

Xiang Yue, Yuansheng Ni, Kai Zhang, and Tianyu Zheng, Ruoqi 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, Wenhui Chen

*CVPR **Oral**, 2024.*

### **TIGERScore: Towards Building Explainable Metric for All Text Generation Tasks**

Dongfu Jiang\*, Yishan Li\*, Ge Zhang, Wenhao Huang, Bill Yuchen Lin, Wenhui Chen

*Transactions on Machine Learning Research, 2024.*

### **LLM-Blender: Ensembling Large Language Models with Pairwise Ranking and Generative Fusion**

Dongfu Jiang, Xiang Ren, Bill Yuchen Lin

*ACL, 2023.*

## Experience

### **Research Intern at Allen Institute of AI**

Advisors: (Bill) Yuchen Lin, Qianhui Qin

June 2024 – Sep 2024

### **Research associate at SeaAI**

Advisors: Qian Liu, Tianyu Pang, Chao Du

Feb 2024 – Today

MANTIS: Interleaved Multi-Image Instruction Tuning

### **Research Intern at University of Southern California**

Advisors: Xiang Ren, (Bill) Yuchen Lin

Mar 2022 – Mar 2023

LLM-Blender: Ensembling Large Language Models with Pairwise Ranking and Generative Fusion

Honors and  
Scholarships

National Scholarship (ZJU, <i>top 1%</i> )	2020
Zhejiang University First-Class Scholarship (ZJU, <i>top 3%</i> )	2020,2021
Xiaomi Scholarship (ZJU)	2022