Ruochen Wang (王若宸)

Home page: https://www.turningpoint-ai.com https://ruocwang.github.io/ • ruocwang@ucla.edu

Research Focus

Harnessing the power of Multimodal Agents, with the goal of building tools that are helpful, reliable, and fully automated.

- LLM-era: **(M)LLM** - Instruction and agent finetuning, compositional agent, trustworthy, prompt optimization, reasoning. **Diffusion Model** - controllable generation, alignment, and analytical study.
- Pre-LLM-era: Efficient and automated ML, including AutoML and Dataset Compression.

Education

01/2020 to present — **University of California at Los Angeles (UCLA)** — **U.S.**
Computer Science Department
- Ph.D. in Computer Science; Advisor: Prof. Cho-Jui Hsieh
- M.S. in Computer Science; GPA=4.0/4.0; Advisor: Prof. Cho-Jui Hsieh

09/2015 to 08/2019 — **The University of Michigan-Ann Arbor (UMich)** — **U.S.**
Department of Electrical Engineering and Computer Science (EECS)
- B.S. in Computer Science & B.S. in Statistics; GPA: 4.0/4.0

09/2013 to 06/2015 — **(Transferred) Shanghai University of Finance and Economics (SUFE)** — **China**
School of Finance
- Financial Experimental Class; GPA: 3.93/4.0; Program Rank: 1/30

Selected Honors

- Outstanding Graduate Student (for Master's degree, 1 per department), ***UCLA CS Department***, 05/2022.
- Outstanding Paper Award, ***ICLR 2021***, 04/2021.
- Award of Excellence (10%), ***Microsoft Research Asia (MSRA)***, 09/2019.
- Highest Distinction Graduate Award, ***The University of Michigan***, 08/2019.
- Berkeley Fung's Excellence Scholarship, ***UC Berkeley Graduate Admission Committee***, 03/2019.
- Outstanding Intern Award, ***SenseTime***, 01/2019.
- James B. Angell Scholar, ***The University of Michigan***, 2017-2019.
- Shanghai City Scholarship (0.6%), ***Shanghai City Government***, 09/2014.

Publications (* denote equal contribution)

(M)LLMs:

- MOSSBench: Is Your Multimodal Language Model Oversensitive to Safe Queries? (*TurningPoint AI, 2024*)
Xirui Li*, Hengguang Zhou*, **Ruochen Wang**, Tianyi Zhou, Minhao Cheng, Cho-Jui Hsieh.

- Large Language Models are Interpretable Learners. (*Google, 2024*)
Ruochen Wang, Si Si, Felix Yu, Dorothea Wiesmann, Cho-Jui Hsieh, Inderjit Dhillon.

- Solving for X and Beyond: Can Large Language Models Solve Complex Math Problems with More-Than-Two Unknowns?
Kuei-Chun Kao, **Ruochen Wang**, Cho-Jui Hsieh (2024)

- DrAttack: Prompt Decomposition and Reconstruction Make Powerful LLM Jailbreakers. (*TurningPoint AI, 2024*)
Xirui Li, **Ruochen Wang**, Ting Liu, Cho-jui Hsieh, Boqing Gong.

- One prompt is not Enough: Automated Construction of a Mixture-of-Expert Prompts. (*TurningPoint AI, ICML 2024*)
Ruochen Wang*, Sohyun An*, Minhao Cheng, Tianyi Zhou, Sung Ju Hwang, Cho-jui Hsieh.

Diffusion Models:

- MuLan: Multimodal-LLM Agent for Progressive Multi-Object Diffusion. (*TurningPoint AI, 2024*)
Sen Li, **Ruochen Wang**, Cho-jui Hsieh, Minhao Cheng, Tianyi Zhou.

- The Crystal Ball Hypothesis in Diffusion Models: Anticipating Object Positions from Initial Noise. (*TurningPoint AI, 2024*)
Yuanhao Ban, **Ruochen Wang**, Tianyi Zhou, Boqing Gong, Cho-Jui Hsieh, Minhao Cheng.

- Understanding the Impact of Negative Prompts: When and How Do They Take Effect? (*TurningPoint AI, ECCV 2024*)
Yuanhao Ban, **Ruochen Wang**, Tianyi Zhou, Minhao Cheng, Cho-jui Hsieh.

- On the Discrete Prompt Optimization for Text-to-Image Diffusion Models. (*Google, ICML 2024*)
Ruochen Wang, Ting Liu, Cho-jui Hsieh, Boqing Gong.

Dataset Compression:

- FedDM: Iterative Distribution Matching for Communication-Efficient Federated Learning. (*CVPR 2023*)
Yuanhao Xiong*, **Ruochen Wang***, Minhao Cheng, Cho-Jui Hsieh.

- Mitigating Bias in Dataset Distillation. (*ICML 2024*)
Justin Cui, **Ruochen Wang**, Yuanhao Xiong, Cho-Jui Hsieh.
- Scaling Up Dataset Distillation to ImageNet-1K with Constant Memory. (*ICML 2023*)
Justin Cui, **Ruochen Wang**, Si Si, Cho-Jui Hsieh.
- DC-BENCH: Dataset Condensation benchmark. (*NeurIPS 2022*)
Justin Cui, **Ruochen Wang**, Si Si, Cho-Jui Hsieh.

AutoML:
- Efficient Non-Parametric Optimizer Search for Diverse Tasks. (*NeurIPS 2022*)
Ruochen Wang, Yuanhao Xiong, Minhao Cheng, Cho-Jui Hsieh.
- Learning to Schedule Learning Rate with Graph Neural Networks. (*ICLR 2022*)
Yuanhao Xiong, Li-Cheng Lan, Xiangning Chen, **Ruochen Wang**, Cho-Jui Hsieh.
- Generalizing Few-Shot NAS with Gradient Matching. (*ICLR 2022*)
Shoukang Hu*, **Ruochen Wang***, Lanqing Hong, Zhenguo Li, Cho-Jui Hsieh, Jiashi Feng.
- RANK-NOSH: Efficient Predictor-Based Architecture Search via Non-Uniform Successive Halving. (*ICCV 2021*)
Ruochen Wang, Xiangning Chen, Minhao Cheng, Xiaocheng Tang, Cho-Jui Hsieh.
- Rethinking architecture selection in differentiable NAS. (*ICLR 2021*) **Outstanding Paper Award**.
Ruochen Wang, Minhao Cheng, Xiangning Chen, Xiaocheng Tang, Cho-Jui Hsieh.
- DrNAS: Dirichlet Neural Architecture Search. (*ICLR 2021*)
Xiangning Chen*, **Ruochen Wang***, Minhao Cheng*, Xiaocheng Tang, Cho-Jui Hsieh.

Research Experience

2023 to Present — **Google Research - Ads ML** — **U.S.**
Student Researcher, with Prof. Indejri Dellion, Dr. Felix Yu, and Dr. Si Si.
- Designing instruction finetuning tasks for improving the (M)LLM's ability to reflect.
- Demonstrate that Multimodal Large Language Models (MLLMs) are interpretable learners and can be used to implement accurate and interpretable Neural-Symbolic Programs.

09/2023 to Present — **TurningPoint AI Research** — **U.S.**
Founder & Principal
Advisory Board: Tianyi Zhou (Prof), Cho-Jui Hsieh (Prof), Minhao Cheng (Prof)
Researchers: Xirui Li, Hengguang Zhou, Yuanhao Ban, Yihang Chen, Sohyun An, Licheng Lan, Andrew Bai
- TPAI is a compact and hardcore research team focused on harnessing the power of Multimodal Agents.
- Serve as the Lead Initiator and Principal Investigator of the team, setting research directions, and overseeing daily operations and projects.

05/2022 to 2023 — **Google Research - Perception Team** — **U.S.**
Student Researcher on Diffusion Models and Transformers, with Dr. Boqing Gong and Dr. Ting Liu
- Controllable generation of Text-to-Image Diffusion Models
- Efficient CNN-Transformer hybrid architecture with linear attention for video processing.

05/2019 to 09/2019 — **Microsoft Research** — **China**
Research Intern
- Resource-constrained neural architecture search.
- Improving the optimization of the Proximal Policy Gradient via Interior Point methods.

09/2018 to 03/2019 — **SenseTime** — **China**
Research Intern
- Adversarial robustness in the frequency domain.

09/2017 to 04/2018 — **UMich College of Engineering** — **U.S.**
Research Assistant with Prof. Honglak Lee
- Worked on debasing a language-queried object detection model trained on datasets with sparse annotations.

Services

- Co-organizer of the 1st Dataset Distillation Workshop at CVPR 2024
- Reviewer for ICML Since 2021, NeurIPS Since 2021, ICLR Since 2022, TMLR, CVPR Since 2023, ICCV Since 2023, ECV 2023, NeuroComputing, e.t.c.