# Henry Hengyuan Zhao

+65-9135-0337 | hengyuan.z@u.nus.edu | https://zhaohengyuan1.github.io/

Google Scholar | zhaohengyuan1 | Linkedin

Singapore

## Self Introduction

I am generally interested in multimodal reasoning and Human-AI Interaction. Recently, I am excited about building intelligent AI systems to solve real-world problems and exploring the potential role of current AI models.

## Education

- **National University of Singapore** — *2022 - 2025*
  *Ph.D., Electrical Computer Engineering* — Singapore
- **Nanjing University of Posts and Telecommunications** — *2016 - 2020*
  *B.S., Communication Engineering* — China

## Experience

- **Microsoft Research GenAI Team** — *Aug 2025 - present*
  *Research Intern, Multimodal Model Modeling; Video VLM* — remote
- **Sea AI Lab** — *June 2023 - March 2024*
  *Research Intern, Multimodal Large Language Models: Data Generation and Training Paradigm* — Singapore
- **Alibaba DAMO Academy** — *Dec. 2021 - March 2023*
  *Research Intern, Parameter-Efficient Fine-tuning for Large Vision Transformers* — Singapore
- **MIG, SenseTime Inc** — *June 2021 - October 2021*
  *Research Intern* — China
- **Vision Technology (VIS), Baidu Inc** — *Dec. 2020 - June 2021*
  *Research Intern* — China
- **Shenzhen Institutes of Advanced Technology (SIAT), CAS** — *Sep. 2019 - Dec. 2021*
  *Research Intern* — China

## Publications

- Jiahao Tang*, **Henry Hengyuan Zhao***†, Lijian Wu, Yifei Tao, Dongxing Mao, Wanyang, Jingru Tan, Min Zeng, Min Li, Alex Jinpeng Wang, "From Charts to Code: A Hierarchical Benchmark for Multimodal Models", **Submission of ICLR 2026**. *: Euqali first author. †: Project Lead.
- **Henry Hengyuan Zhao**, Difei Gao, and Mike Zheng Shou, "WorldGUI: An Interactive Benchmark for Desktop GUI Automation from Any Starting Point", **Submission of ICLR 2026**.

  Benchmark: An early work for testing GUI agents in a dynamic setting. Agent: An effective and universal agent framework for GUI automation building uppn critic-thinking philosophy.
- **Henry Hengyuan Zhao**, Wenqi Pei, Yifei Tao, Mike Zheng Shou, "InterFeedback: Unveiling Interactive Intelligence of Large Multimodal Models via Human Feedback", **EMNLP 2025 Findings**.

  Summary: Can Large Multimodal Models evolve through Interactive Human Feedback? We found that (1) Accuracy may not fully reflect the models's intelligence; (2) LMMs may cater humans; (3) Low-quality feedback can degrade performance more than simply providing binary (0/1) feedback.
- **Henry Hengyuan Zhao**, Pan Zhou, Difei Gao, Zechen Bai, Mike Zheng Shou, "LOVA$^3$: Learning to Visual Question Answering, Asking and Assessment", **NeurIPS, 2024**.

  Only answering questions? Let's think about asking and assessing questions when training MLLMs? Without hyperparameter tuning or additional data annotation, consistent performance improvements are achieved!
- **Henry Hengyuan Zhao**, Pan Zhou, Mike Zheng Shou, "Genixer: Empowering Multimodal Large Language Model as a Powerful Data Generator", **ECCV, 2024**.

  How MLLMs perform in data generation? Take a look at using MLLMs to generate diverse multimodal data and observe the performance improvements.
- **Henry Hengyuan Zhao**, Pichao Wang, Yuyang Zhao, Hao Luo, Fan Wang, Mike Zheng Shou, "SCT: A Simple Baseline for Parameter-Efficient Fine-Tuning via Salient Channels", **IJCV, 2023**.

  We found that tuning only a small number of task-specific channels, referred to as salient channels, is sufficient. This work represents a remarkable reduction of 780x in parameter costs compared to its full fine-tuning counterpart.
- **Henry Hengyuan Zhao**, Hao Luo, Yuyang Zhao, Pichao Wang, Fan Wang, Mike Zheng Shou, "Revisit Parameter-Efficient Transfer Learning: A Two-Stage Paradigm", **Arxiv, 2023**.
- Yihao Liu, **Hengyuan Zhao**, Jinjin Gu, Yu Qiao, Chao Dong, "Evaluating the Generalization Ability of Super-resolution Networks", **TPAMI, 2023**.

- Yihao Liu*, **Hengyuan Zhao***, Kelvin CK Chan, Xintao Wang, Chen Change Loy, Yu Qiao and Chao Dong, "Temporally Consistent Video Colorization with Deep Feature Propagation and Self-regularization Learning", **CVM, 2023**.
- Xiangtao Kong, **Hengyuan Zhao**, Qiao Yu and Chao Dong, "ClassSR: A General Framework to Accelerate Super-Resolution Networks by Data Characteristic", *IEEE Conference on Computer Vision and Pattern Recognition*, **CVPR, 2021**.
- **Hengyuan Zhao**, Xiangtao Kong, Jingwen He, Yu Qiao and Chao Dong, "Efficient Image Super-Resolution using Pixel Attention", *European Conference on Computer Vision Workshop* (**ECCV Workshop, 2020**).
- **Hengyuan Zhao**, Wenze Shao, Bingkun Bao and Haibo Li, "A Simple and Robust Deep Convolutional Approach to Blind Image Denoising", *International Conference on Computer Vision Workshop* (**ICCV Workshop, 2019**).
- **Hengyuan Zhao***, Wenhao Wu*, Yihao Liu*, Dongliang He, "Color2Embed: Fast Exemplar-Based Image Colorization using Color Embeddings", **Arxiv, 2021**.

## TALKS

- I gave an Oral Presentation in the ICLR 2025 Bi-Align Workshop.
- I gave a talk about "Memory Efficient Tecniques" at Show Lab internal meeting.
- I gave a talk on AIM 2020 Efficient Super-Resolution Challenges.

## RESEARCH SERVICE

**Journal Reviewer**: TIP, IJCV
**Conference Reviewer**: ICLR 2026, NeurIPS 2025, ACM MM 2025, ACL 2025, WACV 2024, 2025, AAAI 2025.