

Hao Zhao (for pronunciation: /how-jow/) [in](#) [Linkedin](#) [g](#) [Google Scholar](#)

My main research interests include:

- Understanding and Improving LLM Alignment
- Robustness and Generalization of AI Models

Email: haozh510@gmail.com

Mobile: +81 188 8892 0951

Github: [MarcellusZhao](#)

Site: marcelluszhao.github.io

EDUCATION

- **École Polytechnique Fédérale de Lausanne** Lausanne, Switzerland
MSc in Automatic Systems (Nomination for Outstanding Master's Thesis). GPA: 5.40/6.00 Sept. 2021 – Feb. 2024
- **Zhejiang University** Hangzhou, China
BSc in Mechanical Engineering (with honors). GPA: 3.80/4.00 Sept. 2017 – July. 2021

PUBLICATION

(† stands for equal contribution.)

- *[Hao Zhao](#), [Maksym Andriushchenko](#), [Francesco Croce](#), [Nicolas Flammarion](#) , "Is In-Context Learning Sufficient for Instruction Following in LLMs?". *Under review*, submitted to the *13th International Conference on Learning Representations (ICLR 2025)*. Decision date: Jan 22, 2025. A short version was abridged in NeurIPS 2024 Workshop on Adaptive Foundation Models. Featured by [MIT Technology Review \(China\)](#). [\[Paper\]](#) [\[Code\]](#)
- *[Hao Zhao](#), [Maksym Andriushchenko](#), [Francesco Croce](#), [Nicolas Flammarion](#) , "Long Is More for Alignment: A Simple but Tough-to-Beat Baseline for Instruction Fine-Tuning". In *Proceedings of the 41th International Conference on Machine Learning (ICML 2024)*. 60674-60703. International. Poster Presentation. A short version was abridged in ICLR 2024 Workshop on Data-centric Machine Learning Research. [\[Paper\]](#) [\[Code\]](#)
- Haobo Song†, *[Hao Zhao](#)†, Soumajit Majumder, [Tao Lin](#), "Increasing Model Capacity for Free: A Simple Strategy for Parameter Efficient Fine-tuning". In *Proceedings of the 12th International Conference on Learning Representations (ICLR 2024)*. International. Poster Presentation. [\[Paper\]](#) [\[Code\]](#)
- *[Hao Zhao](#)†, [Yuejiang Liu](#)†, [Alexandre Alahi](#), [Tao Lin](#), "On Pitfalls of Test-time Adaptation". In *Proceedings of the 40th International Conference on Machine Learning (ICML 2023)*. 42058-42080. International. Poster Presentation. A short version was abridged in ICLR 2023 Workshop on Domain Generalization (**Spotlight**). [\[Paper\]](#) [\[Code\]](#)
- *[Hao Zhao](#), Weifei Hu, Zhenyu Liu, Jianrong Tan, "A CapsNet-Based Fault Diagnosis Method for a Digital Twin of a Wind Turbine Gearbox". In *Proceedings of the ASME 2021 Power Conference*. Vol. 85109, p. V001T09A016. International. Oral Presentation. (**Best Paper Award** in Renewable Energy Systems track) [\[Paper\]](#)

RESEARCH EXPERIENCE

- **Aligning Base LLMs Using Many-shot In-Context Learning** Advisor: [Prof. Nicolas Flammarion](#)
Research Assistant, Theory of Machine Learning Lab, EPFL April. 2024 - Present
 - **Uncover factors behind the empirical success of many-shot ICL:** We discover the crucial role of decoding parameters and find that many-shot ICL can indeed be improved by adding high-quality data in context.
 - **Fair comparisons between many-shot ICL and IFT:** We rigorously compare the performance of many-shot ICL to that of instruction fine-tuning (IFT) and give evidence showing that many-shot ICL underperforms IFT on the established benchmark MT-Bench, especially with more capable base LLMs, but ICL could be a viable alternative to IFT in the low data regime.
 - **Outcome:** A first-author paper is under review, and a short version paper was accepted to NeurIPS 2024 Workshop on Adaptive Foundation Models. This project was featured by [MIT Technology Review \(China\)](#).
- **Efficient Data Selection for Instruction Fine-tuning LLMs** Advisor: [Prof. Nicolas Flammarion](#)
Master Thesis Student, Theory of Machine Learning Lab, EPFL Oct. 2023 - Feb. 2024

- **Response length is a surprisingly good metric for efficient data selection:** We found that using reply length as a heuristic can effectively select a few examples from large-scale IFT datasets such as Alpaca-52k, on which the fine-tuned LLM outperforms fine-tuning on full datasets.
- **A lightweight LLM-based refining step further improves the abilities of fine-tuned LLMs:** We propose a new instruction refinement method that prompts a powerful LLM to rewrite demonstrations via introspection.
- **Outcome:** A first-author paper was accepted to ICML 2024, and a short version paper was abridged in ICLR 2024 Workshop on Data-centric Machine Learning Research.

Parameter-Efficient Fine-tuning

Advisor: [Prof. Tao Lin](#)

Research Assistant (remote), Learning and Inference Systems Lab, Westlake University

Mar. 2023 - Sep. 2023

- **Boost model capacity without additional costs:** We propose CapaBoost, a simple strategy that effectively increases the rank of injected weight matrices without adding new parameters, thereby enhancing the model capacity for free, by leveraging low-rank updates through interconnected parallel weight modules in target layers.
- **Outcome:** A co-first author paper was accepted to ICLR 2024.

On Pitfalls of Test-Time Adaptation

Advisor: [Prof. Alexandre Alahi](#)

Semester Project Student, Visual Intelligence for Transportation Lab, EPFL

Mar. 2022 - Febr. 2023

- **The first TTA Benchmark:** We build the first ever comprehensive TTA benchmark, TTAB, which encompasses over 10 state-of-the-art algorithms, a diverse array of distribution shifts, and two evaluation protocols.
- **Towards understanding the empirical success of TTA:** We identify 3 primary pitfalls in prior efforts, including hyperparameter sensitiveness, disparities among different base models trained using various strategies, and vulnerability to certain distribution shifts.
- **Outcome:** A co-first author paper was accepted to ICML 2023, and a short version paper was abridged in ICLR 2023 Workshop on Domain Generalization (**Spotlight**).

Wind Turbine Gearbox Fault Diagnosis

Advisor: [Prof. Weifei Hu](#)

Research Assistant, Institute of Design Engineering, Zhejiang University

Sept. 2020 - Mar. 2021

- **Build a new fault diagnosis pipeline:** We proposed a novel method, based on signal analysis in the frequency domain, that enables to label training data without human supervision, followed by training a neural network classifier, achieving better fault recognition accuracy and earlier warning time than prior works.
- **Outcome:** A first-author paper was accepted to the ASME 2021 Power Conference, which received a **Best Paper Award** in Renewable Energy Systems track.

INDUSTRY INTERNSHIP

Neural Radiance Fields for Thermal Novel View Synthesis

Advisor: [Dr. Malcolm Mielle](#)

Machine Learning Engineer Intern, Schindler Elevator AG, Switzerland

Mar. 2023 - Aug. 2023

- **Multimodal scene reconstruction:** I developed a novel multimodal approach based on Neural Radiance Fields, capable of rendering new RGB and thermal views of a scene jointly, and helped release a new RGB+thermal dataset for scene reconstruction.

HONORS AND AWARDS

- Nomination for EPFL Outstanding Master's Thesis (Institutional. Research. \$0) 2024
- Best Paper Award at ASME 2021 Power Conference (International. Research. \$0) 2021
- Outstanding Graduate of Zhejiang University (top 20%. Institutional. Academic. \$0) 2021
- Excellent Design Award for Undergraduate Thesis (top 10%. Institutional. Research. \$0) 2021
- Academic Scholarship of Zhejiang University (top 10%. Institutional. Academic. \$1000 each) 2018-2020

MISCELLANEA

- **Reviewer for conference:** [ICLR 2025](#)
- **Reviewer for workshop:** [FITML@NeurIPS'24](#), [ICL@ICML'24](#), [WANT@ICML'24](#), [DMLR@ICLR'24](#)
- **Media coverage.** My first-author paper, [Is In-Context Learning Sufficient for Instruction Following in LLMs?](#), was featured by [MIT Technology Review \(China\)](#)
- **English proficiency.** TOEFL iBT: 108 (R30/W30/L25/S23); GRE: 325 (Q170/V155)
- **Other interests.** I enjoy strength training, swimming, and hiking.