Suozhi Huang

Graduate Student

suozhi.huang@princeton.edu—suozhi.huang@gmail.com — +86-178-1022-7930 +1-201-600-4936
—https://hsz0403.github.io/ —

EDUCATION

Princeton University, Princeton, NJ, USA   Sept 2025 — Now

Electrical and Computer Engineering
>   PhD student

Tsinghua University, Beijing, China   Sept 2020 — Jun 2025

Yao Class, Institute for Interdisciplinary Information Sciences
>   Bachelor of Engineering in Computer Science and Technology (majored in Artificial Intelligence)

New York University, New York, USA   May 2023 — Sept 2023

Tandon school of Engineering
>   Exchange visiting student

INTERNSHIPS

Code LLM Pretraining   Beijing, China

Inf tech   April 2024 — Jun 2024

- Participated in code llm pretraining, especially data processing(deduplication) and evaluate with multiple benchmarks
- Built verified code benchmarks with models (like deepseek coder, qwen coder) and aligned the benchmarks to our own model.

RESEARCH INTERESTS

LLMs agents, LLMs inference.

PUBLICATIONS

Zihan Qiu, Zekun Wang, Bo Zheng, Zeyu Huang, Kaiyue Wen, Songlin Yang, Rui Men, Le Yu, Fei Huang, **Suozhi Huang**, Dayiheng Liu, Jingren Zhou, Junyang Lin 'Gated Attention for Large Language Models: Non-linearity, Sparsity, and Attention-Sink-Free', Paper(Neurips 2025 Best paper)

Suozhi Huang, Peiyang Song, Robert Joseph George, Anima Anandkumar. 'LeanProgress: Guiding Search for Neural Theorem Proving via Proof Progress Prediction', Paper

Zijian Wu*, **Suozhi Huang***, Zhejian Zhou, Huaiyuan Ying, Jiayu Wang, Dahua Lin, Kai Chen. 'InternLM2.5-StepProver: Advancing Automated Theorem Proving via Expert Iteration on Large-Scale LEAN Problems', Tech report(ICML 2025 AI4MATH Best Paper Runner Up)

YuChen Gao, YuHang Yuan, **Suozhi Huang**, Nan Yao, Legeng Yu, YaoPeng Chen, Professor Qiang Zhang, Prof Xiang Chen. 'A Knowledge-Data Dual-Driven Framework for Predicting the Molecular Properties of Rechargeable Battery Electrolytes', accepted by Angewandte Chemie. Paper

Suozhi Huang*, Juexiao Zhang*, Yiming Li, Chen Feng. 'ActFormer: Scalable Collaborative Perception via Active Queries', accepted by ICRA 2024. Webpage

RESEARCH EXPERIENCE

Anima AI+Science Lab, Caltech   Caltech, USA

Lean steps prediction analyst with external models

Advisor: Prof. Anima Anandkumar   Aug 2024— Present

- Using Reprover(a lean tactic generator) generate searching tree dataset
- Train and validate steps preditor and progress prediction model with formal math problems' search data
- Integrate supplemental functions for LeanCopilot(an interactive tool for theorem proving) to provide additional predictions for a user-friendly environment
- Submitted to ICML 2025 and accepted by ICLR 2025 verifai-workshop.

REAP Lab, UofT   Toronto, Canada

Program Verification Copilot(Coq) training

Advisor: Dr. Kaiyu Yang(Meta) , Prof. Xujie Si   Sep 2024— Present

- Built Best First Search for Coq and Python interface
- Construct the first Coq finetuning dataset using Coqgym(an interactive tool for theorem proving) for both whole-proof and step-proof
- Trained the first Coq finetuned model
- Integrate VSCode Built-in link with our own model and plan to submit to ICML 2025

Shanghai AI Lab   Shanghai, China
Theorem proving(formal math) LLM training
Advisor: Dr. Zheng Yuan, Prof. Dahua Lin   Jun 2024—Oct 2024

- Built critic model for expert iteration on Lean dataset(Lean workbook & Lean github)
- Replaced Best first search method with critic model guided search in tactic selection
- Construct a new comprehensive dataset by expert iteration searching
- Finished with SOTA performance on minif2f(a common formal math dataset) in LLM for theorem proving Tech report

Department of Chemistry, Tsinghua.   Beijing, China
Machine Learning Study of Flash Point for Designing Highly Safe Battery Electrolytes (GNN for Molecules)
Advisor: Prof. Chen Xiang, Prof. Zhang Qiang   Sept 2023 — Feb 2024

- Completed prediction of safety properties of electrolytes on pretrained molecular property prediction model Uni-Mol
- Implemented quantitative interpretation module for molecules' subgraph by applying MCTS on each's Shapley value to minimize the mutual information
- Run the baseline of electrolyte vector database, which contains billions of molecules with various predicted scientific molecular properties e.g. HOMO,LUMO
- Accepted by Angewandte Chemie as coauthor Paper

Tandon School of Engineering, NYU   New York, USA
Scalable Collaborative Camera-only Perception based on Active BEV Queries (Vision Perception)
Advisor: Prof. Chen Feng, Dr. Yiming Li   Feb 2023 — Sept 2023

- Enhanced the general applicability of the collaborative autonomous driving dataset V2X-Sim by adapting it to the mmdet 3d framework
- Proposed a scalable, active paradigm in query-based camera-only collaboration task, which reduced the cost in feature transfer
- Implemented a pose-guided selection network in deformable attention in camera-based perception, which greatly reduced the number of BEV queries with effective selection
- Completed all experiments in multi-agent dataset, analyzed the results, and improved scalability
- Accepted by ICRA 2024 as first author Webpage

AWARDS

Gold Award(Top 30) in 36th Chinese Physics Olympiad,   HangZhou, China, 2019
Freshman Scholarship at Tsinghua Univeristy   Beijing, China, 2020

SKILLS

- **Programming:** C++, Python
- **ML/AI:** Pytorch, ffmpeg, opencv, mmdet, unicore, etc.
- **Finance:** Quantitative Finance, Financial Modeling, Risk Management
- **Misc:** Quantum mechanics & Quantum computing