Jiajun Zhu

3200106048@zju.edu.cn · (+86) 173-009-89120 · zhuconv.github.io

Education

The University of Texas at Austin *Sept. 2025 - Present*

Ph.D. in Electrical and Computer Engineering

Zhejiang University *Sept. 2020 - Jun. 2025*

B.S. in Mathematics

- Major GPA: 3.90/4.00 (91.3/100, Top 1 of 64)

Publication

1. **Jiajun Zhu**, Yuehao Wang, Peihao Wang, Ruisi Cai, Siqi Miao, Liliang Ren, Pan Li, Jason D. Lee, Zhangyang Wang. Self Interpretable, Robust SSMs with Stochastic Selective Memory, *The Fourteenth International Conference on Learning Representations (ICLR)*, 2026. Submitted.
2. **Jiajun Zhu**, Peihao Wang, Ruisi Cai, Jason D. Lee, Pan Li, Zhangyang Wang. Rethinking Addressing in Language Models via Contextualized Equivariant Positional Encoding, *The Forty-Second International Conference on Machine Learning (ICML)*, 2025.
3. **Jiajun Zhu**, Siqi Miao, Rex Ying, Pan Li. Towards Understanding Sensitive and Decisive Patterns in Explainable AI: A Case Study of Model Interpretation in Geometric Deep Learning, *Nature Machine Intelligence*, 2025.
4. Peihao Wang, Ruisi Cai, Yuehao Wang, **Jiajun Zhu**, Pragya Srivastava, Zhangyang Wang, Pan Li, Understanding Bottlenecks of State Space Models through the Lens of Recency and Over-smoothing, *The Thirteenth International Conference on Learning Representations (ICLR)*, 2025.

Selected Research

Improved, Robust and Interpretable Mamba with Stochastic Memory *Jan. 2025 - Jul. 2025*

Supervised by Prof. Zhangyang Wang *University of Texas at Austin*

- Proposed injecting noise based on the information bottleneck (IB) principle to enhance the interpretability and robustness of the selective state-space model against attacks.
- Implemented SSIB with two types of stochastic models within the S6 framework: Gaussian-SSIB and Bernoulli-SSIB.
- Demonstrated that both SSIB variants offer inherent interpretability, increased robustness to corrupted input, and improved generation performance on standard language benchmarks.

Long-Context Transformer with Equivariant Position Encoding *Jun. 2024 - Sept. 2024*

Supervised by Prof. Zhangyang Wang *University of Texas at Austin*

- Introduced TAPE, a novel framework for dynamically layer-updated positional embeddings in transformers, adapting to content and surpassing the limitations of fixed long-term decay in traditional positional embeddings.
- Proposed the principles of permutation invariance and orthogonal equivariance to enhance the generalization of positional embeddings, and designed an enhanced Transformer with modules that integrate positional information into both the attention and feedforward layers in alignment with these principles.
- Demonstrated that TAPE excels in language modeling and downstream tasks such as arithmetic reasoning and long-context retrieval, achieving strong performance in both pretraining from scratch and parameter-efficient fine-tuning.

Experience

Interactive Reasoning of Visual Language Models *Dec. 2023 - Apr. 2024*

Supervised by Prof. Yaochu Jin *Westlake University*

- Proposed a paradigm enabling interaction with visual language models (VLMs) through visual referencing, specifically utilizing "click and segment" actions to improve interactivity and reference accuracy.
- Enhanced the reasoning capabilities of VLMs for image segmentation by fine-tuning them on a custom-built multimodal dataset.

Skills

- Programming languages: Python, C/C++, CUDA.
- Software & Frameworks: LaTeX, Git, PyTorch, PyTorch Geometric, Transformers, Triton.