

Jucheng Shen

 [juchengshen.github.io](https://github.com/juchengshen) |  [juchengshen](https://twitter.com/juchengshen) |  js237@rice.edu |  +1 (346) 448-4031

EDUCATION

Rice University B.A. in Computer Science, Math, Econ GPA: 3.67/4.0 Expected May 2027

Selected Coursework: Algorithms and Data Structures, Discrete Mathematics, Computer Systems, Linear Algebra, Real Analysis, Probability and Statistics, Honors ODE

RESEARCH EXPERIENCE

Princeton University, Research Intern Nov 2025 – Present

- Investigating data filtering and dataset merging strategies for large-scale text-to-image models, with a focus on improving generalization and training efficiency.
- Studying the role of image-text alignment by analyzing how VAE feature-caption consistency influences representation learning in open text-to-image foundation models.

The University of Texas at Austin, Research Intern May 2025 – Sep 2025

- Proposed One-Shot Dynamic Thresholding (OSDT), a training-free adaptive decoding method for diffusion LLMs achieving 24–50% higher throughput without accuracy loss.
- Directed large-scale experiments and analyses, including all main results, ablations, and figure production for publication.
- Led a collaboration with Intel Labs, proposing CadLLM, a unified adaptive decoding system that jointly tunes step size, block size, vocabulary, and thresholds, achieving up to $2.3\times$ speedup on diffusion LLMs.

Rice University, Research Intern Apr 2025 – Aug 2025

- Implemented *SuperGen*, a novel framework enabling generation of ultra-high resolution (2-4K) videos with limited GPU resources, on two baseline model architectures.
- Conducted end-to-end and ablations experiments, assisted with figure production and paper writing for publication.

PUBLICATIONS AND PREPRINTS

- [1] **Jucheng Shen**, Gaurav Sarkar, Yeonju Ro, Sharath Nittur Sridhar, Zhangyang Wang, Aditya Akella, Souvik Kundu. *Improving the Throughput of Diffusion-based Large Language Models via a Training-Free Confidence-Aware Calibration*. Preprint, 2025. [arXiv:2512.07173](https://arxiv.org/abs/2512.07173).
- [2] **Jucheng Shen**, Yeonju Ro. *Beyond Static Cutoffs: One-Shot Dynamic Thresholding for Diffusion Language Models*. NeurIPS 2025 Efficient Reasoning Workshop. [arXiv:2511.02077](https://arxiv.org/abs/2511.02077).
- [3] Fanjiang Ye, Zepeng Zhao, Yi Mu, **Jucheng Shen**, Renjie Li, Kaijian Wang, Desen Sun, Saurabh Agarwal, Myungjin Lee, Triston Cao, Aditya Akella, Arvind Krishnamurthy, T. S. Eugene Ng, Zhengzhong Tu, Yuke Wang. *SuperGen: An Efficient Ultra-High-Resolution Video Generation System with Sketching and Tiling*. Preprint, 2025. [arXiv:2508.17756](https://arxiv.org/abs/2508.17756).

AWARDS AND HONORS

- Frank Liu Jr. Prize for Creative Innovations in Music, Fashion, and the Arts — Top 11 of 111 teams in Rice Launch Challenge.

LEADERSHIP

Board Member (Tracks and Workshops), HackRice

Apr 2025 – Oct 2025

- Hosted 2 technical workshops for 100+ attendees; ran Q&A and live coding demos.
- Co-designed and shipped the official starter code used by 500+ hackers ([repo](#)).
- Designed competition tracks and judging criteria for HackRice 15.

Co-founder & CEO, NudgeBridge Nonprofit

Jul 2022 – Oct 2024

- Direct 25+ team members to provide educational support for 200+ children; raised \$25,000+ through strategic partnerships.
- Designed and managed long-term mentorship curricula emphasizing emotional support and academic motivation, impacting 40+ mentors and students, respectively.

TECHNICAL SKILLS

PyTorch, Python, Bash/Shell scripting; LLM evaluation & benchmarking; large-scale experiment management (HPC/Slurm); academic writing; rebuttal and peer-review response.