Guangzhi (Allen) Su

+86 182 5130 9825 | gs285@duke.edu | https://guangzhisu.github.io/

EDUCATION

Duke University & Duke Kunshan University Dual Degree Undergraduate Program May 2026
B.S. in Computer Science and Technology Kunshan, China
B.S. in Interdisciplinary Studies (Subplans: Applied Math & Computational Science; Computer Science) Durham, NC, USA

- Cumulative GPA: 3.85/4.0 (DKU) | 3.84/4.0 (Duke)
- Relevant Coursework: Applied Computer Vision (Graduate), Computer Engineering Machine Learning (ML) and Deep Neural Nets (Graduate), ML in Adversarial Settings (Graduate), Numerical Analysis, Probability and Statistics

CONFERENCE PUBLICATIONS

- **G. Su**, S. Huang, Y. Ke, Z. Liu, L. Qian, K. Huang. **SmoothGuard: Defending Multimodal Large Language Models (MLLMs) with Noise Perturbation and Clustering Aggregation.** In *Proceedings of the IEEE International Conference on Data Mining (ICDM)*, Washington D.C., USA, 2025
- Y. Liu, J. Sun, Y. Lin, J. Zhang, M. Yin, Q. Wang, J. Zhang, **G. Su**, H. Li, Y. Chen. **Keyframe-oriented vision token pruning: Enhancing efficiency of large vision language models on long-form video processing.** In *Proceedings of the International Conference on Computer Vision (ICCV)*, Honolulu, Hawaii, 2025
- (In Preparation) **G. Su**, H. Ye, Y. Liu, Y. Lin, H. Li, Y. Chen. **ParaFrame: A Parallel Agent-Based Framework for Efficient Long-Video Question Answering.**
- (In Preparation) **G. Su**, Y. Hu, N. Gong. **ARGuard: A circuit-breaker-powered defense for Autoregressive Generative models.**

CONFERENCE PRESENTATION

- **SmoothGuard: Defending MLLMs with Noise Perturbation and Clustering Aggregation.** Presented at the *1st International Workshop on Realistic Robustness and Generalization in Data Mining (RRoG-DM) at ICDM 2025*. Oral Presentation: Nov 12, 2025
- **SmoothGuard: Defending MLLMs with Noise Perturbation and Clustering Aggregation.** Presented at the *DKU 2025 SRS/SELF Poster Session.* Poster Presentation: Oct 15, 2025
- **StockGuard: Evaluating Membership Inference Vulnerabilities in Stock-Prediction Models.** Presented at the *Duke ECE 661: Computer Engineering Machine Learning and Deep Neural Nets Poster Session*, Poster Presentation: Dec 10, 2024
- **Context-Aware Token Pruning for MLLMs.** Presented at the *DKU 2024 SRS/SELF Poster Session*. Poster Presentation: Nov 8, 2024

AWARDS, HONORS, AND FELLOWSHIPS

Student Experiential Learning Fellowship (USD 2.5K) Summer 2024 and 2025
Dean's List Spring 2024 and 2025
Dean's List with Distinction Spring 2023, Fall 2023, and Fall 2024
Semester Research Program @ Duke (USD 1.1K) Fall 2024
Genscript Corporation Scholarship (USD 1K) Summer 2024
United Nations Millennium Fellowship Fall 2023

RESEARCH EXPERIENCE

Research Assistant Dec 2024 – Present
Department of Electrical and Computer Engineering (ECE), Duke University Durham, NC, USA
Supervisor: Dr. Neil Gong, Associate Professor of ECE and Computer Science
Project: A circuit-breaker powered defense for Image Autoregressive Generative Model

- Pioneered a feature-specific safety framework for autoregressive image generators by designing an end-to-end defense pipeline on the *Infinity* model to mitigate prompt-based vulnerabilities using a circuit-breaker mechanism.
- Conducted harmfulness probing experiments by checking features in latent space to identify the optimal finetuning parameter settings for balancing safety and generation quality.
- Trained a large-scale parameter generation module to enhance user accessibility, finetuning the Infinity model with a self-curated harmful-retained dataset using by model parallelism for efficient large-scale training.

Undergraduate Researcher, Signature Work (Senior Thesis) Sep 2024 – Present
Division of Natural and Applied Sciences, Duke Kunshan University Kunshan, China
Supervisor: Dr. Kaizhu Huang, Professor of ECE

Project: Defending MLLMs with Noise Perturbation and Clustering Aggregation

- Proposed SmoothGuard, a randomized smoothing–based defense for MLLMs that injects Gaussian noise to disrupt adversarial prompts and employs sentiment-aware clustering aggregation to stabilize outputs.
- Achieved a 30% reduction in attack success rate (ASR) with minimal utility loss across POPE, MM-SafetyBench, and LLaVA-Bench (In-the-Wild) benchmarks.

Research Assistant Jun 2024 – Present
Center for Computational Evolutionary Intelligence, Duke University Durham, NC, USA
Supervisors: Dr. Yiran Chen, John Cocke Distinguished Professor of ECE; Dr. Hai Li, Marie Foote Reel E'46 Distinguished Professor of ECE

Project 1: Agent-Based Framework for Efficient Long-Video Understanding Jun 2025 – Present

- Proposed and implemented a parallel multi-agent system for long-video understanding, integrating segmentation, keyframe selection, caption and reasoning agents to process extended video sequences with high efficiency.
- Built the system using Huggingface pipeline with multi-processing for asynchronous execution, achieving 60% accuracy on EgoSchema and Video-MME benchmarks, demonstrating competitive performance and substantial speedups.

Project 2: Context-Aware Token Pruning for MLLMs Jun 2024 – May 2025

- Proposed a context-aware token pruning framework for MLLMs, selecting high-attention vision tokens conditioned on textual prompts to improve efficiency.
- Extended the method to long-video tasks via Keyframe-oriented Vision Token Pruning, reducing token usage by 80% and Floating-Point Operations Per Second by 64% while maintaining or improving accuracy on relevant benchmarks.

Project 3: Adverse Events Detection for Surgical Videos Jul 2024 – Dec 2024

- Developed a deep learning pipeline leveraging InceptionV1 to detect and localize severe bleeding in surgical videos, integrating frame selection, anti-vibration filtering, and severity-aware classification for temporal intensity prediction.
- Curated 1,000+ surgical videos via Selenium web-scraping on Youtube and annotated in collaboration with Duke University School of Medicine, achieving 90% accuracy on short-video benchmarks and aiming to deliver an end-to-end post-operative report system to accelerate clinical review.

INTERNSHIP EXPERIENCE

Technology Research and Development Intern Dec 2023 – Jan 2024
Microsoft-INESA AI Innovation Center Shanghai, China
Supervisor: Jianzhi Liu, Senior Algorithm Engineer

- Drafted analytical reports on frontier AI research and industry trends, covering topics such as AI-generated content detection, LLM training and fine-tuning techniques, and emerging developments in China and abroad.
- Investigated deployment workflows for large language models, exploring fine-tuning methods including Retrieval-Augmented Generation, Docker-based containerization, and scalable cloud infrastructure for inference.

TEACHING & MENTORING EXPERIENCE

Undergraduate Peer Tutor Aug 2025 – Present
Office of Undergraduate Advising, Duke Kunshan University Kunshan, China

- Leading weekly 1-hour drop-in and 1:1 sessions for Principles of Machine Learning (24 students) and Computer Vision (5 students), covering course material and homework Q&A, exam reviews, and advice on final project topics.
- Undergoing bi-weekly 2-hour training sessions for College Reading and Learning Association Level 1 certification and maintaining a 4.85/5 average in student feedback.

Peer Mentor Aug 2025 – Present
Office of Undergraduate Advising, Duke Kunshan University Kunshan, China

- Mentoring 15 first-year students via regular lunch meetings, text check-in, and ice-breaker events, advising on community involvement, intercultural communication, and academic planning.
- Leading the "AI + X" initiative to help students adapt to interdisciplinary study in the AI era, coordinating career sharing alumni talks, AI tools hands-on workshops, and peer panels.

Teaching Assistant Aug 2024 – Mar 2025
Division of Natural and Applied Sciences, Duke Kunshan University Kunshan, China

- Designed and taught weekly 2-hour lab sessions, held drop-in office hours, and graded daily assignments for Discrete Math for Computer Science (21 students) and Introduction to Computer Science (60 students).
- Received a 4.23/5 average in student course evaluations and an 'A' rating in the professor's teaching assistant evaluation.

LEADERSHIP & COMMUNITY ENGAGEMENT

Chair of Student Advocacy Committee Dec 2024 – Present
Student Leaders Board (SLB), Duke Kunshan University Kunshan, China

- Appointed as part of SLB's inaugural cohort through campus-wide endorsement. Led a committee of six to represent student voices and conduct weekly meetings with the university leadership team.
- Led initiatives including raising emergency funds for Myanmar humanitarian relief, refining the Student Handbook's policy on racial and ethnic discrimination by defining violation classes, and contributing to curriculum and major review processes.

Founder & Co-President Jul 2023 – Aug 2025
DKU Computer Science (CS) Club, Duke Kunshan University Kunshan, China

- Founded DKU's CS Club, growing the community to 600+ followers and delivering 30+ events, including the *Technology for Sustainability Symposium* (400+ attendees; invited speakers from Intel, Microsoft, and IEEE) and the *Technology Innovation for Social Good Panel* (150+ attendees; co-hosted with Microsoft Asia Pacific R&D Group).
- Chaired the Program Committee for DKU's 1st and 2nd annual hackathons (400+ attendees; sponsored by Microsoft and AWS). Organized interdisciplinary tracks in finance, biotech, and environment featuring workshops, keynotes, mentoring, and final pitch competitions.

SKILLS

Programming & Scripting	Python, Java, R, HTML, CSS, JavaScript
ML Frameworks & libraries	PyTorch, Scikit-learn, OpenCV, Ollama
Tools & Platforms	Huggingface, Docker, Linux, Git, Selenium
Data visualization tools	Numpy, Pandas, Matplotlib
AI expertise	MLLMs, Multi-agent Systems, Computer Vision, NLP, Deep Learning Architecture Design
Languages	English (fluent), Mandarin (native)