Kunjun Li

Singapore | +65 89423624 | kunjun@u.nus.edu | kunjun-li | kunjun

Education

National University of Singapore Aug 2022 - Present
Bachelor of Engineering in Computer Engineering
Cumulative GPA: 4.8 / 5.0 (Top 5%)
Research Interest: Efficient Deep Learning, Edge-AGI

Publications

TinyFusion: Diffusion Transformers Learned Shallow - SOTA Depth Pruning Method for DiTs
Gongfan Fang*, **Kunjun Li***, Xinyin Ma, Xinchao Wang *(Equal-first author)*
Arxiv 2412.01199

PixelGen: Rethinking Embedded Camera Systems for Mixed-Reality
Kunjun Li, Manoj Gulati, Dhairya Shah, Steven Waskito, Shantanu Chakrabarty, Ambuj Varshney
The 30th Annual International Conference on Mobile Computing and Networking (MobiCom '24)

Professional Experience

NUS Learning and Vision Lab Singapore
Undergraduate Research Assistant 07/2024 – 1/2025
Supervisor: Prof. Wang Xinchao

• Developed TinyFusion, a SOTA learnable depth pruning framework for Diffusion Transformers, achieving a 2.86 FID score with halved model parameters and depth, and reducing pre-training costs to under 7%

• Significantly outperformed models of similar sizes, demonstrating substantial improvements in computational efficiency while maintaining high performance across various generative network architectures (Diffusion, Flow-based, and Visual Autoregressive).

• Conducted in-depth analysis on post-pruning recoverability in large-scale Diffusion Transformer models.

• Paper submitted to top-tier computer vision conference and received positive feedback from reviewers

NUS-NCS Joint Laboratory for Cyber Security Singapore
Undergraduate Research Assistant 07/2023 – 05/2024
Supervisor: Prof. Ambuj Varshney

• Proposed PixelGen, an innovative Embedded Camera System integrating Language Models and Diffusion Models, to generate High-Resolution RGB Images from monochrome images and sensor data.

• Paper Won Best Demo Runner Up at ACM/IEEE IPSN 2024 Conference.

Selected Honors

Dean's List, Top 5% of Cohort, NUS School of Computing 2025

IPSN'24 Best Demonstration Runner-Up, ACM/IEEE 2024

Project Experience

Parallel Virus Scanning with CUDA NUS
• Developed and optimized a CUDA parallel program with asynchronous kernel launches and memory transfers, achieving significant performance gains on NVIDIA A100 and H100 GPUs.

High-Performance RISC-V Processor Design NUS
• Designed a pipelined RISC-V CPU in Verilog with dynamic branch prediction, hazard handling, and optimized matrix multiplication using the Karatsuba Algorithm.

More projects can be found at: kunjun-li.github.io/projects