

Kunjun Li

📍 Seattle | 📞 +1 2064966816 | ✉️ kunjun@u.nus.edu | [in kunjun-li](https://www.linkedin.com/in/kunjun-li) | [globe kunjun](https://github.com/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 Generative Models

PUBLICATIONS

TinyFusion: Diffusion Transformers Learned Shallow

Gongfan Fang*, **Kunjun Li***, Xinyin Ma, Xinchao Wang (*Equal-first author*)

IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR '25 Highlight)

CVPR Highlight (3%) | Tiny DiTs at 7% Training Costs | 2x Faster Inference

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

UW Information Processing Lab

Seattle

Undergraduate Research Assistant

01/2025 – Present

Supervisor: Prof. Jenq-Neng Hwang

- Enhanced Segment Anything 2 (SAM2) to process multiple-view camera video data and won **Second Place** at 2025 SkiTB Visual Tracking Challenge.
- Currently exploring efficient Visual Autoregressive model (VAR), I proposed a multi-scale compression policy to reduce the KV cache down to **1%** with competitive image generation quality.

NUS xML Lab

Singapore

Undergraduate Research Assistant

06/2024 – 01/2025

Supervisor: Prof. Xinchao Wang

- Developed TinyFusion, a SOTA learnable depth pruning framework for Diffusion Transformers, achieving comparable performance 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 DiTs.

NUS-NCS Joint Laboratory for Cyber Security

Singapore

Undergraduate Research Assistant

07/2023 – 05/2024

- Proposed PixelGen, an innovative Embedded Camera System integrating Language Models and Diffusion Models, to generate High-Res 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

Second Place, 2025 SkiTB Visual Tracking Challenge

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.