

# ZEKAI CHEN

+1-607-262-1435 | zc542@cornell.edu | linkedin.com/in/zekai-chen | https://zekai-chen.github.io

## Education

### **Cornell University, College of Engineering, Ithaca, NY**

*Master of Engineering in Systems Engineering (Robotics) GPA:4.0*

Dec 2025

**Relevant Coursework:** NLP & LLMs, Robot Learning, Software Engineering, Autonomous Mobile Robots, Robotic MiniBot System, Integrating Human Driving in a VR Testbed, Frontiers of Computer Vision

### **The Chinese University of Hong Kong, Shenzhen, China**

*Bachelor of Electronic Information Engineering - Computer engineering (First Class Honours)*

May 2024

**Relevant Coursework:** Software Engineering, Machine Learning, Machine Intelligence and Applications, Microprocessors and Computer Systems, Computer Architecture, Operating Systems, Data structures, Database system, Digital System Design

## Professional Experience

### **Robotics Research Engineer | Maxinsights Cooperation, Santa Clara, CA**

Dec 2025 – Present

- Architected an end-to-end robotics MLOps pipeline from teleop data collection to π0 fine-tuning and real-world execution on AgileX/Piper for towel-folding, with Dockerized model serving, reproducible data processing and client-side inference.
- Integrated π real-time chunking (RTC) across server- and robot-side inference, reducing control jitter and improving towel-folding success from ~30% to ~55% in comparable real-world trials.
- Led camera/IMU calibration for an ego-capture robotics data rig and authored step-by-step tutorials to ensure reproducible data collection and processing across the team.

### **AI Application Research Intern | Futurewei Technologies Inc., San Jose, CA**

June 2025 – Nov 2025

- **Co-first Author**, *High-Fidelity 4x Neural Reconstruction of Real-time Path Traced Videos. ICCV-AIGENS 2025*.
- Built production-grade automation agents with Google Vertex AI ADK, reducing Marketing's daily semiconductor news preparation from ~2 hours to seconds with standardized outputs.
- Partnered with Linaro to design and validate an automated Cortex-A710 TF-A erratum patching workflow, cutting manual effort by ~90% and improving reproducibility.
- Optimized OpenSSL by implementing Camellia SIMD128 in AArch64 assembly, delivering ~25% throughput gain via an incremental, test-driven, ABI-compliant workflow.
- Evaluated multiple video-generation models on internal and open benchmarks; built a reproducible pipeline (env, preprocessing, metrics, visualization) and authored replication guides.

### **AI Engineer Intern | Nexa AI, Cupertino, CA**

Dec 2024 – May 2025

- Designed and implemented systematic benchmarks for RAG pipelines and LLM rerankers (jina, BGE, Qwen2.5-VL), informing model selection subsequently adopted in the Hyperlink product. [https://hyperlink.nexa.ai].
- Built a cross-platform on-device evaluation testbed (Mac M-series & Windows x86) to profile latency, memory, and power across FP16/Q8/Q4 models with standardized experimental protocols.
- Refactored and simplified open-source model codebases (Kokoro, Wav2Vec2, Jina-v2, PP-OCR-v4) by stripping out training/experimental code and standardizing interfaces, enabling clean, one-command inference pipelines.
- Shipped a Gradio demo interface (nexa run <model> -gr) within the Nexa SDK, enabling one-command model demonstrations for users and partners.

### **Robot Engineer | Cornell Cup Robotics, Ithaca, NY**

Sep 2024 – May 2025

- Implemented autonomy stack on XRP & iRobot Create (EKF SLAM, RRT/PRM, potential-field local control), achieving <15 cm waypoint error in dynamic mazes.
- Re-engineered Disney BB-8 on open-source XRP and produced LEGO-compatible variant, enabling rapid algorithm benchmarking and showcased at the 2025 FIRST Robotics Showcase.

### **Research Assistant | Shenzhen Institute of Artificial Intelligence and Robotics for Society, China**

Sep 2023 – May 2024

- Integrated Whisper + GPT-4 into a Pepper service robot to enable real-time speech-to-response (<2s), deployed in daily interactions at Shenzhen Longgang Station and serving 500+ user queries.

## Leadership

### **Co-Founder & Co-president | Stanford University ASES (GBA Chapter), Hong Kong SAR**

Nov 2023 – Nov 2025

- Ran Asia Launchpad for 80+ students; mentored 4 startups.

### **Founder | CUHK SZ Alumni Association, Shenzhen, China**

Sep 2023 – Present

- Built network of 300+ alumni, launched Entrepreneurship & Zhejiang divisions.

## Skills and Interests

- **Programming & Tools:** Python, C++, Linux, Git, Docker, CUDA, Google Cloud, AWS
- **Robotics & ML:** ROS2, SLAM, Teleop data collection, MediaPipe, OpenCV, Lerobot, Openpi,
- **Deployment:** PyTorch, Hugging Face, Model Benchmarking, RAG, on-device inference, reproducible MLOps pipelines
- **Research Methods:** Benchmark design, ablation studies, reproducible experimentation, data quality validation.
- **Interests:** Hiking and specialty coffee; Active participation in Bay Area robotics meetups and open-source discussions.