

# Bin-Lun Li

 [mike911209](#) |  [my site](#) |  [mike911209@gmail.com](mailto:mike911209@gmail.com) |  +886 982-880-498

## EDUCATION

---

### National Tsing Hua University (NTHU)

Sep 2021 - Jun 2025

B.S. in Computer Science

Advisor: [Jerry Chou](#)

Average GPA 4.2/4.3 on last 4 semesters.

## SKILLS

---

**Programming Languages** C/C++, Python, Verilog, Go

**Tools & Libraries** Git, Unix-like shells, Docker, Kubernetes, CUDA, MPI, OpenMP, Triton

**Languages** TOEIC Listening and Reading Test 915/990

## EXPERIENCE

---

### Machine Learning Engineer Intern - Lasertec Taiwan

Dec 2024 - Present

Pytorch/Triton/CUDA/Linux

- Achieved up to **297x speedup** in morphological operations using parallel programming.
- Developed production-grade GPU software tailored for industrial applications.

### Teaching Assistant, National Tsing Hua University

Operating System/Hardware Design

- **Operating Systems** - Authored specifications for Operating System implementations.
- **Hardware Design and Lab** - Drafted specifications for hardware design projects.

### President of Student Association of Dept. of CS, NTHU

Sep 2022 - Aug 2023

Leadership/Communication

- **Led a team of 10 people**, collaborated with faculty, industry professionals, and student groups.

## AWARDS

---

### Second Prize - 2024 Meichu Hackathon

[Link to Project](#)

Python/Flask/Backend/LLM

- Achieved **2nd place out of 230 contestants**, first prize in Logitech group.
- Multi-agents chatroom, where each agent embodies a distinct personality, enabling dynamic interactions and valuable insights through conversations.
- **Containerized** the backend, ensuring consistent deployment across different environments.

## PROJECTS

---

### Scalable LLM Inference Serving System

Go/Kubernetes/Prometheus/Grafana/GPU sharing/System Design

- Leveraged **GPU sharing technique** to allocate resources, resolved GPU underutilization issue.
- Implemented **dynamic resource allocation** strategies based on LLM inference workload analysis.
- Real-time monitoring with Prometheus and Grafana to visualize resource consumption.

### All-Pairs Shortest Path

C/C++/CUDA/Parallel Programming

- Optimized both SRAM and host memory access using parallel programming techniques.
- Achieved **10th place** out of 110 contestants in the course competition.

### Robotic Arm Color Classification

[Link to Demo](#)

Verilog/Hardware Design/FPGA

- Design novel hardware architecture which leverage hydraulic pressure to operate a robotic arm.