Pingchuan Ma

9500 Gilman Drive, La Jolla, CA, 92093
linkedin.com/in/pingchuan-ma-73531b239/
https://pingchuanma.netlify.app

Email: pima@ucsd.edu
Phone: (+86)18521309832
GitHub: github.com/HensonMa

Education

- **University of California, San Diego** — La Jolla, San Diego, US
  M.S. in Computer Science — *09/2023-03/2025(Expected)*
- **University of Michigan - Shanghai Jiao Tong University Joint Institute** — Shanghai, China
  B.S. in Electrical and Computer Engineering (minor in Computer Science) — *09/2019-08/2023(Expected)*
  - **Grade**: GPA: 3.947/4.0, Rank: 1/162
  - **Related Courseworks**: Probabilistic Methods in Engineering, Linear Algebra, Data Structures and Algorithms, Computer Vision, Artificial Intelligence, Operating Systems, Computer Architecture, Computer Network.
  - **Honors**: 2 times National Scholarship (top 0.2% national-wide), 2020 John Wu & Jane Sun Excellence Scholarship (Full Tuition), 2022 Zhan Jiajun Scholarship (only six recipients a year).
- **Cornell University** — Ithaca, New York, US
  Exchange Program in Engineering — *01/2022-05/2022*
  - **Coursework**: Embedded System, Optimization II, Intro to Engineering Stochastic Process

Skills Summary

- **Programming Languages**: Python, C, C++, Bash, Verilog, MATLAB, YAML, HTML.
- **Frameworks & Tech Stacks**: PyTorch, TensorFlow, Keras, NumPy, Pandas, Larq, Docker, Flannel, CUDA.
- **Operating Systems**: Linux, MacOS, Windows, Minix.

Experience

- **Software Engineering, Machine Learning Platform Department** — Intel, Shanghai
  Internship position — *12/2022-06/2023*
  - Maintained **backend** support of Intel's BigDL-Nano to provide acceleration for **PyTorch** and **TensorFlow** programs.
  - Supported Dictionary and Tuple data type of input samples for PyTorch models' tracing and inferencing by ONNX Runtime and OpenVINO **accelerator framework** in Intel's BigDL-Nano source codes.
  - Applied Intel Arc series GPU support and Intel's BigDL-Nano framework tools to accelerate **Stable Diffusion**'s inferencing on frontend client platform, by **halving** the latency on average using the best combination.
  - Added quantization (bfloat16, int8, and int4) support for **large language models** (LLMs from huggingface, including ChatGLM, GPT) using LLaMA.cpp to optimize their inference time on both Windows and Linux platforms.
  - Benchmarked LLM's performance with Google's Big-bench tools on **distributed** clusters by using Intel's BigDL-Orca.
  - Supported GitHub **CI/CD actions** (including unit test workflows) for Intel's BigDL-LLM new package development.
- **Efficient Machine Learning, Electrical and Computer Engineering Department** — Cornell University
  Research assistant — *03/2022-09/2022*
  - Examined and improved **PokeBNN** (a typical binarized neral network)'s performance including inference latency and classification task accuracy on **embedded processors**.
  - Re-implemented the inference network of PokeBNN using **TensorFlow Keras** and **Larq** framework.
  - Analyzed PokeBNN's inference process by breaking down the overall inference time into partial latencies of high-level functional abstractions and low-level **TFLite** atomic graph operations using **Larq** benchmarking methods.
  - Reduced the inference time of PokeBNN on embedded processors by **61.54%** by replacing heavy Shortcuts and Activation functions with alternative **TensorFlow** and **Larq** efficient **MLOps**.
  - Designed two-phase CIFAR10 training (first float-point then quantized) of PokeBNN with modified structures using **CUDA**, obtaining the best validation accuracy of **87.94**%.
- **AI Interpretability and Computer Vision, John Hopcroft Center** — Shanghai Jiao Tong University
  Research assistant — *09/2021-09/2022*
  - Trained neural networks to lower the interaction order among input variables to improve their adversarial robustness.
  - Generated adversarial image samples from back-propagation of **PGD-20** attack on adversarial-trained **ResNet18**.
  - Added untrained regional fully-connected layers implemented by **PyTorch** modules to do classification task in training, which yields **73.2%** improvement on the **adversarial robustness** of ResNet18 compared with the original one.
  - Designed a new loss function based on **Hessian matrix** to lower the high-order gradient influence between different regions of feature maps in neural networks.
- **Computational Methods and Machine Learning, DCE Laboratory** — Shanghai Jiao Tong University
  Research assistant — *01/2021-11/2021*
  - Analyzed annual data from COSCO SHIPPING Lines by Correlation Analysis and Principal Component Analysis to extract important features using **Pandas and NumPy**.
  - Built a new **time-series** model [Long-short Term Memory framework with Empirical Mode Decomposition and Attention mechanism] using **TensorFlow Keras** to help predict daily shipment prices and market demands, which reduced the error rate by **52.31%** on average compared to the baseline model RNN.

Projects

- **LemonDB Database** — Shanghai Jiao Tong University
  Software Programmer — *10/2022-11/2022*
  - Implemented a multi-threading read/write database supporting multiple instructions written in C++.
  - Designed the overall data storage structure based on **hash table** to realize **O(1) fast data retrieval**.
  - Designed and tested a flexible **thread pool** by using future, lock and conditional variables in C++ STD class to realize **multi-threading** searching, improving the performance by **3x** compared to the single-thread version on servers.
  - Implemented a **shell** interaction interface supporting I/O redirection for inputs of instructions and output results.
- **Mumsh Shell** — Shanghai Jiao Tong University
  Software Programmer — *09/2022-10/2022*
  - Designed and implemented a basic shell program written in C, supporting command/program execution, input/output redirection, pipelining from one program into another, interruption, and background processes.
- **TCP Communication Network** — Shanghai Jiao Tong University
  Software Programmer — *06/2022-07/2022*
  - Designed and implemented the basic **TCP socket** communication network program using **VMware** and **dockers**.
  - Built an **overlay network** between docker containers from two different virtual machines using **Flannel** framework.
  - Implemented a **selective repeat ARQ** with TCP socket program written in **Python** to simulate package transmission between different containers, as well as using **wireshark** to capturing TCP packages at the meantime.
- **Vintage 2D Plane Wars Video Game** — Cornell University
  Embedded System Interaction Programmer — *04/2022-05/2022*
  - Designed a video game supporting interaction between PC and **embedded processor** written by **C** and **PyGame**.
  - Implemented **UART communication** between the FRDM-KL64z board and user's computer.
  - Enabled finger-sliding and pressing on the embedded FRDM-KL64z board to control the plane's position by **TSL** (Touch Sensing Input) module and button controls in **KL46 sub-family**.
- **Litter Collecting Robot WALLE** — Shanghai Jiao Tong University
  Programmer and Designer — *06/2020-08/2020*
  - Apply **Mecanum wheels** to allow the robot to move in four directions. The final tested speed is chosen to be **0.335m/s** for best performance considering both quickness and stopping accuracy (deviation within 2 cm).
  - Use **HC-06 Bluetooth** module and **Arduino** components to control the robot's grabbing and movement. The experimental maximum operating distance is 98.49m for steady connection.
  - Design a phone **APP with convenient GUI** incorporating Bluetooth function for users' control.

Honors and Awards

- **Jiajun Zhan Scholarship** (only six recipients a year), Shanghai Jiao Tong University — Oct 2022
- **National Scholarship**, The Ministry of Education of the People's Republic of China (MOE) — Dec 2021 & Dec 2020
- **John Wu & Jane Sun Merit Scholarship** (Half Tuition), Shanghai Jiao Tong University Joint Institute — Nov 2021
- **Three Good Student Designation**, Shanghai Jiao Tong University — Oct 2021
- **Excellent Organization Award**, Shanghai Jiao Tong University — Oct 2021
- **John Wu & Jane Sun Excellence Scholarship** (Full Tuition), Shanghai Jiao Tong University Joint Institute — Dec 2020
- **Undergraduate Excellent Scholarship**, Shanghai Jiao Tong University — Dec 2020
- **Student Development Scholarship**, Shanghai Jiao Tong University Joint Institute — Jun 2020
- **Freshman Outstanding Scholarship** (Half Tuition), Shanghai Jiao Tong University Joint Institute — Sep 2019

Activities and Leadership

- **Leader of Publicity Group at Rong Chang Youth Organization** — Shanghai Jiao Tong University
  Publicize members' perception in social practice (e.g. commemorative album editor) — *05/2021-04/2022*
- **Media Department Director at Institute Student Union** — Shanghai Jiao Tong University
  Take charge of the student union's publicity, hold workshops for portrait photography. — *07/2020-07/2021*
- **Institute Writing Center Consultant** — Shanghai Jiao Tong University
  Offer guidance for students to polish their academic papers. — *09/2020-04/2021*
- **Volunteer for 2020 Shanghai Marathon** — Shanghai, China
  Provide assistance for material transportation, cheer up for athletes. — *29/11/2020*
- **Volunteer Teaching Program** — Yunnan, China
  Teach local primary students English, conduct field studies of local English education. — *12/2019-01/2020*