

Qianli Ma

mobile (+86)17855801919

wechat Fazzie17855801919

CV View on Browser

mail 1240419984@qq.com

github Fazziekey

blog fazzie-key.cool

EDUCATION BACKGROUND

National University of Singapore Master

2022.8 – Present

- Major: Computer Science

Zhejiang University Bachelor 3.80/4

2018.9 – 2022.7

- Major: Electronic Science and Technology Shannon elite class of Information Science and Electronic Engineer College
- Minor: Intensive training Program of Innovation and Entrepreneurship(ITP)

WORK & INTERNSHIP EXPERIENCE

Huawei 2012 Lab Distributed Parallel Lab

2021.7 – 2021.12

Algorithm Engineering Internship

Hangzhou

- Contributed code to [Mindspore](#), a full-scene deep learning developing framework; Developed three new features for Mindspore Lite.
- Implemented and iterated the Log system of Mindspore on x86 platform based on Glog.
- Accomplished the GPU support for OpenCL of Mindspore, developed GPU operators of Mindspore core.
- Completed Mindspore Lite OpenGL texture passing core code.
- Mindspore Lite partial GPU arithmetic performance optimization.
- Technology Stack:C++,OpenCL,OpenGL,Cmake,Python

SenseTime Large model training

2021.12 – 2022.6

Algorithm Researcher Internship

Hangzhou

- Participated in the development of large-scale distributed machine learning training framework(sensetime spring) of Shangtang Technology, and participated in research related to machine learning systems.
- Support for large model landing for target detection (Vision Transformer, Swin Transformer, etc.), implementation of Shangtang's generic target detection framework POD using pytorch distributed data parallel training and hybrid fine-reading training.
- Involved in MLops related work, machine learning cloud platform development, supporting model lifecycle management database.
- Technology Stack:Python,C++,Cuda,Pytorch,go,Nebula DB

HPC-AI Technology

2022.7 – Present

Machine Learning System Engineer

Singapore

- Participated in the development of [colossal-AI](#),A Unified Deep Learning System for Big Model Era.
- Leading the development of the AIGC Big Model [ColoDiffusion](#) training solution
 - Use zero optimizer, auto chunk, flash-attention and other technologies to break the memory wall, support super batch acceleration training, [Diffusion Pretraining and Hardware Fine-Tuning Can Be Almost 7X Cheaper!](#)
 - As a huggingface external developer, support huggingface diffuser library Dreambloom for Finetune tasks on a consumer GPU with 8G of memory
- Participated in the development of [Fastfold](#)(Optimizing AlphaFold Training and Inference on GPU Clusters)
 - Support the data pre-processing Parallel(Triple the speed) for Fastfold by [ray](#)
 - support the predict the multimer fold for Fastfold
- Technology Stack:Python,C++,Cuda,Pytorch,ray,colossal-AI

RESEARCH PROJECTS & COMPETITION

Data-Driven Beam Tracking based on Deep Learning

2020.1 – 2021.5

Institute of Intelligent Communication Network and Security

Prof. Min Li(ZJU 100 Young Professor)

- Proposed a novel deep learning algorithm to solve the beam tracking in the mmWave communication system.
- Deducted the mathematical function and built the mmWave communication models, generated the beam-index dataset through model inference.
- Proposed an efficient beam tracking algorithm based on transformer, achieved 91% predicting accuracy, 16 percent higher than existing algorithms.

Intel Embedding System Competition Magic mirror based on openpose

2020.7 – 2020.11

- Developed a device equipped with a mirror and a monitor, which can detect user posture for fitness.
- Abstracted the feature of posture with Openpose, build a VGG based model to do posture classification. Deployed on hardware platform AI-Box with Intel Openvino tool.
- Developed web front-end for the mirror, enabled features including voice assistant and entrance posture recognition. Lead the team to plan and execute the whole competition, won National Second Prize.

CLUBS & ORGANISATIONAL EXPERIENCE

Zhejiang University Internet Society Technology department AI lab

2021.10 – Present

String Program Technology department Member of the machine learning subdepartment

2020.7 – Present

KNOWLEDGE & SKILLS

- Program Language: C++, Python, C, Golang, Matlab, Verilog, Dart, HTML/CSS/JavaScript
- AI Full Stack
 - Familiar with Pytorch, Mindspore, Tensorflow and other deep learning frameworks for distributed training, mixed precision training.
 - Familiarity with the end-side AI inference framework Mindsporelite for model transformation, deployment and underlying source code.
 - Familiar with common computer vision and machine learning libraries such as Opencv, Sklearn, ray, etc. Familiar with GPU programming using OpenCL, Cuda.
- Others:Linux,Vim,shell Language Level: IELTS 6.5
- Core Courses:Data Structures and Algorithms, Digital System Design, Operating System, Distributed Systems,Computer version, Artificial Intelligence, Computer Organization, Network of Computer, Edge Calculation, Matrix Theory and Optimize.