

Yilong Zhao (赵恽龙)

E-mail: sjtuzyl@sjtu.edu.cn

Phone: 15221833996

Website: <https://xiaoke0515.github.io/>

EDUCATION

2022.9 – Now	Shanghai Jiao Tong University (SJTU), Shanghai, China <i>Ph.D Student</i> , Major: Computer Science and Technology	
2018.9 – 2021.3	Shanghai Jiao Tong University (SJTU), Shanghai, China <i>M.Eng.</i> , Major: Computer Technology	GPA: 3.49/4.0
2014.9 – 2018.6	Shanghai Jiao Tong University (SJTU), Shanghai, China <i>B.Eng.</i> , Major: Electronic Science and Technology Minor: Business Administration	GPA: 3.51/4.3

PUBLICATIONS & PATENT (Full list can be found [here](#))

- **Yilong Zhao**, Fangxin Liu, Xiaoyao Liang, Mingyu Gao, Naifeng Jing, Chengyang Gu, Qidong Tang, Tao Yang, and Li Jiang, "*STAMP: Accelerating Second-order DNN Training Via ReRAM-based Processing-in-Memory Architecture*," in **APPT'25**
- **Yilong Zhao**, Mingyu Gao, Huanchen Zhang, Fangxin Liu, Gongye Chen, He Xian, Haibing Guan, and Li Jiang, "*PUSHtap: PIM-based In-Memory HTAP with Unified Data Storage Format*," in **ASPLOS'25** (Accepted)
- **Yilong Zhao**, Mingyu Gao, Fangxin Liu, Yiwei Hu, Zongwu Wang, Han Lin, Ji Li, He Xian, Hanlin Dong, Tao Yang, Naifeng Jing, Xiaoyao Liang, and Li Jiang, "*UM-PIM: DRAM-based PIM with Uniform & Shared Memory Space*," in **ISCA'24**
- Weidong Cao, **Yilong Zhao (co-first author)**, Adith Bolor, Yinhe Han, Xuan Zhang, and Li Jiang, "*Neural-PIM: Efficient Processing-In-Memory with Neural Approximation of Peripherals*," in **TC'21**
- **Yilong Zhao**, Zhezhi He, Naifeng Jing, Xiaoyao Liang, and Li Jiang, "*Re2PIM: A Reconfigurable ReRAM-Based PIM Design for Variable-Sized Vector-Matrix Multiplication*," In **GLSVLSI'21**

Patent

- Li Jiang, Yilong Zhao, "Reconfigurable Architecture, Accelerator, Circuit Deployment and Dataflow Methods," Application No. 202010910280.5
- Li Jiang, Yilong Zhao, Xiaosong Cui, Yun Chen, Jianxing Liao, "Neural Network Circuit," Application No. 202010729402.0

SCIENTIFIC RESEARCH EXPERIENCE

Shanghai Jiao Tong University, ACA-IMPACT Lab,

Shanghai Qizhi Institute

Sep 2021 – Now

Unified data format for HTAP database on DRAM PIM

Design a PIM-based database with unified data format that can satisfy both OLTP and OLAP workload.

DRAM-PIM with Shared and Uniform Memory Space

To address the contradiction between PIM and memory interleaving, we design a uniform memory space where CPU memory pages and PIM pages with different interleaving scheme co-exists.

A PIM based Second-Order Training Accelerator

The purpose of the research is to design a neural network second-order optimizer based on the integration of storage and calculation technology.

PIM Implementation Towards Optical Communication Project

The purpose of the research is to realize the receiver of optical communication and wireless communication based on the integrated technology of storage and calculation. I am responsible for the following tasks:

- Design the overall architecture of optical communication and wireless communication systems based on the integration of storage and calculation, including operator splitting and algorithm reconstruction.
- Realize the circuit simulation of some operators.
- Aiming at the high-power calculation module to achieve a lower calculation amount than the existing numerical algorithm under the conditions of allowable error.

Shanghai Jiao Tong University, ACA Lab

Aug 2018 – Mar 2021

A Reconfigurable ReRAM-based DNN Accelerator Architecture

Design a ReRAM-based DNN accelerator which can significantly reduce the peripheral circuit's overhead.

ReRAM-based Efficient and Reliable DNN Accelerator Project

The project investigates the enhancement of computational reliability and the utilizes sparsity to improve energy efficiency in ReRAM-based DNN accelerator. I am responsible for the following work:

- Design and code a cycle-accurate simulator for the ReRAM-based NN accelerator. The simulator is built based on GEM5.
- Rewrite the simulator to evaluate the reliability and performance of architecture for pruned NN, The results of the simulator are used as an important metric for the project evaluation.
- Design a ReRAM-based DNN accelerator for pruned NN.