Zhihe (Kyrie) ZHAO 赵之赫

2nd year Ph.D student, AIoT Lab, The Chinese University of Hong Kong
Homepage: https://kyrie-zhao.github.io/

ACADEMIC INTERESTS

System for AI; DNN Compiler; Efficient Computing for Future IoT

EDUCATION BACKGROUND

B.E. in Computer Science and Technology, **University of Liverpool**	9/2014 – 7/2019
Master in Computer Engineering (Quit Ph.D with MS), **Duke University**, (Advisor: Prof. Maria Gorlatova)	8/2019 – 6/2021
PhD Student, **The Chinese University of Hong Kong**, (Advisor: Prof. Guoliang Xing)	9/2021 – Now

PROJECT EXPERIENCES

Mixed-critical DNN Inference Tasks Coordination on edge GPU, CUHK (Aaron)	Adviser: Prof. Guoliang Xing	2/2022–Now
Cross-device tensor program compiling domain adaptation, CUHK (Moses)	Adviser: Prof. Guoliang Xing	10/2021 – Now
Multi-user real-time object tracking for AR, Duke University	Adviser: Prof. Maria Gorlatova	8/2019 – 3/2020
AutoML framework for efficient inference on Edge, CUHK (EdgeML)	Adviser: Prof. Guoliang Xing	9/2018 – 5/2020
Edge Computing for Real-time Object Tracking, CUHK (ECRT)	Adviser: Prof. Guoliang Xing	6/2018 – 9/2018
Real-Time Data Monitoring System on wind tunnel, UoL	Adviser: Prof. Dawei Liu	5/2017 – 5/2018

PUBLICATIONS

AS FIRST AUTHOR:

- **Zhihe Zhao**, Xian Shuai, Yang Bai, Neiwen Ling, Nan Guan, Zhenyu Yan, Guoliang Xing, *"Moses: Exploiting Cross-device Transferable Features for On-device Tensor Program Optimization"* The Twenty-fourth International Workshop on Mobile Computing Systems and Applications (**ACM HotMobile 2023**)
- **Zhihe Zhao**, Neiwen Ling, Nan Guan, Guoliang Xing, *"Aaron: Compile-time Kernel Adaptation for Multi-DNN Inference Acceleration on Edge GPU"* In Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems (*Poster, **SenSys'22***). Association for Computing Machinery, New York, NY, USA, 394–395. **[Best Poster Award]**
- **Zhihe Zhao**, Kai Wang, Neiwen Ling, and Guoliang Xing *"EdgeML: An AutoML Framework for Real-Time Deep Learning on the Edge."* In Proceedings of the International Conference on Internet-of-Things Design and Implementation (***IoTDI '21***). Association for Computing Machinery, Virtual.
- **Zhihe Zhao**, Zhehao Jiang, Neiwen Ling, Xian Shuai, and Guoliang Xing. *"ECRT: An Edge Computing System for Real-Time Image-based Object Tracking."* In Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems (*Demo Presentaiton, **SenSys '18***). Association for Computing Machinery, New York, NY, USA, 394–395.
- **Zhihe Zhao**, J. Wang, C. Fu, D. Liu and B. Li, *"Demo Abstract: Smart City: A Real-Time Environmental Monitoring System on Green Roof,"* 2018 IEEE/ACM Third International Conference on Internet-of-Things Design and Implementation (*Demo Presentaiton, **IoTDI '18***), 2018, Orlando, FL, USA, pp. 300-301
- **Zhihe Zhao**, J. Wang, C. Fu, D. Liu, B. Li, *"Design of a Smart Sensor Network System for Real-Time Air Quality Monitoring on Green Roof",* Journal of Sensors (Sensing and Data-Driven Control for Smart Building and Smart City Systems (SBSCS)), Hindawi

AS CO-AUTHOR:

- Neiwen Ling, Xuan Huang, **Zhihe Zhao**, Nan Guan, Zhenyu Yan, Guoliang Xing, *"BlastNet: Exploiting Duo-Blocks for Cross-Processor Real-Time DNN Inference"* In Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems (***SenSys '22***). Association for Computing Machinery, New York, NY, USA, 394–395. **[Best Paper Candidate]**
- Zhang Xiangjun, Wu Weiguo, **Zhihe Zhao**, Wang Jinyu, Liu Song, *"MRMDDQN-Learning: Computation offloading algorithm based on dynamic adaptive multi-objective reinforcement learning in Internet of Vehicles"* (In Submission to **IEEE TVT**)
- Xian Shuai, Yulin Shen, Siyang Jiang, **Zhihe Zhao**, Wenhai Lan, Guoliang Xing, *"BalanceFL: Addressing Class Imbalance in Long-tail Federated Learning"* ACM / IEEE International Conference on Information Processing in Sensor Networks (***IPSN'22***), Milan, Italy.

INTERNSHIP EXPERIENCES

Research Intern, ECIL Lab, Huawei Cloud, Shenzhen, China 3/2022-7/2022

Embedded Software Engineer Intern, Rt-Thread Electronic Technology Co. Ltd., Shanghai, China 2/2017-6/2017

Co-founder, YouDu Smart Technology Co., Ltd., Suzhou, China (Raised 5M $, took a gap year in 15-16) 10/2015-3/2017

ACADEMIC SERVICE

TPC: MLSys'23@On-device Intelligence Workshop

Reviewer: AAAI'23@DCAA | IEEE Transactions on Mobile Computing (TMC) | MICCAI'23

SKILLS

Language & Framework & OS:, Python, C/C++, CUDA | PyTorch, TensorFlow, TVM, Android | Linux, RT-Thread OS, Euler

Hardware: GPU, MCU(STM32, S3C2440), WIFI Chip(ESP8266, ESP32, RT5350), NPU(ATLAS500), FPGA(PYNQ)

AWARDS

2022: Best Poster Award, SenSys'22 | Best Paper Candidate, SenSys'22 | Huawei Spark Award (First Place)

2021: BOSCH AIoT Fellowship | CUHK IE Ph.D Fellowship, 2021-2025

Before 2021: Duke ECE Ph.D Fellowship, 2019-2021 | National Scholarship, 2018