JIAN MENG

https://mengjian0502.github.io/
(+1) 503 810 8292 ⋄ jmeng15@asu.edu

EDUCATION

Cornell University *August 2023 - Present*
Doctor of Philosophy
School of Electrical, Computer and Energy Engineering

Arizona State University *August 2019 - May 2023*
Doctor of Philosophy Overall GPA: 3.83/4.0
School of Electrical, Computer and Energy Engineering

Portland State University *September 2015 - June 2019*
Bachelor of Science Overall GPA: 3.61/4.0; Graduate level GPA: 3.75/4.0
Department of Electrical and Computer Engineering

RESEARCH EXPERIENCE

Meta Reality Lab **May 2023 - Aug 2023**

· **Energy-efficient compression algorithm design for Codec Avatar project** Investigate and enhance the rendering quality under the low-precision model that is designed for AR/VR.

Seo Lab, ASU & Cornell **August 2019 - Present**
Advisor: Jae-sun Seo

· **Energy efficient Self-supervised Learning:** Hardware-aware sparse contrastive learning; High-performance contrastive learning algorithm for compact DNN encoder.
· **Robustness Enhancement Algorithm for In-Memory-Computing:** Robustness enhancement algorithm to overcome the temperature-induced variation in RRAM-based in-memory computing.
· **DNN Training Accelerator:** Hardware-compatible sparse training algorithm for a 28nm taped-out training accelerator.
· **FPGA-based Fully on-chip Inference:** Full on-chip, fully-pipelined sparse model inference accelerator with low precision (sub 4-bit) operation.
· **Sparse In-Memory-Computing:** System-level RRAM accelerator design incorporates low precision model with structured sparsity.

Teuscher Lab, Portland State University **January 2018 - June 2019**

· Design an energy-efficient deep neural network for mobile radiation detection platforms.

WORKING EXPERIENCE

Research Scientist: Meta Reality Lab May. 2023 - Aug. 2023

· **Energy-efficient and high-quality compression algorithm design for Codec Avatar model** Investigate the quality enhancement strategy and algorithm with the low-precision model.

Teaching Assistant: Arizona State University Jan. 2022 - May 2022
EEE598: Neuromorphic Computing Hardware Design

System Engineer: Kilby Labs of Texas Instrument Jun. 2021 - Aug. 2021

· **End-to-end compiler design for neural network acceleration.** Design and test a Pytorch-based compiler for deploying low-precision neural networks to the in-memory-computing-based accelerator.

Teaching Assistant: Portland State University Sep. 2018 - Jun. 2019

· ECE 221/2/3 (Circuit Analysis) and ECE 510 (Mathematical Foundation of Machine Learning)

SELECTED PUBLICATIONS

Conference Publications (*=Equal Contribution)

[C11] [**AAAI'24**] **Jian Meng**, Li Yang, Jinwoo Shin, Deliang Fan, and Jae-sun Seo, "Synchronized Contrastive Pruning for Efficient Self-Supervised Learning" (under review) (**Invited & Presented in IBM Research AI Hardware Forum**)

- **[C10]** [**NeurIPS'22**] **Jian Meng***, Li Yang*, Jae-sun Seo, and Deliang Fan, "Get More at Once: Alternating Sparse Training with Gradient Correction," Conference on Neural Information Processing Systems (NeurIPS), 2022.
- **[C9]** [**DATE'22**] Fan Zhang, Li Yang, **Jian Meng**, Jae-sun Seo, Yu Cao and Deliang Fan, "XST: A Crossbar Column-wise Sparse Training for Efficient Continual Learning," IEEE Design, Automation & Test in Europe (DATE) [**Best IP (Interactive Presentations) Paper Award**].
- **[C8]** [**CVPR'22**] **Jian Meng**, Li Yang, Jinwoo Shin, Deliang Fan, and Jae-sun Seo, "Contrastive Dual Gating: Learning Sparse Features With Contrastive Learning," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022. (**Invited & Presented in IBM Research AI Hardware Forum**)

[C7] [**IRPS'22**] **Jian Meng**, Injune Yeo, Wonbo Shim, Li Yang, Deliang Fan, Shimeng Yu, and Jae-sun Seo "Sparse and Robust RRAM-based Efficient In-memory Computing for DNN Inference" (IRPS).

[C6] [**ESSCIRC'22**] Shreyas K. Venkataramanaiah, **Jian Meng**, Han-Sok Suh, Injune Yeo, Jyotishman Saikia, Sai Kiran Cherupally, Yichi Zhang, Zhiru Zhang, and Jae-sun Seo, A 28nm 8-bit Floating-Point Tensor Core based CNN Training Processor with Dynamic Activation/Weight Sparsification, IEEE European Solid-State Circuits Conference (ESSCIRC), 2022.

[C5] [**FPT'21**] Han-sok Suh, **Jian Meng**, Ty Nguyen, Shreyas K. Venkataramanaiah, Vijay Kumar, Yu Cao, and Jae-sun Seo, Algorithm-Hardware Co-Optimization for Energy-Efficient Drone Detection on Resource-Constrained FPGA, IEEE ICFPT, 2021.

[C4] [**FPL'21**] **Jian Meng**, Shreyas Kolala Venkataramanaiah, Chuteng Zhou, Patrick Hansen, Paul Whatmough and Jae-sun Seo, "FixyFPGA: Efficient FPGA Accelerator for Deep Neural Networks with High Element-Wise Sparsity and without External Memory Access", International Conference on Field Programmable Logic and Applications (FPL), 2021.

[C3] [**IRPS'21**] Wonbo Shim, **Jian Meng**, Xiaochen Peng, Jae-sun Seo, and Shimeng Yu, "Impact of Multilevel Retention Characteristics on RRAM based DNN Inference Engine" (IRPS), 2021

[C2] [**DATE'21**] Jyotishman Saikia, Shihui Yin, Bo Zhang, **Jian Meng**, Mingoo Seok and Jae-sun Seo, "Modeling and Optimization of SRAM-based In-Memory Computing Hardware Design," IEEE Design, Automation, and Test in Europe (DATE), 2021.

[C1] [**INTERSPEECH'20**] Deepak Kadetotad, **Jian Meng**, Visar Berisha, Chaitali Chakrabarti, and Jae-sun Seo, Compressing LSTM Networks with Hierarchical Coarse-Grain Sparsity, INTERSPEECH, 2020.

Journal Publications

[J4] [**IEEE SSCM**] Jae-sun Seo, Jyotishman Saikia, **Jian Meng**, Wangxin He, Han-sok Suh, Anupreetham, Yuan Liao, Ahmed Hasssan, and Injune Yeo, Advances in Digital vs. Analog AI Accelerators, IEEE Solid-State Circuits Magazine, 2022

[J3] [**IEEE MICRO**] **Jian Meng**, Wonbo Shim, Li Yang, Deliang Fan, Shimeng Yu, and Jae-sun Seo, Temperature-Resilient RRAM-based In-Memory Computing for DNN Inference, IEEE Micro, 2021 (**Invited & Presented in IBM Research AI Hardware Forum**)

[J2] [**IEEE JETCAS**] Arnab Neelim Mazumder, **Jian Meng**, Hasib-Al Rashid, Utteja Kallakuri, Xin Zhang, Jae-sun Seo, Tinoosh Mohsenin, "A Survey on the Optimization of Neural Network Accelerators for Micro-AI On-Device Inference", IEEE JETCAS, 2021

[J1] [**IEEE TCAS-II**] **Jian Meng**, Li Yang, Xiaochen Peng, Shimeng Yu, Deliang Fan, Jae-sun Seo, "Structured Pruning of RRAM Crossbars for Efficient In-Memory Computing Acceleration of Deep Neural Networks". IEEE TCAS-II, 2021.