Hyunwuk Lee

Ph.D candidate
School of Electrical and Electronic Engineering
Yonsei University
50, Yonsei-ro, Seodaemun-gu, Seoul, Republic of Korea
hyunwuklee0519@gmail.com
Cell phone: +82-10-3122-6169

CAREER EXPERIENCE

Sep. 2024 ~	**Samsung Electronics** DS division (semiconductor), Memory Solution Development Department Solution Platform Development Team	Hwasung, Korea
Mar. 2018 ~ Aug. 2024	**Yonsei University** School of Electrical and Electronic Engineering *Integrated Master and Ph.D. Course* Advisor: Prof. Won Woo Ro *GPA: 4.25 / 4.5*	Seoul, Korea
Mar. 2014 ~ Feb. 2018	**Yonsei University** School of Electrical and Electronic Engineering *B.S. in Electrical and Electronic Engineering* Advisor: Prof. Won Woo Ro *GPA: 3.61 / 4.5*	Seoul, Korea

RESEARCH INTEREST

- **Accelerator Architecture for Neural Network**
  - ✓ NPU architecture and scheduling
  - ✓ Generative AI Acceleration (LLM, Diffusion Model, etc.)
  - ✓ Neural Network Compression (Quantization, Pruning, etc.)
  - ✓ Systems for Machine Learning
- **Graphic Processing Unit**
  - ✓ GPU Memory Systems
  - ✓ Multi-GPU Systems
  - ✓ GPU Device Driver

PUBLICATIONS

1. Sungbin Kim*, **Hyunwuk Lee***, Wonho Cho, Mincheol Park, and Won Woo Ro, "Ditto: Accelerating Diffusion Model via Temporal Value Similarity," 2025 IEEE International Symposium on High-Performance Computer Architecture (HPCA 2025) *Co-first-author
2. Gun Ko, Jiwon Lee, Hongju Kal, **Hyunwuk Lee**, Won Woo Ro, "REC: Enhancing fine-grained cache coherence protocol in multi-GPU systems," Journal of Systems Architecture (2025): 103339.
3. Sungbin Kim, **Hyunwuk Lee**, Sungwoo Kim, Cheolhwan Kim, and Won Woo Ro, "AirGun: Adaptive Granularity Quantization for Accelerating Large Language Models," The 42nd IEEE International Conference on Computer Design (ICCD 2024)

4. Seunghyun Jin, **Hyunwuk Lee**, and Won Woo Ro, "GUMSO: Gating Unnecessary On-Chip Memory Slices for Power Optimization on GPUs," ACM/IEEE International Symposium on Low Power Electronics and Design (ISLPED 2024)
5. Seunghyun Jin, **Hyunwuk Lee**, Jonghyun Lee, Junsung Kim, and Won Woo Ro, "SHREG: Mitigating register redundancy in GPUs." Journal of Systems Architecture (2024): 103152.
6. **Hyunwuk Lee**, Hyungjun Jang, Sungbin Kim, Sungwoo Kim, Wonho Cho, and Won Woo Ro, "Exploiting Inherent Properties of Complex Numbers for Accelerating Complex Valued Neural Networks", The 56th International Symposium on Microarchitecture (MICRO 2023)
7. Seokjin Go, **Hyunwuk Lee**, Junsung Kim, Jiwon Lee, Myung Kuk Yoon, and Won Woo Ro, "Early-Adaptor: An Adaptive Framework for Proactive UVM Memory Management", 2023 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS 2023)

PATENTS

1. **Hyunwuk Lee**, Gun Ko, Ipoom Jeong, and Won Woo Ro "Memory device including a plurality of area having different refresh periods, memory controller controlling the same and memory system including the same", **US Patent No. 11276452,** Korea Application No. 10-2020-0045023, China Patent No. ZL 2020 1 1090146.1
2. Hongju Kal, Cheonjun Park, **Hyunwook Lee**, Ipoom Jeong, Jiwon Lee, and Won Woo Ro "A method for neural network processing including memory-optimization techniques", Korea Application No. 10-2022-0041848
3. Seunghyun Jin, Jonghyun Lee, **Hyunwuk Lee**, and Won Woo Ro "Apparatus and method with register sharing", US Application No. 18/ 315576, Korea Application No. 10-2022-0074653
4. Jiwon Lee, Ipoom Jeong, Hongju Kal, Gun Ko, **Hyunwuk Lee**, and Won Woo Ro "Memory management unit and method of walking page table", US Application No. 18/502058, Korea Application No. 10-2022-0175909, China Application No. 202311700848.0
5. Seokjin Go, Junsung Kim, **Hyunwuk Lee**, Jiwon Lee, and Won Woo Ro "Memory management apparatus and method for UVM", Korea Application No. 10-2023-0076606
6. **Hyunwuk Lee**, Hyungjun Jang, Sungbin Kim, Wonho Cho, Sungwoo Kim, and Won Woo Ro "Polar Form Aware Apparatus and Method for Complex Valued Neural Network and Rectangular Form Conversion Apparatus", Korea Application No. 10-2023-0158538

RESEARCH EXPERIENCES

Graduate Researcher, Embedded Systems and Computer Architecture Lab (eSCaL), at Yonsei University

Feb. 2018 ~ Aug. 2024

- **Project:** "Memory-Centric Architecture using the Reconfigurable PIM Devices", Institute for Information & communication Technology Planning & evaluation (IITP), Korea

Jan. 2024 ~ Aug. 2024

- ✓ Led the project team in eSCaL
- ✓ Developed virtualized multi-PIM API for DNN applications
- ✓ Designed data allocation schemes for multi-PIM architecture

- **Project:** "Developing Data Processing Unit for AI Workloads", Korea Evaluation Institute of Industrial Technology (Keit), Korea

July 2022 ~ Jan. 2024

- ✓ Led the project team in eSCaL
- ✓ Profiled multi-GPU system running neural network workloads
- ✓ Modeled DPU integrated GPU by CPU-GPU simulator

- **Project:** "Analysis on High Performance GPU Workloads and Architecture Design", Samsung Advanced Institute of Technology, Korea

May 2021 ~ Apr. 2022

- ✓ Led the project team in eSCaL
- ✓ Profiled GPU running neural network workloads
- ✓ Designed GPU register architecture sharing data inter warp and intra warp for general matrix multiplications
- ✓ Modeled the register architecture in GPU simulator

- **Project:** "Development of High Performance Multi-GPU Memory System", National Research Foundation of Korea (NRF) , Korea

Mar. 2021 ~ Feb. 2022

- ✓ Researched HW/SW memory management of modern multi-GPU system
- ✓ Modeled the memory management techniques in multi-GPU simulator

- **Project:** "Architectural Exploration of Parallel Execution Processing Units for Supercomputer CPU", National Research Foundation of Korea (NRF), Korea

July 2020 ~ Jan. 2024

- ✓ Led the project team in eSCaL
- ✓ Designed SIMD architecture for supercomputer CPU
- ✓ Modeled SIMD architecture for supercomputer based on RISC-V CPU SIMD extension version in CPU simulator
- ✓ Designed register architecture for SIMD unit in supercomputer CPU

- **Project:** "Performance Analysis of Neural Network Workloads and Development of Energy Efficient Approximate Memory for Neural Networks", SK Hynix, Korea

July 2018 ~ June 2019

- ✓ Profiled GPU running neural network workloads
- ✓ Designed Approximate DRAM architecture for energy efficient neural networks
- ✓ Modeled DRAM architecture in DRAM simulator

HONOR AND AWARDS

- SK Hynix, Industry-Academic Research Project Outstanding Invention

2021

TEACHING EXPERIENCES

- Teaching Assistant, *School of Electrical and Electronic Engineering, Yonsei University*
- ✓ *Computer Architecture* Spring, 2018

SKILLS AND TECHNIQUES

- **Programming Languages**
- ✓ C / C++
- ✓ Python
- ✓ CUDA
- **DNN Framework**

- ✓ PyTorch
- ✓ Tensorflow

- **Simulator**
- ✓ DRAM simulator: DRAMSim3, Ramulator
- ✓ CPU simulator: Gem5
- ✓ GPU simulator: GPGPU-sim, Accel-Sim, MGPUSim
- ✓ NPU simulator: SCALE-Sim
- ✓ Simpy

- **CAD Tools**
- ✓ Synopsys Design Compiler

- **Languages**
- ✓ Korean
- ✓ English

REFERENCE

eSCaL, Electrical and Electronics Engineering, Yonsei University, Seoul

- Advisor: Professor Won Woo Ro, wro@yonsei.ac.kr
- eSCaL home page: escal.yonsei.ac.kr