

Seonggon Kim

Dept. of Computer Science and Engineering, POSTECH, Republic of Korea
sungonuni@postech.ac.kr

RESEARCH INTEREST

I'm currently focusing on Efficient AI, particularly in enhancing **memory efficiency** and **computation acceleration** during the training and inference of various models (Vision, LLM, Video Generation, etc.) via **quantization** and **low-rank approximation**.

KEYWORD

- Fast and Memory-Efficient Training of Vision models *(CVPR 2025, arXiv preprint)*
- Fast and Memory-Efficient Inference of Vision models *(ICML 2025, WACV 2025)*
- Parameter Efficient Fine-tuning of LLMs *(NeurIPS 2025 Under Review)*
- Fast Sampling of Video Generation Diffusion Models
- CUDA Kernel optimization

EDUCATION

POSTECH

Ph.D. in Computer Science and Engineering

Pohang, Korea

Sep. 2023 - Present

KYUNG HEE UNIVERSITY

B.S. in Computer Science and Engineering

Yongin, Korea

Feb. 2017 - Aug. 2023

PUBLICATIONS

HOT: Hadamard-based Optimized Training

Seonggon Kim, Juncheol Shin, Seung-taek Woo, Eunhyeok Park
Computer Vision and Pattern Recognition (**CVPR 2025**), Nashville.

PTQ4VM: Post-training Quantization for Visual Mamba

Younghyun Cho*, Changhun Lee*, **Seonggon Kim**, Eunhyeok Park
Winter Conference on Applications of Computer Vision (**WACV 2025 Oral**), Tucson.

Merge-Friendly Post-Training Quantization for Multi-Target Domain Adaptation

Juncheol Shin, Minsang Seok, **Seonggon Kim**, Eunhyeok Park
International Conference on Machine Learning (**ICML 2025**), Vancouver.

HoLA: Overcoming the full-finetuning with Hadamard-oriented LoRA

Seonggon Kim, Taehyeon Kim, Byeori Kim, Eunhyeok Park

Neural Information Processing Systems (**NeurIPS 2025**, Under review), San Diego.

HLQ: Fast and Efficient Backpropagation via Hadamard Low-rank Quantization

Seonggon Kim, Eunhyeok Park

arXiv 2406.

EXPERIENCE

SOFTWARE ENGINEER INTERN

Spirent Communications

Jul. 2022 - Feb. 2023

San Jose, CA, USA

- C++ backend engineer of 5G testing program 'LandSlide'
- C++ backend engineer of 5G analysis program 'Drools'

SOFTWARE ENGINEER INTERN

Common Computer

Feb. 2022 - Jun. 2022

Seoul, Korea

- Transplanted Dapp 'AI NFT' from Ethereum to Harmony
- Developed smart contract of 'AI NFT' based on ERC-721
- Presented at ETHDenver 2022

RESEARCH INTERN

SI Analytics

Mar. 2021 - Dec. 2021

Daejeon, Korea

- Semantic segmentation model for land cover classification of Satellite imagery
- Unsupervised, Semi-supervised Learning and Domain Adaptation

AWARDS & HONORS

- ETHDenver 2022 Blockchain Hackathon, NFT project, 3rd Prize Feb. 2022
- CVPR 2021 Earthvision workshop, Land Cover Classification Challenge, Jun. 2021
Selected as the final five teams

TEACHING EXPERIENCE

TEACHING ASSISTANT

POSTECH

Mar. 2025 - June. 2025

Pohang, Korea

- CS311: Computer Architecture [2025-Spring]