# Gordon Euhyun Moon

Sogang University Adam Schall Hall 35, Baekbeom-ro, Mapo-gu, Seoul, Republic of Korea ehmoon@sogang.ac.kr https://gordonmoon.github.io

last updated July 2024

### **EXPERIENCE**

Sogang University Seoul, Korea

Assistant Professor September 2022-present

Department of Computer Science and Engineering

**Korea Aerospace University** Goyang, Korea

Assistant Professor March 2021-August 2022

Department of Software and Department of Artificial Intelligence

Sandia National Laboratories Albuquerque, NM October 2019–January 2021

Postdoctoral Researcher Center for Computing Research

### **EDUCATION**

The Ohio State University Columbus, OH Ph.D. in Computer Science & Engineering

Thesis: "Parallel Algorithms for Machine Learning" Advisor: Professor Ponnuswamy Sadayappan

Committee: Professor Eric Fosler-Lussier and Professor Srinivasan Parthasarathy

**Indiana University** Bloomington, IN

M.S. in Computer Science 2013

Seoul, Korea Yonsei University

B.S. in Computer Science & Industrial System Engineering 2011

## **RESEARCH INTERESTS**

Deep Learning, High-Performance Computing, and Deep Learning Accelerators

2019

#### **SELECTED PUBLICATIONS**

#### **Conference Proceedings**

Sejeong Oh, **Gordon E. Moon** and Sungyong Park, "ML-based Dynamic Operator-Level Query Mapping for Stream Processing Systems in Heterogeneous Computing Environments," *To Appear in Proceedings of the IEEE International Conference on Cluster Computing* (CLUSTER), 2024

Eunji Lee, Yoonsang Han and **Gordon E. Moon**, "Accelerated Block-Sparsity-Aware Matrix Reordering for Leveraging Tensor Cores in Sparse Matrix–Multivector Multiplication," *Proceedings of the 30th International European Conference on Parallel and Distributed Computing* (Euro-Par), 2024

Bokyeong Yoon, Yoonsang Han and **Gordon E. Moon**, "Layer-Wise Sparse Training of Transformer via Convolutional Flood Filling,"

Proceedings of the 28th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD), 2024

Sanha Maeng, **Gordon E. Moon** and Sungyong Park, "Chronica: A Data-Imbalance-Aware Scheduler for Distributed Deep Learning,"

*Proceedings of the 23rd IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGrid), 2023* 

**Gordon E. Moon** and Eric C. Cyr, "Parallel Training of GRU Networks with a Multi-Grid Solver for Long Sequences,"

Proceedings of the 10th International Conference on Learning Representations (ICLR), 2022

Eric Qin, Geonhwa Jeong, William Won, Sheng-Chun Kao, Hyoukjun Kwon, Sudarshan Srinivasan, Dipankar Das, **Gordon E. Moon**, Sivasankaran Rajamanickam and Tushar Krishna, "Extending Sparse Tensor Accelerators to Support Multiple Compression Formats," *Proceedings of the 35th IEEE International Parallel & Distributed Processing Symposium* (IPDPS), 2021

**Gordon E. Moon**, J. Austin Ellis, Aravind Sukumaran-Rajam, Srinivasan Parthasarathy and P. Sadayappan, "ALO-NMF: Accelerated Locality-Optimized Non-negative Matrix Factorization,"

*Proceedings of the 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (KDD), 2020 (acceptance rate:  $216/1279 \approx 16.9\%$ , research track, oral and poster presentations)

#### **Journals**

Bokyeong Yoon, Ah-hyun Lee, Jinsung Kim and **Gordon E. Moon**, "Exploring Attention Sparsity to Accelerate Transformer Training on GPUs," *IEEE Access*, Forthcoming, 2024 (accepted on July 2024)

**Gordon E. Moon**, Hyoukjun Kwon, Geonhwa Jeong, Prasanth Chatarasi, Sivasankaran Rajamanickam and Tushar Krishna, "Evaluating Spatial Accelerator Architectures with Tiled

Matrix-Matrix Multiplication,"

IEEE Transactions on Parallel and Distributed Systems (TPDS), 2022

#### **RESEARCH GRANTS**

- Optimizing Distributed Deep Learning and Federated Learning for Accelerating Large-Scale Deep Learning Models, PI, National Research Foundation of Korea (NRF), ₩800,000,000, April 2024–March 2027
- Architecture-aware Parallel Algorithms for Accelerating Training of Deep Neural Networks, PI, National Research Foundation of Korea (NRF), ₩90,000,000, September 2021–February 2024
- Deep Learning based Route Recommender Systems, Co-PI, Industry-Academy Collaboration R&D, Korea Technology and Information Promotion Agency for SMEs, May 2022–December 2022

#### PROFESSIONAL SERVICE

Program Committee Member

- The 40<sup>th</sup> IEEE International Conference on Computer Design (ICCD 2022), October 2022
- Tenth International Workshop on Accelerating Analytics and Data Management Systems Using Modern Processor and Storage Architectures (ADMS 2019), August 2019

#### **INVITED TALKS**

**Gordon E. Moon**, "Optimizing Sparse Matrix Multiplication Using Tensor Cores to Accelerate DNN Training", HPC Society Conference 2024, The Korean Institute of Information Scientists and Engineers (KIISE), July, 2024

**Gordon E. Moon**, "Accelerated Computing for Deep Learning", Newcomers Session, HPC Society Conference 2022, The Korean Institute of Information Scientists and Engineers (KIISE), July, 2022

**Gordon E. Moon**, "Accelerated Computing for Deep Learning", Newcomers Session, KCC 2022, The Korean Institute of Information Scientists and Engineers (KIISE), June, 2022

**Gordon E. Moon**, "Accelerated Computing for Deep Learning", Newcomers Session, Computer System Society Conference 2022, The Korean Institute of Information Scientists and Engineers (KIISE), February, 2022

**Gordon E. Moon**, "Deep Learning based Recommender Systems for Bicycling Route", NVIDIA GTC'21, A31282 - Regional Panel with Top Startups from Korea, Virtual Conference, November, 2021

Eric C. Cyr and **Gordon E. Moon**, "Parallel-in-Time Training of Recurrent Neural Networks", 2021 AMS Fall Western Virtual Sectional Meeting, SS17B - AMS Special Session on Theoretical and Applied perspectives in Machine Learning, II, October, 2021

**Gordon E. Moon** and Eric C. Cyr, "Parallel Training of an LSTM Network with a Multigrid Solver", SIAM Conference on Computational Science and Engineering (CSE'21), Virtual Conference, March, 2021

Siva Rajamanickam and **Gordon E. Moon**, "Mixed-Precision Schemes for Linear Algebra Kernels on GPUs", SIAM Conference on Computational Science and Engineering (CSE'21), Virtual Conference, March, 2021

**Gordon E. Moon**, "Accelerated Computing for Machine Learning", Sandia National Laboratories, Albuquerque, NM, August, 2019

# **CERTIFICATION/SKILLS**

- Proficient in deep learning frameworks such as PyTorch, TensorFlow, etc.
- Proficient in parallel programming using CUDA, OpenMP, MPI, etc.
- Programming Languages Proficiency: C/C++, Java, Python, and MATLAB