

Sunwoo Lee

Department of Electrical and Computer Engineering
Northwestern University
Evanston, IL
+1) 224-999-5923

sunwoolee1.2014@u.northwestern.edu
<http://sites.northwestern.edu/slz839>

Education

- | | |
|---|----------------|
| Northwestern University
Ph.D. in Computer Engineering
Advisors: Prof. Alok Choudhary and Prof. Wei-keng Liao | 2015 – present |
| Hanyang University , Seoul, South Korea
M.S. in Computer Engineering
Advisor: Prof. Minsoo Ryu | Feb 2009 |
| Hanyang University , Seoul, South Korea
B.S. in Computer Engineering | Feb 2007 |

Research Experience

- | | |
|---|---------------------|
| Northwestern University , Research Assistant
Designed parallel deep neural network training algorithms for scientific image regression in collaboration with ANL and ORNL under RAPIDS (https://rapids.lbl.gov/)
Developed pipelined collective I/O in ROMIO under ECP (https://exascaleproject.org/) | Mar 2015 – present |
| Fermi National Laboratory , Summer Research Intern
Designed a parallel data aggregation strategy for High-Energy Physics data analysis under SciDAC (https://www.scidac.gov/) | Jul 2019 – Sep 2019 |
| Argonne National Laboratory , W. J. Cody Associate
Developed a parallel deep learning software framework using DIY, an object-parallel communication library | Jun 2018 – Aug 2018 |

Professional Experience

- | | |
|--|-------------|
| Samsung Electronics , Memory Division
Software researcher at Memory Solutions Laboratory (MSL) | 2013 – 2015 |
| Humax (<i>alternative military service</i>)
Software developer at Software Laboratory | 2009 – 2013 |

Teaching Experience

Northwestern University, Teaching Assistant

Fall 2019

- CE303: Advanced Digital Design

Honors and Awards

Best Paper Award Finalist

2017

IEEE HiPC 2017 best paper finalist for Parallel Deep Convolutional Neural Network Training by Exploiting the Overlapping of Computation and Communication

Travel Grants

IEEE International Conference on BigData 2019 Travel Grant

2019

Northwestern TGS Travel Grant

2016, 2018, 2019

Publications

1. **Sunwoo Lee**, Qiao Kang, Sandeep Madireddy, Prasanna Balaprakash, Ankit Agrawal, Alok Choudhary, Richard Archibald, and Wei-keng Liao. Improving Scalability of Parallel CNN Training by Adjusting Mini-Batch at Run-Time. IEEE International Conference on BigData, December 2019 (19.3%)
2. **Sunwoo Lee**, Qiao Kang, Reda Al-bahrani, Ankit Agrawal, Alok Choudhary, and Wei-keng Liao, Improving Scalability of Parallel CNN Training by Adaptively Adjusting Parameter Update Frequency. (*Under review by IEEE Transactions on Parallel and Distributed Systems*)
3. Sandeep Madireddy, Ji Hwan Park, **Sunwoo Lee**, Prasanna Balaprakash, Shinjae Yoo, Wei-keng Liao, Cory D. Hauck, M. Paul Laiu, and Richard Archibald. In Situ Compression Artifact Removal in Scientific Data Using Deep Transfer Learning. (*Under review by Neural Networks*)
4. Qiao Kang, **Sunwoo Lee**, Kai-yuan Hou, Robert Ross, Ankit Agrawal, Alok Choudhary, and Wei-keng Liao, Improving MPI Collective I/O Performance with Intra-node Request Aggregation. (*Under review by IEEE Transactions on Parallel and Distributed Systems*)
5. Qiao Kang, Alex Sim, Peter Nugent, **Sunwoo Lee**, Wei-keng Liao, Ankit Agrawal, Alok Choudhary, and Kesheng Wu. Predicting Resource Requirement in Intermediate Palomar Transient Factory Workflow. (*Under review by International Symposium on Cluster, Cloud and Internet Computing (CCGRID) 2020*)
6. **Sunwoo Lee**, Ankit Agrawal, Prasanna Balaprakash, Alok Choudhary, and Wei-keng Liao. Communication-Efficient Parallelization Strategy for Deep Convolutional Neural Network Training. In Workshop on Machine Learning in High-Performance Computing Environments (MLHPC), held in conjunction with International Conference for High Performance Computing, Networking, Storage, and Analysis (SC), November 2018

7. **Sunwoo Lee**, Dipendra Jha, Ankit Agrawal, Alok Choudhary, and Wei-keng Liao. Parallel Deep Convolutional Neural Network Training by Exploiting the Overlapping of Computation and Communication. In Proceedings of the 24th International Conference on High-Performance Computing, Data, and Analytics (HiPC), December 2017 (22.8%)
8. **Sunwoo Lee**, Wei-keng Liao, Ankit Agrawal, Nikos Hardavellas, and Alok Choudhary. Evaluation of K-Means Data Clustering Algorithm on Intel Xeon Phi. In Workshop on Advances in Software and Hardware for Big Data to Knowledge Discovery, held in conjunction with the IEEE International Conference on BigData, December 2016
9. Diana Palsetia, William Hendrix, **Sunwoo Lee**, Ankit Agrawal, Wei-keng Liao, and Alok Choudhary. Parallel Community Detection Algorithm Using a Data Partitioning Strategy with Pairwise Subdomain Duplication. In the 31st International Conference on High Performance Computing (ISC), June 2016

Skills and Qualifications

Programming Language:
C/C++, Python

Deep Learning Software Framework:
Caffe, TensorFlow, PyTorch, Horovod

Parallelization Library:
MPI, OpenMP, Pthreads

Compiling, Debugging, and Analyzer:
GNU and Intel compilers, Intel VTune

I/O Library:
MPI-IO (ROMIO), HDF5, PNetCDF

Container
Docker

Contributions to Open-Source Software

[PCNN]: A software framework for distributed Convolutional Neural Network training

[ph5concat]: Developed a parallel HDF5 file concatenating program

[ROMIO]: Developed a pipelined two-phase I/O for lustre parallel file system