

# Sunwoo Lee

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

[sunwoolee1.2014@u.northwestern.edu](mailto:sunwoolee1.2014@u.northwestern.edu)  
<http://sites.northwestern.edu/slz839>

## Education

---

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

## Research Experience

---

<b>Northwestern University</b> , Research Assistant Designed parallel deep neural network training algorithms for scientific image regression in collaboration with ANL and ORNL under RAPIDS ( <a href="https://rapids.lbl.gov/">https://rapids.lbl.gov/</a> ) Developed pipelined collective I/O in ROMIO under ECP ( <a href="https://exascaleproject.org/">https://exascaleproject.org/</a> )	Mar 2015 – present
<b>Lawrence Berkeley National Laboratory</b> , Summer Research Intern Will study how to improve scalability of Deep Learning-based Cosmology data analysis software	Jun 2020 – Aug 2020
<b>Fermi National Laboratory</b> , Summer Research Intern Designed a parallel data aggregation strategy for High-Energy Physics data analysis under SciDAC ( <a href="https://www.scidac.gov/">https://www.scidac.gov/</a> )	Jul 2019 – Sep 2019
<b>Argonne National Laboratory</b> , 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 2013 – 2015  
Software researcher at Memory Solutions Laboratory (MSL)

**Humax** (*alternative military service*) 2009 – 2013  
Software developer at Software Laboratory

## Teaching Experience

---

**Northwestern University**, Teaching Assistant

- CE303: Advanced Digital Design Fall 2019
- CE510: Social Media Mining Spring 2020

## 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. 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. International Symposium on Cluster, Cloud and Internet Computing (CCGrid) 2020
2. **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 Size at Run-Time. IEEE International Conference on BigData, December 2019 (18.7%)
3. **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
4. **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%)

5. **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
6. 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

## Preprints

---

1. **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 Journal of Parallel and Distributed Computing*)
2. **Sunwoo Lee**, Kai-yuan Hou, Kewei Wang, Saba Sehrish, Marc Paterno, Ankit Agrawal, Alok Choudhary, Robert B. Ross, Quincey Koziol, James Kowalkowski, and Wei-keng Liao, A Case Study on Parallel HDF5 Dataset Concatenation for High-Energy Physics Data Analysis. (In preparation)
3. Qiao Kang, **Sunwoo Lee**, Ankit Agrawal, Alok Choudhary, and Wei-keng Liao, Improving All-to-many Personalized Communication in MPI I/O. (*Under review by SC 20*)
4. 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*)
5. 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*)

## Skills and Qualifications

---

Programming Language:  
C/C++, Python

Parallelization Library:  
MPI, OpenMP, Pthreads

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

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

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

Container  
Docker

## Contributions to Open-Source Software

---

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

[ph5concat]: Developed a parallel HDF5 file concatenation program for large-scale scientific data

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

## References

---

### **Alok Choudhary**

Professor

Electrical and Computer Engineering Department

Northwestern University, Evanston, IL

Email: [choudhar@eecs.northwestern.edu](mailto:choudhar@eecs.northwestern.edu)

### **Wei-keng Liao**

Research Professor

Electrical and Computer Engineering Department

Northwestern University, Evanston, IL

Email: [wkliao@eecs.northwestern.edu](mailto:wkliao@eecs.northwestern.edu)

### **Prasanna Balaprakash**

Computer Scientist

Mathematics and Computer Science Division

Argonne National Laboratory, Lemont, IL

Email: [pbalapra@mcs.anl.gov](mailto:pbalapra@mcs.anl.gov)