

# Sunwoo Lee

---

Postdoctoral Researcher  
ECE Department  
University of Southern California

Tel: (+1) 224-999-5923  
Email: [sunwool@usc.edu](mailto:sunwool@usc.edu)  
<https://sites.google.com/view/sunwoolee>

## Education

---

### Northwestern University

Ph.D. in Computer Engineering  
Advisors: Prof. Alok Choudhary and Prof. Wei-keng Liao

Evanston, IL USA

Sep 2020

### Hanyang University

B.S. and M.S. in Computer Engineering  
Advisor: Prof. Minsoo Ryu

Seoul, South Korea

Feb 2009

## Employment

---

### Inha University

Assistant Professor of Computer Engineering

Incheon, South Korea

Sep 2022 – now

### University of Southern California

Postdoctoral Researcher  
Advisor: Prof. Salman Avestimehr

Los Angeles, CA USA

Oct 2020 – Aug 2022

### Samsung Electronics, Memory Solutions Lab.

System Software Researcher

May 2013 – Jan 2015

### Humax

Software Engineer (Alternative Military Service)

Feb 2009 – Mar 2013

## Research Internship

---

### Lawrence Berkeley National Laboratory

Research Intern

Berkeley, CA USA

Jun 2020 – Aug 2020

### Fermi National Accelerator Laboratory

Research Intern

Batavia, IL USA

Jul 2019 – Sep 2019

### Argonne National Laboratory

Research Intern

Lemont, IL USA

Jun 2018 – Aug 2018

## Teaching Experience

---

University of Southern California

- AEOP Scholarship Program in Data Science

Summer 2021

Northwestern University

- CE501: Social Media Mining
- CE303: Advanced Digital Design

Spring 2020

Fall 2019

## Honors & Awards

---

FL-AAAI Workshop Best Paper Award

2022

- SSFL: Tackling Label Deficiency in Federated Learning via Personalized Self-Supervision

IEEE HiPC Best Paper Finalist

2017

- Parallel Deep Convolutional Neural Network Training by Exploiting the Overlapping of Computation and Communication

## Publications

---

1. **Sunwoo Lee**, Jaeyoung Jeon, Kitae Eom, Chaehwa Jeong, Yongsoo Yang, Ji-Yong Park, Chang Beom Eom, and Hyungwoo Lee, Variance-aware Weight Quantization of Multi-level Resistive Switching Devices based on Pt/LaAlO<sub>3</sub>/SrTiO<sub>3</sub> Heterostructures. *Scientific Reports*, **12**, 1-10, (2022)
2. Kewei Wang, **Sunwoo Lee**, Jan Balewski, Alex Sim, Peter Nugent, Ankit Agrawal, Alok Choudhary, Kesheng Wu, and Wei-keng Liao, Using Multi-resolution Data to Accelerate Neural Network Training in Scientific Applications. *International Symposium on Cluster, Cloud and Internet Computing (CCGrid)*, 2022
3. **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. *Journal of Distributed and Parallel Computing*, **159**, 10-23, (2022)
4. **Sunwoo Lee**, Kai-yuan Hou, Kewei Wang, Saba Sehrish, Marc Paterno, James Kowalkowski, Quincey Koziol, Ross Robert, Ankit Agrawal, Alok Choudhary, and Wei-keng Liao, A Case Study on Parallel HDF5 Dataset Concatenation for High-Energy Physics Data Analysis. *Parallel Computing*, **110**, 102877, (2022)
5. Kai-yuan Hou, Qiao Kang, **Sunwoo Lee**, Ankit Agrawal, Alok Choudhary, and Wei-keng Liao, Supporting Data Compression in PnetCDF, *International Conference on BigData*, December 2021 (19.9%)
6. **Sunwoo Lee**, Qiao Kang, Kewei Wang, Jan Balewski, Alex Sim, Kesheng Wu, Ankit Agrawal, Alok Choudhary, Peter Nugent, and Wei-keng Liao, Asynchronous I/O Strategy for Large-Scale Deep Learning Applications. *International Conference on High-Performance Computing, Data, and Analytics (HiPC)*. December 2021 (22.9%)

7. Reda Al-bahrani, Dipendra Jha, Qiao Kang, **Sunwoo Lee**, Zijiang Yang, Wei-keng Liao, Ankit Agrawal, and Alok Choudhary, SIGRNN: Synthetic minority Instances Generation in imbalanced datasets using a Recurrent Neural Network. *International Conference on Pattern Recognition Applications and Methods*, February 2021
8. **Sunwoo Lee**, Qiao Kang, Ankit Agrawal, Alok Choudhary, and Wei-keng Liao, Communication-Efficient Local SGD for Scalable Deep Learning. *International Conference on Big Data*, December 2020 (15.7%)
9. Sandeep Madireddy, Ji Hwan Park, **Sunwoo Lee**, Prasanna Balaprakash, Shinjae Yoo, Wei-keng Liao, Cory Hauck, M. Paul Laiu, and Richard Archibald, In Situ Compression Artifact Removal in Scientific Data Using Deep Transfer Learning and Experience Replay. *Machine Learning: Science and Technology*, 2020
10. Qiao Kang, **Sunwoo Lee**, Ankit Agrawal, Alok Choudhary, and Wei-keng Liao, Improving All-to-many Personalized Communication in MPI I/O. *International Conference for High Performance Computing, Networking, Storage, and Analysis (SC)*, 2020
11. Qiao Kang, **Sunwoo Lee**, Kai-yuan Hou, Robert Ross, Ankit Agrawal, Alok Choudhary, and Wei-keng Liao, Improving MPI Collective I/O for High Volume Non-contiguous Requests with Intra-node Aggregation. *IEEE Transactions on Parallel and Distributed Systems*, 2020
12. 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
13. **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. *International Conference on Big Data*, December 2019 (18.7%)
14. **Sunwoo Lee**, Ankit Agrawal, Prasanna Balaprakash, Alok Choudhary, and Wei-keng Liao. Communication-Efficient Parallelization Strategy for Deep Convolutional Neural Network Training. *Machine Learning in High-Performance Computing Environments (MLHPC)*, November 2018
15. **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. *International Conference on High-Performance Computing, Data, and Analytics (HiPC)*, December 2017 (22.8%)
16. **Sunwoo Lee**, Wei-keng Liao, Ankit Agrawal, Nikos Hardavellas, and Alok Choudhary. Evaluation of K-Means Data Clustering Algorithm on Intel Xeon Phi. *International Conference on Big Data*, December 2016
17. 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. *International Conference on High Performance Computing (ISC)*, June 2016

18. **Sunwoo Lee**, Byung Kwan Jung, Minsoo Ryu, Seungwon Lee, Extending Component-based Approaches for Multi-threaded Design of Multiprocessor Embedded Software. *International Symposium on Object/Component/Service-Oriented Real-Time Distributed Computing*, 2009

## Pre-prints

---

1. **Sunwoo Lee**, Tuo Zhang, Chaoyang He, and Salman Avestimehr, Layer-wise Model Aggregation for Scalable Federated Learning. *arXiv 2021 (Under review in a top-tier ML conference)*
2. **Sunwoo Lee**, Anit Sahu, Chaoyang He, and Salman Avestimehr, Partial Model Aggregation in Federated Learning: Performance Guarantees. *arXiv 2022 (Under review in IEEE Transactions on Neural Networks and Learning Systems)*
3. **Sunwoo Lee**, Chaoyang He, and Salman Avestimehr, Achieving Small-Batch Accuracy with Large-Batch Scalability via Hessian-Aware Learning Rate Adjustment. *(Under review in Neural Networks)*
4. Chaoyang He, Zhengyu Yang, Erum Mushtaq, **Sunwoo Lee**, Mahdi Soltanolkotabi, Salman Avestimehr, SSFL: Tackling Label Deficiency in Federated Learning via Personalized Self-Supervision. *arXiv 2021 (Under review in a top-tier ML conference)*
5. **Sunwoo Lee**, Jaeyoung Jeon, and Hyungwoo Lee, Probing Oxygen Vacancy Distribution in Oxide Heterostructures by Deep Learning-based Spectral Analysis of Charge Fluctuation. *(Under review in Materials Today Physics (MTP))*

## Invited Talks

---

- U.S. Department of Energy, SciDAC, RAPIDS Institute, Tech Talk: Asynchronous I/O Strategy for Large-Scale Deep Learning Applications, 12/01/2021
- HDF5 User Group Meeting: A Case Study on Parallel HDF5 Dataset Concatenation for Scientific Data Analysis, 10/21/2021
- U.S. Department of Energy, SciDAC, RAPIDS Institute, Tech Talk: Communication-Efficient Local SGD for Scalable Deep Learning, 7/7/2021

## Service

---

- Program Chair Committee of Federated Learning for Natural Language Processing (FL4NLP) workshop held in conjunction with Association for Computational Linguistics (ACL) 2022

## **Skills and Qualifications**

---

### **Programming Language**

C/C++, Python

### **Parallelization Libraries**

MPI, OpenMP, Pthreads

### **Deep Learning Software Frameworks**

TensorFlow, PyTorch, Caffe

### **I/O Libraries**

MPI-I/O (ROMIO), HDF5, NetCDF