Sunwoo Lee

Postdoctoral Researcher
Department of Electrical and Computer Engineering
University of Southern California

Tel: +1-224-999-5923
Email: sunwool@usc.edu
https://sites.google.com/view/sunwoolee

Research Interests

1. Scalable distributed optimization algorithms for large-scale Deep Learning
2. Federated Learning on resource-constrained heterogeneous devices
3. Applied Deep Learning for scientific applications

Education

Northwestern University, USA — Sep 2020
Ph.D. in Computer Engineering
Advisors: Prof. Alok Choudhary and Prof. Wei-keng Liao

Hanyang University, Seoul, South Korea — Feb 2009
B.S. and M.S. in Computer Engineering
Advisor: Prof. Minsoo Ryu

Research Experience

University of Southern California — Oct 2020 – Present
Postdoctoral Researcher
Advisor: Prof. Salman Avestimehr

Lawrence Berkeley National Laboratory — Jun 2020 – Aug 2020
Research Intern

Fermi National Laboratory — Jul 2019 – Sep 2019
Research Intern

Argonne National Laboratory — Jun 2018 – Aug 2018
W.J.Cody Associate (Research Intern)

Professional Experience

Samsung Electronics, Memory Solutions Lab. — May 2013 – Jan 2015
Software Researcher

Humax, Software Lab. — Feb 2009 – Mar 2013
Software Engineer (alternative military service)

Teaching Experience

University of Southern California, Mentor
- AEOP Scholarship Program in Data Science — Summer 2021

Northwestern University, Teaching Assistant
- CE303: Advanced Digital Design — Fall 2019
- CE501: Social Media Mining — Spring 2020

Honors & Awards

FL-AAAI Workshop Best Paper Award, 2022
- SSFL: Tackling Label Deficiency in Federated Learning via Personalized Self-Supervision

IEEE HiPC Best Paper Award Finalist, 2017

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

Publications

1. 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

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. *Journal of Distributed and Parallel Computing*, 2022

3. **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*, 2022

4. 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%)

5. **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%)

6. 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

7. **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%)

8. 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

9. 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

10. 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

11. 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

12. **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%)

13. **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

14. **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%)

15. **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

16. 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

17. **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

Preprints

1. **Sunwoo Lee**, Tuo Zhang, Yue Niu, Saurav Prakash, and Salman Avestimehr, Inclusive Federated Learning: Enabling Weak Client Participation via Partial Training. (*Under review in a top-tier ML conference*)

2. **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*)

3. **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)*

4. **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*)

5. 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*)

6. Tuo Zhang, Lei Gao, **Sunwoo Lee**, Chaoyang He, Mi Zheng, and Salman Avestimehr, Towards Adaptive and Efficient Asynchronous Federated Learning for Mobile Devices. (*Under review in a top-tier ML conference*)

7. Yue Niu, Saurav Prakash, **Sunwoo Lee**, and Salman Avestimehr, Federated Learning at the Edge: Overcoming Constraints via Learning Principal Sub-Models. (*Under review in a top-tier ML conference*)

8. Yue Niu, Zalan Fabian, **Sunwoo Lee**, Mahdi Soltanolkotabi, and Salman Avestimehr, SLIM-QN: A Stochastic, Light, Momentumized Quasi-Newton Optimizer for Deep Neural Networks. (*Under review in a top-tier ML conference*)

9. **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/LaAlO3/SrTiO3 Heterostructures for Memristive Applications. (*Revision submitted to Scientific Reports*)

10. **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 Science Advances (AAAS)*)

Workshop Presentations

1. **Sunwoo Lee**, Anit Sahu, Chaoyang He, and Salman Avestimehr, Partial Model Averaging in Federated Learning: Performance Guarantees and Benefits. *International Workshop on Trustable, Verifiable, and Auditable Federated Learning in conjunction with AAAI*, February 2022 (Oral presentation)

2. Chaoyang He, Zhengyu Yang, Erum Mushtaq, **Sunwoo Lee**, Mahdi Soltanolkotabi, and Salman Avestimehr, SSFL: Tackling Label Deficiency in Federated Learning via Personalized Self-Supervision. *International Workshop on Trustable, Verifiable, and Auditable Federated Learning in conjunction with AAAI,* February 2022 (Oral presentation)

3. Yue Niu, Zalan Fabian, **Sunwoo Lee**, Mahdi Soltanolkotabi, and Salman Avestimehr, SLIM-QN: A Stochastic, Light, Momentumized Quasi-Newton Optimizer for Deep Neural Networks. *Beyond first-order methods in ML Systems in conjunction with ICML*, July 2021

4. **Sunwoo Lee**, Ankit Agrawal, Prasanna Balaprakash, Alok Choudhary, and Wei-keng Liao. Communication-Efficient Parallelization Strategy for Deep Convolutional Neural Network Training. *International Workshop on Machine*

Learning in High-Performance Computing Environments (MLHPC) in conjunction with SC, November 2018

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 & Qualifications		
	Programming Languages *C/C++, Python*	Deep Learning Software Frameworks *TensorFlow, PyTorch, Caffe*
	Parallelization Libraries *MPI, OpenMP, Pthreads*	I/O Libraries *MPI-I/O (ROMIO), HDF5, NetCDF*

References

Alok Choudhary
Henry and Isabelle Dever Professor
Department of Electrical and Computer Engineering
Northwestern University, IL, USA
Email: a-choudhary@northwestern.edu
Phone: +1-847-467-4129

Salman Avestimehr
Dean's Professor
Department of Electrical and Computer Engineering
University of Southern California, CA, USA
Email: avestime@usc.edu
Phone: +1-213-740-7326

Wei-keng Liao
Research Professor
Department of Electrical and Computer Engineering
Northwestern University, IL, USA
Email: wkliao@northwestern.edu
Phone: +1-847-491-2916

Minsoo Ryu
Professor
Department of Computer Science
Hanyang University
Email: msryu@hanyang.ac.kr
Phone: +82-2-2220-4242