

Wei Wen

412-944-8906 | weiwen.web@gmail.com | <http://www.pittnuts.com/> | <https://github.com/wenwei202>

EDUCATION

Ph.D. in Electrical and Computer Engineering, Duke University, USA, 08/2014-12/2019 (Expected)

Advisor: Dr. Hai Li

Research Interest: Deep Learning & Machine Learning & Neuromorphic Computing

M.S. in Electronic and Information Engineering, Beihang University, China, 09/2010-01/2013

B.S. in Electronic and Information Engineering, Beihang University, China, 09/2006-07/2010

RESEARCH STATEMENT

My research is Machine Learning and its applications in Computer Vision and Natural Language Processing. My research focuses on automated machine learning, learning algorithm understanding, efficient deep neural networks, and optimization algorithms for distributed deep learning. I was invited to give talks in UC Berkeley, Cornell University, NeurIPS 2017 oral, etc. I worked with Google Brain, Facebook Research, Microsoft Research, Intel Labs and HP Labs, where I sometimes incorporated my research into AI productions.

INDUSTRIAL RESEARCH & ENGINEERING

Google Brain, Research Intern, Mountain View, CA, USA 05/2019-08/2019

Lead: Quoc Le. Mentors: Pieter-Jan Kindermans & Gabriel Bender

- Automated Machine Learning (AutoML), using machine learning to design machine learning models.

Facebook Research, AI Infra and Applied Machine Learning, Research Intern, Menlo Park, CA, USA 05/2018-08/2018

Mentor: Yangqing Jia

- Personalization and Distributed Machine Learning.

Microsoft Research Redmond, Business AI, Research Intern, Redmond, WA, USA 05/2017-07/2017

- Model Compression and Efficient Recurrent Neural Networks.

HP Labs, Platform Architecture Group, Research Intern, Palo Alto, CA, USA 06/2016-09/2016

- Distributed Deep Learning Systems.

Agricultural Bank of China, Software Engineer Employee, Beijing, China 07/2013-07/2014

Microsoft Research Asia, Mobile and Sensing Systems Group, Research Intern, Beijing, China 04/2013-06/2013

SELECTED PUBLICATIONS

- W. Inkawhich, **W. Wen**, Y. Chen, H. Li. "Feature Space Perturbations Yield More Transferable Adversarial Examples." *CVPR*, 2019.
- W. Wen**, Y. He, S. Rajbhandari, M. Zhang, W. Wang, F. Liu, B. Hu, Y. Chen, H. Li, "Learning Intrinsic Sparse Structures within Long Short-Term Memory", *ICLR*, 2018.
- W. Wen**, C. Xu, F. Yan, C. Wu, Y. Wang, Y. Chen, H. Li, "TernGrad: Ternary Gradients to Reduce Communication in Distributed Deep Learning", *NeurIPS*, 2017. (**Oral, 40/3240=1.2%**) (Integrated into PyTorch/Caffe2)
- W. Wen**, C. Xu, C. Wu, Y. Wang, Y. Chen, H. Li, "Coordinating Filters for Faster Deep Neural Networks", *ICCV*, 2017.
- W. Wen**, C. Wu, Y. Wang, Y. Chen, H. Li, "Learning Structured Sparsity in Deep Neural Networks", *NeurIPS*, 2016.
- J. Park, S. Li, **W. Wen**, P. T. P. Tang, H. Li, Y. Chen, P. Dubey, "Faster CNNs with Direct Sparse Convolutions and Guided Pruning", *ICLR*, 2017.
- C. Wu, **W. Wen**, T. Afzal, Y. Zhang, Y. Chen, H. Li, "A Compact DNN: Approaching GoogLeNet-Level Accuracy of Classification and Domain Adaptation", *CVPR*, 2017.
- Y. Wang, **W. Wen**, L. Song, H. Li, "Classification Accuracy Improvement for Neuromorphic Computing Systems with One-level Precision Synapses", *ASP-DAC*, 2017. (**Best Paper Award**)

SELECTED HONORS & AWARDS

- Best Paper Award, Asia and South Pacific Design Automation Conference (ASP-DAC), IEEE 2017
- Best Paper Nomination, Design Automation Conference (DAC), IEEE 2016
- Best Paper Nomination, Design Automation Conference (DAC), IEEE 2015
- ICLR Travel Award 2018
- Graduate Student Conference Travel Fellowship, Duke ECE 2017
- NeurIPS Travel Award 2017
- NeurIPS Travel Award 2016
- National Scholarship (3/233), Ministry of Education China 2009
- Second Prize, National College Physics Competition China 2007

SKILLS

- Machine learning: TensorFlow, PyTorch, Caffe
- Languages: Python, C/C++/CUDA C
- Android Development (with Google Play publications)

TEACHING

- Teach Assistant, CEE 690/ECE 590: Introduction to Deep Learning, Duke University, Fall 2018
- Teach Assistant, STA561/COMPSCI571/ECE682: Probabilistic Machine Learning, Duke University, Spring 2019

TALKS

- UC Berkeley, Scientific Computing and Matrix Computations Seminar, “On Matrix Sparsification and Quantization for Efficient and Scalable Deep Learning”, 10/10/2018
- Cornell University, AI Seminar, “Efficient and Scalable Deep Learning”, 10/05/2018
- NeurIPS 2017, TernGrad: Ternary Gradients to Reduce Communication in Distributed Deep Learning, 12/6/2017
- Alibaba DAMO Academy, “Deep Learning in Cloud-Edge AI Systems”, SunnyVale, CA, 06/28/2018

SERVICE

- Paper reviewer, NeurIPS, ICML, ICLR, CVPR, ICCV, TPAMI, IJCV, TNNLS, TCAD, Neurocomputing, etc.
- Activity volunteer, Machine Learning for Girls, FEMMES (Female Excelling More in Math, Engineering, and Science) Capstone at Duke University, 02/2018
- Conference volunteer, Embedded Systems Week (ESWEEK), Pittsburgh, PA, USA, 10/2016