

Wei WEN

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

EDUCATION

Duke University	Durham, NC, United States	08/2017-12/2019 (Expected)
University of Pittsburgh	Pittsburgh, PA, United States	09/2014-08/2017 (Transferred to Duke U)
<u>Ph.D.</u>	Electrical and Computer Engineering	Supervisor: Dr. Hai Li
	Research Area: Deep Learning & Machine Learning & Neuromorphic Computing	
Beihang University	Beijing, China,	09/2006-07/2010 (B.S.), 09/2010-01/2013(M.S.)
<u>B.S., M.S.</u>	Electronic and Information Engineering	

RESEARCH STATEMENT

My research is Machine Learning and its applications in Computer Vision and Natural Language Processing. Recently, I focus on *efficient deep learning on the edge, optimization algorithms for distributed machine learning, and learning algorithm understanding*. I am also interested in AutoML and interpretable machine learning. I was invited to give talks in UC Berkeley, Cornell University, etc. I closely worked with Facebook Research, Microsoft Research, Intel Labs, and HP Labs, where I incorporated my research into industrial AI productions, including Facebook Applied Machine Learning, Microsoft Bing, Intel Nervana & SkimCaffe, etc. I am a contributor of PyTorch/Caffe2.

SELECTED PUBLICATIONS

- **W. Wen**, Y. Wang, F. Yan, C. Xu, Y. Chen, H. Li, “*SmoothOut: Smoothing Out Sharp Minima for Generalization in Large-Batch Deep Learning*”, **Preprint**, 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*”, **NIPS**, 2017. (**Oral, 40/3240=1.2%**) (Integrated into [PyTorch/Caffe2](#))
- **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, 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*”, **NIPS**, 2016. (Integrated into [Intel AI](#))
- 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.
- 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**)

INDUSTRY EXPERIENCE

Facebook Research, Menlo Park, CA, USA	05/2018-08/2018
Research Intern. Mentor: Yangqing Jia	
• Caffe2, Personalization and Distributed Machine Learning.	
Microsoft Research & Business AI, Redmond, WA, USA	05/2017-07/2017
Research Intern. Mentor: Yuxiong He	
• Machine Reading Comprehension and Recurrent Neural Networks.	
HP Labs, Platform Architecture Group, Palo Alto, CA, USA	06/2016-09/2016
Research Intern. Mentor: Cong Xu & supervisor: Paolo Faraboschi	
• Distributed Deep Learning Systems.	

INDUSTRY EXPERIENCE (Continued)

<i>Agricultural Bank of China, Software Engineer Employee, Beijing, China</i>	<i>07/2013-07/2014</i>
<i>Microsoft Research Asia, Mobile and Sensing Systems Group, Beijing, China</i>	<i>04/2013-06/2013</i>
<i>Tencent Inc., Advertising Platform and Products Division, Beijing, China</i>	<i>07/2012-09/2012</i>

SKILLS

- Machine learning: TensorFlow, PyTorch, Caffe2, Caffe
- Languages: C/C++/CUDA C, Python
- Android Development (with [Google Play](#) publications)

SELECTED HONORS & AWARDS

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

ACADEMIC ACTIVITIES

- Talks in UC Berkeley, Cornell University, NIPS 2017, Facebook Research, Microsoft Research, Alibaba DAMO Academy, HP Labs, etc.
- Paper reviewers of NIPS, TCAD, TNNLS, 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