Huan Zhang

University of California, Los Angeles
http://www.huan-zhang.com

☎ *530-601-3829*
✉ *huanzhang@ucla.edu*
404 Westwood Plaza
Engineering VI
Los Angeles, CA 90095-1596
http://www.huan-zhang.com

Education

2012-present **Ph.D. in Computer Science (in progress)**, UCLA (2018-), UC Davis (2012-2018).
- *Advisor:* Prof. Cho-Jui Hsieh (2015-present) and Prof. Venkatesh Akella (2012-2018)
- *Area:* large-scale machine learning and optimization. Some recent works include exploring the robustness issues and adversarial examples in deep neural networks from an optimization perspective, accelerating gradient boosted decision tree (GBDT) training on GPUs, and designing large-scale asynchronous parallel stochastic gradient descent and coordinate descent optimizers in multi-core and distributed settings especially for deep learning training.

2012-2014 **M.S. in Computer Engineering**, University of California, Davis.
- *Advisor:* Prof. Venkatesh Akella (akella@ucdavis.edu)
- *Area:* computer architecture and parallel computing, especially on improving scalability of parallel workloads on heterogeneous multi-core processors and distributed systems.

2008-2012 **Bachelor of Engineering**, Zhejiang University, China.
- *Major:* Information Engineering

Work Experience

11/2017-3/2018 **Internship in Amazon A9.com Search Lab**, Palo Alto, CA.
- *Mentor:* Inderjit Dhillon (isd@a9.com)

I am working on the design and implementation of a deep-learning based product search suggestion system used on amazon.com production website. Given an incomplete search-bar query, our LSTM-based model provides candidates to complete the query in real-time.

6/2017-11/2017 **Internship in IBM T.J. Watson Research Center**, Yorktown Heights, NY.
- *Mentor:* Jinfeng Yi (jinfengy@us.ibm.com), Pin-Yu Chen (Pin-Yu.Chen@ibm.com)

4/2018-6/2018 I worked on exploring the robustness issues of deep neural networks by crafting adversarial examples in a black-box setting using zeroth order optimization (ZOO) based attack, achieving very high success rates. I also investigated the robustness of image captioning system with recurrent neurons and attention mechanism and proposed the Show-and-Fool attack.

6/2015-8/2015, 6/2013-9/2013 **Internship in Alcatel-Lucent Bell Labs**, Murray Hill, NJ.
- *Mentor:* Noriaki Kaneda and Young-Kai Chen (ykchen@bell-labs.com)

My work involves building a high speed optical communication system on FPGA with real-time coherent OFDM modulation. Results were accepted as a conference and a journal paper.

Skills

- Languages: C/C++, Python, Julia, Matlab, Lua, Bash, Verilog, VHDL
- Parallel programming on CPU/GPUs: POSIX threads, OpenMP, MPI, OpenCL, CUDA
- Deep Learning: TensorFlow, PyTorch, Torch, Theano, Keras
- Program profiling and tuning: gprof, Intel VTune Amplifier, OProfile
- Linux development and embedded system programming

List of Publications

2018 **Efficient Neural Network Robustness Certification with General Activation Functions**, Advances in Neural Information Processing Systems (NIPS), 2018, Huan Zhang*, Tsui-Wei Weng*, Pin-Yu Chen, Cho-Jui Hsieh, Luca Daniel. (* Equal contribution).

2018 **Is Robustness the Cost of Accuracy? Lessons Learned from 18 Deep Image Classifiers**, European Conference on Computer Vision (ECCV), 2018, Dong Su*, Huan Zhang*, Hongge Chen, Jinfeng Yi, Pin-Yu Chen, Yupeng Gao. (* Equal contribution).

2018 **Towards Robust Neural Networks via Random Self-ensemble**, European Conference on Computer Vision (ECCV), 2018, Xuanqing Liu, Minhao Cheng, Huan Zhang, Cho-Jui Hsieh.

2018 **Realtime query completion via deep language models**, SIGIR Workshop On eCommerce, 2018, Po-Wei Wang, Huan Zhang, Vijai Mohan, Inderjit S. Dhillon and J. Zico Kolter.

2018 **Towards Fast Computation of Certified Robustness for ReLU Networks**, International Conference on Machine Learning (ICML), 2018, Tsui-Wei Weng*, Huan Zhang*, Hongge Chen, Zhao Song, Cho-Jui Hsieh, Duane Boning, Inderjit S. Dhillon, Luca Daniel. (* Equal contribution).

2018 **Attacking Visual Language Grounding with Adversarial Examples: A Case Study on Neural Image Captioning**, 56th Annual Meeting of the Association for Computational Linguistics (ACL), 2018, Hongge Chen*, Huan Zhang*, Pin-Yu Chen, Jinfeng Yi and Cho-Jui Hsieh (*** Equal contribution**).

2018 **Evaluating the Robustness of Neural Networks: An Extreme Value Theory Approach**, *International Conference on Learning Representations (ICLR)*, 2018, Tsui-Wei Weng*, Huan Zhang*, Pin-Yu Chen, Jinfeng Yi, Dong Su, Yupeng Gao, Cho-Jui Hsieh, Luca Daniel (*** Equal contribution**).

2017 **ZOO: Zeroth Order Optimization based Black-box Attacks to Deep Neural Networks without Training Substitute Models**, *10th ACM Workshop on Artificial Intelligence and Security*, 2017, Pin-Yu Chen*, Huan Zhang*, Yash Sharma, Jinfeng Yi and Cho-Jui Hsieh (*** Equal contribution**).

2017 **EAD: Elastic-Net Attacks to Deep Neural Networks via Adversarial Examples**, *In AAAI Conference on Artificial Intelligence* (AAAI), 2018, Pin-Yu Chen, Yash Sharma, Huan Zhang, Jinfeng Yi and Cho-Jui Hsieh.

2017 **Towards Robust Neural Networks via Random Self-ensemble**, arXiv:1712.00673, Xuanqing Liu, Minhao Cheng, Huan Zhang, Cho-Jui Hsieh.

2017 **GPU-acceleration for Large-scale Tree Boosting**, arXiv Preprint:1706.08359, 2017, Huan Zhang, Si Si and Cho-Jui Hsieh.

2017 **Can Decentralized Algorithms Outperform Centralized Algorithms? A Case Study for Decentralized Parallel Stochastic Gradient Descent**, *Advances in Neural Information Processing Systems (NIPS)*, 2017, Xiangru Lian, Ce Zhang, Huan Zhang, Cho-Jui Hsieh, Wei Zhang and Ji Liu (**Oral paper**).

2017 **Gradient Boosted Decision Trees for High Dimensional Sparse Output**, *34th International Conference on Machine Learning (ICML)*, 2017, Si Si, Huan Zhang, Sathiya Keerthi, Dhruv Mahajan, Inderjit Dhillon and Cho-Jui Hsieh.

2016 **HogWild++: A New Mechanism for Decentralized Asynchronous Stochastic Gradient Descent**, *IEEE International Conference on Data Mining (ICDM)*, 2016, Huan Zhang, Cho-Jui Hsieh and Venkatesh Akella.

2016 **Fixing the Convergence Problems in Parallel Asynchronous Dual Coordinate Descent**, *IEEE International Conference on Data Mining (ICDM)*, 2016, Huan Zhang and Cho-Jui Hsieh.

2016 **Sublinear Time Orthogonal Tensor Decomposition**, *Advances in Neural Information Processing Systems (NIPS)*, 2016, Zhao Song, David P. Woodruff and Huan Zhang.

2016 **A Comprehensive Linear Speedup Analysis for Asynchronous Stochastic Parallel Optimization from Zeroth-Order to First-Order**, *Advances in Neural Information Processing Systems (NIPS)*, 2016, Xiangru Lian, Huan Zhang, Cho-Jui Hsieh, Yijun Huang and Ji Liu.

2015 **Field Demonstration of 100-Gb/s Real-Time Coherent Optical OFDM Detection**, *Journal of Lightwave Technology*, Vol. 33, No. 7, April 1 2015, Noriaki Kaneda, Timo Pfau, Huan Zhang, Jeffrey Lee, Young-Kai Chen, Chun Ju Youn, Yong Hwan Kwon, Eun Soo Num, and S. Chandrasekhar.

2014 **Burst Mode Processing: An Architectural Framework for Improving Performance in Future Chip Microprocessors**, *Workshop on Managing Overprovisioned Systems, Co-located with ASPLOS-19*, April 2014, Huan Zhang, Rajeevan Amirtharajah, Christopher Nitta, Matthew Farrens and Venkatesh Akella.

2013 **HySIM: Towards a Scalable, Accurate and Fast Simulator for Manycore Processors**, *3rd Workshop on the Intersections of Computer Architecture and Reconfigurable Logic, Co-located with MICRO-46*, December 2013, Kramer Straube, Huan Zhang, Christopher Nitta, Matthew Farrens and Venkatesh Akella.

2013 **Spectral and Spatial 2D Fragmentation-Aware Routing and Spectrum Assignment Algorithms in Elastic Optical Networks**, *IEEE/OSA Journal of Optical Communications and Networking*, Yawei Yin, Huan Zhang, Mingyang Zhang, Ming Xia, Zuqing Zhu, Stefan Dahlfort and S. J. B. Yoo.

Open Source Projects

2017 **LightGBM on GPU**, `https://github.com/huanzhang12/lightgbm-gpu`.
LightGBM is a popular tree boosting package by Microsoft with high efficiency on large-scale datasets. I accelerated its decision tree construction process on GPUs using a novel histogram based algorithm which efficiently exploits the GPU's capability to approximately find the best split. Experimental results show that it is up to 7 to 8 times faster than LightGBM and 20 to 30 times faster than XGBoost comparing with a 28-core CPU. My code (about 5,000 lines C++) reaches production quality and has been merged into the LightGBM official repository.

List of Awards

2018 **IBM PhD Fellowship**, 2018-2019.

Student Travel Award, NIPS 2016, 2017; ICDM 2016; ICLR 2018; ACM CCS 2017.

2010 **Meritorious Winner**, The U.S. Mathematical Contest in Modeling.

2009 **First Prize**, China Undergraduate Mathematical Contest in Modeling.

2009 **Second Prize**, East China Undergraduate Mathematical Modeling Contest.