

Huan Zhang

262 Coordinated Science Laboratory
1308 W Main St
Urbana, IL 61801

✉ huan@huan-zhang.com

🎓 [Google Scholar](#)

🌐 huan-zhang.com

🐙 [Github](#)

PROFESSIONAL APPOINTMENTS

Assistant Professor, University of Illinois Urbana-Champaign (UIUC) Urbana, IL (July 2023 - Present).

Department of Electrical and Computer Engineering

Department of Computer Science (affiliated)

Coordinated Science Laboratory (affiliated)

Post-Doctoral Fellow, Carnegie Mellon University (CMU)

Pittsburgh, PA (Jan 2021 - July 2023).

Supervisor: [J. Zico Kolter](#)

EDUCATION

Ph.D. in Computer Science

UCLA (2020).

Advisor: [Cho-Jui Hsieh](#)

Thesis: Machine Learning with Provable Robustness Guarantees

M.S. in Computer Engineering

UC Davis (2014).

Advisor: [Venkatesh Akella](#)

Area: computer architecture, parallel computing, scalable machine learning.

Bachelor of Engineering

Zhejiang University (2012).

Major: Information & Optical Engineering

EMPLOYMENT

Internship at Google DeepMind

London, UK, Summer 2019.

Mentor: [Krishnamurthy \(Dj\) Dvijotham](#) and [Po-Sen Huang](#)

Internship at Microsoft Research

Redmond, WA, Summer 2018.

Mentor: [Pengchuan Zhang](#) and [Lin Xiao](#)

Internship at Amazon A9.com

Palo Alto, CA, Spring 2018.

Mentor: [Inderjit Dhillon](#)

Internship at IBM T.J. Watson Research Center

Yorktown Heights, NY, Summer 2017, 2018.

Mentor: [Jinfeng Yi](#) and [Pin-Yu Chen](#)

Internship at Nokia Bell Labs

Murray Hill, NJ, Summer 2013, 2015.

Mentor: [Noriaki Kaneda](#) and [Young-Kai Chen](#)

AWARDS AND HONORS

2022 **Schmidt Futures AI2050 Early Career Fellowship**, with a \$300,000 research grant

2023 **First Place**, Forth International Verification of Neural Networks Competition ([VNN-COMP 2023](#))

- 2022 **First Place**, Third International Verification of Neural Networks Competition ([VNN-COMP 2022](#)), I led a multi-institutional team that developed the α, β -CROWN verification toolkit, outperforming 11 competing tools from Stanford, Imperial College London, ETH Zurich, and other prestigious universities on various benchmarks from computer vision, robotics, and aerospace applications.
- 2021 **First Place**, Second International Verification of Neural Networks Competition ([VNN-COMP 2021](#))
- 2021 **Adversarial Machine Learning (AdvML) Rising Star Award**, sponsored by [MIT-IBM Watson AI Lab](#)
- 2018 **IBM PhD Fellowship**, with a stipend of \$60,000
- 2011 **National Merit Scholarship**, Ministry of Education, China, awarded to top 2% students.
- 2010 **Meritorious Winner**, [The U.S. Mathematical Contest in Modeling, 2010](#).
- 2009 **National Merit Scholarship**, Ministry of Education, China, awarded to top 2% students.
- 2009 **First Prize**, China Undergraduate Mathematical Contest in Modeling, 2009.
- 2009 **Second Prize**, East China Undergraduate Mathematical Contest in Modeling, 2009.

PROFESSIONAL SERVICES

Workshop Organization

- 2022 **Lead organizer**, [Trustworthy and Socially Responsible Machine Learning \(TSRML\)](#), co-located with NeurIPS 2022.
- 2022 **Lead organizer**, [1st Workshop on Formal Verification of Machine Learning](#), co-located with ICML 2022.
- 2022 **Lead organizer**, [Queer in AI Workshop](#), co-located with ICML 2022.
- 2022 **Co-organizer**, [Workshop on Socially Responsible Machine Learning](#), co-located with ICLR 2022.
- 2021 **Co-organizer**, [Workshop on Security and Reliability of Machine Learning](#), co-located with 19th International Symposium on Automated Technology for Verification and Analysis (ATVA 2021).

Conference/Journal Reviewing and Journal Editing

- 2022 **Guest Journal Editor**, Special Issue “Black-Box Algorithms and Their Applications”, MDPI Algorithms.
 - 2021 **Guest Journal Editor**, Trustworthy Machine Learning Research Topic, Frontiers in Big Data, 2021.
 - 2021 **Senior Program Committee/Area Chair**, AAAI 2022.
- Conference Paper Reviewer/Program Committee**, NIPS 2016, 2018, 2019, 2020, 2021, 2022; ICML 2019, 2020, 2021, 2022, 2023; ICLR 2019, 2020, 2021, 2022, 2023; AAAI 2020, 2021, 2023; UAI 2020, 2021; AISTATS 2021, 2022; CVPR 2020, 2021. USENIX 2020.
- Journal Reviewer**, Journal of Machine Learning Research (JMLR), IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI), Springer Journal of Machine Learning.

RECENT INVITED TALKS

- 2022 **Caltech**, DOLCIT Seminar Series, title “Formal Verification of Deep Neural Networks: Challenges and Recent Advances”.
- 2022 **Princeton University**, [Virtual Seminars on Security and Privacy in Machine Learning \(SPML\)](#), title “Formal Verification of Deep Neural Networks: Challenges and Recent Advances”.
- 2022 **Johns Hopkins University**, [Institute for Assured Autonomy Seminar Series](#), title “How Can We Trust a Black-box? A Quest for Scalable and Powerful Neural Network Verifiers”.

- 2022 **Carnegie Mellon University (CMU)**, [CMU AI Seminar](#), title “How Can We Trust a Black-box? A Quest for Scalable and Powerful Neural Network Verifiers”.
- 2021 **University of California Santa Barbara (UCSB)**, Computer Science Colloquium, title “How Can We Trust a Black-box? A Quest for Scalable and Powerful Neural Network Verifiers”.
- 2021 **Northeastern University**, Security Seminar, title “How Can We Trust a Black-box? A Quest for Scalable and Powerful Neural Network Verifiers”.
- 2021 **University of Illinois at Urbana-Champaign (UIUC)**, [Computer Science Speakers Series](#), title “How Can We Trust a Black-box? A Quest for Scalable and Powerful Neural Network Verifiers”.
- 2021 **University of Southern California (USC)**, [USC AI Seminar](#), title “How Can We Trust a Black-box? A Quest for Scalable and Powerful Neural Network Verifiers”.
- 2021 **Lorentz Center Workshop on Robust Artificial Intelligence**, title “Robust Reinforcement Learning Against Adversarial Perturbations on State Observations”.
- 2021 **Bosch Center for Artificial Intelligence (BCAI)**, title “Complete and Incomplete Neural Network Verification with Efficient Bound Propagations”.
- 2020 **3rd Workshop on Formal Methods for ML-Enabled Autonomous Systems**, title “Robustness Verification for Ensemble Stumps and Trees”.

— PUBLICATIONS

Peer-reviewed Conference papers (* indicates **co-first** authors)

- [1] **General Cutting Planes for Bound-Propagation-Based Neural Network Verification**
[Huan Zhang*](#), [Shiqi Wang*](#), [Kaidi Xu*](#), [Linyi Li](#), [Bo Li](#), [Suman Jana](#), [Cho-Jui Hsieh](#) and [Zico Kolter](#)
Advances in Neural Information Processing Systems (NeurIPS), 2022 (winner of VNN-COMP 2022)
- [2] **Are AlphaZero-like Agents Robust to Adversarial Perturbations?**
[Li-Cheng Lan](#), [Huan Zhang](#), [Ti-Rong Wu](#), [Meng-Yu Tsai](#), [I-Chen Wu](#), [Cho-Jui Hsieh](#)
Advances in Neural Information Processing Systems (NeurIPS), 2022
- [3] **Efficiently Computing Local Lipschitz Constants of Neural Networks via Bound Propagation**
[Zhouxing Shi](#), [Yihan Wang](#), [Huan Zhang](#), [Zico Kolter](#), [Cho-Jui Hsieh](#)
Advances in Neural Information Processing Systems (NeurIPS), 2022
- [4] **A Branch and Bound Framework for Stronger Adversarial Attacks of ReLU Networks**
[Huan Zhang*](#), [Shiqi Wang*](#), [Kaidi Xu](#), [Yihan Wang](#), [Suman Jana](#), [Cho-Jui Hsieh](#) and [Zico Kolter](#)
International Conference on Machine Learning (ICML), 2022
- [5] **Linearity Grafting: Relaxed Neuron Pruning Helps Certifiable Robustness**
[Tianlong Chen*](#), [Huan Zhang*](#), [Zhenyu Zhang](#), [Shiyu Chang](#), [Sijia Liu](#), [Pin-Yu Chen](#) and [Zhangyang Wang](#)
International Conference on Machine Learning (ICML), 2022
- [6] **δ -SAM: Sharpness-Aware Minimization with Dynamic Reweighting**
[Wenxuan Zhou](#), [Fangyu Liu](#), [Huan Zhang](#), [Muhao Chen](#)
Findings in Empirical Methods in Natural Language Processing (EMNLP), 2022
- [7] **ViP: Unified Certified Detection and Recovery for Patch Attack with Vision Transformers**
[Junbo Li](#), [Huan Zhang](#), [Cihang Xie](#)
European Conference on Computer Vision (ECCV), 2022

- [8] **COPA: Certifying Robust Policies for Offline Reinforcement Learning against Poisoning Attacks**
Fan Wu, Linyi Li, Chejian Xu, Huan Zhang, Bhavya Kailkhura, Krishnaram Kenthapadi, Ding Zhao and Bo Li
International Conference on Learning Representations (ICLR), 2022
- [9] **Beta-CROWN: Efficient Bound Propagation with Per-neuron Split Constraints for Complete and Incomplete Neural Network Verification**
Shiqi Wang*, Huan Zhang*, Kaidi Xu*, Xue Lin, Suman Jana, Cho-Jui Hsieh and Zico Kolter
Advances in Neural Information Processing Systems (NeurIPS), 2021 (winner of VNN-COMP 2021)
- [10] **Training Certifiably Robust Neural Networks with Efficient Local Lipschitz Bounds**
Yujia Huang, Huan Zhang, Yuanyuan Shi, Zico Kolter and Anima Anandkumar
Advances in Neural Information Processing Systems (NeurIPS), 2021
- [11] **Robustness Between the Worst and Average Case**
Leslie Rice, Anna Bair, Huan Zhang, and Zico Kolter
Advances in Neural Information Processing Systems (NeurIPS), 2021
- [12] **Fast Certified Robust Training via Better Initialization and Shorter Warmup**
Zhouxing Shi*, Yihan Wang*, Huan Zhang, Jinfeng Yi and Cho-Jui Hsieh
Advances in Neural Information Processing Systems (NeurIPS), 2021
- [13] **Robust Reinforcement Learning on State Observations with Learned Optimal Adversary**
Huan Zhang*, Hongge Chen*, Duane Boning, Cho-Jui Hsieh
International Conference on Learning Representations (ICLR), 2021
- [14] **Fast and Complete: Enabling Complete Neural Network Verification with Rapid and Massively Parallel Incomplete Verifiers**
Kaidi Xu*, Huan Zhang*, Shiqi Wang, Yihan Wang, Suman Jana, Xue Lin, Cho-Jui Hsieh
International Conference on Learning Representations (ICLR), 2021
- [15] **Double Perturbation: On the Robustness of Robustness and Counterfactual Bias Evaluation**
Chong Zhang, Jieyu Zhao, Huan Zhang, Kai-Wei Chang, Cho-Jui Hsieh
Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 2021
- [16] **Robust Deep Reinforcement Learning against Adversarial Perturbations on State Observations**
Huan Zhang*, Hongge Chen*, Chaowei Xiao, Bo Li, Mingyan Liu, Duane Boning, Cho-Jui Hsieh
Advances in Neural Information Processing Systems (NeurIPS), 2020 (Spotlight paper, 100+ citations)
- [17] **Automatic Perturbation Analysis for Scalable Certified Robustness and Beyond**
Kaidi Xu*, Zhouxing Shi*, Huan Zhang*, Yihan Wang, Minlie Huang, Kai-Wei Chang, Bhavya Kailkhura, Xue Lin, Cho-Jui Hsieh
Advances in Neural Information Processing Systems (NeurIPS), 2020 (100+ citations)
- [18] **An Efficient Adversarial Attack for Tree Ensembles**
Chong Zhang, Huan Zhang, Cho-Jui Hsieh
Advances in Neural Information Processing Systems (NeurIPS), 2020
- [19] **Reducing Sentiment Bias in Language Models via Counterfactual Evaluation**
Po-Sen Huang*, Huan Zhang*, Ray Jiang, Robert Stanforth, Johannes Welbl, Jack Rae, Vishal Maini, Dani Yogatama, Pushmeet Kohli
Findings in Empirical Methods in Natural Language Processing (EMNLP), 2020

- [20] **On ℓ_p -norm Robustness of Ensemble Decision Stumps and Trees**
Yihan Wang, Huan Zhang, Hongge Chen, Duane Boning and Cho-Jui Hsieh
International Conference on Machine Learning (ICML), 2020
- [21] **Towards Stable and Efficient Training of Verifiably Robust Neural Networks**
Huan Zhang, Hongge Chen, Chaowei Xiao, Sven Gowal, Robert Stanforth, Bo Li, Duane Boning, Cho-Jui Hsieh
International Conference on Learning Representations (ICLR), 2020 (**200+ citations**)
- [22] **Robustness Verification for Transformers**
Zhouxing Shi, Huan Zhang, Kai-Wei Chang, Minlie Huang, Cho-Jui Hsieh
International Conference on Learning Representations (ICLR), 2020
- [23] **MACER: Attack-free and Scalable Robust Training via Maximizing Certified Radius**
Runtian Zhai, Chen Dan, Di He, Huan Zhang, Boqing Gong, Pradeep Ravikumar, Cho-Jui Hsieh, Liwei Wang
International Conference on Learning Representations (ICLR), 2020 (**100+ citations**)
- [24] **Robustness Verification of Tree-based Models**
Hongge Chen*, Huan Zhang*, Si Si, Yang Li, Duane Boning, Cho-Jui Hsieh.
Advances in Neural Information Processing Systems (NeurIPS), 2019
- [25] **A Convex Relaxation Barrier to Tight Robustness Verification of Neural Networks**
Hadi Salman, Greg Yang, Huan Zhang, Cho-Jui Hsieh, Pengchuan Zhang
Advances in Neural Information Processing Systems (NeurIPS), 2019 (**100+ citations**)
- [26] **Provably Robust Deep Learning via Adversarially Trained Smoothed Classifiers**
Hadi Salman, Greg Yang, Jerry Li, Pengchuan Zhang, Huan Zhang, Ilya Razenshteyn, Sebastien Bubeck
Advances in Neural Information Processing Systems (NeurIPS), 2019 (**Spotlight, 300+ citations**)
- [27] **The Limitations of Adversarial Training and the Blind-Spot Attack**
Huan Zhang*, Hongge Chen*, Zhao Song, Duane Boning, Inderjit Dhillon, Cho-Jui Hsieh
International Conference on Learning Representations (ICLR), 2019 (**100+ citations**)
- [28] **Query-Efficient Hard-label Black-box Attack: An Optimization-based Approach**
Minhao Cheng, Thong Le, Pin-Yu Chen, Huan Zhang, Jinfeng Yi, Cho-Jui Hsieh
International Conference on Learning Representations (ICLR), 2019 (**300+ citations**)
- [29] **Structured Adversarial Attack: Towards General Implementation and Better Interpretability**
Kaidi Xu, Sijia Liu, Pu Zhao, Pin-Yu Chen, Huan Zhang, Quanfu Fan, Deniz Erdogmus, Yanzhi Wang, Xue Lin
International Conference on Learning Representations (ICLR), 2019
- [30] **Robust Decision Trees Against Adversarial Examples**
Hongge Chen, Huan Zhang, Duane Boning, Cho-Jui Hsieh
International Conference on Machine Learning (ICML), 2019 (**long oral presentation**)
- [31] **Evaluating Robustness of Deep Image Super-Resolution Against Adversarial Attacks**
Jun-Ho Choi, Huan Zhang, Jun-Hyuk Kim, Cho-Jui Hsieh, Jong-Seok Lee
International Conference on Computer Vision (ICCV), 2019
- [32] **Second Rethinking of Network Pruning in the Adversarial Setting**
Shaokai Ye, Kaidi Xu, Sijia Liu, Hao Cheng, Jan-Henrik Lambrechts, Huan Zhang, Aojun Zhou, Kaisheng Ma, Yanzhi Wang, Xue Lin
International Conference on Computer Vision (ICCV), 2019

- [33] **RecurJac: An Efficient Recursive Algorithm for Bounding Jacobian Matrix of Neural Networks and Its Applications**
Huan Zhang, Pengchuan Zhang, Cho-Jui Hsieh
AAAI Conference on Artificial Intelligence (AAAI), 2019
- [34] **AutoZOOM: Autoencoder-based Zeroth Order Optimization Method for Attacking Black-box Neural Networks**
Chun-Chen Tu, Paishun Ting, Pin-Yu Chen, Sijia Liu, Huan Zhang, Jinfeng Yi, Cho-Jui Hsieh, Shin-Ming Cheng
AAAI Conference on Artificial Intelligence (AAAI), 2019 (**200+ citations**)
- [35] **Efficient Neural Network Robustness Certification with General Activation Functions**
Huan Zhang*, Tsui-Wei Weng*, Pin-Yu Chen, Cho-Jui Hsieh, Luca Daniel.
Advances in Neural Information Processing Systems (NIPS), 2018 (**400+ citations**)
- [36] **Towards Fast Computation of Certified Robustness for ReLU Networks**
Tsui-Wei Weng*, Huan Zhang*, Hongge Chen, Zhao Song, Cho-Jui Hsieh, Duane Boning, Inderjit S. Dhillon, Luca Daniel.
International Conference on Machine Learning (ICML), 2018 (**500+ citations**)
- [37] **Evaluating the Robustness of Neural Networks: An Extreme Value Theory Approach**
Tsui-Wei Weng*, Huan Zhang*, Pin-Yu Chen, Jinfeng Yi, Dong Su, Yupeng Gao, Cho-Jui Hsieh, Luca Daniel
International Conference on Learning Representations (ICLR), 2018 (**300+ citations**)
- [38] **Attacking Visual Language Grounding with Adversarial Examples: A Case Study on Neural Image Captioning**
Hongge Chen*, Huan Zhang*, Pin-Yu Chen, Jinfeng Yi, Cho-Jui Hsieh
56th Annual Meeting of the Association for Computational Linguistics (ACL), 2018 (**100+ citations**)
- [39] **Is Robustness the Cost of Accuracy? Lessons Learned from 18 Deep Image Classifiers**
Dong Su*, Huan Zhang*, Hongge Chen, Jinfeng Yi, Pin-Yu Chen, Yupeng Gao.
European Conference on Computer Vision (ECCV), 2018 (**300+ citations**)
- [40] **Towards Robust Neural Networks via Random Self-ensemble**
Xuanqing Liu, Minhao Cheng, Huan Zhang, Cho-Jui Hsieh
European Conference on Computer Vision (ECCV), 2018 (**300+ citations**)
- [41] **EAD: Elastic-Net Attacks to Deep Neural Networks via Adversarial Examples**
Pin-Yu Chen, Yash Sharma, Huan Zhang, Jinfeng Yi and Cho-Jui Hsieh
In AAAI Conference on Artificial Intelligence (AAAI), 2018 (**500+ citations**)
- [42] **GPU-acceleration for Large-scale Tree Boosting**
Huan Zhang, Si Si and Cho-Jui Hsieh
SysML Conference, 2018
- [43] **Gradient Boosted Decision Trees for High Dimensional Sparse Output**
Si Si, Huan Zhang, Sathiya Keerthi, Dhruv Mahajan, Inderjit Dhillon and Cho-Jui Hsieh
34th International Conference on Machine Learning (ICML), 2017
- [44] **Can Decentralized Algorithms Outperform Centralized Algorithms? A Case Study for Decentralized Parallel Stochastic Gradient Descent**
Xiangru Lian, Ce Zhang, Huan Zhang, Cho-Jui Hsieh, Wei Zhang and Ji Liu
Advances in Neural Information Processing Systems (NIPS), 2017 (**Oral, 700+ citations**)

- [45] **HogWild++: A New Mechanism for Decentralized Asynchronous Stochastic Gradient Descent**
Huan Zhang, Cho-Jui Hsieh, Venkatesh Akella
IEEE International Conference on Data Mining (ICDM), 2016
- [46] **Fixing the Convergence Problems in Parallel Asynchronous Dual Coordinate Descent**
Huan Zhang, Cho-Jui Hsieh
IEEE International Conference on Data Mining (ICDM), 2016
- [47] **Sublinear Time Orthogonal Tensor Decomposition**
Zhao Song, David P. Woodruff, Huan Zhang
Advances in Neural Information Processing Systems (NIPS), 2016
- [48] **A Comprehensive Linear Speedup Analysis for Asynchronous Stochastic Parallel Optimization from Zeroth-Order to First-Order**
Xiangru Lian, Huan Zhang, Cho-Jui Hsieh, Yijun Huang and Ji Liu
Advances in Neural Information Processing Systems (NIPS), 2016

Peer-reviewed workshop papers (* indicates co-first authors)

- [49] **Enhancing Certifiable Robustness via a Deep Model Ensemble**
Huan Zhang, Minhao Cheng and Cho-Jui Hsieh
ICLR 2019 Safe Machine Learning Workshop, 2019
- [50] **Realtime Query Completion via Deep Language Models**
Po-Wei Wang, Huan Zhang, Vijai Mohan, Inderjit S. Dhillon and J. Zico Kolter
SIGIR Workshop On eCommerce, 2018
- [51] **ZOO: Zeroth Order Optimization based Black-box Attacks to Deep Neural Networks without Training Substitute Models**
Pin-Yu Chen*, Huan Zhang*, Yash Sharma, Jinfeng Yi and Cho-Jui Hsieh
10th ACM Workshop on Artificial Intelligence and Security, 2017 (*Best paper finalist, 1300+ citations*)
- [52] **Burst Mode Processing: An Architectural Framework for Improving Performance in Future Chip Microprocessors**
Huan Zhang, Rajeevan Amirtharajah, Christopher Nitta, Matthew Farrens and Venkatesh Akella
Workshop on Managing Overprovisioned Systems, Co-located with ASPLOS-19, 2014
- [53] **HySIM: Towards a Scalable, Accurate and Fast Simulator for Manycore Processors**
Kramer Straube, Huan Zhang, Christopher Nitta, Matthew Farrens and Venkatesh Akella
3rd Workshop on the Intersections of Computer Architecture and Reconfigurable Logic, Co-located with MICRO-46, 2013