Linyi Li

3107 Siebel Center. 201 N. Goodwin Ave. Urbana, IL 61801, USA

(+1) 217 205 6216 • linyi2@illinois.edu • linyil.com

Ph.D. candidate in computer science with research interests in **machine learning** and **computer security**, especially in building **certifiably trustworthy** deep learning systems

- by providing rigorous guarantees of various trustworthy properties (robustness, fairness, reliabiltiy, etc) for a given deep neural network model;
- by improving such guaranteed trustworthiness for machine learning via strategic architecture design, dataset building, model training, post-processing, etc.

Education

- **University of Illinois Urbana-Champaign**
  Ph.D. Candidate (since May 2020) in Computer Science *Aug 2018 – Jul 2023 (expected)*
  - Advisor: *Prof.* Bo Li Co-advisor: *Prof.* Tao Xie
  - Thesis proposal: Enabling large-scale certifiably trustworthy deep learning systems

- **Tsinghua University** **Beijing, China**
  Bachelor of Computer Science and Technology *Aug 2014 – Jul 2018*
  - GPA: Major: 91.6/100 Overall: 90.1/100
  - Advisor: *Prof.* Xiaoying Bai
  - Thesis: Model-Based Automated Web API Test Generation.
  - Tsinghua University Outstanding Undergraduate, Class of 2018
  - Excellent Undergraduate, Department of Computer Science and Technology

Selected Publications

(* stands for equal contribution)

1. **Linyi Li**, Tao Xie, Bo Li. SoK: Certified Robustness for Deep Neural Networks. *IEEE Symposium on Security and Privacy (SP) 2023.*
2. Mintong Kang*, **Linyi Li***, Maurice Weber, Yang Liu, Ce Zhang, Bo Li. Certifying Some Distributional Fairness with Subpopulation Decomposition. *Advances in Neural Information Processing Systems (NeurIPS) 2022.*
3. **Linyi Li**, Jiawei Zhang, Tao Xie, Bo Li. Double Sampling Randomized Smoothing. *International Conference on Machine Learning (ICML) 2022.*
4. Fan Wu*, **Linyi Li***, Chejian Xu, Huan Zhang, Bhavya Kailkhura, Krishnaram Kenthapadi, Ding Zhao, Bo Li. COPA: Certifying Robust Policies for Offline Reinforcement Learning against Poisoning Attacks. *International Conference on Learning Representations (ICLR) 2022.*
5. Zhuolin Yang*, **Linyi Li***, Xiaojun Xu, Bhavya Kailkhura, Tao Xie, Bo Li. On the Certified Robustness for Ensemble Models and Beyond. *International Conference on Learning Representations (ICLR) 2022.*
6. Zhuolin Yang*, **Linyi Li***, Xiaojun Xu*, Shiliang Zuo, Qian Chen, Benjamin Rubinstein, Ce Zhang, Bo Li. TRS: Transferability Reduced Ensemble via Encouraging Gradient Diversity and Model Smoothness. *Advances in Neural Information Processing Systems (NeurIPS) 2021.*
7. Jiawei Zhang*, **Linyi Li***, Huichen Li, Xiaolu Zhang, Shuang Yang, Bo Li. Progressive-Scale Boundary

Blackbox Attack via Projective Gradient Estimation. *International Conference on Machine Learning (ICML) 2021.*

8. **Linyi Li***, Maurice Weber*, Xiaojun Xu, Luka Rimanic, Bhavya Kailkhura, Tao Xie, Ce Zhang, Bo Li. TSS: Transformation-Specific Smoothing for Robustness Certification. *ACM Conference on Computer and Communications Security (CCS) 2021.*
9. Huichen Li*, **Linyi Li***, Xiaojun Xu, Xiaolu Zhang, Shuang Yang, Bo Li. Nonlinear Projection Based Gradient Estimation for Query Efficient Blackbox Attacks. *International Conference on Artificial Intelligence and Statistics (AISTATS) 2021.*
10. **Linyi Li***, Zexuan Zhong*, Bo Li, Tao Xie. Robustra: Training Provable Robust Neural Networks over Reference Adversarial Space. *International Joint Conference on Artificial Intelligence (IJCAI) 2019.*

Selected Awards

- 2022 AdvML Rising Star Award *2022*
- 1st Place, 3rd International Verification of Neural Networks Competition (VNN-COMP'22) *2022*
- Qualcomm Innovation Fellowship Finalist (among 44 in North America) *2022*
- Two Sigma PhD Fellowship Finalist (among 13 worldwide) *2022*
- ACM CCS Travel Conference Award *2021*
- 2nd Place, ICPC Mid-Central USA Regional Contest *2019*
- 3rd Place, ICPC Mid-Central USA Regional Contest *2018*
- Wing Kai Cheng Fellowship *2018*
- Tsinghua University Outstanding Undergraduate, Class of 2018 (301 of 3555) *2018*
- Excellent Undergraduate, Department of Computer Science and Technology at Tsinghua *2018*
- Academic Excellence Award with HUAWEI Scholarship *2017*
- "Sogou Cup" Artificial Intelligence Programming Contest Top 16 *2015*
- Top 0.03% in the National College Entrance Exam *2014*
- National Olympics in Informatics, Bronze Medal *2013*
- National Olympics in Informatics in Provinces, First Prize *2012*

Invited Talks

- Webinar at TrustML Young Scientist Seminars, RIKEN-AIP *Aug 2022*
- Talk at 4th Workshop on Adversarial Learning Methods for Machine Learning and Data Mining (AdvML), KDD 2022 *Aug 2022*
- Virtual Talk at AI Time platform *Aug 2022*
- Webinar at Jiangmen platform *Feb 2022*
- Virtual Talk at Visual Informatics Group, University of Texas at Austin *Oct 2021*
- Webinar at Safe AI, Bilibili *Mar 2021*
- Virtual Talk at Safe AI Lab, Carnegie Mellon University *Mar 2021*
- Virtual Talk at Workshop on Robust Artificial Intelligence, Lorentz Center *Jan 2021*

Teaching and Mentorship Experiences

- **Logic and AI (Graduate Level)** — **Lead Teaching Assistant**
  University of Illinois Urbana-Champaign — *Aug 2021 – Dec 2021*
  - Lead the course project design and grading.
  - Setup infrastructure and help the lecture design of the new course.

- **Data Structure (Undergraduate Level)** — **Teaching Assistant**
  Tsinghua University — *Sept 2015 – Jan 2016*
  - Host two seminar for homework problem discussions.
  - Contribute several original problems for assignments and exams.

- **Undergraduate Research Intern Co-Mentorship**
  Mentored students:
  - Mintong Kang — *Nov 2021 – May 2022*
    Paper published at NeurIPS 2022 on certified fairness. Now a PhD student at UIUC.
  - Chenhui Zhang — *Dec 2021 – May 2022*
    Paper submitted on ensemble pruning for certified robustness. Now an undergraduate student at UIUC.
  - Jiawei Zhang — *Aug 2020 – Feb 2021*
    Paper published at ICML 2021 on black-box neural network attacks. Now a master student at UIUC.
  - Wenda Chu — *Nov 2021 – Feb 2022*
    Paper published at ICML 2022 on certification of point cloud models. Now an undergraduate student at Tsinghua University.

Student mentor for new PhD students in computer science at UIUC. *Fall 2022*
Graduate ambassador for prospective PhD students in computer science at UIUC. *Spring 2021*

Internship Experiences

- **Microsoft Research Lab - New England** — **Cambridge, MA**
  Research Intern mentored by Dr. Adam Kalai — *May 2022 – Aug 2022*
  - Program synthesis by finetuning from large language models with a handcrafted distributed training framework.

- **Fujitsu Laboratories of America** — **Remote**
  Research Intern mentored by Dr. Mukul Prasad — *May 2021 – Aug 2021*
  - Program Synthesis for AutoML based on learning from mined corpus and static analysis based data augmentation.
  - Lead to a paper accepted by ICSE 2022.

- **Microsoft** — **Redmond, WA**
  Data Scientist Intern mentored by Dr. Neel Sundaresan — *Jun 2019 – Aug 2019*
  - Build an efficient search engine for PR comments and commits.
  - Utilize transformer models for unsupervised commit classification and code change pattern extraction.

- **Carnegie Mellon University** — **Pittsburgh, PA**
  Undergrad Research Intern mentored by Prof. Matt Fredrikson — *Jun 2017 – Sept 2017*
  - Apply integrated gradients to explain and visualize convolutional neural networks.
  - Develop an automatic method to capture vital lesions for diabetic retinopathy diagnosis, leading to a paper accepted by ITC 2018.

- **Sogou Inc.** — **Beijing, China**
  Back-end Engineer Intern — *Aug 2015 – Oct 2015*
  - Design the interfaces between back-end and front-end for a tutor ordering platform.
  - Implement an efficient and advanced tutor search module that supported multiple keys.

Selected Open-Source Projects

- Developer of leaderboard on provable training and verification approaches for DNNs. ≈ 100 stars
  `https://github.com/AI-secure/Provable-Training-and-Verification-Approaches-Towards-Robust-Neural-Networks`
  `https://github.com/sokcertifiedrobustness/sokcertifiedrobustness.github.io`
- Developer of VeriGauge: unified toolbox for representative robustness verification approaches for deep neural networks. ≈ 80 stars
  `https://github.com/AI-secure/VeriGauge`
  `https://github.com/sokcertifiedrobustness/certified-robustness-benchmark`
  - Over 20 verification approaches are reliably reproduced.
  - Accompanying paper published at SP 2023.
- Developer of TSS: transformation-specific smoothing-based robustness certification against geometric perturbations. ≈ 20 stars
  `https://github.com/AI-secure/semantic-randomized-smoothing`
  - State-of-the-art verification approach for robustness against geometric perturbations.
  - Accompanying paper published at CCS 2021.
- Key contributor of α-β-CROWN (alpha-beta-CROWN), a scalable neural network verifier. ≈ 100 stars
  `https://github.com/huanzhang12/alpha-beta-CROWN`
  - 2x winner of International Verification of Neural Networks Competition (VNN-COMP'21, '22).
  - Accompanying paper published at NeurIPS 2022.

Services

- NeurIPS 2022, Workshop on Trustworthy and Socially Responsible Machine Learning — Organizer
- NeurIPS (2021-) — PC Member
- ICML (2022-) — PC Member
- ICLR (2021-) — PC Member
- AAAI (2022-) — PC Member
- UAI (2021-) — PC Member
- AISTATS (2021-) — PC Member
- TPAMI — Reviewer
- TMLR — Reviewer
- Neurocomputing — Reviewer
- ICML 2022, Workshop on Formal Verification of Machine Learning — PC Member
- KDD (2020-), Workshop on Adversarial Learning Methods for Machine Learning and Data Mining — PC Member
- ICML 2019, Workshop on the Security and Privacy of Machine Learning (SPML) — PC Member
- CVPR 2019, Workshop on Adversarial Machine Learning in Real-World Computer Vision Systems (AdvMLCV) — PC Member

Full Publication List

(* stands for equal contribution)

1. **Linyi Li**, Tao Xie, Bo Li. SoK: Certified Robustness for Deep Neural Networks. *IEEE Symposium on Security and Privacy (SP) 2023.*

2. Mintong Kang*, **Linyi Li***, Maurice Weber, Yang Liu, Ce Zhang, Bo Li. Certifying Some Distributional Fairness with Subpopulation Decomposition. *Advances in Neural Information Processing Systems (NeurIPS) 2022.*

3. Xiaojun Xu, **Linyi Li**, Bo Li. LOT: Layer-wise Orthogonal Training on Improving l2 Certified Robustness. *Advances in Neural Information Processing Systems (NeurIPS) 2022.*

4. Bhaskar Ray Chaudhury, **Linyi Li**, Mintong Kang, Bo Li, Ruta Mehta. Fairness in Federated Learning via Core-Stability. *Advances in Neural Information Processing Systems (NeurIPS) 2022.*

5. Huan Zhang*, Shiqi Wang*, Kaidi Xu*, **Linyi Li**, Bo Li, Suman Jana, Cho-Jui Hsieh, J. Zico Kolter. General Cutting Planes for Bound-Propagation-Based Neural Network Verification. *Advances in Neural Information Processing Systems (NeurIPS) 2022.*

6. Zhuolin Yang*, Zhikuan Zhao*, Boxin Wang, Jiawei Zhang, **Linyi Li**, Hengzhi Pei, Bojan Karlaš, Ji Liu, Heng Guo, Ce Zhang, Bo Li. Improving Certified Robustness via Statistical Learning with Logical Reasoning. *Advances in Neural Information Processing Systems (NeurIPS) 2022.*

7. Hanjiang Hu, Zuxin Liu, **Linyi Li**, Jiacheng Zhu, Ding Zhao. Robustness Certification of Visual Perception Models via Camera Motion Smoothing. *6th Annual Conference on Robot Learning (CoRL 2022).*

8. **Linyi Li**, Jiawei Zhang, Tao Xie, Bo Li. Double Sampling Randomized Smoothing. *International Conference on Machine Learning (ICML) 2022.*

9. Wenda Chu, **Linyi Li**, Bo Li. TPC: Transformation-Specific Smoothing for Point Cloud Models. *International Conference on Machine Learning (ICML) 2022.*

10. Maurice Weber, **Linyi Li**, Boxin Wang, Zhikuan Zhao, Bo Li, Ce Zhang. Certifying Out-of-Domain Generalization for Blackbox Functions. *International Conference on Machine Learning (ICML) 2022.*

11. Fan Wu*, **Linyi Li***, Chejian Xu, Huan Zhang, Bhavya Kailkhura, Krishnaram Kenthapadi, Ding Zhao, Bo Li. COPA: Certifying Robust Policies for Offline Reinforcement Learning against Poisoning Attacks. *International Conference on Learning Representations (ICLR) 2022.*

12. Zhuolin Yang*, **Linyi Li***, Xiaojun Xu, Bhavya Kailkhura, Tao Xie, Bo Li. On the Certified Robustness for Ensemble Models and Beyond. *International Conference on Learning Representations (ICLR) 2022.*

13. Fan Wu, **Linyi Li**, Zijian Huang, Yevgeniy Vorobeychik, Ding Zhao, Bo Li. CROP: Certifying Robust Policies for Reinforcement Learning through Functional Smoothing. *International Conference on Learning Representations (ICLR) 2022.*

14. Ripon Saha, Akira Ura, Sonal Mahajan, Chenguang Zhu, **Linyi Li**, Yang Hu, Hiroaki Yoshida, Sarfraz Khurshid, Mukul R. Prasad. SapientML: Synthesizing Machine Learning Pipelines by Learning from Human-Written Solutions. *International Conference on Software Engineering (ICSE) 2022.*

15. Zhuolin Yang*, **Linyi Li***, Xiaojun Xu*, Shiliang Zuo, Qian Chen, Benjamin Rubinstein, Ce Zhang, Bo Li. TRS: Transferability Reduced Ensemble via Encouraging Gradient Diversity and Model Smoothness. *Advances in Neural Information Processing Systems (NeurIPS) 2021.*

16. Jiawei Zhang*, **Linyi Li***, Huichen Li, Xiaolu Zhang, Shuang Yang, Bo Li. Progressive-Scale Boundary Blackbox Attack via Projective Gradient Estimation. *International Conference on Machine Learning (ICML) 2021.*

17. **Linyi Li***, Maurice Weber*, Xiaojun Xu, Luka Rimanic, Bhavya Kailkhura, Tao Xie, Ce Zhang, Bo Li. TSS: Transformation-Specific Smoothing for Robustness Certification. *ACM Conference on Computer and Communications Security (CCS) 2021.*

18. Huichen Li*, **Linyi Li***, Xiaojun Xu, Xiaolu Zhang, Shuang Yang, Bo Li. Nonlinear Projection Based Gradient Estimation for Query Efficient Blackbox Attacks. *International Conference on Artificial Intelligence and Statistics (AISTATS) 2021.*

19. **Linyi Li**, Zhenwen Li, Weijie Zhang, Jun Zhou, Pengcheng Wang, Jing Wu, Guanghua He, Xia Zeng, Yuetang Deng, Tao Xie. Clustering Test Steps in Natural Language toward Automating Test Automation. *ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE) 2020, Industry Track.*

20. **Linyi Li***, Zexuan Zhong*, Bo Li, Tao Xie. Robustra: Training Provable Robust Neural Networks over Reference Adversarial Space. *International Joint Conference on Artificial Intelligence (IJCAI) 2019.*

21. Klas Leino, Shayak Sen, Anupam Datta, Matt Fredrikson, **Linyi Li**. Influence-Directed Explanations for Deep Convolutional Networks. *International Test Conference (ITC) 2018.*

22. Junyi Wang, Xiaoying Bai, **Linyi Li**, Zhicheng Ji, Haoran Ma. A Model-Based Framework For Cloud API Testing. *Computer Software and Applications Conference (COMPSAC) 2017.*

23. Junyi Wang, Xiaoying Bai, Haoran Ma, **Linyi Li**, Zhicheng Ji. Cloud API Testing. *IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW) 2017.*

Preprints

1. Jiawei Zhang, **Linyi Li**, Ce Zhang, Bo Li. CARE: Certifiably Robust Learning with Reasoning via Variational Inference. *arXiv: 2209.05055.*

2. Wenda Chu*, Chulin Xie*, Boxin Wang, **Linyi Li**, Lang Yin, Han Zhao, Bo Li. FOCUS: Fairness via Agent-Awareness for Federated Learning on Heterogeneous Data. *arXiv: 2207.10265.*

3. Zhangheng Li, Tianlong Chen, **Linyi Li**, Bo Li, Zhangyang Wang. Can pruning improve certified robustness of neural networks. *arXiv: 2206.07311.*

4. Zhonghan Niu, Zhaoxi Chen, **Linyi Li**, Yubin Yang, Bo Li, Jinfeng Yi. On the Limitations of Denoising Strategies as Adversarial Defenses. *arXiv: 2012.09384.*

Last updated: Sept 25, 2022