Avery Ma

W1140-108 College St
Toronto, ON M5G 0C6
averyma.com
ama@cs.toronto.edu
Last update: Feb 2025

Research Statement: My research integrates theoretical and empirical approaches to robustness and generalization in deep learning, with the goal of enabling more reliable AI systems. Currently, I focus on the safety and security of LLMs, advancing research that drives practical solutions for AI deployment.

Education

Ph.D in Computer Science **Toronto ON**
University of Toronto, Vector Institute *2018 – 2024*
- Thesis: Understanding Adversarial Robustness in Deep Learning
- Supervisors: Amir-massoud Farahmand and Richard Zemel

M.A.Sc. in Systems Design Engineering **Waterloo ON**
University of Waterloo, Vision and Image Processing Lab *2016 – 2018*
- Thesis: Computational Depth from Defocus via Active Quasi-random Pattern Projections
- Supervisors: Alexander Wong and David Clausi

B.A.Sc. in Mechatronics Engineering *with Distinction*, Honours, Co-op Program **Waterloo ON**
University of Waterloo *2011 – 2016*
- Capstone project: All Terrain Personal Transportation Device

Research Experience

Research Intern **Toronto ON**
Samsung - Samsung AI Center (Host: Afsaneh Fazly) *May 2021 – Aug 2022*
- Developed a data augmentation method to improve generalization in multi-modal learning (Patented)

Research Intern **Kitchener ON**
Christie Digital - Advanced Technologies Group (Host: Mark Lamm) *May 2016 – Apr 2017*
- Led the research and development of real-time super-resolution techniques for projectors (Patented)

Selected Publications

- **Avery Ma**, Yangchen Pan, Amir-massoud Farahmand (2025). PANDAS: Improving Many-shot Jailbreaking via Positive Affirmation, Negative Demonstration, and Adaptive Sampling. *Under Review.*
- **Avery Ma**, Amir-massoud Farahmand, Yangchen Pan, Philip Torr, Jindong Gu (2024). Improving Adversarial Transferability via Model Alignment. *ECCV'24: European Conference on Computer Vision.*
- Jindong Gu, Xiaojun Jia, Pau de Jorge, Wenqian Yu, Xinwei Liu, **Avery Ma**, Yuan Xun, Anjun Hu, Ashkan Khakzar, Zhijiang Li, Xiaochun Cao, Philip Torr (2023). A Survey on Transferability of Adversarial Examples Across Deep Neural Networks. *TMLR: Transactions on Machine Learning Research.*
- **Avery Ma**, Yangchen Pan, Amir-massoud Farahmand (2023). Understanding the robustness difference between stochastic gradient descent and adaptive gradient methods. *TMLR: Transactions on Machine Learning Research (**Featured Certification (Top 3%)**, ICLR'24 Journal-to-Conference*).

- **Avery Ma**, Nikita Dvornik, Ran Zhang, Leila Pishdad, Konstantinos Derpanis, Afsaneh Fazly (2022). SAGE: Saliency-Guided Mixup with Optimal Rearrangements. *BMVC'22: British Machine Vision Conference.*
- **Avery Ma**, Aladin Virmaux, Kevin Scaman, Juwei Lu (2021). Improving Hierarchical Adversarial Robustness of Deep Neural Network. *arXiv preprint arXiv: 2102.09012.*
- **Avery Ma**, Fartash Faghri, Nicolas Papernot, Amir-massoud Farahmand (2020). SOAR: Second-Order Adversarial Regularization. *arXiv preprint arXiv: 2004.01832* .

Full list available at Google Scholar.

Patents

- **Bojie Ma**, Nikita Dvornik, Ran Zhang, Konstantinos Derpanis, Afsaneh Fazly (2023). Saliency-guided mixup with optimal re-arrangements for efficient data augmentation. Patent App.: 18/201,521
- **Bojie Ma**, Ahmed Gawish, Alexander Wong, Paul Fieguth, Mark Lamm (2018). Real-time spatial-based resolution enhancement using shifted superposition. Patent No.: US10009587 B1

Honors and Awards

- DAAD AInet Fellowship for the Postdoc-NeT-AI Program on Safety and Security in AI *Apr 2024*
- NeurIPS'23 Top Reviewer *Dec 2023*
- NSERC Canada Graduate Scholarship - Doctoral (CGS-D) *Sept 2018 – Dec 2022*
- University of Waterloo Alumni Gold Medal (Department Nomination) *Sept 2018*
- Ontario Graduate Scholarship *May 2017 – Apr 2018*
- University of Waterloo President's Graduate Scholarship *May 2017 – Apr 2018*
- University of Waterloo Provost Graduate Scholarship *May 2016 – Apr 2017*

Invited Talks

- **Ludwig Maximilian University of Munich, Mathematisches Institut** **Nov 2024**
  "Understanding generalization, robustness, and adversarial transferability."
- **Ludwig Maximilian University of Munich, Tresp Lab** **Nov 2024**
  "Understanding generalization, robustness, and adversarial transferability.""
- **University of Toronto, CSC413: Neural Networks and Deep Learning (Guest Lecturer) Apr 2024**
  "Is Your Neural Network at Risk? The Pitfall of Adaptive Gradient Optimizers"

Professional Activities and Services

- International Conference on Learning Representations (**ICLR**) (2023, 2025)
- Conference on Neural Information Processing Systems (**NeurIPS**) (2023, 2024)
- International Conference on Machine Learning (**ICML**) (2023, 2024, 2025)
- Computer Vision and Image Understanding (**CVIU**) (2022)
- Artificial Intelligence and Statistics (**AISTATS**) (2022, 2025)
- Transactions on Machine Learning Research (**TMLR**)
- Graduate application assistance program for prospective students in groups underrepresented in Computer Science, University of Toronto (2021, 2022, 2023, 2024)
- Graduate admissions committee at the Department of Computer Science, University of Toronto (2020)