Yiwei Lu

(226)7476966
yiwei.lu@uottawa.ca
Yiwei Lu's homepage

Professional Appointment

2025-present **University of Ottawa**, *Tenure-track Assistant Professor*.
School of Engineering Design and Teaching Innovation
Faculty of Engineering

2025-present **Vector Institute**, *Faculty Affliate*.

Education

2020-2025 **University of Waterloo**, *Ph.D. in Computer Science*.
Waterloo AI Lab and Vector Institute, Supported by David R. Cheriton Graduate Scholarship
Supervisor: Prof. Yaoliang Yu and Dr. Sun Sun
Thesis: Trustworthy Machine Learning with Data in the Wild

2018-2020 **University of Manitoba**, *M.Sc. in Computer Science*.
Computer Vision Lab, Supported by University of Manitoba Graduate Funding
Supervisor: Prof. Yang Wang
Thesis: Anomaly Detection in Surveillance Videos using Deep Learning

2013-2017 **University of Electronic Science and Technology of China**, *B.E. in Electronic Engineering*.
Statistical Machine Intelligence and Learning Lab
Supervisor: Prof. Zenglin Xu and Prof. Zhao Kang
Thesis and patent: Design of a Phase-Reconfigurable Antenna Unit

2016 **University of California, Santa Barbara**, *Electrical Computer Engineering*.
Exchange Program, Funded by China Scholarship Council

Award and Honors

2024 Apple Scholars PhD fellowship in AI/ML Nomination (1 of 3), University of Waterloo
2024 Google PhD Fellowship Nomination (1 of 4), University of Waterloo
2024 Vector Institute Travel Grant
2024 IEEE Computer Society Honorarium for SaTML (with grant)
2024-2025 David R. Cheriton Graduate Scholarship, University of Waterloo
2023 NeurIPS 2023 Top Reviewer (with grant)
2023 Conference Funding, University of Waterloo
2023 Graduate Student Research Dissemination Award, University of Waterloo
2020-2025 Vector Research Grant
2019-2020 University of Manitoba Graduate Funding (UMGF)
2019 University of Manitoba Graduate Studies Travel Award
2018-2019 University of Manitoba International Graduate Student Entrance Scholarship
2013-2017 People's Scholarship of China
2016 China Scholarship Council Scholarship for Studying Abroad
2016 UESTC Exchange Award

Professional Services

2024-2025 **Equity, Diversity, and Inclusion (EDI) Committee**, *University of Waterloo*.
Graduate student representative.

2024 **School Advisory Committee on Appointments (SACA)**, *University of Waterloo*.
Graduate student representative and student host for faculty candidate meetings.

2024 **Graduate Recruiting Committee**, *University of Waterloo*.
Graduate student representative and an invited panelist for the grad student visiting.

2021-2025 **Conference Reviewer**, *NeurIPS (2022-2025), ICML (2022-2025), ICLR (2024-2025), AAAI (2021-2022, 2025), AISTATS (2025)*.

2021-2025 **Journal Reviewer**, *Transactions on Machine Learning Research, Neural Networks, IEEE Transactions on Multimedia, IEEE Transactions on Circuits and Systems for Video Technology*.

Publications

Journal Papers

[1] **MUC: Machine Unlearning for Contrastive Learning with Black-box Evaluation**, *Yihan Wang*, **Yiwei Lu***, Guojun Zhang, Franziska Boenisch, Adam Dziedzic, Yaoliang Yu, Xiao-Shan Gao*, in *Transactions on Machine Learning Research (TMLR), 2025* (also appeared in *ICML 2024 Next Generation of AI Safety Workshop (Oral)*).

[2] **f-MICL: Understanding and Generalizing InfoNCE-based Contrastive Learning**, ***Yiwei Lu****, *Guojun Zhang*, Sun Sun, Hongyu Guo, Yaoliang Yu*, in *Transactions on Machine Learning Research (TMLR), 2023* (also appeared in *NeurIPS 2021 Workshop on Self-supervised Learning (Contributed Talk)*).

[3] **Indiscriminate Data Poisoning Attacks on Neural Networks**, ***Yiwei Lu***, *Gautam Kamath, Yaoliang Yu*, in *Transactions on Machine Learning Research (TMLR), 2022* (also appeared in *NeurIPS 2022 ML Safety Workshop and Trustworthy and Socially Responsible Machine Learning (TSRML) Workshop*).

[4] **AdaCrowd: Unlabeled Scene Adaptation for Crowd Counting**, *Mahesh Kumar Krishna Reddy, Mrigank Rochan, **Yiwei Lu**, Yang Wang*, in *IEEE Transactions on Multimedia, Volume 24, Page 1008-1019, 2022*.

[5] **Structure Learning with Similarity Preserving**, *Zhao Kang, Xiao Lu, **Yiwei Lu**, Chong Peng, Wenyu Chen, and Zenglin Xu*, in *Neural Networks, 2020, Volume 129, Page 138-148*.

Conference Papers

[6] **Demystifying Foreground-Background Memorization in Diffusion Models**, *Jimmy Z Di*, **Yiwei Lu***, Yaoliang Yu, Gautam Kamath, Adam Dziedzic, Franziska Boenisch, in AAAI 2026*.

[7] **BridgePure: Revealing the Fragility of Black-box Data Protection**, *Yihan Wang*, **Yiwei Lu***, Xiao-Shan Gao, Yaoliang Yu, Gautam Kamath, in NeurIPS 2025*.

[8] **Machine Unlearning Fails to Remove Data Poisoning Attacks** , *Martin Pawelczyk, Jimmy Z. Di, **Yiwei Lu**, Gautam Kamath, Ayush Sekhari, Seth Neel*, in *ICLR 2025* (also appeared in *ICML 2024 Generative AI and Law Workshop (Spotlight)*).

[9] **Disguised Copyright Infringement of Latent Diffusion Models**, ***Yiwei Lu****, *Matthew Y.R Yang*, Zuoqiu Liu*, Gautam Kamath, Yaoliang Yu*, in *ICML 2024*.

[10] **Indiscriminate Data Poisoning Attacks on Pre-trained Feature Extractors**, ***Yiwei Lu***, *Matthew Y.R Yang, Gautam Kamath, Yaoliang Yu*, in *IEEE SaTML 2024*.

[11] **Understanding Neural Network Binarization with Forward and Backward Proximal Quantizers**, ***Yiwei Lu***, *Yaoliang Yu, Xinlin Li, Vahid Partovi Nia*, in *NeurIPS 2023*.

[12] **Exploring the Limits of Model-Targeted Indiscriminate Data Poisoning Attacks**, ***Yiwei Lu***, *Gautam Kamath, Yaoliang Yu*, in *ICML 2023*.

[13] **Few-shot Scene-Adaptive Anomaly Detection (Spotlight)**, ***Yiwei Lu***, *Frank Yu, Mahesh Kumar K and Yang Wang*, in *ECCV 2020*.

[14] **Future Frame Prediction Using Convolutional VRNN for Anomaly Detection**, ***Yiwei Lu***, *Mahesh Kumar K, Seyed shahabeddin Nabavi and Yang Wang.*, in *IEEE AVSS 2019*.

[15] **Similarity Learning via Kernel Preserving Embedding**, *Zhao Kang,* ***Yiwei Lu****, Yuanzhang Su, Changsheng Li, Zenglin Xu*, in *AAAI 2019* .

[16] **Homoglyph Attack Detection with Unpaired Data**, ***Yiwei Lu***, *Mahesh Kumar K, Noman Mohammed, Yang Wang*, in *ACM/IEEE Symposium on Edge Computing 2019*.

Workshop Papers

[17] **On the Robustness of Neural Networks Quantization against Data Poisoning Attacks**, ***Yiwei Lu***, *Yihan Wang, Guojun Zhang, Yaoliang Yu*, in *ICML 2024 Next Generation of AI Safety Workshop*.

[18] **CM-GAN: Stabilizing GAN Training with Consistency Models**, *Haoye Lu,* ***Yiwei Lu****, Dihong Jiang, Spencer Ryan Szabados, Sun Sun, Yaoliang Yu*, in *ICML 2023 Workshop on Structured Probabilistic Inference & Generative Modeling*.

[19] **Semantic Segmentation in Compressed Videos**, *Ang Li*,* ***Yiwei Lu****, Yang Wang*, in *IEEE International Workshop on Multimedia Signal Processing 2019*.

Preprint

[20] **Not All Samples Are Equal: Quantifying Instance-level Difficulty in Targeted Data Poisoning**, *William Xu*,* ***Yiwei Lu****, Yihan Wang, Matthew YR Yang, Zuoqiu Liu, Gautam Kamath, Yaoliang Yu, Preprint (2025)*.

Talks (Selected)

2025 **Vector Machine Learning Security and Privacy Workshop**, *Adversarial Perturbation for Data Protection: Limitations and Pitfalls*.

2025 **University of Ottawa**, *Trustworthy Machine Learning with Data in the Wild*.

2025 **Concordia University**, *Trustworthy Machine Learning with Data in the Wild*.

2025 **University of Texas at Arlington**, *Trustworthy Machine Learning with Data in the Wild*.

2025 **University of Victoria**, *Trustworthy Machine Learning with Data in the Wild*.

2025 **Western University**, *Trustworthy Machine Learning with Data in the Wild*.

2024 **University of Waterloo (ECE)**, *Copyright Issues in Generative Models and Countermeasures*.

2024 **UESTC**, *Trustworthy Machine Learning with Data in the Wild*.

2024 **Vector Machine Learning Security and Privacy Workshop**, *Disguised Copyright Infringement of Latent Diffusion Models*.

2024 **IEEE SaTML**, *Indiscriminate Data Poisoning Attacks on Pre-trained Feature Extractors*.

2023 **Fudan University**, *Exploring the Limit of Indiscriminate Data Poisoning Attacks*.

2022 **National Research Council Canada**, *f-Mutual Information Contrastive Learning*.

Teaching

2024 **Open Course Development**, *University of Waterloo & Cybersecurity and Privacy Institute*.
Artificial Intelligence (AI) Literacy: module on "Data Poisoning Attacks and Defenses".

2020-2024 **Teaching Assistant**, *University of Waterloo*.
CS480/680: Introduction to Machine Learning
CS245: Logic and Computation
CS116: Introduction to Computer Science 2

2018-2020 **Teaching Assistant**, *University of Manitoba*.
COMP4550: Real-time Systems
COMP4190: Advanced Artificial Intelligence

Work Experiences

2022 **Research Internship**, *Huawei Noah's Ark Lab, Montreal*.
Machine Learning Researcher on Neural Network Quantization
Published a paper in NeurIPS 2023.

2021 **Research Internship**, *National Research Council Canada, Waterloo*.
Machine Learning Researcher on Contrastive Learning
Published a paper in NeurIPS 2021 Workshop and TMLR 2023.

Supervision of Research Students

2025-present **Jack Anderson**, *University of Chicago*, Vector Internship.

2024-2025 **William Xu**, *University of Waterloo*, Undergraduate Researcher, Now software engineer at Waabi.

2023-2024 **Robert Liu**, *University of Waterloo*, Undergraduate Researcher, Now machine learning engineer at Google.

2023-2024 **Matthew Y.R Yang**, *University of Waterloo*, Undergraduate Researcher, Now Master's student at CMU.

2019-2020 **Frank Yu**, *University of Manitoba*, Undergraduate Researcher, Now research engineer at Meta.

2019-2020 **Ang Li**, *University of Manitoba*, Undergraduate Researcher, Now machine learning engineer at Primate Labs Inc.