

# Wenxiao Wang

✉ [wwx@umd.edu](mailto:wwx@umd.edu) | PhD student at UMD

## Education

### University of Maryland, College Park

PH.D. IN COMPUTER SCIENCE

Advisor: Prof. [Soheil Feizi](#)

*College Park, US*

*Sept. 2021 - present*

### Institute for Interdisciplinary Information Sciences (a.k.a. Yao class), Tsinghua University

B.ENG. IN COMPUTER SCIENCE AND TECHNOLOGY.

*Peking, China*

*Sept. 2016 - June. 2020*

## Experience

### Sony AI

RESEARCH INTERN MENTORED BY **WEIMING ZHUANG** AND **LINGJUAN LYU**

model selection

*Remote, US*

*May. 2023 - Aug. 2023*

### Bytedance

RESEARCH INTERN MENTORED BY **LINJIE YANG**, **HENG WANG** AND **YU TIAN**

self-supervised pre-training of visual models

*Remote, US*

*June. 2022 - Nov. 2022*

### Institute for Interdisciplinary Information Sciences, Tsinghua University

RESEARCH ASSISTANT MENTORED BY PROF. **HANG ZHAO**

self-supervised learning

*Peking, China*

*Sep. 2020 - Aug. 2021*

### University of California, Berkeley

VISITING STUDENT RESEARCHER ADVISED BY **XINYUN CHEN**, **RUOXI JIA** AND PROF. **DAWN SONG**

watermarking of neural networks & differentially private deep learning

*Berkeley, US*

*Apr. 2019 - Aug. 2019*

### Bytedance AI Lab

INTERN IN VISUAL SEARCH GROUP MENTORED BY **YI HE** AND **LEI LI**

deep representation learning for large scale near duplicate video retrieval

*Peking, China*

*May. 2018 - Nov. 2018*

## Preprints

### Robustness of AI-Image Detectors: Fundamental Limits and Practical Attacks [\[url\]](#)

MEHRDAD SABERI, VINU SANKAR SADASIVAN, KEIVAN REZAEI, AOUNON KUMAR, ATOOSA CHEGINI, **WENXIAO WANG**,

SOHEIL FEIZI

Media Coverage: [\[Wired\]](#) [\[MIT Tech Review\]](#) [\[Bloomberg News\]](#) [\[The Register\]](#)

*2023*

### Can AI-Generated Text be Reliably Detected? [\[url\]](#)

VINU SANKAR SADASIVAN, AOUNON KUMAR, SRIRAM BALASUBRAMANIAN, **WENXIAO WANG**, SOHEIL FEIZI

Media Coverage: [\[Washington Post\]](#) [\[Wired\]](#) [\[New Scientist\]](#) [\[The Register\]](#) [\[TechSpot\]](#) [\[UMD Science\]](#)

*2023*

## On Practical Aspects of Aggregation Defenses against Data Poisoning Attacks [\[url\]](#)

WENXIAO WANG, SOHEIL FEIZI

2023

## DRSM: De-Randomized Smoothing on Malware Classifier Providing Certified Robustness [\[url\]](#)

SHOUMIK SAHA, WENXIAO WANG, YIGITCAN KAYA, SOHEIL FEIZI, TUDOR DUMITRAS

2023

## Publications

---

### Temporal Robustness against Data Poisoning [\[url\]](#)

WENXIAO WANG, SOHEIL FEIZI

2023

Conference on Neural Information Processing Systems (NeurIPS)

### Spuriousity Rankings: Sorting Data for Spurious Correlation Robustness [\[url\]](#)

MAZDA MOAYERI, WENXIAO WANG, SAHIL SINGLA, SOHEIL FEIZI

2023

Conference on Neural Information Processing Systems (NeurIPS)[\[spotlight\]](#)

### Lethal Dose Conjecture on Data Poisoning [\[url\]](#)

WENXIAO WANG, ALEXANDER LEVINE, SOHEIL FEIZI

2022

Conference on Neural Information Processing Systems (NeurIPS)

### Improved Certified Defenses against Data Poisoning with (Deterministic) Finite Aggregation [\[url\]](#)

WENXIAO WANG, ALEXANDER LEVINE, SOHEIL FEIZI

2022

International Conference on Machine Learning (ICML)

### On Feature Decorrelation in Self-Supervised Learning [\[url\]](#)

TIANYU HUA\*, WENXIAO WANG\*, ZIHUI XUE, SUCHENG REN, YUE WANG, HANG ZHAO

2021

(\*EQUAL CONTRIBUTION)

International Conference on Computer Vision (ICCV)[\[oral\]](#)

### DPLis: Boosting Utility of Differentially Private Deep Learning via Randomized Smoothing [\[url\]](#)

WENXIAO WANG, TIANHAO WANG, LUN WANG, NANQING LUO, PAN ZHOU, DAWN SONG, RUOXI JIA

2021

Privacy Enhancing Technologies Symposium (PETS)

### REFIT: A Unified Watermark Removal Framework For Deep Learning Systems With Limited Data [\[url\]](#)

XINYUN CHEN\*, WENXIAO WANG\*, YIMING DING, CHRIS BENDER, RUOXI JIA, BO LI, DAWN SONG

2021

(\*EQUAL CONTRIBUTION)

ACM Asia Conference on Computer and Communications Security (AsiaCCS)

### The Secret Revealer: Generative Model Inversion Attacks Against Deep Neural Networks [\[url\]](#)

YUHENG ZHANG\*, RUOXI JIA\*, HENGZHI PEI, WENXIAO WANG, BO LI, DAWN SONG

2020

(\*EQUAL CONTRIBUTION)

Conference on Computer Vision and Pattern Recognition (CVPR)[\[oral\]](#)

## Leveraging Unlabeled Data for Watermark Removal of Deep Neural Networks [\[url\]](#)

XINYUN CHEN\*, WENXIAO WANG\*, YIMING DING, CHRIS BENDER, RUOXI JIA, BO LI, DAWN SONG

(\*EQUAL CONTRIBUTION)

ICML2019 Workshop on Security and Privacy of Machine Learning

2019

## Talks

---

- **Temporal Robustness against Data Poisoning**, AI TIME Youth PhD Talk, November 2023.
- **Lethal Dose Conjecture: From Few-shot Learning to Potentially Nearly Optimal Defenses against Data Poisoning**, TMLR Group, Hong Kong Baptist University, December 2022.
- **Lethal Dose Conjecture on Data Poisoning**, AI TIME Youth PhD Talk, November 2022.
- **Improved Certified Defenses against Data Poisoning with (Deterministic) Finite Aggregation**, AI TIME Youth PhD Talk, August 2022.

## Services

---

Program Committee / Reviewer of:

- TPAMI
- NeurIPS 2022, 2023
- ICML 2022, 2023 (Outstanding Reviewer in 2022)
- ICLR 2023
- ICCV 2023
- Workshop on Adversarial Robustness In the Real World (ICCV 2021, ECCV 2022)
- Workshop on Socially Responsible Machine Learning (ICML 2021)
- Workshop on Adversarial Machine Learning in Real-World Computer Vision Systems and Online Challenges (CVPR 2021)
- Workshop on Security and Safety in Machine Learning Systems (ICLR 2021)

## Awards

---

Gold Medal( <b>4th place</b> )	National Olympiad in Informatics	2015
Gold Medal( <b>1st place</b> )	Asia and Pacific Informatics Olympiad in China District	2015
Gold Medal( <b>10th place</b> )	China Team Selection Competition	2015
Gold Medal	National Olympiad in Informatics	2014
Bronze Medal	Asia and Pacific Informatics Olympiad in China District	2014
Silver Medal	China Team Selection Competition	2014

## Teaching

---

- Teaching Assistant of *CMSC828W: Foundations of Deep Learning*, Fall 2022, University of Maryland, College Park.
- Teaching Assistant of *CMSC422: Introduction to Machine Learning*, Spring 2022, University of Maryland, College Park.
- Teaching Assistant of *CMSC351: Algorithms*, Fall 2021, University of Maryland, College Park.