

JAECHUL (Harry) Roh

Ph.D. in Computer Science, University of Massachusetts Amherst
jroh@umass.edu · [Personal Website](#) · [Github](#) · [Google Scholar](#)

EDUCATION

University of Massachusetts Amherst

Ph.D. in Computer Science

Advisor: Prof. [Amir Houmansadr](#)

GPA: 4.0/4.0

September 2023 – Present

Amherst, Massachusetts, USA

Hong Kong University of Science and Technology

B.Eng. in Computer Engineering, School of Engineering

Final Year Thesis Advisor: Prof. [Jun Zhang](#)

September 2017 – May 2023

Clear Water Bay, Hong Kong, HK

RESEARCH INTERESTS

My research focuses on the realms of **Privacy & Security in AI** and **Trustworthy ML**. Specifically, I am fascinated by the complexities of adversarial attacks and the methods involved in adversarial training, which play crucial roles in improving the resilience of models across diverse domains. I am also interested in exploring related areas of study such as robustness of federated learning and the dynamics of backdoor attacks and defenses. Presently, I am actively researching on the trustworthiness of generative models under the supervision of Prof. Amir Houmansadr.

PUBLICATIONS

- OSLO: One-Shot Label-Only Membership Inference Attacks**
Yuefeng Peng, **Jaechul Roh**, Subhransu Maji, Amir Houmansadr
NeurIPS 2024
[\[paper\]](#)
- Backdooring Bias into Text-to-Image Models**
Ali Naseh, **Jaechul Roh**, Eugene Bagdasaryan, Amir Houmansadr
Under Review
[\[paper\]](#) [\[code\]](#)
- Memory Triggers: Unveiling Memorization in Text-To-Image Generative Models through Word-Level Duplication**
Ali Naseh, **Jaechul Roh**, Amir Houmansadr
The 5th AAAI Workshop on Privacy-Preserving Artificial Intelligence
[\[paper\]](#)
- Understanding (Un)Intended Memorization in Text-to-Image Generative Models**
Ali Naseh, **Jaechul Roh**, Amir Houmansadr
Preprint at arXiv
[\[paper\]](#)
- Robust Smart Home Face Recognition under Starving Federated Data**
Jaechul Roh, Yajun Fang
Oral Presentation in the IEEE International Conference on Universal Village (IEEE UV2022)
[\[paper\]](#)[\[code\]](#)[\[slides\]](#)[\[video\]](#)
- MSDT: Masked Language Model Scoring Defense in Text Domain**
Jaechul Roh, Minhao Cheng, Yajun Fang
Oral Presentation in the IEEE International Conference on Universal Village (IEEE UV2022)
[\[paper\]](#)[\[code\]](#)[\[slides\]](#)[\[video\]](#)
- Impact of Adversarial Training on the Robustness of Deep Neural Networks**
Jaechul Roh
2022 IEEE 5th International Conference on Information Systems and Computer Aided Education (ICISCAE)
[\[paper\]](#)[\[code\]](#)

RESEARCH / WORK EXPERIENCE

BAID: Backdoor Attack for Gradient Inversion Defense

Final Year Thesis, Supervisor: Prof. [Jun Zhang](#)

August 2022 – May 2023

Clear Water Bay, Hong Kong

- Proposed novel text domain defense method against gradient inversion attack in the context of federated learning.

IEEE International Conference on Universal Village 2022

May 2022 – October 2022

Student Research Program, Supervisor: Dr. [Yajun Fang](#)

Cambridge, Massachusetts

- Experimented the robustness of federated learning in smart home face recognition system.

MSDT: Masked Language Model Scoring Defense in Text Domain

December 2021 – May 2022

Independent Work Research, Supervisor: Prof. [Minhao Cheng](#)

Clear Water Bay, Hong Kong

- Proposed a novel improved textual defense method against backdoor attack on pre-trained language models.

Personal Research Project

January 2022 – March 2022

Topic: “Impact of Adversarial Training on the Robustness of Deep Neural Networks”

- Experimented the effectiveness of various methods of adversarial training on improving the robustness of neural networks against classifying perturbed histopathological images.

Super Chain AI (Conard International)

June 2021 – August 2021

NLP Engineer Intern

Kowloon Bay, Hong Kong

- In charge of topic modeling and semantic analysis based on customer reviews and assigning specific semantics to the topics extracted.
- Competitors’ analysis through web-scraping customer reviews from other drop-shipping websites.

Military Service at Head Quarter of 12th Infantry Division

July 2018 – March 2020

Sergeant of Republic of Korea Army

Injae, Kang Won Do, Republic of Korea

- Officer Administrative Clerk Specialist
- Squad Leader of the Head Quarter

PROJECTS

Histopathological Scan Cancer Detection

December 2021 - January 2022

2022 Personal Winter Project, Supervisor: Prof. [Mark Vogelsberger](#) (MIT)

- Demonstrated a user-friendly application that aids to classify whether a histopathologic scan contains metastatic cancer using modified Convolutional Neural Network and modified ResNet-18.
- In charge of implementing the neural networks for the classification task.

Presentation Project on “Adversarial Attack”

September 2021 – November 2021

Machine Learning course Final Project, Instructor: Prof. [Dit-Yan YEUNG](#)

Clear Water Bay, Hong Kong

- 30-minute video presentation on the topic of “Adversarial Attack”
[slides] [video]

SKILLS / LANGUAGES

Programming Language: Python

Languages: Korean (Native), English (Native), Chinese (Fluent)