JAECHUL (Harry) Roh

Ph.D. in Computer Science, University of Massachusetts Amherst
+1 (470) 915 - 1137 · jroh@umass.edu · Personal Website · Github · Google Scholar

EDUCATION

University of Massachusetts Amherst — **September 2023 – Present**
Ph.D. in Computer Science — Amherst, Massachusetts, USA
Advisor: Prof. Amir Houmansadr
GPA: 4.0/4.0

Hong Kong University of Science and Technology — **September 2017 – May 2023**
B.Eng. in Computer Engineering, School of Engineering — Clear Water Bay, Hong Kong, HK
Final Year Thesis Advisor: Prof. Jun Zhang

RESEARCH INTERESTS

My research is centered on **Privacy & Security in AI** and **Trustworthy Machine Learning**. Currently, I am investigating the trustworthiness of multimodal generative models across various domains, including **text-to-image** and **audio-based** modalities, under the supervision of Prof. Amir Houmansadr.

PUBLICATIONS

1. **OSLO: One-Shot Label-Only Membership Inference Attacks**
   Yuefeng Peng, **Jaechul Roh**, Subhransu Maji, Amir Houmansadr
   NeurIPS 2024
   [paper]

2. **Backdooring Bias into Text-to-Image Models**
   Ali Naseh, **Jaechul Roh**, Eugene Bagdasaryan, Amir Houmansadr
   Under Review
   [paper] [code]

3. **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]

4. **Understanding (Un)Intended Memorization in Text-to-Image Generative Models**
   Ali Naseh, **Jaechul Roh**, Amir Houmansadr
   Preprint at arXiv
   [paper]

5. **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]

6. **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]

7. **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 — **August 2022 – May 2023**
Final Year Thesis, Supervisor: Prof. Jun Zhang — 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-scrapping 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)