

Xiao Li

Sancai building 906 or FIT building 1-508, Tsinghua University, Beijing, P.R. China, 100084

☎ (+86) 188-1096-3683 | ✉ xiaoli.cst@gmail.com | 🏠 lixiaothu.github.io | 🎓 Google Scholar

Education

Tsinghua University

Ph.D. Student Department of Computer Science and Technology

Beijing, China

2020/09 - Present

- Advisor: Prof. Xiaolin Hu and Prof. Bo Zhang.
- TSAIL Group (directed by Prof. Bo Zhang and Prof. Jun Zhu).
- **Research interests:** My researches aim to build up trustworthy AI systems, hopefully bring AI closer to human-level intelligence. With this goal, I have explored topics including adversarial machine learning, representation learning, brain-inspired learning, scalable multimodal learning, and generative models. My current research interests focus on *the security of foundation models*, including large text-to-image and text-to-video generative models, and large language models.

Tsinghua University

B.ENG. Department of Computer Science and Technology

Beijing, China

2016/09 - 2020/06

- Minor in Statistics. in Center for Statistical Science, Tsinghua University.

Selected Publications

See more publications on Google Scholar.

Preprint & Under review

- **Xiao Li**, Wenxuan Sun, Huanran Chen, Qiongxiu Li, Yining Liu, Yingzhe He, Jie Shi, Xiaolin Hu. *Adversarial Diffusion Bridge Model for Reliable Adversarial Purification*. Under review.
- **Xiao Li**, Yining Liu, Na Dong, Sitian Qin, Xiaolin Hu. *PartImageNet++ Dataset: Scaling up Part-based Models for Robust Recognition*. Under review.
- **Xiao Li**, Hang Chen, Xiaolin Hu. *On the Importance of Backbone to the Adversarial Robustness of Object Detectors*. Under review & Preprint. arXiv:2305.17438.
- Hang Chen, Chufeng Tang, **Xiao Li**, Lesi Wei, Xiaolin Hu. *Improving Neuron Segmentation in Electron Microscopy with Affinity-Guided Queries*. Under review.
- Wei Zhang, Zhanhao Hu, **Xiao Li**, Xiaopei Zhu, Xiaolin Hu. *Adversarial Patch Defenses Give a False Sense of Security for Physical Defense: Circumventing Defenses with a Single Set of Clothes*. Under review.

Published & Accepted

- **Xiao Li**, Wei Zhang, Yining Liu, Zhanhao Hu, Xiaolin Hu. *Language-Driven Anchors for Zero-Shot Adversarial Robustness*. CVPR 2024.
- **Xiao Li**, Qiongxiu Li, Zhanhao Hu, Xiaolin Hu. *On the Privacy Effect of Data Enhancement via the Lens of Memorization*. IEEE TIFS 2024.
- **Xiao Li**, Ziqi Wang, Bo Zhang, Fuchun Sun, Xiaolin Hu. *Recognizing Object by Components with Human Prior Knowledge Enhances Adversarial Robustness of Deep Neural Networks*. IEEE TPAMI 2023.
- Xiaolin Hu, Chufeng Tang, Hang Chen, **Xiao Li**, Jianmin Li, Zhaoxiang Zhang. *Improving Image Segmentation with Boundary Patch Refinement*. IJCV 2022.
- Chufeng Tang, Hang Chen, **Xiao Li**, Jianmin Li, Zhaoxiang Zhang, Xiaolin Hu. *Look closer to segment better: Boundary patch refinement for instance segmentation*. CVPR 2021.
- Xiaopei Zhu, **Xiao Li**, Jianmin Li, et al. *Fooling thermal infrared pedestrian detectors in real world using small bulbs*. AAAI 2021.

Experience

Tsinghua University

Teaching Assistant

Beijing, China

2021/08-2023/02

- 2021/08 - 2023/02. *Training Camp of Deep Learning*, instructed by Prof. Xiaolin Hu.
- 2022 Fall. *Neural and Cognitive Computation* (THU-80240642), instructed by Prof. Xiaolin Hu.
- 2021 Fall. *Introduction to Deep Learning* (THU-00240332), instructed by Prof. Xiaolin Hu.

Momenta

Research Intern

Beijing, China

2018/08-2018/12

- Adviser: Prof. Xiang Li
- Research topic: appearance-based gaze direction estimation, multi-task learning, continual learning, attention mechanism, etc.

Presentation

- *Insights into Security Risks of the Diffusion (Text-to-Image) Generative Models*. Shield Laboratory, Singapore. 2023/09.
- *Adversarial Suffix Prompt Attacks and Defenses on Large Language Models*. Shield Laboratory, Singapore. 2024/03.

Service

- Conference Reviewer: AAAI 2023, CVPR 2023, NeurIPS 2023, AAAI 2024, CVPR 2024, ECCV 2024.
- Journal Reviewer: IEEE TPAMI, IEEE TIP, IEEE TIFS.