Xia Song

Email: sxia002@e.ntu.edu.sg

EDUCATION

MSc in signal processing, Nanyang Technological University 2021 - 2022
Supervised by Professor Yap Kim Hui.
Dissertation project: Certified Defense and Robust AI
Average score:

BSc in Communication Engineering, Wuhan University 2016 - 2020
Supervised by Professor Hao Jiang.
Dissertation project: The Generation and Defense of Adversarial Examples Based on Reinforcement Learning
Average score: 84/100

RESEARCH EXPERIENCE

Research on certified defense method and watermark attack in deep neural networks

Jun. 2020- Mar. 2021

- Work as a research assistance in Zhejiang Lab for three months, work as a research assistance in Jiang Hao's research group for six months.
- My work is conducting research about adversarial defense and attack methods in deep neural networks, including enhancing the certified robustness for deep learning classification models by maximize Randomized Smoothing theory margin and researching the rotation feature for watermark attack to convolution neural networks.

Research on interactive adversarial attack methods in deep learning

Oct. 2019- May 2020

- Work as a part of Professor Jiang Hao's research group in the field of robust AI
- My work is mainly about the black box attack methods in image classification models. A reinforcement-based interactive attack method is proposed to attack the classification models in MNIST dataset.

Research on the data analysis

Jan. 2019- Jun. 2019

- Work as a part of Professor Jiang Hao's research group in the field of big data analysis
- My work is analyzing the mobile phone signal data for guiding the internet advertisement delivery, which uses the Multi-state models, Hierarchical clustering, and Tucker decomposition to mine the potential behavioral preferences of people.

PUBLICATIONS

Published:

- Hao Jiang; Jintao Yang; Guang Hua; Lixia Li; Ying Wang; Shenghui Tu; **Song Xia**. FAWA: Fast Adversarial Watermark Attack. In *IEEE Transactions on Computers*, 2021.
- Wenli Xiao, Hao Jiang, and **Song Xia**. A new black box attack generating adversarial examples based on reinforcement learning. In *2020 Information Communication Technologies Conference (ICTC).* Pages 141-146, 2020.
- **Song Xia**, Hao Jiang, Yi Zhang, Duo Peng. Internet advertising investment analysis based on Beijing and Jinhua signaling data. *In 2019 IEEE International Conference on Computational Science and*

Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC), pages 419-426, 2019.

Submitted:

- **Song Xia**, Hao Jiang, Haijun Shan, Jintao Yang. SODM: Maximize Certified Robustness in Randomized Smoothing by Solvable Optimization Via Differentiable Mapping, *In 2021 IEEE conference on computer vision and pattern recognition*, 2021.

PROGRAMMING SKILL

Pytroch, Python, Matlab, $C^{\#}$.