Yanyun Wang 王彦云

DOB: 15/05/2000
Tel: +86 18678857129
Email: yynwang@connect.hku.hk
Homepage: https://yywang.netlify.app

Education

M.Phil. Artificial Intelligence, The Hong Kong University of Science and Technology (Guangzhou) **09/2024 – Now**
- **GPA:** 3.70/4.3 - **Supervisor:** Prof. Li Liu

M.Sc. Computer Science, The University of Hong Kong **09/2022 – 10/2023**
- **GPA:** 3.35/4.3
- **Degree Project:** *Towards Robust Speaker Recognition through Crafting Imperceptible Adversarial Speech Samples* (Grade: A+)

B.Eng. Software Engineering, East China Normal University (985 / Double First Class) **09/2018 - 06/2022**
- **GPA:** 3.52/4.0 - **Ranking:** 30/186 (16.13%)
- **Degree Thesis:** *RNN Adversarial Samples Generation Approach based on Weighted Finite Automaton Abstraction* (Grade: A)

Publications & Preprints

As The 1st Author:

New Paradigm of Adversarial Training: Breaking Inherent Trade-Off between Accuracy and Robustness via Dummy Classes
Under review (ICLR'25 6665 rej -> IJCAI'25 first round passed) https://arxiv.org/abs/2410.12671 (arXiv preprint)

TSFool: Crafting Highly-Imperceptible Adversarial Time Series through Multi-Objective Attack
27th European Conference on Artificial Intelligence (ECAI'24) **Oral**
https://ebooks.iospress.nl/doi/10.3233/FAIA240644

Meta Pattern Concern Score: A Novel Evaluation Measure with Human Values for Multi-classifiers
2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC'23)
https://ieeexplore.ieee.org/abstract/document/10394380

As Co-Author:

BackdoorDM: A Comprehensive Benchmark for Backdoor Learning in Diffusion Model
Under review (IJCAI'25 first round passed) https://arxiv.org/abs/2502.11798 (arXiv preprint)

"Yes, My LoRD." Guiding Language Model Extraction with Locality Reinforced Distillation
Under review (ACL Rolling Review - 2025 February) https://arxiv.org/abs/2409.02718 (arXiv preprint)

Efficient Adversarial Sequence Generation for RNN with Symbolic Weighted Finite Automata
SafeAI Workshop @ 36th AAAI Conference on Artificial Intelligence (AAAI'22)
http://ceur-ws.org/Vol-3087/paper_19.pdf Best Paper Award Nomination

Research & Intern Experience

Research Assistant, The Hong Kong Polytechnic University **10/2023 - 06/2024**
- **Laboratory:** Applied Security, Trust And Privacy Lab for Enterprise (ASTAPLE) - **Supervisor:** Prof. Haibo Hu

Research Assistant (part-time), East China Normal University **12/2021 - 09/2023**
- **Laboratory:** Shanghai Key Laboratory of Trustworthy Computing - **Supervisor:** Prof. Dehui Du

Algorithm Engineer, Ping An Technology Co., Ltd **08/2021 - 11/2021**

- **Department:** NLP Innovation Research and Development Department, OLATOP Knowledge Graph Team

Software Development Engineer, Dareway Software Co., Ltd **07/2020 - 08/2020**

Extracurricular Practice

Summer Workshop Student, School of Computing, National University of Singapore **05/2021 - 07/2021**

- **Topic:** AI/ML for Financial Services - **Grade:** A+
- **Project:** *Portfolio Management - Based on LSTM Models and Optimal Combination*

Participant, Shanghai College Students' Innovation and Entrepreneurship Training Program **10/2019 - 10/2020**

- **Project:** *Jiangnan Mengxun* – a 2.5D indie word adventure game - **Grade:** Provincial Level - B

Associate Director, Human Resource Center of Students' Union **06/2019 - 06/2020**

Honors & Awards

Excellent Graduate Award, East China Normal University **2022**

Excellent Bachelor's Degree Thesis Award, Software Engineering Institute, East China Normal University **2022**

People's Choice Award, School of Computing Summer Workshop, National University of Singapore **2021**

Excellent Undergraduate Student, East China Normal University **2021**

Second-class Scholarship, East China Normal University **2021**

Excellent Undergraduate Student, East China Normal University **2020**

First-class Scholarship, East China Normal University **2020**

Services

Conference Reviewer: CVPR (2025, 2024), ICCV (2025), AAAI (2025)

Journal Reviewer: TKDE