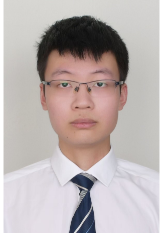
# Hanling Tian

Master student, Shanghai Jiao Tong University, Shanghai, China
Supervisor: Prof. Xiaolin Huang
hanlingtian@sjtu.edu.cn — (+86) 18905258081 — BlueBlood6.github.io



**Research interest**: Machine Learning, Generative models, LLM Agent Safety

## EDUCATION

**Shanghai Jiao Tong University**, Shanghai, China — 09 2023 — 03 2026
Master student in Automation Science and Engineering — GPA: 3.83/4.00

**Xi'an JiaoTong University**, Xi'an, China — 09 2019 — 06 2023
Bachelor of Engineering: Automation — GPA: 4.11/4.30, Average Grade: 94.15/100, Rank: 1/191
Young Gifted Program, Qian Xuesen Honors College, Outstanding Graduate
National Scholarship (0.2%): Ministry of Education, the People's Republic of China

## PUBLICATIONS

### Featured Publications

**H. Tian**, Y. Liu, M. He, Z. He, Z. Huang, R. Yang & X. Huang (2025). **Simulating Training Dynamics to Reconstruct Training Data from Deep Neural Networks**. ICLR 2025. https://openreview.net/forum?id=ZJftXKy12x

- We propose SimuDy to successfully reconstruct training data from a trained ResNet's parameters for the first time.
- We consider trained parameters as accumulation of gradients throughout the dynamical training process and formulate dataset reconstruction into a high-level gradient inversion attack.
- We show that indeed there is memorization in DNNs, providing a promising tool for investigating deep learning memory.

**H. Tian**, Z. Sha, J. Wang, Y. Hang, Z. Huang & X. Huang (2025). **InjecMEM: Memory Injection Attack on LLM Agent Memory Systems**. Submitted to ICLR 2026.

- We identify and formalize the core vulnerability of agent memory systems.
- We propose an injection attack that interacts with agents using crafted prompt and causes subsequent harmful outputs.
- We indirectly inject poisoned memory through other subsystems and show the black-box transferability of our method.

### Collaborative Publications

Z. Sha, **H. Tian**, Z. Xu, S. Cui, C. Meng & W. Wang (2025). **Agent Safety Alignment via Reinforcement Learning**. ArXiv. https://arxiv.org/pdf/2507.08270

Q. Xiao, **H. Tian**, Z. Huang & X. Huang (2025). **GradCFG: Gradient Inversion of Classifier-Free Guidance Diffusion Models**. Submitted to ICLR 2026.

M. He, R. Yang, **H. Tian**, Y. Qiu & X. Huang (2025). **Primphormer: Efficient Graph Transformers with Primal Representations**. ICML 2025. https://openreview.net/forum?id=fMAihjfJij

Z. Huang, Y. Hang, B. Lin, Y. Lou, Z. He, **H. Tian**, T. Li & X. Huang (2025). **RAIN-Merging: A Gradient-Free Method to Enhance Instruction Following in Large Reasoning Models with Preserved Thinking Format**. Submitted to ICLR 2026.

Z. Huang, Y. Hang, Y. Lou, Z. He, M. He, W. Zhou, **H. Tian**, T. Li, K. Li, Z. Huang & X. Huang (2025). **T2I-ConBench: Text-to-Image Benchmark for Continual Post-training**. Submitted to ICLR 2026.

D. Huang, J. Guo, S. Sun, **H. Tian**, J. Lin, Z. Hu, C. Lin, J. Lou & D. Zhang (2023). **A Survey for Graphic Design Intelligence**. ArXiv. https://arxiv.org/pdf/2309.01371.

## PROFESSIONAL EXPERIENCE

**Microsoft Research Asia** (MSRA), Beijing, China
Intern of Data, Knowledge, and Intelligence Group. Mentor: Shizhao Sun — 07 2022 — 06 2023
**Pre-training of Graphic Layout Generation** & **Design Image Generation with Text Constrains**

**Ant Group**, Shanghai, China
Intern of Security and Risk Management Group - LLM Safety. Mentor: Changhua Meng — 05 2025 — 08 2025
**Agent Safety Alignment via Reinforcement Learning** & **Attacks to Agent Memory System**