Weitao Feng

weitaofeng@mail.ustc.edu.cn Weitao Feng github.com/Georgefwt

Education

University of Science and Technology of China **Sep. 2022 – Now**
Master of Engineering in Cyber Scurity *Hefei, Anhui*

Microsoft Research AI4Science **Nov. 2022 – Jan. 2024**
Research Intern *Beijing*

University of Science and Technology of China **Sep. 2018 – May 2022**
Bachelor of Engineering in Information Science *Hefei, Anhui*

Publications

Toward White-box Protection for Customized Stable Diffusion Models via Watermark LoRA
Weitao Feng**, Wenbo Zhou, Jiyan He, Jie Zhang, Tianyi Wei, Guanlin Li, Tianwei Zhang, etal.* ***ICML *2024*

- Point out the necessity for white-box protection to Stable Diffusion model.
- Propose the Scaling Matrix for LoRA and prior-preserving fine-tuning.
- Release the first implementation of a watermark by integrating the watermark into the U-Net structure.

Catch You Everything Everywhere: Guarding Textual Inversion via Concept Watermarking
***Weitao Feng**, Jiyan He, Jie Zhang, Tianwei Zhang, Wenbo Zhou, Weiming Zhang, Nenghai Yu* *2023*

- Point out guarding against concept sharing is necessary, with novel concept watermarking proposed as a solution.
- Utilize a progressive training strategy is adopted to enhance concept watermarking's fidelity.
- Concept watermarking shows inherent robustness against pixel-level distortions.

Control Risk for Potential Misuse of Artificial Intelligence in Science
Jiyan He,* ***Weitao Feng****, Yaosen Min*, Jingwei Yi*, etal.* *2023*

- Identify and classify potential risks associated with AI models in science.
- Propose building safeguarded scientific AI systems, and present SciGuard.
- Build Sci-Safetybench, a red-teaming dataset to assess the safety risks of scientific AI systems

Predicting Equilibrium Distributions for Molecular Systems with Deep Learning
Shuxin Zheng, Jiyan He*, Chang Liu*, Yu Shi*, Ziheng Lu*,* ***Weitao Feng**, etal.* ***Nature Machine Intelligence*** *2024*

- Introduction of Distributional Graphormer (DiG) for predicting molecular systems' equilibrium distributions.
- DiG employs deep neural networks for efficient conformation generation and state density estimation.
- Application of DiG in tasks like protein conformation and ligand structure sampling, demonstrating its versatility.

Recent Projects

Face Landmark ControlNet | *Stable Diffusion, ControlNet* **Mar 2023**

- Train a ControlNet based on 68 face landmarks to achieve finer face control. This project is finished a few weeks after ControlNet's first release.
- Generate faces with identical poses and expressions; Control the facial expressions and poses of generated images.
- Available both on GitHub and Huggingface.

Human Perception Conformed Authenticity Assessment for Deepfake | *Deepfake, IQA* **Oct 2022**

- Conducted a crowdsourced quality assessment to create a deepfake quality assessment dataset
- Propose a novel deepfake authenticity assessment metric consistent with human perception.

Selected Awards

- Best Paper at the 3rd CSIG China Media Forensics and Security Conference, 2023
- The 4th place in 2022 CVPR biometric workshop face forgery detection track
- 2021 China Artificial Intelligence Competition Multimedia Forgery Generation Track A Level
- Gold Medal, Scholarship for Outstanding Students of USTC, 2021
- Wang Xiaomo Scholarship for Excellence in Science and Technology, 2020
- The second prize of Anhui Province in the 2020 National College Student Mathematics Competition

Activities

Laboratory | *Laboratory Cluster Administrator* **Oct 2022 - Now**

- Cluster storage structure modification and expansion, from ZFS to Ceph.
- Replacement of cluster compute nodes and routine maintenance.

Information Security Design and Practice Class | *Teaching Assistant* **Mar 2022 - Jun 2022**

- Teach web browser security and designing the experiments.

USTC Nebula CTF Team | *Team Member* **Jun 2020 - Dec 2021**

- Won 2nd place in 2021 TencentCTF.
- Participated in the organizing committee for Hackergame 2021 (with 4k registered participants).

Student Union | *Vice president of Propaganda Department* **Jun 2020 - Jun 2021**

- Run the Student Union's WeChat Official Account.

Linux User Group | *Group Member* **Oct 2019 - Now**