Qidong Huang

Ph.D, University of Science and Technology of China

Building No.7, USTC West Campus
Hefei, Anhui, China
(+86) 13085060686
hqd0037@mail.ustc.edu.cn

Short Biography

Qidong Huang is a PhD student at University of Science and Technology of China. He has published more than 10 papers at top1-tier conferences and journals, such as CVPR/ICCV/AAAI/TIP/TCSVT. His research interest focus on multi-modal LLMs and trustworthy/efficient AI, including explainable VLMs, parameter-efficient fine-tuning, AI privacy and robustness. He is the reviewer of many top conferences and journals (e.g., CVPR, ICCV, ECCV) and top journals (e.g., TNNLS, TIP, PR).

Education

09/2020–present **PhD of Cyberspace Security**, *University of Science and Technology of China*, Hefei, China, CAS Key Laboratory of Electromagnetic Space Information. Supervised by Prof. Weiming Zhang.

09/2016–06/2020 **Bachelor of Information Security**, *School of Information Science and Technology, University of Science and Technology of China*, Hefei, China.

Skills

- **Expertise in multi-modal LLMs :** My recent research mainly focuses on multi-modal large language models, where I have published one paper (first author) on top-tier computer vision conferences. The most recent work is OPERA, which investigates the deep reason for the hallucination of multi-modal LLMs, and gives a information attenuation explanation. Based on this, we propose the over-trust logit penalty and retrospection-allocation mechanism to mitigate the hallucination issue.
- **Expertise in trustworthy AI :** I have been researching the trustworthy system for supervised/unsupervised vision models, where I published four paper (first author) on top-tier computer vision conferences. Another one is RobustMAE, which reveals the flaw of masked-autoencoder-style vision pretraining on adversarial robustness, and improve it with test-time frequency-domain prompting. Moreover, I also dedicated the early time of my PhD career to other topics, such as adversarial attack/defense for 3D models and anti-DeepFake (where we are the first to propose the concept of "initiative defense" against DeepFakes by proactively protecting users' facial privacy before the manipulation, unlike previous ex-post countermeasures like DeepFake detection.)
- **Expertise in efficient AI :** I have been researching the prompt learning for large-scale vision pretrained models and published one paper on top-tier computer vision conferences, in which I propose DAM-VP, a data diversity-aware method for efficient and adaptive vision prompt learning. This work addresses the mismatch between vision prompts and downstream data diversity.

Publications (First Author)

- **Qidong Huang**, Xiaoyi Dong, Pan Zhang, Bin Wang, Conghui He, Jiaqi Wang, Dahua Lin, Weiming Zhang, Nenghai Yu. OPERA : Alleviating Hallucination in Multi-Modal Large Language Models via Over-Trust Penalty and Retrospection-Allocation. *Conference on Computer Vision and Pattern Recognition (CVPR), 2024.* *(Highlight, 2.8% of submissions)*
- **Qidong Huang**, Xiaoyi Dong, Dongdong Chen, Hang Zhou, Weiming Zhang, Kui Zhang, Gang Hua, Nenghai Yu. PointCAT : Contrastive Adversarial Training for Robust Point Cloud Recognition. *IEEE Transactions on Image Processing (TIP), 2024.*

- ⋆ **Qidong Huang**, Xiaoyi Dong, Dongdong Chen, Yinpeng Chen, Lu Yuan, Gang Hua, Weiming Zhang, Nenghai Yu. Improving Adversarial Robustness of Masked Autoencoders via Test-time Frequency-domain Prompting. *International Conference on Computer Vision (ICCV), 2023.*
- ⋆ **Qidong Huang**, Xiaoyi Dong, Dongdong Chen, Weiming Zhang, Feifei Wang, Gang Hua, Nenghai Yu. Diversity-Aware Meta Visual Prompting. *Conference on Computer Vision and Pattern Recognition (CVPR), 2023.*
- ⋆ **Qidong Huang**, Xiaoyi Dong, Dongdong Chen, Hang Zhou, Weiming Zhang, Nenghai Yu. Shape-invariant 3D Adversarial Point Clouds. *Conference on Computer Vision and Pattern Recognition (CVPR), 2022.*
- ⋆ **Qidong Huang***, Jie Zhang*, Wenbo Zhou, Weiming Zhang, Nenghai Yu. Initiative Defense against Facial Manipulation. *AAAI Conference on Artificial Intelligence (AAAI), 2021.* (*Qidong Huang and Jie Zhang contribute equally.)

Publications (Collaborate)

- ⋆ Feifei Wang, Zhentao Tan, Tianyi Wei, Yue Wu, **Qidong Huang**†. SimAC : A Simple Anti-Customization Method against Text-to-Image Synthesis of Diffusion Models. *Conference on Computer Vision and Pattern Recognition (CVPR), 2024.* *(† Corresponding author)*
- ⋆ Kui Zhang, Hang Zhou, Jie Zhang, **Qidong Huang**, Weiming Zhang, Nenghai Yu. Ada3Diff : Defending against 3D Adversarial Point Clouds via Adaptive Diffusion. *ACM International Conference on Multimedia (MM), 2023*
- ⋆ Han Fang, Dongdong Chen, **Qidong Huang**, Jie Zhang, Zehua Ma, Weiming Zhang and Nenghai Yu. Deep Template-based Watermarking. *IEEE Transactions on Circuits and Systems for Video Technology (TCSVT), 2020.*
- ⋆ Jie Zhang, Dongdong Chen, **Qidong Huang**, Jing Liao, Weiming Zhang, Huamin Feng, Gang Hua, Nenghai Yu. Poison ink : Robust and invisible backdoor attack. *IEEE Transactions on Image Processing (TIP), 2022.*

Services

- ⋆ Reviewer for CVPR 2022, 2023, 2024
- ⋆ Reviewer for ICCV 2023
- ⋆ Reviewer for ECCV 2022, 2024
- ⋆ Reviewer for NeurIPS 2024
- ⋆ Reviewer for ACCV 2024
- ⋆ Reviewer for ICPR 2022
- ⋆ Reviewer for IEEE Transactions on Neural Networks and Learning Systems (TNNLS)
- ⋆ Reviewer for IEEE Transactions on Image Processing (TIP)
- ⋆ Reviewer for Pattern Recognition (PR)

Awards & Honors

2021 China National Scholarship

2023 "Internet +" Innovation and Entrepreneurship Competition, Provincial Bronze Award

2023 Anheng Information Scholarship