

# SHAO, Rui



## Professor

School of Computer Science and Technology,  
Harbin Institute of Technology (Shenzhen)

**Homepage:** <https://rshaojimmy.github.io/>

**GitHub:** <https://github.com/rshaojimmy>

**Email:** shaorui@hit.edu.cn; rshaojimmy@gmail.com

(Birthday: 1992.12)

## ❖ Education

---

- **Doctor of Philosophy**  
Hong Kong Baptist University  
—— Department of Computer Science //2017-2021
- **Bachelor of Science**  
University of Electronic Science and Technology of China  
—— School of Information and Communication Engineering // 2011-2015
- **High School**  
Shenzhen Foreign Languages School // 2008-2011

## ❖ Research

---

- **Nanyang Technological University, Singapore** // 2021 – 2023  
**Research fellow**  
**Advisor:** [Prof. Ziwei Liu](#), [Prof. Chen Change Loy](#)  
([Prof. Ziwei Liu serves as an Associate Editor of International Journal of Computer Vision](#);  
[Prof. Chen Change Loy serves as an Associate Editor of International Journal of Computer Vision](#)  
[and IEEE Transactions on Pattern Analysis and Machine Intelligence](#))
- **Hong Kong Baptist University, Hong Kong, China** //2017-2021  
**Ph.D**  
**Advisor:** [Prof. Pong C. Yuen](#)  
([Chair Professor; IAPR Fellow](#))
- **Johns Hopkins University, Baltimore, U.S.** // 2020.02 – 2020.07  
**Visiting scholar**  
**Advisor:** [Prof. Vishal M Patel](#)  
([Associate Editor of IEEE Transactions on Pattern Analysis and Machine Intelligence](#))
- **SenseTime, Shenzhen, China** // 2021.07 – 2021.11  
**Researcher**

## Research Direction: Multimodal Large Language Models and Embodied AI

(\*denotes corresponding authors)

### First Author/Corresponding Author Papers:

#### Journal Papers:

- **Rui Shao**, Tianxing Wu, Jianlong Wu, Liqiang Nie, Ziwei Liu. "Detecting and Grounding Multi-Modal Media Manipulation and Beyond". *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2024. (SCI Q1; Impact Factor: 23.6; CCF-A; No. 1 Journal in Computer Vision and Artificial Intelligence)
- **Rui Shao**, Tianxing Wu, Liqiang Nie, Ziwei Liu. "DeepFake-Adapter: Dual-Level Adapter for DeepFake Detection". *International Journal of Computer Vision (IJCV)*, 2025. (JCR Q1; Impact Factor: 13.369; CCF-A; Top Journal in Computer Vision)
- **Rui Shao**, Tianxing Wu, Ziwei Liu. "Robust Sequential DeepFake Detection". *International Journal of Computer Vision (IJCV)*, 2025. (JCR Q1; Impact Factor: 13.369; CCF-A; Top Journal in Computer Vision)
- **Rui Shao**, Pramuditha Perera, Pong C. Yuen, Vishal M. Patel. "Open-set Adversarial Defense with Clean-Adversarial Mutual Learning". *International Journal of Computer Vision (IJCV)*, 2022. (JCR Q1; Impact Factor: 13.369; CCF-A; Top Journal in Computer Vision)
- **Rui Shao**, Pramuditha Perera, Pong C. Yuen, Vishal M. Patel. "Federated Generalized Face Presentation Attack Detection". *IEEE Transactions on Neural Networks and Learning Systems (TNNLS)*, 2022. (SCI Q1; Impact Factor: 14.255; CCF-B; Top Journal in Artificial Intelligence)
- **Rui Shao**, Xiangyuan Lan, Pong C. Yuen. "Joint Discriminative Learning of Deep Dynamic Textures for 3D Mask Face Anti-spoofing". *IEEE Transactions on Information Forensics and Security (TIFS)*, 2019. (SCI Q1; Impact Factor: 7.231; CCF-A; Top Journal in Security of Artificial Intelligence)
- **Rui Shao**, Xiangyuan Lan. "Adversarial Auto-encoder for Unsupervised Deep Domain Adaptation". *IEEE Image Processing*, 2019. (SCI Q3; Impact Factor: 2.373; CCF-C)

#### Conference Papers:

- Bin Xie, **Rui Shao\***, Gongwei Chen, Kaiwen Zhou, Yinchuan Li, Jie Liu, Min Zhang, Liqiang Nie. "GUI-explorer: Autonomous Exploration and Mining of Transition-aware Knowledge for GUI Agent". *The 63rd Annual Meeting of the Association for Computational Linguistics (ACL)*, 2025. (CCF-A; Top Conference in Natural Language Processing)
- Hao Li, Qi Lv, **Rui Shao\***, Xiang Deng, Yinchuan Li, Jianye HAO, Liqiang Nie. "STAR: Learning Diverse Robot Skill Abstractions through Rotation-Augmented Vector Quantization". *International Conference on Machine Learning (ICML)*, 2025. **Spotlight**. (CCF-A; Top Conference in Machine Learning)
- Wei Li, Bing Hu, **Rui Shao\***, Leyang Shen, Liqiang Nie. "LION-FS: Fast & Slow Video-Language Thinker as Online Video Assistant". *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025. (CCF-A; Top Conference in Computer Vision)
- Zaijing Li, Yuquan Xie, **Rui Shao\***, Gongwei Chen, Dongmei Jiang, Liqiang Nie\*. "Optimus-2: Multimodal Minecraft Agent with Goal-Observation-Action Conditioned Policy". *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025. (CCF-A; Top Conference in Computer Vision)
- Jingxuan Chen, Derek Yuen, Bin Xie, Yuhao Yang, Gongwei Chen, Zhihao Wu, Li Yixing, Xurui Zhou, Weiwen Liu, Shuai Wang, Kaiwen Zhou, **Rui Shao\***, Liqiang Nie, Yasheng Wang, Jianye Hao, Jun Wang, Kun Shao\*. "SPA-Bench: A Comprehensive Benchmark for SmartPhone Agent Evaluation". *The Thirteenth International Conference on Learning Representations (ICLR)*, 2025. **Spotlight**. (TH-CPL-A; Top Conference in Machine Learning)
- Gongwei Chen, Leyang Shen, **Rui Shao\***, Xiang Deng, Liqiang Nie\*. "LION: Empowering Multimodal Large Language Model with Dual-Level Visual Knowledge". *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024. (CCF-A; Top Conference in Computer Vision; SOTA performance in 13

### benchmarks of Multimodal Large Language Model)

- Zaijing Li, Yuquan Xie, **Rui Shao\***, Gongwei Chen, Dongmei Jiang, Liqiang Nie\*. "Optimus-1: Hybrid Multimodal Memory Empowered Agents Excel in Long-Horizon Tasks". *Thirty-eighth Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2024. (CCF-A; Top Conference in Machine Learning)
- Leyang Shen, Gongwei Chen, **Rui Shao\***, Weili Guan, Liqiang Nie\*. "MoME: Mixture of Multimodal Experts for Generalist Multimodal Large Language Models". *Thirty-eighth Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2024. (CCF-A; Top Conference in Machine Learning)
- **Rui Shao**, Tianxing Wu, Ziwei Liu. "Detecting and Grounding Multi-Modal Media Manipulation". *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023. (CCF-A; Top Conference in Computer Vision)
- **Rui Shao**, Tianxing Wu, Ziwei Liu. "Detecting and Recovering Sequential DeepFake Manipulation". *European Conference on Computer Vision (ECCV)*, 2022. (TH-CPL-A; CCF-B; Top Conference in Computer Vision)
- **Rui Shao**, Pramuditha Perera, Pong C. Yuen, Vishal M. Patel. "Open-set Adversarial Defense". *European Conference on Computer Vision (ECCV)*, 2020. (TH-CPL-A; CCF-B; Top Conference in Computer Vision)
- **Rui Shao**, Xiangyuan Lan, Pong C. Yuen. "Regularized Fine-grained Meta Face Anti-spoofing". *Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI)*, 2020. (CCF-A; Top Conference in Artificial Intelligence)
- **Rui Shao**, Xiangyuan Lan, Jiawei Li, Pong C. Yuen. "Multi-adversarial Discriminative Deep Domain Generalization for Face Presentation Attack Detection". *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019. (CCF-A; Top Conference in Computer Vision)
- **Rui Shao**, Xiangyuan Lan, Pong C. Yuen. "Feature Constrained by Pixel: Hierarchical Adversarial Deep Domain Adaptation". *ACM International Conference on Multimedia (ACM MM)*, 2018. (CCF-A; Top Conference in Multimedia)
- **Rui Shao**, Bochao Zhang, Pong C. Yuen, Vishal M. Patel. "Federated Test-Time Adaptive Face Presentation Attack Detection with Dual-Phase Privacy Preservation". *IEEE International Conference on Automatic Face and Gesture Recognition (FG)*, 2021. (TH-CPL-B; CCF-C; Top Conference in Biometrics)
- **Rui Shao**, Xiangyuan Lan, Pong C. Yuen. "Deep Convolutional Dynamic Texture Learning with Adaptive Channel-discriminability for 3D Mask Face Anti-spoofing". *International Joint Conference on Biometrics (IJB)*, 2017. (TH-CPL-B; CCF-C; Top Conference in Biometrics)

### Co-Author Papers:

- Qi Lv, Hao Li, Xiang Deng, **Rui Shao**, Yinchuan Li, Jianye HAO, Longxiang Gao, MICHAEL YU WANG, Liqiang Nie. "Spatial-Temporal Graph Diffusion Policy with Kinematics Modeling for Bimanual Robotic Manipulation." *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025. (CCF-A; Top Conference in Computer Vision)
- Qilang Ye, Zitong Yu, **Rui Shao**, Xinyu Xie, Philip Torr, Xiaochun Cao, "CAT: Enhancing Multimodal Large Language Model to Answer Questions in Dynamic Audio-Visual Scenarios". *European Conference on Computer Vision (ECCV)*, 2024. (TH-CPL-A; CCF-B; Top Conference in Computer Vision)
- Qi Lv, Hao Li, Xiang Deng, **Rui Shao**, Michael Yu Wang, Liqiang Nie. "RoboMP: A Robotic Multimodal Perception-Planning Framework with Multimodal Large Language Models". *International Conference on Machine Learning (ICML)*, 2024. (CCF-A; Top Conference in Machine Learning)
- Xinyu Hou, Liming Jiang, **Rui Shao**, Chen Change Loy. "Video Infilling with Rich Motion Prior". *British Machine Vision Conference (BMVC)*, 2023. (CCF-C; Famous Conference in Computer Vision)
- Xiangyuan Lan, Mang Ye, **Rui Shao**, Bineng Zhong, Pong C. Yuen, Huiyu Zhou. "Learning Modality-Consistency Feature Templates: A Robust RGB-Infrared Tracking System". *IEEE Transactions on Industrial Electronics (TIE)*, 2019. (SCI Q1; Impact Factor: 8.236)
- Chong Yin, Siqi Liu, **Rui Shao**, Pong C. Yuen. "Focusing on Clinically Interpretable Features: Selective Attention Regularization for Liver Biopsy Image Classification". *Medical Image Computing and Computer Assisted Interventions (MICCAI)*, 2021. (CCF-B; Top Conference in Medical Imaging)

### Pre-prints

- Renshan Zhang, **Rui Shao\***, Gongwei Chen, Kaiwen Zhou, Weili Guan, Liqiang Nie\*. "FALCON: Resolving Visual Redundancy and Fragmentation in High-resolution Multimodal Large Language Models via Visual

Registers.” *arXiv preprint*, 2025.

- Renshan Zhang, Yibo Lyu, **Rui Shao\***, Gongwei Chen, Weili Guan, Liqiang Nie\*. “Token-level Correlation-guided Compression for Efficient Multimodal Document Understanding”. *arXiv preprint*, 2024.
- Zaijing Li, **Rui Shao\***, Gongwei Chen, Dongmei Jiang, Liqiang Nie\*. “Enhancing Emotional Generation Capability of Large Language Models via Emotional Chain-of-Thought”. *arXiv preprint*, 2024.

## ❖ Services

---

- **Conference Chairs**

ICMR 2025 Panel Co-Chairs; ACM Multimedia 2024, BMVC 2024 2025 Area Chairs

- **Program Committee Member:**

AAAI Conference on Artificial Intelligence (AAAI), 2021, 2022, 2023

Workshop on Audiovisual Deepfake Generation and Detection, ACM MM 2021

- **Invited Reviewer:**

**Journal:** IEEE Transactions on Pattern Analysis and Machine Intelligence

International Journal of Computer Vision

IEEE Transactions on Image Processing

IEEE Transactions on Neural Networks and Learning Systems

IEEE Transactions on Information Forensics and Security

IEEE Transactions on Circuits and Systems for Video Technology

Journal of Selected Topics in Signal Processing

Neural Networks

Pattern Recognition

**Conference:** CVPR, ICCV, ECCV, NeurIPS, ICML, ICLR, ACM MM, AAAI, IJCAI, ICPR, ICB

- **Student Volunteer:** IAPR/IEEE Winter School on Biometrics 2017

## ❖ Awards

---

- Huawei Talent Funding
- 2019-2020 Computer Science Department RPg Performance Award
- 2018-2019 Computer Science Department RPg Performance Award
- 2017-2018 Computer Science Department RPg Performance Award

## ❖ Competition

---

- **Ranked 1st in TRECVID, 2015 and 2016**

Zheng Wang, Yang Yang, Shuosun Guan, Chenxia Han, Jiamei Lan, **Rui Shao**, Jinqiao Wang, Chao Liang. “WHU-NERCMS at TRECVID2016: Instance Search Task”. *Participant Notebook Paper, TRECVID, 2016*.

- **Ranked 8th in AI Singapore Trusted Media Challenge, 2021.**