

YUJUN CAI

Email: vanora.caiyj@gmail.com ◇ [webpage](#)

[Google Scholar](#) ◇ [LinkedIn](#)

WORK EXPERIENCE

Professor at Southeast University	2026 - Present
Lecturer at The University of Queensland (QS #42)	2024 - 2026
Research Scientist at Meta Reality Labs (Seattle)	2022 - 2024

EDUCATION

Doctor of Philosophy in Computer Science	2017 - 2021
Nanyang Technological University	
Bachelor of Science in Information Technology	2013 - 2017
Southeast University, GPA: 4.341/4.8; Rank: 1/229	

PUBLICATIONS

- [1] Honghao Fu, Jiacheng Chen, Manxi Lin, Junjun Zheng, Xiangheng Kong, Yiwei Wang, Xin Yu, Miao Xu, Yuning Jiang, **Yujun Cai**. AdaM-Rec: Adaptive Modality Routing for Multimodal Recommendation, *NeurIPS 2026*. (CCF-A)
- [2] Haonan Ge, Yiwei Wang, Hang Wu, **Yujun Cai**. What Should a Streaming Video Model Remember?, *NeurIPS 2026*. (CCF-A)
- [3] Zehao Li, Hao Jiang, **Yujun Cai**, Jianing Chen, Baolong Bi, Shuqin Gao, Honglong Zhao, Yiwei Wang, Tianlu Mao, Zhaoqi Wang. STDR: Spatio-Temporal Decoupling for Real-Time Dynamic Scene Rendering, *ACM MM 2026*. (CCF-A)
- [4] Puwei Lian, **Yujun Cai**, Songze Li, Bingkun Bao. Noise as a Probe: Membership Inference Attacks on Diffusion Models Leveraging Initial Noise, *ACM MM 2026*. (CCF-A)
- [5] Xingjian Tao, Yiwei Wang, **Yujun Cai**, Yifan Song, Jing Tang. ViewFusion: Structured Spatial Thinking Chains for Multi-View Reasoning, *ECCV 2026*. (ICORE A*)
- [6] Lexin Wang, Shenghua Liu, Yiwei Wang, **Yujun Cai**, Yuyao Ge, Jiayu Yao, Jiafeng Guo, Xueqi Cheng. HighlightBench: Benchmarking and Diagnosing Markup-Driven Table Reasoning in Scientific Documents, *ECCV 2026*. (ICORE A*)
- [7] Yu Li, **Yujun Cai**, Chi Zhang. CRAFT-LoRA: Content-Style Personalization via Rank-Constrained Adaptation and Training-Free Fusion, *CVPR 2026*. (CCF-A)
- [8] Xiangyan Qu, Zhenlong Yuan, Jing Tang, Rui Chen, Datao Tang, Meng Yu, Lei Sun, Yancheng Bai, Xiangxiang Chu, Gaopeng Gou, Gang Xiong, **Yujun Cai**. From Scale to Speed: Adaptive Test-Time Scaling for Image Editing, *CVPR 2026*. (CCF-A)

- [9] Samyak Rawlekar, Amitabh Swain, **Yujun Cai**, Yiwei Wang, Ming-Hsuan Yang, Narendra Ahuja. Finding Distributed Object-Centric Properties in Self-Supervised Transformers, *CVPR 2026*. (CCF-A)
- [10] Bowen Sun, **Yujun Cai**, Ming-Hsuan Yang, Hang Wu, Yiwei Wang. PAS: A Training-Free Stabilizer for Temporal Encoding in Video LLMs, *CVPR 2026*. (CCF-A)
- [11] Qianxun Xu, Chenxi Song, **Yujun Cai**, Chi Zhang. SwitchCraft: Training-Free Multi-Event Video Generation with Attention Controls, *CVPR 2026*. (CCF-A)
- [12] Yiyang Yang, Wei Cheng, Sijin Chen, Honghao Fu, Xianfang Zeng, **Yujun Cai**, Gang Yu, Xingjun Ma. OmniLottie: Generating Vector Animations via Parameterized Lottie Tokens, *CVPR 2026*. (CCF-A)
- [13] Tong Guan, Sheng Pan, Johan Barthelemy, Zhao Li, **Yujun Cai**, Cesare Alippi, Ming Jin, Shirui Pan. TimeOmni-VL: Unified Models for Time Series Understanding and Generation, *ICML 2026*. (CCF-A)
- [14] Zehao Li, Hongwei Yu, Hao Jiang, Qiang Sheng, Yilong Xu, Baolong Bi, Yang Li, Zhenlong Yuan, **Yujun Cai**, Zhaoqi Wang. FactGuard: Agentic Video Misinformation Detection via Reinforcement Learning, *ICML 2026*. (CCF-A)
- [15] Puwei Lian, **Yujun Cai**, Songze Li, Bingkun Bao. Enhancing Membership Inference Attacks on Diffusion Models from a Frequency-Domain Perspective, *ICML 2026*. (CCF-A)
- [16] Yixin Song, Guangchi Liu, Xiangyu Xu, Shaofeng Li, Zhen Ling, Yiwei Wang, **Yujun Cai**. MIMO-LP: A Multi-Input Multi-Output Framework for Subgraph-based Link Prediction, *ICML 2026*. (CCF-A)
- [17] Siqian Tong, Li Xuan, Yiwei Wang, Baolong Bi, **Yujun Cai**, Shenghua Liu, Yuchen He, Hao Chengpeng. AuTAgent: A Reinforcement Learning Framework for Tool-Augmented Audio Reasoning, *ICML 2026*. (CCF-A)
- [18] Honghao Fu, Miao Xu, Yiwei Wang, Dailing Zhang, Jun Liu, **Yujun Cai**. VideoStir: Understanding Long Videos via Spatio-Temporally Structured and Intent-Aware RAG, *ACL 2026*. (CCF-A)
- [19] Boyan Han, Yiwei Wang, Yi Song, **Yujun Cai**, Chi Zhang. Dynamic Infilling Anchors for Format-Constrained Generation in Diffusion Large Language Models, *ACL 2026*. (CCF-A)
- [20] Zheng He, Yiwei Wang, Hongxing Wang, **Yujun Cai**. ReCon: Active Defense against Large Vision-Language Model Jailbreaks via Reverse Safety Concept Injection, *ACL 2026*.
- [21] Nan Sun, Jing Tang, Lei Sun, Rui Chen, Yuxing Lu, Xiangxiang Chu, Hefei Ling, **Yujun Cai**. Beyond Query Bias: Candidate-Aware Iterative Refinement for Zero-Shot Composed Image Retrieval, *ACL 2026*.
- [22] Xingjian Tao, Yiwei Wang, **Yujun Cai**, Yihong Luo, Kai Han, Jing Tang. Rectifying Coordinate Drift in MLLMs via Counterfactual Positional Guidance, *ACL 2026*.
- [23] Zhetong Zhang, Honghao Fu, Miao Xu, Yiwei Wang, **Yujun Cai**. V-DEAL: Diagnosing Video Safety De-Calibration as an Understanding–Refusal Coupling Failure, *COLM 2026*.
- [24] Jianing Chen, Zehao Li, **Yujun Cai**, Hao Jiang, Shuqin Gao, Honglong Zhao, Tianlu Mao, Yucheng Zhang. From Tokens to Nodes: Semantic-Guided Motion Control for Dynamic 3D Gaussian Splatting, *ICLR 2026*. (CCF-A)
- [25] Honghao Fu, Yuan Ouyang, Kai-Wei Chang, Yiwei Wang, Zi Huang, **Yujun Cai**. ContextNav: Towards Agentic Multimodal In-Context Learning, *ICLR 2026*. (CCF-A)

- [26] Xingjian Tao, Yiwei Wang, **Yujun Cai**, Zhicheng Yang, Jing Tang. Are LLMs Really Not Knowledgeable? Mining the Submerged Knowledge in LLMs’ Memory, *ICLR 2026*. (CCF-A)
- [27] Yike Wu, Yiwei Wang, **Yujun Cai**. ChainMPQ: Interleaved Text-Image Reasoning Chains for Mitigating Relation Hallucinations, *ICLR 2026*. (CCF-A)
- [28] Zhen Xiong, **Yujun Cai**, Zhecheng Li, Yiwei Wang. Unveiling the Potential of Diffusion Large Language Model in Controllable Generation, *ICLR 2026*. (CCF-A)
- [29] Haojin Yang, Rui Hu, Zequn Sun, Rui Zhou, **Yujun Cai**, Yiwei Wang. WavefrontDiffusion: Dynamic Decoding Schedule for Improved Reasoning, *ICLR 2026*. (CCF-A)
- [30] Zhenlong Yuan, Chengxuan Qian, Jing Tang, Rui Chen, Zijian Song, Lei Sun, Xiangxiang Chu, **Yujun Cai**, Dapeng Zhang, Shuo Li. AutoDrive-R²: Incentivizing Reasoning and Self-Reflection Capacity for VLA Model in Autonomous Driving, *ICLR 2026*. (CCF-A)
- [31] Zhenlong Yuan, Xiangyan Qu, Chengxuan Qian, Rui Chen, Jing Tang, Lei Sun, Xiangxiang Chu, Dapeng Zhang, Yiwei Wang, **Yujun Cai**, Shuo Li. Video-STAR: Reinforcing Open-Vocabulary Action Recognition with Tools, *ICLR 2026*. (CCF-A)
- [32] Haodong Xie, **Yujun Cai**, Rahul Singh Maharjan, Yiwei Wang, Federico Tavella, Angelo Cangelosi. Hierarchical, Interpretable, Label-Free Concept Bottleneck Model, *BMVC 2026*.
- [33] Litian Liang, Jingxi Xu, Xinda Qi, **Yujun Cai**, Houzhu Ding, Luqi Wang, Zhixin Sun, Jyh-Herng Chow, Ming Yang, Mark Cutkosky. Universal Manipulation Exoskeleton: Learning Compliant Whole-body Policies with Real-time Torque Feedback, *CoRL 2026*.
- [34] Baolong Bi, Yuyao Ge, Shenghua Liu, Yuchen He, Siqian Tong, Lizhe Chen, Lingrui Mei, Zehao Li, Yiwei Wang, **Yujun Cai**, Ming-Hsuan Yang, Xueqi Cheng. PromptCD: Test-Time Behavior Enhancement via Polarity-Prompt Contrastive Decoding, *TPAMI 2026*. (CCF-A) (JCR Q1)
- [35] Jianing Chen, Zehao Li, **Yujun Cai**, Hao Jiang, Chengxuan Qian, Juyuan Kang, Shuqin Gao, Honglong Zhao, Tianlu Mao, Yucheng Zhang. HAIF-GS: Hierarchical and Induced Flow-Guided Gaussian Splatting for Dynamic Scene, *NeurIPS 2025*. (CCF-A)
- [36] Honghao Fu, Junlong Ren, Qi Chai, Deheng Ye, **Yujun Cai**, Hao Wang. VistaWise: Building Cost-Effective Agent with Cross-Modal Knowledge Graph for Minecraft, *EMNLP 2025*. (ICORE A*)
- [37] Haonan Ge, Yiwei Wang, Ming-Hsuan Yang, **Yujun Cai**. MRFD: Multi-Region Fusion Decoding with Self-Consistency for Mitigating Hallucinations in LVLMS, *EMNLP 2025*. (ICORE A*)
- [38] Shuyang Hao, Yiwei Wang, Bryan Hooi, Jun Liu, Muhao Chen, Zi Huang, **Yujun Cai**. Making every step effective: Jailbreaking large vision-language models through hierarchical kv equalization, *EMNLP 2025*. (ICORE A*)
- [39] Sifan Li, **Yujun Cai**, Yiwei Wang. SemVink: Advancing VLMs’ Semantic Understanding of Optical Illusions via Visual Global Thinking, *EMNLP 2025*. (ICORE A*)
- [40] Xingjian Tao, Yiwei Wang, **Yujun Cai**, Zhicheng Yang, Jing Tang. Understanding GUI Agent Localization Biases through Logit Sharpness, *EMNLP 2025*. (ICORE A*)
- [41] Hang Wu, Hongkai Chen, **Yujun Cai**, Chang Liu, Qingwen Ye, Ming-Hsuan Yang, Yiwei Wang. DiMo-GUI: Advancing Test-time Scaling in GUI Grounding via Modality-Aware Visual Reasoning, *EMNLP 2025*. (ICORE A*)

- [42] Zhen Xiong, **Yujun Cai**, Zhecheng Li, Yiwei Wang. Mapping the Minds of LLMs: A Graph-Based Analysis of Reasoning LLM, *EMNLP 2025*. (ICORE A*)
- [43] Shuyang Hao, Bryan Hooi, Jun Liu, Kai-Wei Chang, Zi Huang, **Yujun Cai**. Exploring Visual Vulnerabilities via Multi-Loss Adversarial Search for Jailbreaking Vision-Language Models, *CVPR 2025*. (CCF-A)
- [44] Muchen Li, Sammy Christen, Chengde Wan, **Yujun Cai**, Renjie Liao, Leonid Sigal, Shugao Ma. LatentHOI: On the Generalizable Hand Object Motion Generation with Latent Hand Diffusion, *CVPR 2025*. (CCF-A)
- [45] Zhecheng Li, Yiwei Wang, Bryan Hooi, **Yujun Cai**, Nanyun Peng, Kai-Wei Chang. DRS: Deep Question Reformulation With Structured Output, *ACL 2025*. (CCF-A)
- [46] Zhecheng Li, Yiwei Wang, Bryan Hooi, **Yujun Cai**, Zhen Xiong, Nanyun Peng, Kai-wei Chang. Vulnerability of LLMs to Vertically Aligned Text Manipulations, *ACL 2025*. (CCF-A)
- [47] Zhecheng Li, Guoxian Song, **Yujun Cai**, Zhen Xiong, Junsong Yuan, Yiwei Wang. Texture or Semantics? Vision-Language Models Get Lost in Font Recognition, *COLM 2025*.
- [48] Zhaochen Wang, Bryan Hooi, Yiwei Wang, Ming-Hsuan Yang, Zi Huang, **Yujun Cai**. Text Speaks Louder than Vision: ASCII Art Reveals Textual Biases in Vision-Language Models, *COLM 2025*.
- [49] Chunxue Xu, Yiwei Wang, Bryan Hooi, **Yujun Cai**, Songze Li. How does Watermarking Affect Visual Language Models in Document Understanding?, *COLM 2025*.
- [50] Cheng Wang, Yiwei Wang, **Yujun Cai**, Bryan Hooi. Tricking Retrievers with Influential Tokens: An Efficient Black-Box Corpus Poisoning Attack, *NAACL 2025*.
- [51] Cheng Wang, Yiwei Wang, Bryan Hooi, **Yujun Cai**, Nanyun Peng, Kai-Wei Chang. Con-ReCall: Detecting Pre-training Data in LLMs via Contrastive Decoding, *COLING 2025*.
- [52] Haoxuan Qu, Zhuoling Li, Hossein Rahmani, **Yujun Cai**, Jun Liu. DisC-GS: Discontinuity-aware Gaussian Splatting, *NeurIPS 2024*. (CCF-A)
- [53] Sasha Salter, Richard Warren, Collin Schlager, Adrian Spurr, Shangchen Han, Rohin Bhasin, **Yujun Cai**, Peter Walkington, Anuoluwapo Bolarinwa, Robert Wang, others. emg2pose: A Large and Diverse Benchmark for Surface Electromyographic Hand Pose Estimation, *NeurIPS 2024*. (CCF-A)
- [54] Yihong Luo, Siya Qiu, Xingjian Tao, **Yujun Cai**, Jing Tang. Energy-Calibrated VAE with Test Time Free Lunch, *ECCV 2024*. (ICORE A*)
- [55] Haoxuan Qu, **Yujun Cai**, Jun Liu. Llms are good action recognizers, *CVPR 2024*. (CCF-A)
- [56] Li Xu, Haoxuan Qu, **Yujun Cai**, Jun Liu. 6d-diff: A keypoint diffusion framework for 6d object pose estimation, *CVPR 2024*. (CCF-A)
- [57] Zhecheng Li, Yiwei Wang, Bryan Hooi, **Yujun Cai**, Zhen Xiong, Nanyun Peng, Kai-Wei Chang. Vulnerability of LLMs to Vertically Aligned Text Manipulations, *ACL 2024*. (CCF-A)
- [58] Kenrick Kin, Chengde Wan, Ken Koh, Andrei Marin, Necati Cihan Camgoz, Yubo Zhang, **Yujun Cai**, Fedor Kovalev, Moshe Ben-Zacharia, Shannon Hoople, others. STMG: A Machine Learning Microgesture Recognition System for Supporting Thumb-Based VR/AR Input, *CHI 2024*. (CCF-A)
- [59] Yiwei Wang, **Yujun Cai**, Muhao Chen, Yuxuan Liang, Bryan Hooi. Primacy effect of chatgpt, *EMNLP 2023*. (ICORE A*)

- [60] Haoxuan Qu, **Yujun Cai**, Lin Geng Foo, Ajay Kumar, Jun Liu. A Characteristic Function-Based Method for Bottom-Up Human Pose Estimation, *CVPR 2023*. (CCF-A)
- [61] Julian Tanke, Linguang Zhang, Amy Zhao, Chengcheng Tang, **Yujun Cai**, Lezi Wang, Po-Chen Wu, Juergen Gall, Cem Keskin. Social diffusion: Long-term multiple human motion anticipation, *ICCV 2023*. (CCF-A)
- [62] Yiwei Wang, Bryan Hooi, Fei Wang, **Yujun Cai**, Yuxuan Liang, Wenxuan Zhou, Jing Tang, Manjuan Duan, Muhao Chen. How Fragile is Relation Extraction under Entity Replacements?, *CONLL 2023*.
- [63] Haoxuan Qu, Li Xu, **Yujun Cai**, Lin Geng Foo, Jun Liu. Heatmap distribution matching for human pose estimation, *NeurIPS 2022*. (CCF-A)
- [64] Mingfei Chen, Jianfeng Zhang, Xiangyu Xu, Lijuan Liu, **Yujun Cai**, Jiashi Feng, Shuicheng Yan. Geometry-guided progressive nerf for generalizable and efficient neural human rendering, *ECCV 2022*. (ICORE A*)
- [65] Yiwei Wang, Muhao Chen, Wenxuan Zhou, **Yujun Cai**, Yuxuan Liang, Bryan Hooi. Graphcache: Message passing as caching for sentence-level relation extraction, *NAACL 2022*.
- [66] Yiwei Wang, Muhao Chen, Wenxuan Zhou, **Yujun Cai**, Yuxuan Liang, Dayiheng Liu, Baosong Yang, Juncheng Liu, Bryan Hooi. Should we rely on entity mentions for relation extraction? debiasing relation extraction with counterfactual analysis, *NAACL 2022*.
- [67] Yiwei Wang, **Yujun Cai**, Yuxuan Liang, Henghui Ding, Changhu Wang, Bryan Hooi. Time-Aware Neighbor Sampling on Temporal Graphs, *2022 International Joint Conference on Neural Networks (IJCNN) 2022*.
- [68] Ziyin Xin, **Yujun Cai**, Louis T Dang, Hannah MS Burke, Jerico Revote, Natalie Charitakis, Denis Bienroth, Hieu T Nim, Yuan-Fang Li, Mirana Ramialison. MonaGO: a novel gene ontology enrichment analysis visualisation system, *BMC Bioinformatics 2022*. (JCR Q1)
- [69] Shangchen Han, Po-chen Wu, Yubo Zhang, Beibei Liu, Linguang Zhang, Zheng Wang, Weiguang Si, Peizhao Zhang, **Yujun Cai**, Tomas Hodan, others. UmeTrack: Unified multi-view end-to-end hand tracking for VR, *SIGGRAPH Asia 2022*. (ICORE A*)
- [70] Chi Zhang, **Yujun Cai**, Guosheng Lin, Chunhua Shen. Deepemd: Differentiable earth mover’s distance for few-shot learning, *TPAMI 2022*. (CCF-A) (JCR Q1)
- [71] Yiwei Wang, **Yujun Cai**, Yuxuan Liang, Henghui Ding, Changhu Wang, Siddharth Bhatia, Bryan Hooi. Adaptive data augmentation on temporal graphs, *NeurIPS 2021*. (CCF-A)
- [72] Jianfeng Zhang, **Yujun Cai**, Shuicheng Yan, Jiashi Feng, others. Direct multi-view multi-person 3d pose estimation, *NeurIPS 2021*. (CCF-A)
- [73] **Yujun Cai**, Yiwei Wang, Yiheng Zhu, Tat-Jen Cham, Jianfei Cai, Junsong Yuan, Jun Liu, Chuanxia Zheng, Sijie Yan, Henghui Ding, others. A unified 3d human motion synthesis model via conditional variational auto-encoder, *ICCV 2021*. (CCF-A)
- [74] Yiwei Wang, Wei Wang, Yuxuan Liang, **Yujun Cai**, Bryan Hooi. Progressive supervision for node classification, *ECML 2021*.
- [75] Yiwei Wang, **Yujun Cai**, Yuxuan Liang, Henghui Ding, Changhu Wang, Bryan Hooi. Time-aware neighbor sampling for temporal graph networks, *The 2022 International Joint Conference on Neural Networks (IJCNN) 2021*.

- [76] Yiwei Wang, Wei Wang, Yuxuan Liang, **Yujun Cai**, Bryan Hooi. Curgraph: Curriculum learning for graph classification, *WWW 2021*. (CCF-A)
- [77] Yiwei Wang, Wei Wang, Yuxuan Liang, **Yujun Cai**, Bryan Hooi. Mixup for node and graph classification, *WWW 2021*. (CCF-A)
- [78] **Yujun Cai**, Lin Huang, Yiwei Wang, Tat-Jen Cham, Jianfei Cai, Junsong Yuan, Jun Liu, Xu Yang, Yiheng Zhu, Xiaohui Shen, others. Learning progressive joint propagation for human motion prediction, *ECCV 2020*. (ICORE A*)
- [79] Chi Zhang, **Yujun Cai**, Guosheng Lin, Chunhua Shen. Deepemd: Few-shot image classification with differentiable earth mover’s distance and structured classifiers, *CVPR 2020*. (CCF-A)
- [80] Yiwei Wang, Wei Wang, Yujun Ca, Bryan Hooi, Beng Chin Ooi. Detecting implementation bugs in graph convolutional network based node classifiers, *ISSRE 2020*.
- [81] Yiwei Wang, Wei Wang, Yuxuan Liang, **Yujun Cai**, Juncheng Liu, Bryan Hooi. Nodeaug: Semi-supervised node classification with data augmentation, *KDD 2020*. (CCF-A)
- [82] **Yujun Cai**, Liuhao Ge, Jianfei Cai, Nadia Magnenat Thalmann, Junsong Yuan. 3D hand pose estimation using synthetic data and weakly labeled RGB images, *TPAMI 2020*. (CCF-A) (JCR Q1)
- [83] **Yujun Cai**, Liuhao Ge, Jun Liu, Jianfei Cai, Tat-Jen Cham, Junsong Yuan, Nadia Magnenat Thalmann. Exploiting spatial-temporal relationships for 3d pose estimation via graph convolutional networks, *ICCV 2019*. (CCF-A)
- [84] **Yujun Cai**, Liuhao Ge, Jianfei Cai, Junsong Yuan. Weakly-supervised 3d hand pose estimation from monocular rgb images, *ECCV 2018*. (ICORE A*)
- [85] Liuhao Ge, **Yujun Cai**, Junwu Weng, Junsong Yuan. Hand pointnet: 3d hand pose estimation using point sets, *CVPR 2018*. (CCF-A)

ASSOCIATE EDITOR

Neurocomputing

Pattern Recognition (PR)

AREA CHAIR AND SENIOR PROGRAM COMMITTEE

International Conference on Learning Representations (ICLR)	2025
IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)	2026
Conference on Neural Information Processing Systems (NeurIPS)	2025
Conference on Empirical Methods in Natural Language Processing (EMNLP)	
ACM Multimedia	2025
The AAAI Conference on Artificial Intelligence (AAAI)	2025
International Conference on Computer Vision (ICCV)	2025
International Conference on Machine Learning (ICML)	2025

International Joint Conferences on Artificial Intelligence(IJCAI) (SPC)	
The Web Conference (WWW) (SPC)	2025
CIKM (SPC)	2025
Conference on Language Modeling (COLM)	2025
IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)	2024, 2025
International Joint Conference on Neural Network (IJCNN)	2025

REVIEWING

ACM SIGGRAPH

IEEE International Conference on Multimedia and Expo (ICME)	2022-2024
IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)	
IEEE Transactions on Image Processing (TIP)	
IEEE Transactions on Image Processing(TIP)	
IEEE Transactions on Multimedia (TMM)	
IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)	
ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)	
IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)	2018-2025
International Conference on Computer Vision (ICCV)	2019-2024
European Conference on Computer Vision (ECCV)	2018-2024
Conference on Empirical Methods in Natural Language Processing (EMNLP)	2024, 2025
Annual Meeting of the Association for Computational Linguistics (ACL)	2024, 2025
North American Chapter of the Association for Computational Linguistics (NAACL)	2024, 2025
International Conference on Learning Representations (ICLR)	2023, 2024
Conference on Neural Information Processing Systems (NeurIPS)	2022-2024
The AAAI Conference on Artificial Intelligence (AAAI)	2021-2024
International Joint Conference on Artificial Intelligence (IJCAI)	2019-2024

CONFERENCE TUTORIALS

From Perception to Simulation: The Emergence of World Models in Multi-modal Reasoning ECCV 2026	
From Perception to Simulation: The Emergence of World Models in Multi-modal Reasoning CVPR 2026	

Towards Comprehensive Reasoning in Vision-Language Models
ICCV 2025

TEACHING

Instructor, ELEC4630 Deep Learning and Computer Vision

UQ, 2025 Semester 1

Instructor, COMP1100/7110 Introduction to Software Innovation

UQ, 2025 Semester 2

Teaching Assistant, DM6190: Recent Advances in Image Segmentation and Its Applications

NTU, 2018-2019

Teaching Assistant, CE7491: Digital Image Processing

NTU, 2019-2020

Teaching Assistant, Engineering Mathematics II

NTU, 2018

HONORS AND AWARDS

Amazon Research Award	2026
National Scholarship, Ministry of Education of the People's Republic of China	2014
National Scholarship, Ministry of Education of the People's Republic of China	2016