

Yuanhao Zhai

E-mail: yuanhaozhai@gmail.com ◇ Phone: +1 (716)-580-0896 ◇ Homepage: yhzhai.com

RESEARCH INTEREST

Video understanding/generation, human action analysis/synthesis, multi-modal learning, weakly-supervised learning, image manipulation detection, and Deepfakes video detection.

EDUCATION

State University of New York at Buffalo, Buffalo, NY, USA Feb 2021 - Jun 2025 (Expected)

- PhD student in Computer Science & Engineering.
- Advisors: Prof. David Doermann and Prof. Junsong Yuan.

Xi'an Jiaotong University, Xi'an, China Sep 2016 - July 2020

- Bachelor of Engineering in Computer Science.
- Advisors: Ziyi Liu and Prof. Le Wang.
- Outstanding Graduate Award

University of California, Berkeley, Berkeley, CA, USA Jan 2019 - May 2019

- Exchange student of the Berkeley International Study Program.

INTERNSHIP RESEARCH EXPERIENCE

Amazon, New York, NY, USA May 2024 - Nov 2024

Applied scientist intern

Topic: Text-to-outfit generation using large language models

Microsoft, Remote Feb 2024 - May 2024

Research intern

Topic: Video diffusion distillation

Microsoft, Remote May 2023 - Aug 2023

Research intern

Topic: Human-centric joint video-depth generation with diffusion models

Wormpex AI Research, Bellevue, WA, USA May 2022 - Aug 2022

Research intern

Topic: Human action analysis and open-set recognition

Institute for Artificial Intelligence and Data Science, University at Buffalo, Buffalo, NY, USA

Research assistant

Aug 2021 - May 2024

Topic: Image manipulation detection and Deepfakes detection

Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, Xi'an, China

Research intern

Feb 2018 - July 2021

Project: Weakly-supervised temporal action localization

PUBLICATIONS

<https://scholar.google.com/citations?user=AeU2HdYAAAAJ>

First-author Publications

- NeurIPS 2024 [Yuanhao Zhai](#), Kevin Lin, Zhengyuan Yang, Linjie Li, Jianfeng Wang, Chung-Ching Lin, David Doermann, Junsong Yuan, and Lijuan Wang, “Motion Consistency Model: Accelerating Video Diffusion with Disentangled Motion-Appearance Distillation”, in Conference on Neural Information Processing Systems, 2024. [[Project](#)] [[Paper](#)] [[Code](#)]
- ECCV 2024 [Yuanhao Zhai](#), Kevin Lin, Linjie Li, Chung-Ching Lin, Jianfeng Wang, Zhengyuan Yang, David Doermann, Junsong Yuan, Zicheng Liu, and Lijuan Wang, “IDOL: Unified Dual-Modal Latent Diffusion for Human-Centric Joint Video-Depth Generation”, in European Conference on Computer Vision, 2024. [[Project](#)] [[Paper](#)] [[Code](#)]
- ACM MM 2023 [Yuanhao Zhai](#), Mingzhen Huang, Tianyu Luan, Lu Dong, Ifeoma Nwogu, Siwei Lyu, David Doermann, and Junsong Yuan, “Language-guided Human Motion Synthesis with Atomic Actions”, in ACM Multimedia, 2023. [[Paper](#)] [[Code](#)]
- ICCV 2023 [Yuanhao Zhai](#), Ziyi Liu, Zhenyu Wu, Yi Wu, Chunluan Zhou, David Doermann, Junsong Yuan, and Gang Hua, “SOAR: Scene-debiasing Open-set Action Recognition”, in IEEE International Conference on Computer Vision, 2023. [[Paper](#)] [[Code](#)]
- ICCV 2023 [Yuanhao Zhai](#), Tianyu Luan, David Doermann, and Junsong Yuan, “Towards Generic Image Manipulation Detection with Weakly-Supervised Self-Consistency Learning”, in IEEE International Conference on Computer Vision, 2023. [[Paper](#)] [[Code](#)]
- TPAMI 2022 [Yuanhao Zhai](#), Le Wang, Wei Tang, Qilin Zhang, Nanning Zheng, David Doermann, Junsong Yuan, and Gang Hua, “Adaptive Two-Stream Consensus Network for Weakly-Supervised Temporal Action Localization”, in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022. [[Paper](#)]
- TMM 2022 [Yuanhao Zhai](#), Le Wang, Wei Tang, Qilin Zhang, Nanning Zheng, and Gang Hua, “Action Coherence Network for Weakly-Supervised Temporal Action Localization”, in IEEE Transactions on Multimedia, vol. 24, pp. 1857-1870, 2022. [[Paper](#)]
- ECCV 2020 Spotlight [Yuanhao Zhai](#), Le Wang, Wei Tang, Qilin Zhang, Junsong Yuan, and Gang Hua, “Two-Stream Consensus Network for Weakly-Supervised Temporal Action Localization”, in European Conference on Computer Vision, pp. 37-54, 2020. [[Paper](#)]
- ICIP 2019 [Yuanhao Zhai](#), Le Wang, Ziyi Liu, Qilin Zhang, Gang Hua, and Nanning Zheng, “Action Coherence Network for Weakly Supervised Temporal Action Localization”, in IEEE International Conference on Image Processing, pp. 918-922, 2019. [[Paper](#)]

Collaboration Publications

- arXiv 2024 Yangcen Liu, Ziyi Liu, [Yuanhao Zhai](#), Wen Li, David Doerman, and Junsong Yuan, “STAT: Towards Generalizable Temporal Action Localization”, arXiv preprint arXiv:2404.13311, 2024. [[Paper](#)]
- ICML 2024 Fabing Li, [Yuanhao Zhai](#), Shuangyu Cai, and Mingyu Gao, “Seesaw: Compensating for Nonlinear Reduction with Linear Computations for Private Inference”, in International Conference on Machine Learning, 2024. [[Paper](#)]
- CVPR 2024 Tan Wang, Linjie Li, Kevin Lin, [Yuanhao Zhai](#), Chung-Ching Lin, Zhengyuan Yang, Hanwang Zhang, Zicheng Liu, and Lijuan Wang, “DisCo: Disentangled Control for Referring Human Dance Generation in Real World”, in IEEE Computer Vision and Pattern Recognition, 2024. [[Paper](#)]
- CVPR 2023 Tianyu Luan, [Yuanhao Zhai](#), Jingjing Meng, Zhong Li, Zhang Chen, Yi Xu, and Junsong Yuan, “High Fidelity 3D Hand Shape Reconstruction via Scalable Graph Frequency Decomposition”, in IEEE Computer Vision and Pattern Recognition, 2023. [[Paper](#)]

TIP 2021 Le Wang, Rizhi Ding, Yuanhao Zhai, Qilin Zhang, Wei Tang, Nanning Zheng, and Gang Hua, “Giant Panda Identification”, IEEE Transactions on Image Processing (TIP), vol. 30, pp. 2837-2849, 2021. [\[Paper\]](#)

HONORS & AWARDS

Runner Up , CVPR’2021 HACS Weakly-Supervised Temporal Action Localization Challenge	2021
Ji Ke Prize (8-th/763 teams), Alibaba Tianchi Temporal Action Localization Competition	2020
Outstanding Graduate (Top 15%), Xi’an Jiaotong University	2020
Excellent Student (Top 15%), Xi’an Jiaotong University	2017, 2018, 2019
Peng Kang Scholarship (Top 3%), Xi’an Jiaotong University	2017

PROFESSIONAL SERVICES

Conference Reviewer

- IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) 2022, 2024, 2025
- IEEE International Conference on Computer Vision (ICCV) 2023, 2025
- European Conf. on Computer Vision (ECCV) 2022, 2024
- Neural Information Processing Systems (NeurIPS) 2024, 2025
- International Conference on Learning Representations (ICLR) 2025
- Annual AAAI Conference on Artificial Intelligence (AAAI) 2025
- ACM Multimedia (MM) 2024, 2025
- Asian Conference on Computer Vision (ACCV) 2024
- IEEE International Conf. on Automatic Face and Gesture Recognition (FG) 2021, 2023, 2025

Journal Reviewer

- IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)
- International Journal of Computer Vision (IJCV)
- IEEE Trans. on Image Processing (TIP)
- IEEE Trans. on Circuits and Systems for Video Technology (TCSVT)
- IEEE Trans. on Multimedia (TMM)
- IEEE Transactions on Neural Networks and Learning Systems (TNNLS)
- IEEE Intelligent Systems
- International Journal of Machine Learning and Cybernetics
- Journal of Visual Communication and Image Representation
- Machine Vision and Applications
- The Journal of Supercomputing