

Weiming Ren

Email	w2ren@uwaterloo.ca	Homepage	cs.uwaterloo.ca/~w2ren/
LinkedIn	linkedin.com/in/wmren/	GitHub	github.com/wren93

Research Interests

Generative Modelling: developing novel algorithms based on diffusion models for controllable video generation, image and video editing, and image restoration.

Multimodal Understanding: designing novel model architectures, data curation pipelines and evaluation benchmarks to enhance large multimodal models (LMMs) for image and video understanding.

Education

Sep 2023 - Present **Ph.D. in Computer Science** - University of Waterloo
GPA: 4.0/4.0, Advisor: Prof. Wenhui Chen
Related Courses: Neural Networks, Software Engineering for Big Data and AI

Sep 2021 - Dec 2022 **MSc. in Applied Computing** - University of Toronto
GPA: 4.0/4.0
Related Courses: Machine Learning for Healthcare, Neural Networks and Deep Learning, Machine Learning for Machine Vision as Inverse Graphics

Jul 2019 - Jun 2021 **Bachelor of Advanced Computing (Honours)** - The Australian National University
GPA: 6.938/7.0, Graduated with First-class Honours

Aug 2017 - Jun 2021 **B.S. in Computer Science and Technology** - Beijing Institute of Technology
GPA: 90.28/100

Publications

TUNA: Taming Unified Visual Representations for Native Unified Multimodal Models

Zhiheng Liu*, Weiming Ren*, Haozhe Liu, Zijian Zhou, Shoufa Chen, Haonan Qiu, Xiaoke Huang, Zhaochong An, Fanny Yang, Aditya Patel, Viktor Atliha, Tony Ng, Xiao Han, Chuyan Zhu, Chenyang Zhang, Ding Liu, Juan-Manuel Perez-Rua, Sen He, Jürgen Schmidhuber, Wenhui Chen, Ping Luo, Wei Liu, Tao Xiang, Jonas Schult, Yuren Cong

**Equal First-Authors. Arxiv Preprint.*

Hallucination Score: Towards Mitigating Hallucinations in Generative Image Super-Resolution

Weiming Ren*, Raghav Goyal*, Zhiming Hu*, Tristan Ty Aumentado-Armstrong*, Iqbal Mohamed, Alex Levinshtein

**Equal First-Authors. Arxiv Preprint.*

VideoEval-Pro: Robust and Realistic Long Video Understanding Evaluation

Wentao Ma*, Weiming Ren*, Yiming Jia, Zhuofeng Li, Ping Nie, Ge Zhang, Wenhui Chen

**Equal First-Authors. Arxiv Preprint.*

Pixel Reasoner: Incentivizing Pixel-Space Reasoning with Curiosity-Driven Reinforcement Learning

Alex Su*, Haozhe Wang*, Weiming Ren, Fangzhen Lin, Wenhui Chen

Conference on Neural Information Processing Systems (NeurIPS), 2025.

SAMPL: Self-Attention Modelled Patch Learning for Efficient Visual Understanding

Zhiming Hu, Salar Hosseini Khorasgani, Weiming Ren, Iqbal Mohamed

ACM Multimedia Systems Conference (MMSys), 2025.

Vamba: Understanding Hour-Long Videos with Hybrid Mamba-Transformers

Weiming Ren, Wentao Ma, Huan Yang, Cong Wei, Ge Zhang, Wenhu Chen
International Conference on Computer Vision (ICCV), 2025.

VISTA: Enhancing Long-Duration and High-Resolution Video Understanding by Video Spatiotemporal Augmentation

Weiming Ren, Huan Yang, Jie Min, Cong Wei, Wenhu Chen
IEEE / CVF Computer Vision and Pattern Recognition Conference (CVPR), 2025.

OmniEdit: Building Image Editing Generalist Models Through Specialist Supervision

Cong Wei*, Zheyang Xiong*, Weiming Ren, Xinrun Du, Ge Zhang, Wenhu Chen
International Conference on Learning Representations (ICLR), 2025.

MMLU-Pro: A More Robust and Challenging Multi-Task Language Understanding Benchmark

Yubo Wang, Xueguang Ma, Ge Zhang, Yuansheng Ni, Abhranil Chandra, Shiguang Guo, Weiming Ren, Aaran Arulraj, Xuan He, Ziyang Jiang, Tianle Li, Max Ku, Kai Wang, Alex Zhuang, Rongqi Fan, Xiang Yue, Wenhu Chen
*NeurIPS Dataset and Benchmark Track **Spotlight***, 2024.

Video Diffusion Models: A Survey

Andrew Melnik, Michal Ljubicjanac, Cong Lu, Qi Yan, Weiming Ren, Helge Ritter
Transactions on Machine Learning Research (TMLR), 2024.

AnyV2V: A Plug-and-Play Framework for Any Video-to-Video Editing Tasks

Max Ku*, Cong Wei*, Weiming Ren*, Huan Yang, Wenhu Chen
Equal First-Authors. Transactions on Machine Learning Research (TMLR) **Reproducibility Certification*, 2024.

StructLM: Towards Building Generalist Models for Structured Knowledge Grounding

Alex Zhuang, Ge Zhang, Tianyu Zheng, Xinrun Du, Junjie Wang, Weiming Ren, Stephen W Huang, Jie Fu, Xiang Yue, Wenhu Chen
Conference on Language Modeling (COLM), 2024.

ConsistI2V: Enhancing Visual Consistency for Image-to-Video Generation

Weiming Ren, Huan Yang, Ge Zhang, Cong Wei, Xinrun Du, Wenhao Huang, Wenhu Chen
Transactions on Machine Learning Research (TMLR), 2024.

MMM: A Massive Multi-discipline Multimodal Understanding and Reasoning Benchmark for Expert AGI

Xiang Yue, Yuansheng Ni, Kai Zhang, Tianyu Zheng, Ruoqi Liu, Ge Zhang, Samuel Stevens, Dongfu Jiang, Weiming Ren, Yuxuan Sun, Cong Wei, Botao Yu, Ruibin Yuan, Renliang Sun, Ming Yin, Boyuan Zheng, Zhenzhu Yang, Yibo Liu, Wenhao Huang, Huan Sun, Yu Su, Wenhu Chen
*IEEE / CVF Computer Vision and Pattern Recognition Conference (CVPR) **Oral, Best Paper Finalist***, 2024.

HiCu: Leveraging Hierarchy for Curriculum Learning in Automated ICD Coding

Weiming Ren, Ruijing Zeng, Tongzi Wu, Tianshu Zhu, Rahul G. Krishnan
Machine Learning for Healthcare Conference (MLHC), 2022.

Experiences

Jun 2025 - **Meta MGenAI**, London, United Kingdom

Oct 2025 *Research Scientist Intern*
Mentor: Dr. Jonas Schult

- Worked on pretraining unified multimodal models for unified image/video understanding, image/video generation, and image editing. Published one preprint (TUNA).

Sep 2024 - **Samsung AI Centre Toronto**, Toronto, Canada

Mar 2025 *Research Intern*

Mentors: Dr. Iqbal Mohamed, Dr. Alex Levinshtein

- Worked on identifying and mitigating hallucinations for generative image super-resolution. Published one preprint (Hallucination Score).

Aug 2023 - **01.AI**, Beijing, China

Feb 2025 *Research Intern*

Mentor: Dr. Huan Yang

- Worked on post-training foundational models for video generation and understanding. Published four papers (ConsistI2V, AnyV2V, VISTA, Vamba) that have been accepted to TMLR 2024, CVPR 2025 and ICCV 2025.

May 2022 - **Samsung AI Centre Toronto**, Toronto, Canada

Apr 2023 *Research Intern*

Mentor: Dr. Iqbal Mohamed

- Worked on self-supervised learning methods for data-efficient video action recognition. Published one paper (SAMPL) that has been accepted to MMSys 2025.

Awards

Sep 2024 David R. Cheriton Graduate Scholarship, University of Waterloo

Sep 2023 David R. Cheriton Graduate Scholarship, University of Waterloo

May 2021 Chancellor's Letter of Commendation, Australian National University

Apr 2020 Outstanding Winner, COMAP's Mathematical Contest in Modeling (MCM)

Mar/Sep 2018, Mar/Sep 2019 The Second-class Scholarship (4 times), Beijing Institute of Technology

Skills

Programming: Python (fluent), C/C++ (proficient), Arduino, Java, HTML/CSS/Javascript (prior experience)

Libraries & Tools: PyTorch, NumPy, Scikit-learn, Huggingface Transformers/Diffusers, Pandas, Matplotlib, Git