Ye Liu

Ph.D. @ The Hong Kong Polytechnic University

CONTACT INFORMATION

Department of Computing
The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong

Email: coco.ye.liu@connect.polyu.hk
Google Scholar: `bit.ly/2eiaBcsAAAAJ`
Homepage: `yeliu.dev`

RESEARCH INTERESTS

Vision-Language Models: E.T. Bench, VisionMath, UniPixel, VideoMind, Seed 2.0/2.1 Series
Computer Vision: ConsNet, CATNet, UMT, R^2-Tuning, VTG LLM Survey
Generative Models: MADD, ReTake
Urban Computing: TPG, POIFormer

EDUCATION

The Hong Kong Polytechnic University — *Hong Kong SAR*
Ph.D. in Computer Science — *Sep. 2021 – Jan. 2026*
Advisor: Prof. Chang Wen Chen

Wuhan University — *Wuhan, China*
B.S. in Geographical Information Science & B.E. in Computer Science — *Sep. 2016 – Jun. 2020*
Advisor: Prof. Huifang Li & Prof. Gui-Song Xia

Imperial College London — *London, UK*
Summer School in Machine Learning, Robotics and Sensor Networks — *Jul. 2018 – Aug. 2018*
Advisor: Prof. Guang-Zhong Yang & Dr. Benny Lo

The University of Hong Kong — *Hong Kong SAR*
Summer School in Mathematical Laboratory and Modeling — *Jul. 2017 – Aug. 2017*
Advisor: Dr. Chi-Wing Wong

RESEARCH EXPERIENCE

ByteDance Seed — *Shenzhen, China*
Research Intern *(Multimodal Interaction & World Model Team)* — *Dec. 2025 – Present*
Contributor of Seed 2.0/2.1 series:
- Large-scale long video pre-training: scaling from 128K to 1M context length.
- Post-training optimization: video temporal grounding, tracking, counting, captioning, etc.
- Video agent: Seed Video Editor (presented as Seed2.1 Pro official demo) Click to Preview

ARC Lab, Tencent PCG — *Shenzhen, China*
Research Intern *(Host: Dr. Yang Wu & Dr. Zhongang Qi)* — *Jul. 2021 – Apr. 2025*
- Audio-visual moment retrieval and highlight detection (CVPR 2022)
- Event-level and time-sensitive video understanding benchmark (NeurIPS 2024)
- Unified pixel-level video reasoning (NeurIPS 2025)
- These projects saved production costs by ∼¥600,000,000 in Tencent's video platform.

Show Lab, National University of Singapore — *Singapore*
Visiting Scholar *(Host: Prof. Mike Z. Shou)* — *Mar. 2024 – Jul. 2024*
Agentic video-language understanding (ICLR 2026)

Visual Computing Group, Harvard University — *Boston, USA*
Associate *(Host: Prof. Hanspeter Pfister)* — *Jun. 2023 – Mar. 2024*
Parameter- and memory-efficient transfer learning (ECCV 2024)

LIESMARS, Wuhan University *Wuhan, China*
Research Assistant *(Host: Prof. Huifang Li)* *Mar. 2020 – Jun. 2021*
Dense prediction in remote sensing images (TNNLS 2025)

SUNY Buffalo & The Chinese University of Hong Kong, Shenzhen *Buffalo, USA*
Research Intern *(Host: Prof. Junsong Yuan & Prof. Chang Wen Chen)* *Jul. 2019 – Dec. 2019*
Zero-shot human-object interaction detection (ACM MM 2020)

GNSS Research Center, Wuhan University *Wuhan, China*
Research Assistant *(Host: Prof. Chi Guo & Prof. Weiping Jiang)* *Sep. 2018 – Jun. 2019*
Hardware-efficient video object tracking.

The Hamlyn Centre, Imperial College London *London, UK*
Summer Research Intern *(Host: Prof. Guang-Zhong Yang)* *Jul. 2018 – Aug. 2018*
Voice-controlled visual servoing systems.

PUBLICATIONS

Highlighted Publications

ByteDance Seed. "Seed2.1 Model Card: Agentic Intelligence for Productivity". *Tech Report*, 2026.

ByteDance Seed. "Seed2.0 Model Card: Towards Intelligence Frontier for Real-World Complexity". *Tech Report*, 2026.

Ye Liu*, Kevin Qinghong Lin*, Chang Wen Chen†, and Mike Zheng Shou†. "VideoMind: A Chain-of-LoRA Agent for Temporal-Grounded Video Reasoning". In *International Conference on Learning Representations* (**ICLR**), 2026.

Ye Liu, Zongyang Ma, Junfu Pu, Zhongang Qi, Yang Wu, Ying Shan, and Chang Wen Chen†. "UniPixel: Unified Object Referring and Segmentation for Pixel-Level Visual Reasoning". In *Advances in Neural Information Processing Systems* (**NeurIPS**), 2025.

Ye Liu, Zongyang Ma, Zhongang Qi†, Yang Wu, Ying Shan, and Chang Wen Chen†. "E.T. Bench: Towards Open-Ended Event-Level Video-Language Understanding". In *Advances in Neural Information Processing Systems* (**NeurIPS**), 2024.

Ye Liu, Jixuan He, Wanhua Li†, Junsik Kim, Donglai Wei, Hanspeter Pfister, and Chang Wen Chen†. "R^2-Tuning: Efficient Image-to-Video Transfer Learning for Video Temporal Grounding". In *Proceedings of the European Conference on Computer Vision* (**ECCV**), 2024.

Ye Liu, Huifang Li†, Chao Hu, Shuang Luo, Yan Luo, and Chang Wen Chen. "Learning to Aggregate Multi-Scale Context for Instance Segmentation in Remote Sensing Images". *IEEE Transactions on Neural Networks and Learning Systems* (**TNNLS**), 2024.

Ye Liu, Siyuan Li, Yang Wu†, Chang Wen Chen, Ying Shan, and Xiaohu Qie. "UMT: Unified Multi-modal Transformers for Joint Video Moment Retrieval and Highlight Detection". In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (**CVPR**), 2022.

Ye Liu, Junsong Yuan, and Chang Wen Chen. "ConsNet: Learning Consistency Graph for Zero-Shot Human-Object Interaction Detection". In *Proceedings of the ACM International Conference on Multimedia* (**ACM MM**), 2020.

Other Publications

Amol Harsh, Zongyan Han†, Jean Lahoud, **Ye Liu**, Rao Muhammad Anwer, Hisham Cholakkal, Salman Khan, and Fahad Shahbaz Khan. "Ground3D-LMM: Fine-Grained 3D Point Grounding and Spatial Reasoning with LMM". In *Proceedings of the European Conference on Computer Vision* (**ECCV**), 2026.

Junchen Zhou*, Zhixian He*, Yaosi Hu, **Ye Liu**, and Chang Wen Chen. "Retake: A Training-free Framework for Physically Plausible Video Generation". In *Physics-Aware Video Generation and Restoration Workshop* @ **ICPR**, 2026.

Jianlong Wu†, Wei Liu, **Ye Liu**, Meng Liu, Liqiang Nie, Zhouchen Lin, and Chang Wen Chen. "A Survey on Video Temporal Grounding with Multimodal Large Language Model". *IEEE Transactions on Pattern Analysis and Machine Intelligence* (**TPAMI**), 2025.

Jixuan He*, Wanhua Li*, **Ye Liu**, Junsik Kim, Donglai Wei, and Hanspeter Pfister. "Affordance-Aware Object Insertion via Mask-Aware Dual Diffusion". In *P13N: Personalization in Generative AI Workshop* @ **ICCV**, 2025.

Zongyang Ma, Yuxin Chen, Ziqi Zhang, Zhongang Qi, Chunfeng Yuan, Shaojie Zhu, Chengxiang Zhuo, Bing Li, **Ye Liu**, Zang Li, Ying Shan, Weiming Hu. "VisionMath: Vision-Form Mathematical Problem-Solving". In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (**ICCV**), 2025.

Yan Luo, **Ye Liu**†, Fu-lai Chung, Yu Liu, and Takahiro Yabe. "End-to-End Personalized Next Location Recommendation via Contrastive User Preference Modeling". *IEEE Transactions on Computational Social Systems* (**TCSS**), 2025.

Jingyun Xu*, Yang Liu*, **Ye Liu***, Junnan Yu, Yukang Wang, Bobo Li, and Yi Cai†. "Dynamically Incorporating Syntactic and Semantic Knowledge for Sentiment Quadruple Extraction in Dialogues". *Under Review*, 2024.

Yan Luo, Haoyi Duan, **Ye Liu**†, and Fu-lai Chung. "Timestamps as Prompts for Geography-Aware Location Recommendation". In *Proceedings of the ACM International Conference on Information and Knowledge Management* (**CIKM**), 2023.

* *Equal contribution* † *Corresponding authors*

PATENTS

Chang Wen Chen, **Ye Liu**, Zongyang Ma, Junfu Pu, Zhongang Qi, Yang Wu, Ying Shan. "Multimedia Processing Method, Apparatus, Device, Storage Medium, and Program Product". *CN Patent (202510942903.X)*, 2025.

Chang Wen Chen, **Ye Liu**, Zongyang Ma, Zhongang Qi, Yang Wu, Ying Shan. "Video-Language Processing Method, Model Training Method, and Device". *CN Patent (202410866001.8)*, 2024.

Chang Wen Chen, **Ye Liu**, Ying Shan, Siyuan Li, Yang Wu, Xiaohu Qie. "A Video Recognition Method, System, Storage Medium, and Server". *CN Patent (CN114581821A)*, 2022.

TEACHING EXPERIENCE

Teaching Assistant

COMP1004: Introduction to AI and Data Analytics	*Fall 2022, Fall 2024, Spring 2025*
COMP3122: Information Systems Development	*Spring 2023*
COMP3334: Computer Systems Security	*Spring 2022*
COMP1001: Problem Solving Methodology in Information Technology	*Fall 2021*

Final Year Project Mentoring

Multi-modal Tokenization for Video Large Language Models	*2024*
Audio-Visual Video Temporal Understanding	*2024*
Video Moment Retrieval using Deep Learning Models (**IEEE CIS Best Project Award**)	*2023*
Rotated Object Detection in Remote Sensing Images	*2023*

HONORS & AWARDS

Hong Kong RGC Junior Research Fellow Scheme (Top-30 in HK, HK$823,200)	*2026*
PolyU Top 10 Research & Innovation Stories of the Year	*2026*
PolyU Distinguished Postdoctoral Fellowship (Top-3 at PolyU, HK$785,000)	*2025*
NeurIPS Scholar Award	*2025*
PolyU International Collaborative Research Fellowship (HK$55,000)	*2024*
HKSTP Ideation Grant (HK$100,000)	*2023, 2024*
PolyU Student Innovation and Entrepreneurship Scholarship (HK$80,000)	*2022, 2023, 2024*
PolyU Research Student Attachment Programme (HK$55,000)	*2023*
The Most Appreciated Teaching Assistant Award (Top-1 at Department, HK$50,000)	*2022*
WHU Outstanding Scholarship for Outgoing Exchange Students (¥30,000)	*2017, 2018, 2019*
Best Overall Project Award, Imperial College London	*2018*

SERVICES

Journal Reviewer

IEEE TPAMI, IEEE TIP, IEEE TMM, IEEE TCSVT

Conference Reviewer

NeurIPS *(2023 – 2025)*, ICLR *(2024 – 2026)*, ICML *(2024 – 2026)*, CVPR *(2023 – 2026)*, ICCV *(2023 – 2025)*, ECCV *(2024 – 2026)*, ACM MM *(2025 – 2026)*, AAAI *(2025 – 2027)*, BMVC *(2026)*, ACCV *(2024)*, ICIP *(2023)*

SKILLS

Dev Tools: Python, PyTorch, Git, Slurm, Docker, Spark, LaTeX
Languages: English *(fluent)*, Chinese *(native)*, Cantonese *(native)*