Shuquan (Lillie) Ye

Email: shuquanye@gmail.com
Homepage: https://shuquanye.com/
Phone: +852 93623804

Introduction

My research interests focus on computer vision and machine learning, especially for multimodal (text, 2D, 3D, and 4D), creative, and human-aligned AI. Specifically, I am exploring the integration of human-aligned priors, learning strategies, multi-modal knowledge, and scalable-generalizable guidance, to both examine and enhance these models.

I have focused on these topics recently:

- AI for Science & Medical AI
- Learning on Imperfect 3D and Multimodal Data
- Embodied AI
- Human-aligned Perception, Reasoning, Benchmarks

Education

Feb 2023 - Oct 2023	Remote Visiting Student in Computer Science, New York University (NYU) Advisor: Prof. Saining Xie
Sep 2019 - Aug 2023	Ph.D. in Computer Science, City University of Hong Kong (CityUHK) Thesis: 3D Point Cloud Learning with Imperfect and Limited Data. Advisor: Prof. Jing Liao
Sep 2015 - Apr 2019	B.Eng. in Computer Science and Technology (Elite Program), South China University of Technology Thesis: Arbitrary-length Face Video Generation. Advisor: Prof. Shengfeng He

Working Experiences

Assistant Professor, AI for Science and Engineering Center, Shenzhen Loop Area Institute — Mar 2026 - Now

Postdoctoral Researcher, The Chinese University of Hong Kong, w/ Prof. Wanli Ouyang — Sept 2025 - Feb 2026

Research Fellow & Postdoctoral Researcher, City University of Hong Kong — Oct 2023 - Feb 2025

Research Intern, Microsoft Research, Redmond, w/ Dr. Dongdong Chen, Yujia Xie, Yichong Xu — Aug 2022 - Nov 2022

Selected Research Experiences

Human-aligned Multimodal LLM — Project lead @ CityU, Feb 2024 - Present

- Construct **the first** benchmark to evaluate how MLLMs follow the human vision system.
- Reveal new challenges for cutting-edge MLLMs, offer HVS-alignment insights, and propose methods to enhance them.

Animating 3D Gaussian Splatting Scenes — Project lead @ CityU, Oct 2023 - Present

- Propose **the first** method generating general text-conditioned 4D animation for arbitrary static 3DGS scenes.
- The method leverages single-view motion knowledge from video diffusion models and ensures multi-view consistency.

Robust Medical Image Segmentation with Noisy Labels — Project lead @ CityU, Aug 2022 - Apr 2024

- Propose **the first** single-network robust learning framework, by reformulating multi-networks to a memory bank.
- The first to explicitly solve the coarse boundary label problem, by proposing an edge detection branch and a thinning loss.

Inconspicuous Object Segmentation — CityU, May 2024 - Present

- Propose **the first** in-the-wild inconspicuous object benchmark and methods based on RGB-D data.

Open-vocabulary 3D Perception — CityU, Nov 2023 - Present

- Propose **the first** benchmark and method for generalized open-vocabulary 3D understanding. Supervise junior Ph.D..

3D Diffusion Model for Generic 3D Representation Synthesis — Project lead @ NYU, Feb 2023 - Oct 2023

- **The first** to generate general 3D latent (triplane) by learning 3D diffusion on large-scale in-the-wild single-view image.

Analyzing and Improving Commonsense of VL-Models — Project lead @ Microsoft Research, Aug 2022 - Nov 2022

- **The first** to evaluate and enhance VL-models' commonsense, by scalable contrastive learning on 447M+ image-text pairs.
- We generate scalable image-riddle pairs from knowledge graphs, offer benchmark, reveal gaps, and drive advancements.

3D Question Answering (3DQA) — Project lead @ CityU, Jan 2021 - Nov 2022

- **The first** to explore 3D question answering task, proposing the first dataset, with 10K+ fully human-annotated questions.
- We propose a novel transformer with dual encoders and 3D-L Bert for joint appearance-geometry understanding.

Robust Point Cloud Segmentation with Noisy Labels — Project lead @ CityU, Dec 2020 - Dec 2022

- **The first** to learn with noisy labels in real-world point cloud datasets, with both instance-level and boundary-level noise.
- We present a noise-rate blind, cluster-wise label correction framework robust to 60% label noise, and a cleaned ScanNetV2.

Arbitrary-Scale Point Cloud Upsampling Project lead @ CityU, Sep 2019 - Dec 2020

- **The first** to support continuously upsampling point cloud in arbitrary scale factor, by proposing a Meta-PU.

3D Portrait Stylization Project lead @ CityU, Nov 2019 - May 2022

- **The first** one-shot 3D portrait style transfer generating exaggerated geometry, stylized texture and preserved identity.

Arbitrary-length Face Video Generation Project lead @ SCUT, Apr 2017 - Apr 2019

- **The first** to generate arbitrary-length portrait video while maintaining both face identity and inter-frame coherence.

Publications

Journals

[1] Learning A Single Network for Robust Medical Image Segmentation with Noisy Labels
S. Ye, Y. Xu, D. Chen, S. Han, and J. Liao,
in IEEE Transactions on Medical Imaging (**TMI**, **JCR Q1**) 2024

[2] 3D Question Answering
S. Ye, D. Chen, S. Han, and J. Liao,
in IEEE Transactions on Visualization and Computer Graphics (**TVCG**, **CCF A, JCR Q1**) 2022

[3] Robust Point Cloud Segmentation with Noisy Annotations
S. Ye, D. Chen, S. Han, and J. Liao,
in IEEE Transactions on Pattern Analysis and Machine Intelligence (**TPAMI**, **CCF A, JCR Q1, IF: 24.314**) 2022

[4] Exemplar-Based 3D Portrait Stylization
F. Han*, ***S. Ye** **(*joint first authors)**, M. He, M. Chai, and J. Liao,
in *IEEE Transactions on Visualization and Computer Graphics (**TVCG**, **CCF A, JCR Q1**) 2021*

[5] Meta-PU: An Arbitrary-Scale Upsampling Network for Point Cloud
S. Ye, D. Chen, S. Han, Z. Wan, and J. Liao,
in *IEEE Transactions on Visualization and Computer Graphics (**TVCG**, **CCF A, JCR Q1**) 2021*

Conferences

[1] Do MLLMs Exhibit Human-like Perceptual Behaviors? HVSBench: A Benchmark for MLLM Alignment with Human Perceptual Behavior
J. Lin*, ***S. Ye** **(*joint first authors)**, D. Xu, W. Ouyang, R. Lau,
in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**, CCF A), Findings, 2026*

[2] OpenScan: A Benchmark for Generalized Open-Vocabulary 3D Scene Understanding
Y. Zhao, J. Lin, **S. Ye**, Q. Pang, and R. Lau,
in *The Association for the Advancement of Artificial Intelligence (**AAAI**, CCF A), 2026*

[3] Language-Guided Salient Object Ranking
F. Liu, Y. Liu, K. Xu, **S. Ye**, G.P. Hancke, and R. Lau,
in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**, CCF A), 2025*

[4] Leveraging RGB-D Data with Cross-Modal Context Mining for Glass Surface Detection
J. Lin*, Y. H. Yeung*, **S. Ye**# **(#corresponding author)**, and R. Lau,
in *The Association for the Advancement of Artificial Intelligence (**AAAI**, CCF A), 2025*

[5] Improving Commonsense in Vision-Language Models via Knowledge Graph Riddles **Highlight Paper, (2.6% acceptance rate)**
S. Ye, Y. Xie, D. Chen, Y. Xu, L. Yuan, C. Zhu and J. Liao,
in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**, CCF A), 2023*

[6] Learning with Noisy Labels for Robust Point Cloud Segmentation **Oral Presentation, (3.4% acceptance rate)**
S. Ye, D. Chen, S. Han, and J. Liao,
in *IEEE International Conference on Computer Vision (**ICCV**, CCF A), 2021*

[7] Coherence and Identity Learning for Arbitrary-length Face Video Generation
S. Ye, J. Lin, S. He, C. Han, G. Han and J. Qin,
in *25th International Conference on Pattern Recognition (ICPR), 2020*

[8] Two-dimensional-reduction Random Forest
S. Ye, Z. Yu, J. Lin, K. Yang, D. Dan, Z. Zhan, W. Chen, and Z. Zhang,
in *Eighth International Conference on Information Science and Technology*, 2018.

Manuscripts

[1]Do Multimodal Large Language Models See Like Humans?
S. Ye*, J. Lin*, D. Xu, W. Ouyang, and R. Lau,

[2] From Static Spectra to Operando Infrared Dynamics: Physics Informed Flow Modeling and a Benchmark
S. Ye, B. Fei, H. Xu, J. Lin, W. Ouyang

Mentoring and Student Supervision

2023-	Youjun Zhao, junior Ph.D.@ CityU, co-supervised with Prof. Rynson Lau and Dr. Jiaying Lin
2023-	Fang Liu, junior Ph.D.@ CityU, co-supervised with Prof. Rynson Lau and Dr. Ke Xu
2023-2024	Qianshi Pang, undergraduate @ SCUT, co-supervised with Dr. Jiaying Lin

Teaching

2022-2023	**Teaching**, Database System (CS3402), Undergraduate, CityU.
2022-2023	**Online Judge System Developer**, Database System (CS3402), Undergraduate, CityU.
2020-2023	**Teaching Assistant**, Database System (CS3402), Undergraduate, CityU, a total of eight semesters.

Invited Talks

Jul 2023	**Improving Commonsense in Vision-Language Models**, City University of Hong Kong.
Sep 2022	**3D Question Answering**, Beijing, with 2000 audiences.

Awards & Scholarships

2022, 2023	Postgraduate Studentship (Non-UGC funds - Central)
2022, 2023	Research Tuition Scholarship
2021, 2022	Research Tuition Scholarship
2020, 2022	Postgraduate Studentship (UGC-allocated funds and TA Scheme)
2019, 2020	Postgraduate Studentship (UGC-related research projects)
2017, 2018	Student Scholarship in South China University of Technology
2017, 2018	Merit Student in South China University of Technology

Academic Services

Reviewer for international conferences:

- IEEE Conference on Computer Vision and Pattern Recognition (**CVPR**) 2022, 2023, 2024, 2025
- International Conference on Computer Vision (**ICCV**) 2023, 2025
- European Conference on Computer Vision (**ECCV**) 2022,
- ...

Reviewer for international journals:

- IEEE Transactions on Pattern Analysis and Machine Intelligence (**TPAMI**)
- ACM Transactions on Graphics (**TOG**)
- International Journal of Computer Vision (**IJCV**)
- IEEE Transactions on Visualization and Computer Graphics (**TVCG**)
- IEEE Transactions on Multimedia (**TMM**)
- IEEE Journal of Biomedical and Health Informatics (**JBHI**)
- Pattern Recognition (**PR**)
- IEEE Geoscience and Remote Sensing Letters (**GRSL**)
- IEEE Access (**Access**)
- ...

Skills

Languages	English, Mandarin.
Programming Languages	Python, C++, C, Java, Matlab, C#.
Deep Learning Frameworks	Pytorch, Tensorflow, Keras.
Misc	CUDA, Large-scale Multi-GPU training.