Shiyu Hu (胡世宇)

Research Fellow, Nanyang Technological University (NTU), Singapore

1995.10.29 shiyu.hu@ntu.edu.sg hushiyu1995 huuuuusy

https://huuuuusy.github.io/ http://viig.aitestunion.com/

Work Experience

2024.08 - Now **Research Fellow, School of Physical and Mathematical Sciences (SPMS), Nanyang Technological University (NTU)**
- **Direction:** AI4Science, Computer Vision
- **PI:** Assoc Prof. Kanghao Cheong (IEEE Senior Member)

2018.03 - 2018.11 **Research Assistant, University of Hong Kong (HKU)**
- **Direction:** High Performance Computing, Heterogeneous Computing
- **PI:** Prof. Choli Wang

2016.08 - 2016.09 **Research Intern, Institute of Electronics, Chinese Academy of Sciences (CASIE)**

Teaching Experience

2024.08 - 2024.12 **Teach Assistant in Nanyang Technological University (NTU)**
- **Course:** SC1123 (Math 1: Linear Algebra and Calculus for Computing)
- **Course Author:** Assoc Prof. Kanghao Cheong

2018.01 - 2018.04 **Teach Assistant for Experimental Course in University of Hong Kong (HKU)**
- **Course:** COMP7305 (Cluster and Cloud Computing)
- **Course Author:** Prof. Choli Wang

Education Background

2019.09 - 2024.01 **Ph.D, Institute of Automation, Chinese Academy of Sciences (CASIA)**
- **Major:** Computer Applied Technology
- **Supervisor:** Prof. Kaiqi Huang (IAPR Fellow, IEEE Senior Member)
- **Co-supervisor:** Prof. Xin Zhao (IEEE Senior Member)
- **Thesis title:** *Research of Intelligence Evaluation Techniques for Single Object Tracking*
- **Thesis committee:** Prof. Jianbin Jiao, Prof. Yuxin Peng, Prof. Yao Zhao (IEEE Fellow, IET Fellow), Prof. Yunhong Wang (IEEE Fellow, IAPR Fellow, CCF Fellow), Prof. Ming Tang
- **Thesis defense grade:** Excellent

2017.09 - 2019.07 **M.Sc., Department of Computer Science, University of Hong Kong (HKU)**
- **Major:** Computer Science
- **Supervisor:** Prof. Choli Wang
- **Thesis title:** *NightRunner: Deep Learning for Autonomous Driving Cars after Dark*
- **Thesis defense grade:** A+

2013.09 - 2017.07 **B.E., Elite Class in School of Information and Electronics, Beijing Institute of Technology (BIT)**
- **Major:** Information Engineering
- **Thesis title:** *Text Sentiment Analysis Based on Deep Neural Network*
- **Thesis defense grade:** Excellent

2015.07 - 2015.08 **Summer Semester, University of California, Berkeley (UCB)**
- **Major:** New Media
- **Course grade:** A

Research Foundation & Interests

Data-centric AI	Research on construction strategy of single-modal and multi-modal datasets incorporating human knowledge structure [A1], [A2].
	Research on evaluation mechanisms for robustness, generalization, and safety [A3], [A4].
Visual Turing Test	Design of a human-machine universal visual ability evaluation framework [A5], [P1].
	Benchmarking the performance of algorithms based on human abilities in perceptual, cognitive, inferential, etc. Analyzing the bottlenecks of algorithms and human subjects in depth, providing guidance for research on human-like modeling, human-machine collaboration, and human-machine integration [A1], [A6].
Video Understanding	Exploring using Large Language Models (LLMs) and Large Vision Models (LVMs) for long video understanding [R1].
Visual Object Tracking	Research on object tracking algorithms in general scenes and specific scenarios (e.g., unmanned aerial vehicles) [A4], [A6], [O1], [R2].
Visual Language Tracking	Research on multi-modal tracking, video understanding, and visual reasoning tasks based on long video sequences [A7]–[A9], [R3], [R4], [P2].
Human-machine Interaction	Exploring human-computer interaction patterns in video sequences with various proxy tasks [A10], [P3].
AI4Science	**Education:** Research on human-computer interaction (HCI) technology for education scenarios, including designing an intelligent education framework from a multidisciplinary perspective, investigating HCI technology, conducting qualitative and quantitative analysis [R5], [R6], [O2].
	Cognitive Science: Visual task design, environment construction, and human-machine capability analysis based on human-like modeling principles [R7], [R8].
	Medical Science: Research on medical image processing techniques based on artificial intelligence technologies (e.g., cell segmentation and tracking) [A11].
	Psychology: Development of gamified assessment systems targeting psychological dimensions such as anxiety, depression, and obsession, along with research on intelligent psychological evaluation technologies. Exploring using LLMs and LVMs for visual comprehension with psychological elements [A12]–[A14], [R9].

Research Publications

Acceptance

1. **S. Hu**, X. Zhao, L. Huang, and K. Huang, "Global instance tracking: Locating target more like humans," *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, vol. 45, no. 1, pp. 576–592, 2023. DOI: 10.1109/TPAMI.2022.3153312.
2. **S. Hu**, D. Zhang, M. Wu, X. Feng, X. Li, X. Zhao, and K. Huang, "A multi-modal global instance tracking benchmark (mgit): Better locating target in complex spatio-temporal and causal relationship," in *The 37th Conference on Neural Information Processing Systems (NeurIPS, Poster)*, vol. 36, 2023, pp. 25 007–25 030.
3. **S. Hu**, X. Zhao, and K. Huang, "Sotverse: A user-defined task space of single object tracking," *International Journal of Computer Vision (IJCV)*, vol. 132, pp. 872–930, 2024. DOI: 10.1007/s11263-023-01908-5.
4. X. Zhao ✉, **S. Hu** ✉, Y. Wang, J. Zhang, Y. Hu, R. Liu, H. Ling, Y. Li, R. Li, K. Liu, and J. Li, "Biodrone: A bionic drone-based single object tracking benchmark for robust vision," *International Journal of Computer Vision (IJCV)*, vol. 132, pp. 1659–1684, 2024. DOI: 10.1007/s11263-023-01937-0.

5. **S. Hu**, X. Zhao, and K. Huang, "Visual intelligence evaluation techniques for single object tracking: A survey," *Journal of Images and Graphics* (《中国图象图形学报》), 2023.

6. D. Zhang, **S. Hu**, X. Feng, X. Li, M. Wu, J. Zhang, and K. Huang, "Beyond accuracy: Tracking more like human via visual search," *The 38th Conference on Neural Information Processing Systems (NeurIPS, Poster)*, 2024.

7. X. Feng, X. Li, **S. Hu**, D. Zhang, M. Wu, J. Zhang, X. Chen, and K. Huang, "Memvlt: Visual-language tracking with adaptive memory-based prompts," *The 38th Conference on Neural Information Processing Systems (NeurIPS, Poster)*, 2024.

8. X. Feng, D. Zhang, **S. Hu**, X. Li, M. Wu, J. Zhang, X. Chen, and K. Huang, "Enhancing vision-language tracking by effectively converting textual cues into visual cues," *The 50th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP, Poster)*, 2025.

9. X. Li, X. Feng, **S. Hu**, M. Wu, D. Zhang, J. Zhang, and K. Huang, "Dtllm-vlt: Diverse text generation for visual language tracking based on llm," *The 3rd Workshop on Vision Datasets Understanding and DataCV Challenge in The IEEE/CVF Conference on Computer Vision and Pattern Recognition 2024 (CVPRW, Oral, Best Paper Honorable Mention)*, 2024.

10. M. Wu, K. Huang, Y. Cai, **S. Hu**, Y. Zhao, and W. Wang, "Finger in camera speaks everything: Unconstrained air-writing for real-world," *IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)*, 2024.

11. J. Zhang, T. Zhao, **S. Hu**, and X. Zhao, "Robust single-particle cryo-em image denoising and restoration," in *The 49th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP, Poster)*, 2024.

12. M. Wu, Y. Kang, X. Li, **S. Hu**, X. Chen, Y. Kang, W. Wang, and K. Huang, "Vs-llm: Visual-semantic depression assessment based on llm for drawing projection test," *The 7th Chinese Conference on Pattern Recognition and Computer Vision (PRCV, Poster)*, 2024.

13. X. Feng, **S. Hu**, X. Chen, and K. Huang, "A hierarchical theme recognition model for sandplay therapy," in *The 6th Chinese Conference on Pattern Recognition and Computer Vision (PRCV, Poster)*, 2023, pp. 241–252. DOI: 10.1007/978-981-99-8462-6_20.

14. K. Huang, Y. Kang, C. Yan, **S. Hu**, L. Wang, T. Tao, and W. Gao, "A review of intelligent psychological assessment based on interactive environment," *Chinese Mental Health Journal* (《中国心理卫生杂志》), 2024.

15. Y. Wang, **S. Hu**, and X. Zhao, "Rethinking similar object interference in single object tracking," in *The 7th International Conference on Computer Science and Artificial Intelligence (CSAI, Oral)*, 2023, pp. 251–258.

16. K. Huang, X. Zhao, Q. Li, and **S. Hu**, "Visual turing: The next development of computer vision in the view of human-computer gaming," *Journal of Graphics* (《图学学报》), vol. 42, no. 3, p. 339, 2021. DOI: 10.11996/JG.j.2095-302X.2021030339.

17. Y. Zhang, C. Liu, W. Chen, X. Xu, F. Wang, H. Li, **S. Hu**, and X. Zhao, "Revisiting instance search: A new benchmark using cycle self-training," *Neurocomputing (Neu)*, vol. 501, pp. 270–284, 2022. DOI: 10.1016/j.neucom.2022.06.027.

Preprint

1. **S. Hu**, X. Zhao, Y. Wang, Y. Shan, and K. Huang, *Nearing or surpassing: Overall evaluation of human-machine dynamic vision ability*, 2023. URL: https://openreview.net/forum?id=LGbzYw_pnsc.

2. X. Li, **S. Hu**, X. Feng, D. Zhang, M. Wu, J. Zhang, and K. Huang, "Dtvlt: A multi-modal diverse text benchmark for visual language tracking based on llm," *arXiv preprint arXiv:2410.02492*, 2024.

3. X. Li, **S. Hu**, X. Feng, D. Zhang, M. Wu, J. Zhang, and K. Huang, "Visual language tracking with multi-modal interaction: A robust benchmark," *arXiv preprint arXiv:2409.08887*, 2024.

Under Review

1. **S. Hu**∗, X. Li∗, X. Li, J. Zhang, Y. Wang, X. Zhao, and K. Cheong, "Can lvlms describe videos like humans? a five-in-one video annotations benchmark for better human-machine comparison," *The 13th International Conference on Learning Representations (ICLR, Under Review)*, 2024.
2. S. Jia, Y. Cao, **S. Hu**, X. Lu, X. Cheng, and X. Lu, "Distance-guided interaction learning for robust multi-object tracking," *The IEEE/CVF Conference on Computer Vision and Pattern Recognition 2025 (CVPR, Under Review)*, 2025.
3. X. Li∗, **S. Hu**∗, X. Feng, D. Zhang, M. Wu, J. Zhang, and K. Huang, "How texts help? a fine-grained evaluation to reveal the role of language in vision-language tracking," *The IEEE/CVF Conference on Computer Vision and Pattern Recognition 2025 (CVPR, Under Review)*, 2025.
4. X. Feng, **S. Hu**, X. Li, D. Zhang, M. Wu, J. Zhang, X. Chen, and K. Huang, "Atctrack: Aligning target-context cues with dynamic target states for robust vision-language tracking," *The IEEE/CVF Conference on Computer Vision and Pattern Recognition 2025 (CVPR, Under Review)*, 2025.
5. Y. Ma∗, **S. Hu**∗, X. Li, Y. Wang, S. Liu, and K. Cheong, "Students rather than experts: A new ai for education pipeline to model more human-like and personalised early adolescences," *The 13th International Conference on Learning Representations (ICLR, Under Review)*, 2024.
6. L. Tan, **S. Hu**, D. Yao, and K. Cheong, "Artificial intelligence-enabled adaptive learning platforms: A review," *Computers & Education: Artificial Intelligence*, 2024.
7. D. Shang, M. Yao, **S. Hu**, K. Wang, J. Zhang, B. Xu, and G. Li, "Adaptive continual learning through proactive detection of transfer and interference," *The 13th International Conference on Learning Representations (ICLR, Under Review)*, 2024.
8. M. Wu, D. Shang, Z. Song, **S. Hu**, B. Zhang, G. Li, K. Huang, and W. Wang, "Dspikeair: Efficient dynamic spike-driven unconstrained air-writing recognition with balanced energy and accuracy," *The IEEE/CVF Conference on Computer Vision and Pattern Recognition 2025 (CVPR, Under Review)*, 2025.
9. Y. Ren, X. Feng, **S. Hu**, Y. Kang, C. Yan, Y. Zeng, L. Wang, and K. Huang, "Intelligent psychological assessment with sandplay based on evidence-centered design theory," *Acta Psychologica Sinica (《心理学报》, Under Review)*, 2024.

Ongoing Research

1. X. Zhao, **S. Hu**, and X. Yin, *Visual Object Tracking - An Evaluation Perspective*. Springer, 2025.
2. Y. Ma, Z. Yang, Y. Kang, X. Xue, and **S. Hu**, "From collaboration to integration: Research on the dilemma and countermeasures of human-computer interaction in intelligent education," *Intelligent Education PhD Research Fund, supported by the Institute of AI Education Shanghai and East China Normal University (In Progress)*, 2024.

Skills

Languages	Mandarin Chinese (native speaker) and English.
Coding	Python, Java, Matlab, C, LaTeX.
Development	Android, Flask, SQLite.
Linux	Shell, OS virtualization.
Misc.	Academic research, leadership, presentation.

Awards and Honors

2024 **Best Paper Honorable Mention**, the 3rd Workshop on Vision Datasets Understanding and DataCV Challenge in CVPR 2024.

Beijing Outstanding Graduates, Beijing Municipal Education Commission (Top 5%).

2023 **China National Scholarship**, Ministry of Education of the People's Republic of China (Top 1%).

First Prize of Climbing Scholarship, Institute of Automation, Chinese Academy of Sciences.

2022 **Merit Student**, University of Chinese Academy of Sciences.

2017 **Academic Scholarship**, Beijing Institute of Technology.

Excellent Innovative Student, Beijing Institute of Technology.

2016 **College Scholarship**, Chinese Academy of Sciences.

Academic Scholarship, Beijing Institute of Technology.

Excellent League Member on Youth Day Competition, Beijing Institute of Technology.

2015 **National First Prize**, Contemporary Undergraduate Mathematical Contest in Modeling (Top 1%).

Academic Scholarship, Beijing Institute of Technology.

First Prize of Mathematics Modeling Competition, Beijing Institute of Technology.

Outstanding Individual on Summer Social Practice, Beijing Institute of Technology.

Second Prize on Summer Social Practice, Beijing Institute of Technology (Team Leader).

Outstanding Student Cadre, Beijing Institute of Technology.

Outstanding League Cadre on Youth Day Competition, Beijing Institute of Technology.

Outstanding Youth League Branch, Beijing Institute of Technology (Team Leader).

Top 10 Activities on Youth Day Competition, Beijing Institute of Technology (Team Leader).

2014 **Academic Scholarship**, Beijing Institute of Technology.

Outstanding Student, Beijing Institute of Technology.

2013 **Academic Scholarship**, Beijing Institute of Technology.

Academic Activities and Services

Tutorial

31th IEEE International Conference on Image Processing (ICIP)
- **Title:** An Evaluation Perspective in Visual Object Tracking: from Task Design to Benchmark Construction and Algorithm Analysis
- **Date & Location:** 9:00-12:30, 27th October, 2024, Abu Dhabi, United Arab Emirates
- **Duration:** Half-day (Three Hours)

27th International Conference on Pattern Recognition (ICPR)
- **Title:** Visual Turing Test in Visual Object Tracking: A New Vision Intelligence Evaluation Technique based on Human-Machine Comparison
- **Date & Location:** 14:30-18:00, 1st December, 2024, Kolkata, India
- **Duration:** Half-day (Three Hours)

17th Asian Conference on Computer Vision (ACCV)
- **Title:** From Machine-Machine Comparison to Human-Machine Comparison: Adapting Visual Turing Test in Visual Object Tracking
- **Date & Location:** 9:00-12:00, 9th December, 2024, Hanoi, Vietnam
- **Duration:** Half-day (Three Hours)

Associate Editor **Journal:** Innovation and Emerging Technologies

Reviewer **Conference:** NeurIPS, ICML, ICLR, CVPR, ECCV, AAAI, ACMMM, AISTATS, etc.

Academic Activities and Services (continued)

- **Journal:** IEEE Transactions on Image Processing, SCIENCE CHINA Information Sciences, Scientific Reports, IEEE Access, Journal of Computational Science, Journal of Electronic Imaging, Digital Signal Processing, etc.

Assisted Student Supervision

Ph.D. Student

- **Meiqi Wu**, 2022.08-Now, University of Chinese Academy of Sciences (Computer Vision & Human-computer Interaction)
- **Xiaokun Feng**, 2023.04-Now, Institute of Automation, Chinese Academy of Sciences (Visual Object Tracking & Visual Language Tracking)
- **Yiping Ma**, 2023.08-Now, East China Normal University (Intelligent Education Technique & Human-computer Interaction)
- **Dailing Zhang**, 2023.08-Now, Institute of Automation, Chinese Academy of Sciences (Visual Object Tracking & Visual Turing Test & AI Agent)
- **Yipei Wang**, 2024.08-Now, Southeast University (Multimodal Large Language Models & Visual Object Tracking)
- **Xuchen Li**, 2024.08-Now, Institute of Automation, Chinese Academy of Sciences (Visual Language Tracking & Multimodal Large Language Models & AI4Science)
- **Hong-jui Shen**, 2024.10-Now, School of Physical & Mathematical Sciences, Nanyang Technological University (NTU), Singapore (Large Language Models & AI4Science)
- **Kainan Li**, 2024.10-Now, School of Physical & Mathematical Sciences, Nanyang Technological University (NTU), Singapore (Computer Vision & AI4Science)

M.S. Student

- **Yiping Ma**, 2022.05-2023.07, Nanjing Normal University (Intelligent Education Technique & Speech Emotion Recognition)
- **Yipei Wang**, 2022.08-2024.07, Southeast University (Visual Object Tracking & LLM for Recommendation System)
- **Yuqi Cui**, 2024.07-Now, University of Science and Technology Beijing (Visual Object Tracking)
- **Panxi Xu**, 2024.09-Now, University of Science and Technology Beijing (Large Language Models & Visual Object Tracking)

B.E. Student

- **Junyou Zhu**, 2022.09-2023.08, University of Chinese Academy of Sciences (Visual Turing Test)
- **Lihang Hu**, 2022.09-2023.08, University of Chinese Academy of Sciences (Visual Object Tracking)
- **Dailing Zhang**, 2022.09-2023.08, Southeast University (Visual Object Tracking)
- **Xuchen Li**, 2023.04-2024.07, Beijing University of Posts and Telecommunications (Visual Object Tracking & Visual Language Tracking)
- **Zi Ye**, 2024.09-Now, University of Science and Technology Beijing (Large Language Models & Visual Object Tracking)
- **Leying Tan**, 2024.09-Now, Undergrad Student (DSAI) Batch, Nanyang Technological University (NTU), Singapore (Large Language Models & AI4Science)
- **Shunya Hirashima**, 2024.10-Now, Undergrad Student (CSC) Batch, Nanyang Technological University (NTU), Singapore (Large Language Models & AI4Science)
- **Jinlin Ma**, 2024.10-Now, Undergrad Student (CE) Batch, Nanyang Technological University (NTU), Singapore (AI for Social Science)
- **Yuxiao Li**, 2024.10-Now, Undergrad Student (MACS) Batch, Nanyang Technological University (NTU), Singapore (AI for Social Science)

References

Professors Kaiqi Huang and Xin Zhao served as my Ph.D. supervisor and co-supervisor, respectively, with whom I collaborated on research in computer vision. Additionally, Prof. Choli Wang oversaw my M.Sc. studies at HKU, and I had the privilege of working with him on high-performance computing projects. Currently, I am lucky to work with Prof. Kanghao Cheong at NTU.

Prof. Kaiqi Huang
Professor, IAPR Fellow, IEEE Senior Member, 10,000 Talents Program - Leading Talents
Director of Center for Research on Intelligent Systems and Engineering (CRISE)
National Laboratory of Pattern Recognition (NLPR)
Institute of Automation, Chinese Academy of Sciences (CASIA)
95 Zhongguancun East Road, Beijing, China
kqhuang@nlpr.ia.ac.cn

Prof. Choli Wang
Honorary Professor
Department of Computer Science, University of Hong Kong (HKU)
Pokfulam, Hong Kong, China
clwang@cs.hku.hk
choliwang@gmail.com

Prof. Xin Zhao
Professor, IEEE Senior Member, Beijing Science Fund for Distinguished Young Scholars
School of Computer and Communication Engineering, University of Science and Technology Beijing (USTB)
30 Xueyuan Road, Beijing, China
xinzhao@ustb.edu.cn
xzhaopersonal@foxmail.com

Prof. Kanghao Cheong
Associate Professor, IEEE Senior Member
Assistant Dean, School of Physical & Mathematical Sciences
Assistant Dean, College of Science
Nanyang Technological University (NTU)
50 Nanyang Avenue, Singapore
kanghao.cheong@ntu.edu.sg