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/

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. Kang Hao 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. Kang Hao 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 Ba	ickground
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 (CAAI/CIE/CSIG Fellow), Prof. Yao Zhao (IEEE/IET Fellow), Prof. Yunhong Wang (IEEE/IAPR/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 2015.07 - 2015.08	 B.E., Elite Class in School of Information and Electronics, Beijing Institute of Technology (BIT) Major: Information Engineering Diploma project supervisor: Prof. Senlin Luo Thesis title: Text Sentiment Analysis Based on Deep Neural Network Thesis defense grade: Excellent 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 in- corporating human knowledge structure.
	Research on evaluation mechanisms for robustness, generalization, and safety.
Visual Object Tracking	Research on object tracking algorithms in general scenes and specific scenarios (e.g., unmanned aerial vehicles).
Visual Language Tracking	Research on multi-modal tracking, video understanding, and visual reasoning tasks based on long video sequences.
Video Understanding	Exploring using Large Language Models (LLMs) and Large Vision Models (LVMs) for long video understanding.
Visual Turing Test	Design of a human-machine universal visual ability evaluation framework .
	Benchmarking the performance of algorithms based on human abilities in per- ceptual, cognitive, inferential, etc. Analyzing the bottlenecks of algorithms and human subjects in depth, providing guidance for research on human-like mod- eling, human-machine collaboration, and human-machine integration.
Human-machine Interaction	Exploring human-computer interaction patterns in video sequences with various proxy tasks.
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, and conducting qualitative and quantitative analysis.
	Cognitive Science: Visual task design, environment construction, and human- machine capability analysis based on human-like modeling principle.
	Medical Science: Research on medical image processing techniques based on artificial intelligence technologies (e.g., cell segmentation and tracking).
	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.

Research Publications

Book

X. Zhao, S. Hu, and X. Yin, Visual Object Tracking - An Evaluation Perspective. Springer, 2025, ISBN: 978-981-96-4558-9.

Acceptance

- **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, CCF-A)*, vol. 45, no. 1, pp. 576–592, 2023. *O* DOI: 10.1109/TPAMI.2022.3153312.
- **S. Hu**, X. Zhao, and K. Huang, "Sotverse: A user-defined task space of single object tracking," *International Journal of Computer Vision (IJCV, CCF-A)*, vol. 132, pp. 872–930, 2024. *O* DOI: 10.1007/s11263-023-01908-5.
- 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, CCF-A)*, vol. 132, pp. 1659–1684, 2024. *I* DOI: 10.1007/s11263-023-01937-0.
- 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 *Conference on Neural Information Processing Systems (NeurIPS, CCF-A, Poster)*, vol. 36, 2023, pp. 25 007–25 030.
 Last update: 2025-5-28 Page-2

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》, CCF-B), 2023. Y. Wang, J. Zhang, Y. Wang, S. Hu 🖾 , B. Shen, Z. Hou, and W. Zhou, "Improved sar aircraft detection algorithm based on visual state space models," IET Computer Vision (IET-CVI, CCF-C), 2025. X. Feng, D. Zhang, S. Hu, X. Li, M. Wu, J. Zhang, X. Chen, and K. Huang, "Cstrack: Enhancing rgb-x tracking via compact spatiotemporal features," in International Conference on Machine Learning (ICML, CCF-A, Poster), 2025. D. Zhang, S. Hu, X. Feng, X. Li, M. Wu, J. Zhang, and K. Huang, "Beyond accuracy: Tracking more like human via visual search," in Conference on Neural Information Processing Systems (NeurIPS, CCF-A, Poster), 2024. 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," in Conference on Neural Information Processing Systems (NeurIPS, CCF-A, Poster), 2024. 10 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," in IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshop (CVPRW, CCF-A Workshop, Oral, Best Paper Honorable Mention), 2024. 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," in IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP, CCF-B, Poster), 2025. J. Zhang, T. Zhao, S. Hu, and X. Zhao, "Robust single-particle cryo-em image denoising and restoration," 12 in IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP, CCF-B, Poster), 2024. 13 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*, *CCF-B*), 2024. M. Wu, Y. Kang, X. Li, S. Hu, X. Chen, Y. Kang, W. Wang, and K. Huang, "Vs-llm: Visual-semantic 14 depression assessment based on llm for drawing projection test," in Chinese Conference on Pattern Recognition and Computer Vision (PRCV, CCF-C, Poster), 2024. X. Feng, S. Hu, X. Chen, and K. Huang, "A hierarchical theme recognition model for sandplay therapy," in 15 Chinese Conference on Pattern Recognition and Computer Vision (PRCV, CCF-C, Poster), 2023, pp. 241–252. 𝔗 DOI: 10.1007/978-981-99-8462-6_20. 16 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, CCF-C)*, vol. 501, pp. 270–284, 2022. *P* DOI: 10.1016/j.neucom.2022.06.027. K. Huang, X. Zhao, Q. Li, and S. Hu, "Visual turing: The next development of computer vision in the view 17 of human-computer gaming," Journal of Graphics (《图学学报》, CCF-C), vol. 42, no. 3, p. 339, 2021. 𝔗 DOI: 10.11996/JG. j. 2095−302X. 2021030339. 18 L. Tan, S. Hu, D. J. Yeo, and K. H. Cheong, "Artificial intelligence-enabled adaptive learning platforms: A review," Computers & Education: Artificial Intelligence (C&E:AI), 2025. 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》), 2025. Y. Wang, S. Hu, and X. Zhao, "Rethinking similar object interference in single object tracking," in International Conference on Computer Science and Artificial Intelligence (CSAI, EI, Oral), 2023, pp. 251–258.

Preprint

- **S. Hu***, X. Li*, X. Li, J. Zhang, Y. Wang, X. Zhao, and K. H. Cheong, "Fiova: A multi-annotator benchmark for human-aligned video captioning," *arXiv preprint arXiv:2410.15270*, 2024.
- 2 Y. Ma*, **S. Hu***, X. Li, Y. Wang, Y. Chen, S. Liu, and K. H. Cheong, "When llms learn to be students: The soei framework for modeling and evaluating virtual student agents in educational interaction," *arXiv preprint arXiv:2410.15701*, 2024.
- 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," *arXiv preprint arXiv:2411.15600*, 2024.
- **4** 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.
- 5 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.
- 6 S. Hu, X. Zhao, Y. Wang, Y. Shan, and K. Huang, Nearing or surpassing: Overall evaluation of human-machine dynamic vision ability, 2023. *O* URL: https://openreview.net/forum?id=LGbzYw_pnsc.

Projects

Framework

2018.03 - 2018.11	Darknet-Cross: Light-weight Deep Learning Framework for Heterogeneous Com- puting
	https://github.com/huuuuusy/Darknet-Cross
	Darknet-Cross supports algorithm acceleration processing on various platforms (e.g.,
	Android and Ubuntu) and various GPUs (e.g., Nvidia GTX1070 and Adreno 630).
	\checkmark The work is a part of my master's thesis at HKU (thesis defense grade: A+).
Platform (Deve	elopment & Maintenance)
2019.11 - Now	📕 VideoCube / MGIT Platform
	http://videocube.aitestunion.com
	VideoCube / MGIT is the supporting platform for research accepted by IEEE TPAMI 2023 and NeurIPS 2023.
	\checkmark As of Sept. 2024, the platform has received 440k+ page views, 1.2k+ downloads, 420+ trackers from 220+ countries and regions worldwide.
2021.07 - Now	SOTVerse / VLTVerse Platform
	http://metaverse.aitestunion.com
	SOTVerse is the supporting platform for research accepted by IJCV 2024.
	As of Sept. 2024, the platform has received 126k+ page views from 150+ countries and regions worldwide.
2022.05 - Now	📕 BioDrone Platform
	http://biodrone.aitestunion.com/
	BioDrone is the supporting platform for research accepted by IJCV 2024.
	✓ As of Sept. 2024, the platform has received 170k+ page views from 200+ countries and regions worldwide
Platform (Main	itenance)
r tationn (rian	
2020.07 - Now	GOT-10k Platform
	<pre> http://got-10k.aitestunion.com/</pre>
	GOT-10k is the supporting platform for research accepted by IEEE TPAMI 2021.

✓ As of Sept. 2024, the platform has received 3.92M+ page views, 7.5k+ downloads, 21.5k+ trackers from 290+ countries and regions worldwide.

Challenge (Organizer)

2023.05 - 2023.11 Hislopvision Challenge

http://hislopvision.aitestunion.com/

Hislopvision Challenge supports the 3rd High-speed and Low-power Visual Understanding Challenge in the 5th Chinese Conference on Pattern Recognition and Computer Vision. ✓ The participating teams include researchers from Tsinghua University, Beijing Institute of Technology, Jilin University, etc.

Challenge (Participant)

2021.01 - 2021.04 Cell Tracking Challenge

https://celltrackingchallenge.net/

✓ This project was submitted to the Cell Tracking Challenge in Mar. 2021, and maintains the second place in the Fluo-C2FL-MSC+ dataset and the third place in the Fluo-C2FL-Huh7 dataset (statistics by Oct. 2023).

Grant (Participant)

2023.12 - Now Research on the Dilemma and Countermeasures of Human-Computer Interaction in Intelligent Education

> ✓ The project is funded by the 2023 Intelligent Education PhD Research Fund, supported by the Institute of AI Education Shanghai and East China Normal University, and is currently in progress.

Academic Activities and Services

Tutorial	 34th International Joint Conference on Artificial Intelligence (IJCAI) Title: Human-Centric and Multimodal Evaluation for Explainable AI: Moving Beyond Benchmarks Date & Location: 16th-22nd August, 2025, Montreal, Canada
	 28th European Conference on Artificial Intelligence (ECAI) Title: From Benchmarking to Trustworthy AI: Rethinking Evaluation Methods Across Vision and Complex Systems Date & Location: 25th-30th October, 2025, Bologna, Italy
	 2025 IEEE International Conference on Systems, Man, and Cybernetics (SMC) Title: The Synergy of Large Language Models and Evolutionary Optimization on Complex Networks Date & Location: 5th-8th October, 2025, Vienna, Austria
	 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
	 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
	 31st 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
Guest Editor	Journal: Electronics (Special Issue: Techniques and Applications of Multimodal Data Fusion)

Academic Activities and Services (continued)



Development Android, Flask, SQLite.

Linux 📕 Shell, OS virtualization.

Misc. 📕 Academic research, leadership, presentation.

Awards and Honors



Assisted Student Supervision

Ph.D. Student 📕 Meiqi Wu, 2022.08-Now, University of Chinese Academy of Sciences

- Xiaokun Feng, 2023.04-Now, Institute of Automation, Chinese Academy of Sciences
- Yiping Ma, 2023.08-Now, East China Normal University

Assisted Student Supervision (continued)



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. Kang Hao 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) kqhuang@nlpr.ia.ac.cn

Prof. Choli Wang

Honorary Professor

Department of Computer Science, University of Hong Kong (HKU)

🗹 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)

- 🖂 xinzhao@ustb.edu.cn
- ☑ xzhaopersonal@foxmail.com

Prof. Kang Hao Cheong

Associate Professor, IEEE Senior Member Assistant Dean, School of Physical & Mathematical Sciences

Assistant Dean, College of Science

- Nanyang Technological University (NTU)
- 🖂 kanghao.cheong@ntu.edu.sg