

胡世宇

Research Fellow, 新加坡南洋理工大学

1995.10.29

shiyu.hu@ntu.edu.sg

hushiyu1995

huuuuusy

https://huuuuusy.github.io/

工作经历

2024.08 - 至今 ■ Research Fellow, 南洋理工大学 (NTU), 物理与数学科学学院

- 方向: AI4Science, 计算机视觉
- PI: Assoc Prof. Kanghao Cheong (IEEE Senior Member)

2018.03 - 2018.11 ■ Research Assistant, 香港大学 (HKU), 工程学院, 计算机科学系

- 方向: 高性能计算, 异构计算
- PI: Prof. Choli Wang

2016.08 - 2016.09 ■ Research Intern, 中国科学院, 电子学研究所 (CASIE)

教学经历

2024.08 - 2024.12 ■ 助教, 南洋理工大学 (NTU)

- 课程: SC1123 (Math 1: Linear Algebra and Calculus for Computing)
- 授课教师: Assoc Prof. Kanghao Cheong

2018.01 - 2018.04 ■ 实验课助教, 香港大学 (HKU)

- 课程: COMP7305 (Cluster and Cloud Computing)
- 授课教师: Prof. Choli Wang

教育经历

2019.09 - 2024.01 ■ 博士, 中国科学院, 自动化研究所 (CASIA)

- 专业: 计算机应用技术
- 导师: 黄凯奇研究员 (IAPR会士, 万人领军)
- 合作导师: 赵鑫教授 (北京市杰青)
- 论文题目: 单目标跟踪中的智能评估技术研究
- 答辩委员会: 焦建彬教授, 彭宇新教授 (国家杰青), 赵耀教授 (IEEE会士, IET会士, 国家杰青), 王蕴红教授 (IEEE会士, IAPR会士, CCF会士), 唐明研究员
- 论文答辩成绩: 全优

2017.09 - 2019.07 ■ 硕士, 香港大学 (HKU), 工程学院, 计算机科学系

- 专业: 计算机科学
- 导师: Prof. Choli Wang
- 论文题目: NightRunner: Deep Learning for Autonomous Driving Cars after Dark
- 论文答辩成绩: A+

2013.09 - 2017.07 ■ 学士, 北京理工大学 (BIT), 信息与电子学院, 信息工程菁英班

- 专业: 电子信息工程
- 毕业设计导师: 罗森林教授
- 论文题目: 基于深度神经网络的文本情感分类研究
- 论文答辩成绩: 全优

2015.07 - 2015.08 ■ 暑期学期, 加州大学伯克利分校 (UCB)

- 专业: 新媒体
- 课程成绩: A

研究基础 & 研究兴趣

数据为中心的AI ■ 融合人类知识结构的大规模单模态、多模态数据集构建策略研究。

研究基础 & 研究兴趣 (continued)






- 面向视觉鲁棒性、泛化性、安全性的评估机制研究和评价指标设计。
- 视觉目标跟踪 ■ 通用场景、专用场景（例如，无人机）下的单目标跟踪算法研究。
- 视觉语言跟踪 ■ 基于长视频序列的多模态跟踪、视频理解和视觉推理研究。
- 视频理解 ■ 结合大语言模型（LLMs）和基础视觉模型（LVMs）的长视频理解研究。
- 视觉图灵 ■ 人机通用的视觉智能评估框架设计。
- 以人为基准度量算法在感知、认知、推理等能力维度上的表现，并深入分析算法和人类被试的瓶颈，为类人建模、人机协作、人机融合等研究提供引导。
- 人机交互 ■ 以多种视觉任务为代理的人机交互模式探究。
- AI4Science ■ 教育学：面向智能教育场景的人机交互技术研究，包括多学科交叉视角下的智能教育框架设计、人机交互技术研究以及定性和定量分析。
- 认知神经科学：基于类人建模的视觉任务设计、环境构建和人机能力分析。
- 医学：结合人工智能的医学图像处理技术研究（例如，细胞分割与跟踪，低温电子显微镜图像去噪）。
- 心理学：面向焦虑、抑郁、强迫等心理维度的游戏化评测系统开发及智能心理评估技术研究。结合大语言模型（LLMs）和基础视觉模型（LVMs）的心理场景下视觉理解任务研究。

论文成果

专著 (Book)

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

已发表/接收 (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**, 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.
- 3 **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.
- 4 **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.
- 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)*, vol. 132, pp. 1659–1684, 2024.  DOI: 10.1007/s11263-023-01937-0.
- 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 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.
- 10 X. Li, X. Feng, **S. Hu**, M. Wu, D. Zhang, J. Zhang, and K. Huang, "Dtlm-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.
- 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 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.
- 13 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.
- 14 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.
- 15 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.
- 16 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.
- 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.

预印本 (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 **S. Hu***, X. Li*, X. Li, J. Zhang, Y. Wang, X. Zhao, and K. H. Cheong, "Can lvlms describe videos like humans? a five-in-one video annotations benchmark for better human-machine comparison", *arXiv preprint arXiv:2410.15270*, 2024.
- 3 Y. Ma*, **S. Hu***, X. Li, Y. Wang, S. Liu, and K. H. Cheong, "Students rather than experts: A new ai for education pipeline to model more human-like and personalised early adolescences", *arXiv preprint arXiv:2410.15701*, 2024.
- 4 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.
- 5 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.
- 6 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.

项目





框架

- 2018.03 - 2018.11  **Darknet-Cross: Light-weight Deep Learning Framework for Heterogeneous Computing**
 <https://github.com/huuuuusy/Darknet-Cross>
 Darknet-Cross支持多种平台（例如Android和Ubuntu）和多种GPU（例如Nvidia GTX1070和Adreno 630）的算法加速处理。
 该项目是我在香港大学硕士论文的一部分（论文答辩成绩：A+）。





平台（开发和维护）

- 2019.11 - 至今  **VideoCube / MGIT 平台**
 <http://videocube.aitestunion.com>
 VideoCube/MGIT 是两项发表在 IEEE TPAMI 2023 和 NeurIPS 2023 的研究的支持平台。
 截至 2024 年 9 月，该平台已获得来自全球 220 多个国家和地区的 440k+ 页面浏览、1.2k+ 下载量和 420+ 跟踪器提交。
- 2021.07 - 至今  **SOTVerse / VLTVerse 平台**
 <http://metaverse.aitestunion.com>
 SOTVerse 是一项发表在 IJCV 2024 的研究的支持平台。
 截至2024年9月，该平台已获得来自全球150多个国家和地区的126k+页面浏览。
- 2022.05 - 至今  **BioDrone 平台**
 <http://biodrone.aitestunion.com/>
 BioDrone 是一项发表在 IJCV 2024 的研究的支持平台。
 截至2024年9月，该平台已获得来自全球200多个国家和地区的17万+页面浏览。




平台（维护）

- 2020.07 - 至今  **GOT-10k 平台**
 <http://got-10k.aitestunion.com/>
 GOT-10k 是一项发表在 IEEE TPAMI 2021 的研究的支持平台。
 截至 2024 年 9 月，该平台已获得来自全球 290 多个国家和地区的 392 万+次页面浏览量、7.5k+次下载量和 21.5k+个跟踪器。



挑战赛（主办方）

- 2023.05 - 2023.11  **Hislopvision 挑战赛**
 <http://hislopvision.aitestunion.com/>
 Hislopvision Challenge 是第五届中国模式识别与计算机视觉学术会议第三届高速低功耗视觉理解挑战赛。
 参赛队伍包括来自清华大学、北京理工大学、吉林大学等科研单位的研究者。

挑战赛（参与者）

- 2021.01 - 2021.04  **Cell Tracking Challenge**
 <https://celltrackingchallenge.net/>
 该项目于2021年3月提交至Cell Tracking Challenge，并在Fluo-C2FL-MSc+数据集保持第二名，在Fluo-C2FL-Huh7数据集保持第三名（2023年10月统计）。

基金（参与者）

- 2023.12 - Now  **智能教育人机交互困境及对策研究**
 相关项目获得上海智能教育研究院与华东师范大学2023年智能教育博士生科研基金资助，处于研究中。

学术活动和服务

- 讲习班 ■ 31st IEEE International Conference on Image Processing (ICIP)
- 题目：An Evaluation Perspective in Visual Object Tracking: from Task Design to Benchmark Construction and Algorithm Analysis
 - 时间和地点：9:00-12:30, 27th October, 2024, Abu Dhabi, United Arab Emirates
- 27th International Conference on Pattern Recognition (ICPR)
- 题目：Visual Turing Test in Visual Object Tracking: A New Vision Intelligence Evaluation Technique based on Human-Machine Comparison
 - 时间和地点：14:30-18:00, 1st December, 2024, Kolkata, India
- 17th Asian Conference on Computer Vision (ACCV)
- 题目：From Machine-Machine Comparison to Human-Machine Comparison: Adapting Visual Turing Test in Visual Object Tracking
 - 时间和地点：9:00-12:00, 9th December, 2024, Hanoi, Vietnam
- 客座编辑 ■ 期刊：Electronics（特刊：Techniques and Applications of Multimodal Data Fusion）
- 副主编 ■ 期刊：Innovation and Emerging Technologies
- 审稿 ■ 会议：NeurIPS, ICML, ICLR, CVPR, ICCV, ECCV, AACL, IJCAI, ACMMM, AISTATS等
- 期刊：IEEE Transactions on Image Processing, SCIENCE CHINA Information Sciences, Scientific Reports, IEEE Access, Journal of Computational Science, Journal of Electronic Imaging, Digital Signal Processing等
- 会员 ■ 学会：电气电子工程师学会 (IEEE, No.97803543), 中国图象图形学会 (CSIG, No.E651129499M), 中国人工智能学会 (CAAI, No.E660120827A)

技能

- 语言 ■ 汉语普通话（母语）和英语。
- 编程 ■ Python, Java, Matlab, C, \LaTeX .
- 开发 ■ Android, Flask, SQLite.
- Linux ■ Shell, OS虚拟化。
- 其他 ■ 学术研究，团队协作能力，领导力，学术报告。

奖励及荣誉

- 2024 ■ 最佳论文提名，CVPR第三届视觉数据理解研讨会。
- 北京市优秀毕业生，北京市教育委员会（获奖率：5%）。
- 2023 ■ 国家奖学金，中华人民共和国教育部（获奖率：1%）。
- 攀登一等奖学金，中国科学院自动化研究所。
- 2022 ■ 优秀学生，中国科学院大学。
- 2017 ■ 优秀创新学生，北京理工大学。
- 2016 ■ 大学生奖学金，中国科学院。
- 优秀团员，北京理工大学。
- 2015 ■ 国家一等奖，全国大学生数学建模竞赛（获奖率：1%）。
- 数学建模竞赛校第一名，北京理工大学。
- 暑期社会实践优秀个人，北京理工大学。
- 暑期社会实践校二等奖，北京理工大学（队长）。
- 优秀学生干部，北京理工大学。
- 优秀团干部，北京理工大学。
- 优秀团支部，北京理工大学（团支书）。

奖励及荣誉 (continued)

■ 十佳团日活动，北京理工大学（团支书）。

2014 ■ 优秀学生，北京理工大学。

2013-2017 ■ 学业奖学金，北京理工大学。

协助指导学生

博士生 ■ 武美奇，2022.08-至今，中国科学院大学

■ 丰效坤，2023.04-至今，中国科学院自动化研究所

■ 马一平，2023.08-至今，华东师范大学

■ 张岱凌，2023.08-至今，中国科学院自动化研究所

■ 王怡沛，2024.08-至今，东南大学

■ 李旭宸，2024.08-至今，中国科学院自动化研究所

■ Hong-jui Shen，2024.10-至今，新加坡南洋理工大学

■ Kainan Li，2024.10-至今，新加坡南洋理工大学

■ Buyuan Zhu，2025.01-至今，新加坡南洋理工大学

硕士生 ■ 马一平，2022.05-2023.07，南京师范大学

■ 王怡沛，2022.08-2024.07，东南大学

■ 李旭钊，2024.07-至今，北京理工大学

■ 崔雨祺，2024.07-至今，北京科技大学

■ 徐攀曦，2024.09-至今，北京科技大学

■ Nguyen Khanh Truong，2025.02-至今，新加坡南洋理工大学

本科生 ■ 朱峻佑，2022.09-2023.07，中国科学院大学

■ 胡力杭，2022.09-2023.07，中国科学院大学

■ 张岱凌，2022.09-2023.07，东南大学

■ 李旭宸，2023.04-2024.07，北京邮电大学

■ 叶子，2024.09-至今，北京科技大学

■ Leying Tan，2024.09-至今，新加坡南洋理工大学

■ Jinlin Ma，2024.10-至今，新加坡南洋理工大学

■ Yuxiao Li，2024.10-至今，新加坡南洋理工大学

■ 孟肆烺，2024.12-至今，北京科技大学

■ 姜一卉，2025.01-至今，北京科技大学

推荐人

黄凯奇研究员和赵鑫教授分别是我的博士导师和合作导师，我非常荣幸和他们一起开展计算机视觉领域的研究。王卓立教授是我的硕士导师，我非常荣幸在他的指导下开展高性能计算研究。目前，我在南洋理工大学和Kanghao Cheng教授一起工作。

黄凯奇 研究员

研究员, IAPR会士, IEEE高级会员, 万人领军
中国科学院自动化研究所(CASIA)
中关村东路95号, 北京, 中国

✉ kqhuang@nlpr.ia.ac.cn

王卓立 教授

名誉教授

香港大学(HKU), 计算机科学系
薄扶林, 香港特别行政区, 中国

✉ clwang@cs.hku.hk

✉ choliwang@gmail.com

赵鑫 教授

教授, IEEE高级会员, 北京市杰青
北京科技大学(USTB), 计算机与通信工程学院
学院路30号, 北京, 中国

✉ xinzha@ustb.edu.cn

✉ xzhaopersonal@foxmail.com

Kanghao Cheng 教授

副教授, IEEE高级会员
物理与数学科学学院助理院长
科学学部助理院长

南洋理工大学(NTU)

南洋大道50号, 新加坡

✉ kanghao.cheong@ntu.edu.sg