

胡世宇，博士

📅 1995.10.29 ✉ hushiyu2019@ia.ac.cn ✉ hushiyu199510@gmail.com
🌐 hushiyu1995 🐙 huuuuusy 🌐 https://huuuuusy.github.io/

个人概述

- 教育背景优秀 本硕博均毕业于顶尖大学/科研机构，学位论文答辩成绩均为优秀。
荣获多项奖励及荣誉，博士阶段荣获国家奖学金（获奖率：1%）。
- 科研基础扎实 博士阶段共计发表论文10篇。其中一作论文4篇，包含国际顶级期刊IEEE TPAMI（CCF-A期刊，IF=23.6）、IJCV（CCF-A期刊，IF=19.5）、国际顶级会议NeurIPS（CCF-A会议）和《中国图象图形学报》（CCF-B期刊）。
负责搭建并维护的科研平台获得全球130+国家和地区28.7万次访问。
- 交流合作广泛 担任CVPR, AAAI, ACM MM等顶级会议的审稿人，将在ICIP 2024会议上举办讲习班（题目：An Evaluation Perspective in Visual Object Tracking: from Task Design to Benchmark Construction and Algorithm Analysis）。
自2022年9月起，发起并组织基于计算机视觉的跨领域研讨会（40+次，10+学校，20+人），讨论主题涉及计算机视觉、认知神经科学、人智交互等。
协助指导本硕博生近10人开展科研工作，研究方向涉及视觉目标跟踪、视觉语言跟踪、视觉图灵测试、人智交互技术等。

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 Fellow)
 - 论文题目：单目标跟踪中的智能评估技术研究
 - 论文答辩成绩：全优
- 2017.09 - 2019.06 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：王卓立 教授
 - 论文题目：NightRunner: Deep Learning for Autonomous Driving Cars after Dark
 - 论文答辩成绩：A+
- 2013.09 - 2017.06 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

研究基础 & 研究兴趣

- 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。
面向单目标跟踪算法的鲁棒性、泛化性、安全性等研究。
- 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。
以视觉语言跟踪为代理任务的长视频序列下人智交互模式探究。
- 视觉基准建模 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。
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。
- 智能评估技术 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。

研究基础 & 研究兴趣 (continued)

- AI4Science ■ 认知神经科学：基于类人建模思想的视觉任务设计、实验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。

研究经历


主要研究（负责：思路制定 & 方法设计 & 实验分析 & 论文撰写 & 平台开发）


- 2018.03 - 2018.11 ■ **Darknet-Cross: 适配于异构计算的轻量化深度学习框架**
- ✎ 高性能计算 | 异构计算 | 深度学习框架
 - ✎ Darknet-Cross是一个轻量化的深度学习框架，主要基于开源的深度学习算法库Darknet和yolov2_light进行二次开发，并通过交叉编译成功移植到手机端，实现利用移动GPU开展高效的算法推理。
 - ✎ 该框架支持多种平台（例如，Android和Ubuntu）和多种型号GPU（例如，Nvidia GTX1070和Adreno 630）下的算法加速处理。
 - ✓相关工作是我在香港大学硕士毕业论文的一部分（论文答辩成绩：A+）。
- 2019.11 - 2022.02 ■ **VideoCube: 大规模多维度全局实例跟踪任务智能评估平台**
- ✎ 视觉目标跟踪 | 大规模基准构建 | 智能评估技术
 - ✎ 从类人建模出发对单目标跟踪进行拓展，并提出全局实例跟踪(global instance tracking, GIT)任务，使其应用范围拓展至对抗场景中。
 - ✎ 提出基于影视学理论的视频叙事内容解耦框架，并基于此构建了一个包含746万视频帧的大规模多维度的全局实例跟踪任务智能评估平台VideoCube，该平台是目前规模最大的单目标跟踪基准。
 - ✎ 从人机对抗出发，首次在单目标跟踪任务中引入人类被试，旨在以人类为基准度量算法的视觉智能。
 - ✓相关工作于2022年2月被CCF-A期刊TPAMI接收 [A1]，截止2024年1月，平台获得全球130+国家和地区28.7万次访问，1000余次下载，380余次算法测试。
- 2023.03 - 2023.09 ■ **MGIT: 基于层级化语义框架的多模态全局实例跟踪任务基准**
- ✎ 视觉语言跟踪 | 长视频理解与推理 | 层级化语义信息
 - ✎ 在GIT任务和VideoCube基准的基础上，构建一个具有复杂视频叙事关系的多模态基准MGIT，旨在充分体现长视频所刻画的复杂时空关系和因果关系。
 - ✎ 结合人类认知的层级化结构，创新性地提出一种多粒度的语义信息标注策略，旨在为算法提供高质量的语义信息，并通过实验验证该策略的有效性。
 - ✎ 提出一套适配于多模态单目标跟踪任务的评测机制，并基于该评测机制对不同结构的算法进行详细的实验分析，从而深入挖掘算法的性能瓶颈。
 - ✓相关工作于2023年9月被CCF-A会议NeurIPS接收 [A2]。
- 2021.07 - 2023.09 ■ **SOTVerse: 用户自定义的单目标跟踪任务空间**
- ✎ 视觉目标跟踪 | 动态开放环境建模 | 视觉评估技术
 - ✎ 从高效利用研究资源开展评估的视角出发，提出3E范式作为任务分析框架，并从实验环境、评估系统和任务对象三个维度对单目标跟踪进行梳理。
 - ✎ 对代表性单目标跟踪数据集进行整合，将原本静态封闭的数据空间融合为一个包含1256万帧的动态开放数据空间SOTVerse，并提出子空间构建算法，旨在帮助研究者依据研究目标快速定位高挑战性序列形成专用的实验环境。
 - ✎ 提出全新的评测系统，使其能够适配于不同的评测需求和任务对象，并对23个代表性算法开展细粒度分析，有效挖掘出不同架构算法的性能瓶颈。
 - ✓相关工作于2023年9月被CCF-A期刊IJCV接收 [A3]。


研究经历 (continued)

2022.05 - 2023.10  **BioDrone: 基于仿生无人机的单目标跟踪基准和鲁棒视觉研究平台**

 视觉目标跟踪 | 无人机视角跟踪 | 视觉鲁棒性研究


 构建了首个基于仿生扑翼无人机的单目标跟踪基准数据集BioDrone，扑翼无人机特有的空气动力学结构会导致镜头出现剧烈抖动，为鲁棒视觉研究提供一个高质量的环境。


 在BioDrone上复现并测试了20种代表性单目标跟踪算法，并针对算法的鲁棒性瓶颈开展细致分析。


 设计了一个名为UAV-KT的基线算法，通过实验分析验证了该方法可以有效提取视觉特征，并在无人机场景的挑战因素下依旧保持鲁棒性。


✓相关工作于2022年5月-10月作为竞赛数据支撑举办第三届高速低功耗视觉理解挑战赛（在第五届中国模式识别与计算机视觉大会上举办），并于2023年10月被CCF-A期刊IJCV接收 [A4]。


2022.04 - 至今  **基于视觉图灵的单目标跟踪智能评估技术**

 视觉目标跟踪 | 智能评估技术 | AI4Science

 结合图灵测试思想，提出“视觉图灵”评估范式——以人类的视觉能力为基准，全面度量算法的视觉智能。

 构建了一个公平对比人机动态视觉能力的实验环境，该环境可以涵盖任务对象在执行动态视觉任务时需要调用的感知和认知能力。


 挑选合适的任务对象开展人机动态视觉能力测试，包括了20个代表性算法和15名受试者，并基于严格的流程开展实验。


 设计出一套人机通用的多粒度动态视觉任务评估框架，基于中心点距离从三个粒度（帧级、序列级和组级）开展评估和分析。


✓相关工作于2021年和2023年发表中文综述2篇 [A5], [A6]，实验内容及主要结论正在完善 [A7]，准备投稿至cell patterns期刊 [P1]。

独立开发（负责：平台维护 & 升级）

2020.07 - 2024.01  **GOT-10k: 大规模广覆盖单目标视觉跟踪评估平台**

 视觉目标跟踪 | 智能评估技术 | 平台开发维护


 GOT-10k是用来度量算法在未知目标类别和运动模式上泛化性的评估平台。该平台提供了1万段视频序列，包含563类目标类别和87种运动模式，总计150万个矩形框标注。


 和已有的单目标跟踪基准相比，GOT-10k在目标类别的丰富程度上要远超过现有工作，其提出的训练-测试类别不重合的评测思路也成为单目标跟踪评估评测领域的研究规范。


✓相关工作于2019年11月被CCF-A期刊TPAMI接收；截止2024年1月，本平台获得全球150+国家和地区282万次访问，6千余次下载，1.7万余次算法测试。和维护初期相比，有18倍的用户增长。

合作研究（负责：思路讨论 & 实验分析 & 论文修改）

2019.05 - 2019.10  **无色卡场景下的肤色检测系统研究**

 颜色恒常性 | 肤色检测算法 | 环境光照估计

 在18种环境光照条件和4部手机的参数组合下，对110位用户进行肤色数据采集，并以德国CK公司的MPA9肤色检测仪的测试结果作为用户肤色的真值，构建包含7920张图片的肤色数据集。

 基于椭圆皮肤模型得到图像中的关键皮肤区域，并利用开源颜色恒常性模型FC⁴对环境光照进行恢复，最后利用SVR回归计算用户的肤色检测结果。

✓相关工作已成功部署在华为手机官方应用《镜子》的AI测肤功能中。

研究经历 (continued)

2020.11 - 2021.03 ■ 基于深度学习方法的细胞跟踪项目研究

✎ 医学图像处理 | AI4Science | 细胞分割与跟踪

✎ 参考基于检测的跟踪范式，将每一帧的细胞分割结果进行整合，并利用孪生神经网络进行帧间细胞位置匹配，从而完成跟踪过程。

✓ 相关工作于2021年3月提交至ISBI细胞跟踪挑战赛，在2023年10月，依旧保持Fluo-C2FL-MSD+数据集第二名和Fluo-C2FL-Huh7数据集第三名。

2023.08 - 2023.10 ■ 面向单粒子低温电磁场图像的去噪与恢复算法

✎ 医学图像处理 | AI4Science | 扩散模型

✎ 采样自低温电子显微镜的信息通过图像重建可以展示接近原子尺度的生物分子信息，但重建过程需要克服低信噪比和复杂的噪声结构。本工作提出一种带有后处理模块的扩散模型，以有效地去噪和恢复单粒子低温电镜图像。

✎ 在模拟数据集和真实数据集上的实验结果验证了方法的有效性。

✓ 相关工作于2023年12月被CCF-B会议ICASSP接收 [A8]。

2023.01 - 2024.01 ■ 面向相似物体干扰挑战的单目标跟踪任务研究

✎ 视觉目标跟踪 | 相似物体干扰 | 高效数据利用

✎ 基于主流跟踪算法的运行原理及失败案例分析，重新定义了算法视角下的相似物体干扰挑战，并重点分析人机在面对相似物体干扰时的认知差异。

✎ 提出一种数据挖掘算法，在无人工干预的情况下从代表性单目标跟踪数据集中挖掘出存在相似物体干扰的序列，形成TrackingSOI数据集。

✎ 提出一种有效应对相似物体干扰的算法TransKT，其可以将目标与表现信息相似的候选物进行区分，从而获得鲁棒的视觉目标跟踪能力。

✓ 相关工作的初始版于2023年11月被EI会议CSAI接收为oral论文 [A9]，完整版于2024年1月投稿至CCF-B期刊TCSVT，处于审稿中 [R1]。

2022.10 - 至今 ■ AWCV-100k: 面向真实应用场景的无约束空中手写基准

✎ 空中手写技术 | 视觉基准构建 | 人机交互

✎ 基于空中手写任务构建了一个大规模高质量的空中汉字手写数据集AWCV-100k，旨在为人机交互研究提供一种更加自然通用的实验环境。

✎ AWCV-100k包含880万视频帧，涉及多种环境、光照组合，全面覆盖中文一级字符库GB2312-80的3755个字符，是目前规模最大、覆盖范围最全面的中文空中手写视频数据集。

✎ 提出了基于视频的空中手写字符识别算法VCTec，该基线算法可以从稀疏的视觉线索中提取指尖特征，并使用一个时空序列模块进行分析。

✎ 在AWCV-100k数据集上复现并测试了代表性的空中手写算法和基线算法VCTec，实验结果验证了VCTec的鲁棒性和有效性。

✓ 相关工作于2023年12月投稿至CCF-B期刊TCSVT，处于审稿中 [R2]；后续工作将基于该基准开展人机交互技术研究。

2023.02 - 至今 ■ 基于电子沙盘的智能心理评测系统

✎ 心理评测系统 | 游戏化评测 | AI4Science

✎ 针对传统心理评测窥视感强、被试缺乏内在动机等缺陷，构建了基于电子沙盘的智能心理评测系统，可以对被试开展焦虑、抑郁、强迫等多个心理维度的评测和分析。

✎ 和传统的量表评测和游戏化评测相比，基于电子沙盘的智能心理评测系统结合了人工智能技术，可以挖掘游戏内动态过程数据中耦合的心理测量证据，能够进一步提升测量真实性和可解释性。

✎ 在教育场景和公共安全场景招募被试，并开展信效度验证和个案分析，实验结果证实了电子沙盘的有效性。

✓ 相关工作所采用的沙盘主题识别模型于2023年9月被CCF-C会议PRCV接收 [A10]；作为一个三学科交叉的系统性工作（心理学、游戏设计、人工智能），其技术路线发展脉络和技术框架将分别以综述和技术论文的形式投稿至中国顶级心理学期刊 [P2], [R3]。

研究经历 (continued)

2023.07 - 至今 ■ 基于类人记忆机制建模的鲁棒视觉语言跟踪算法

✎ 视觉语言跟踪 | 类人记忆建模 | 自适应提示

✎ 针对视觉语言跟踪算法在长视频中鲁棒性不足的情况，从类人记忆建模角度出发，设计了一个鲁棒的多模态跟踪算法MemTrack。

✎ MemTrack受认知领域的互补学习理论的启发，包含记忆存储、交互和引导三个模块，旨在模拟人类记忆在海马体和新皮层之间的复杂调制过程。

✎ 在多个代表性视觉语言跟踪基准上的实验验证了MemTrack的有效性。

✓ 相关工作于2024年1月投稿至CCF-A会议IJCAI，处于审稿中 [R4]。

论文成果

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, CCF-A Journal, IF=23.6)*, 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, CCF-A Conference, Poster)*, 2023.
- 3 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 Journal, IF=19.5)*, pp. 1–59, 2023. 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, CCF-A Journal, IF=19.5)*, 2023. 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 (单目标跟踪中的视觉智能评估技术综述)", *Journal of Images and Graphics (《中国图象图形学报》, CCF-B Chinese Journal)*, 2023.
- 6 K. Huang, X. Zhao, Q. Li, and S. Hu, "Visual turing: The next development of computer vision in the view of human-computer gaming (视觉图灵：从人机对抗看计算机视觉下一步发展)", *Journal of Graphics (《图学学报》, CCF-C Chinese Journal)*, vol. 42, no. 3, p. 339, 2021. DOI: 10.11996/JG.j.2095-302X.2021030339.
- 7 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.
- 8 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, CCF-B Conference, Poster)*, 2024.
- 9 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, EI Conference, Oral)*, 2023.
- 10 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, CCF-C Conference, Poster)*, 2023, pp. 241–252. DOI: 10.1007/978-981-99-8462-6_20.
- 11 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 Journal, IF=6)*, vol. 501, pp. 270–284, 2022. DOI: 10.1016/j.neucom.2022.06.027.

在审稿 (Review)

Last update: 2024-2-4

- 1 Y. Wang, **S. Hu**, D. Zhang, M. Wu, T. Yao, Y. Wang, L. Chen, and X. Zhao, "Target or distractor? rethinking similar object interference in single object tracking", *IEEE Transactions on Circuits and Systems for Video Technology (TCSVT, CCF-B Journal, IF=8.4, Under Review)*, 2024.
- 2 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 Journal, IF=8.4, Under Review)*, 2024.
- 3 Y. Ren, X. Feng, **S. Hu**, et al., "Intelligent psychological assessment with sandplay based on evidence-centered design theory (基于证据中心设计理论的智能心理沙盘测评系统)", *Acta Psychologica Sinica (《心理学报》, CSSCI Journal, Top-1 Psychological Journal in China, Under Review)*, 2024.
- 4 X. Feng, X. Li, **S. Hu**, D. Zhang, M. Wu, X. Chen, and K. Huang, "Remembering target more like humans: A robust visual-language tracker with adaptive prompts", *the 33rd International Joint Conference on Artificial Intelligence (IJCAI, CCF-A Conference, Under Review)*, 2024.

待投稿 (Preparation)

- 1 **S. Hu**, J. Zhu, Y. Wang, X. Zhao, and K. Huang, "Vt³: A visual tracking turing test of human-machine dynamic vision ability", *Cell Patterns (In Preparation)*, 2024.
- 2 K. Huang, Y. Kang, C. Yan, **S. Hu**, et al., "Intelligent game-based psychological assessment(基于游戏的智能化心理测评系统)", *Chinese Mental Health Journal (《中国心理卫生杂志》, CSSCI Journal, In Preparation)*, 2024.

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

- 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 ■ 学业奖学金，北京理工大学。
■ IE入学奖学金，香港中文大学。
■ 优秀创新学生，北京理工大学。
- 2016 ■ 学业奖学金，北京理工大学。
■ 大学生奖学金，中国科学院。
■ 优秀团员，北京理工大学。
- 2015 ■ 学业奖学金，北京理工大学。
■ 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）。

奖励及荣誉 (continued)

- 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）。
- 十佳团日活动，北京理工大学（团支书）。
- 2014 ■ 学业奖学金，北京理工大学。
- 优秀学生，北京理工大学。
- 2013 ■ 学业奖学金，北京理工大学。
- 2012 ■ 国家二等奖，第27届全国青少年科技创新竞赛，中国科协。
- 2010 ■ 三好学生，郑州市，河南省。

推荐人

黄凯奇 研究员
研究员，IAPR会士，IEEE高级会员
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

赵鑫 教授
教授
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
✉ xinzhao@ustb.edu.cn
✉ xzhaopersonal@foxmail.com

王卓立 教授
名誉教授（已退休）
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