



北京信息科学与技术国家研究中心

Beijing National Research Center for Information Science and Technology

简 报

办公室编印

2026 年 6 月刊

2026 年 6 月 30 日

本期导读

科研动态

- 清华大学与中国电信携手，实现 6G 算网融合关键技术突破
- 吴及课题组两篇论文被 MICCAI 2026 录用

交流合作

- 北京信息科学与技术国家研究中心系列交叉论坛（第 116 期）举办
- 北京信息科学与技术国家研究中心系列交叉论坛（第 117 期）举办

重点成果介绍

- 信息国家研究中心团队研制出多响应类视网膜传感器

◆ 科研动态

清华大学与中国电信携手，实现 6G 算网融合关键技术突破

近日，清华大学信息国家研究中心智慧天网交叉创新群体与中国电信研究院依托中国电信云网融合中试平台，联合开展面向算网融合感知调度技术的前瞻理论研究与创新平台试验，实现中断敏感业务的用户无感平滑迁移，业务资源开销最高降幅达 68.22%，为 6G 低时延通算融合业务提供全场景适配的算力服务支撑和连续性保障。



面向无人机动态目标跟踪场景，针对低空链路时变不稳、业务中断频发、资源调度适配性不足等痛点，双方联合提出 MSO 动态依赖感知调度算法，可自适应匹配低空信道动态变化特性，实现算网资源协同优化部署；针对复杂计算任务与多维算力资源难解耦的问题，双方共同研制端网云算力资源感知仿真平台，创新提出基于超图模型的多维资源表征方案，支持多种资源状态感知模式，实现感知效率与系统开销的最优平衡。依托云网融合中试平台的算力及融合组网验证能力，构建端网云算网融合试验环境，完成上述算法及方案验证。试验数据表明，上述方案可实现网内数十毫秒级算力资源感知，结合先迁后断机制与动态数据同步能力，可保障算网上下文平滑无感知迁移，为 6G UIN（全域智慧）- ADN（AI 定义网络）提供坚实的算网融合底座。相关研究成果已发表于 IEEE Internet of Things Journal 等 SCI 期刊，作为 ADN 技术成果在 2026 年世界移动通信大会(MWC)上海展公开展示。

未来，信息国家研究中心智慧天网交叉创新群体将继续聚焦下一代移动信息网络服务的痛点难点问题，以业务需求为导向，联合中国电信开展以 6G 算网融合技术为代表的新技术研究验证与创新试验，以前沿技术创新推动移动网络快速发展。

吴及课题组两篇论文被 MICCAI 2026 录用

近日，信息国家研究中心数基生命系统交叉创新群体吴及教授课题组论文“Lesion-Centric Structured Report Generation for 3D CT with Multimodal Large Language Models via Hierarchical Evidence Tokens”“Localization-Grounded Supervision: Revisiting Vanilla SFT of Large Vision-Language Models for Medical Image Analysis”被 MICCAI 2026 录用。MICCAI (Medical



Image Computing and Computer Assisted Intervention) 是由国际医学图像计算和计算机辅助干预协会举办的综合性学术会议, 是医学图像分析领域国际公认的最具影响力的学术会议。

论文 “Lesion-Centric Structured Report Generation for 3D CT with Multimodal Large Language Models via Hierarchical Evidence Tokens” (基于多模态大语言模型与分层证据 Token 的 3D CT 病灶中心结构化报告生成), 针对医学影像结构化报告生成中存在的描述不规范、信息遗漏及推理不充分等问题, 我们提出一种基于分层证据 Token 的病灶中心式框架。该方法首先将 3D CT 切片与病灶 mask 输入视觉编码器, 结合全局证据槽查询向量提取多尺度视觉特征; 随后将全局证据词元与细粒度图像词元共同注入大语言模型, 在注意力约束机制引导下完成跨模态对齐与推理学习。实验结果表明, 注意力约束与全局证据槽的联合策略在 LIDC-IDRI 数据集上取得最佳性能, 在不同训练比例下均显著优于基线方法; 同时, 全局证据槽的置零或打乱显著降低了模型效果, 证明其携带了真实、有效的任务相关信息。该研究为医学影像大模型提供了更稳定、可解释且高效的结构化报告生成新范式。

论文 “Localization-Grounded Supervision: Revisiting Vanilla SFT of Large Vision-Language Models for Medical Image Analysis” (定位锚定的监督: 重新审视用于医学图像分析的大型视觉语言模型的常规 SFT), 针对多模态大模型在医学影像分析中常用的监督微调 (SFT) 方法难以同时实现多模态对齐和高效任务学习的问题, 我们提出定位引导的多模态监督微调 (LGS) 方法。该方法借助预训练获得的 2D 空间对齐能力, 利用 2D 和 3D 定位数据在多任务中同步增强空间与语义对齐能力, 并结合 DePass-VL 计算不同区域对回答的贡献度。实验表明, LGS 在不同架构与参数规模下均稳定优于传统 SFT, 显著提升 2D 和 3D 任务性能与注意力定位的准确性, 为医学影像大模型的高效监督微调提供了新思路。

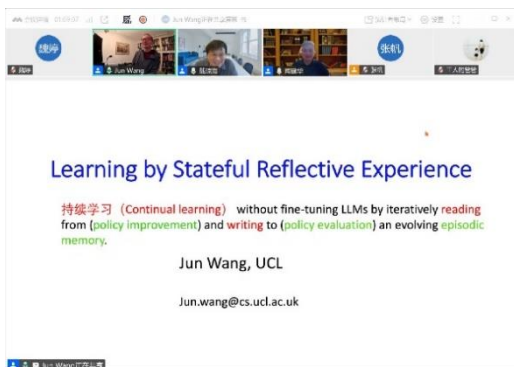
◆ 交流合作

北京信息科学与技术国家研究中心系列交叉论坛 (第 116 期) 举办

6 月 4 日晚, 北京信息科学与技术国家研究中心系列交叉论坛 (第 116 期) 通过线上会议和直播的形式举办, 本次论坛邀请了伦敦大学学院计算机科学系教授汪军作题为 “Learning by Stateful Reflective Memory” 的报告。论坛由清华大学信息学院院长、信息国家研究中心主任戴琼海院士和信息国家研究中心副主任陶建华教授共同主持。信息国家研究中心党政联席会成员、群体负责人以及校内外师生等 80 余人通过腾讯会议在线参加论坛, 累计约 23 万人次通过直播、

新浪、百度、视频号等直播平台在线观看。

报告中,汪军聚焦大型语言模型(LLM)智能体持续经验学习难题,以反思机制为核心:智能体依托历史经验优化后续决策,全程无需改动模型参数。研究借鉴基于案例推理与 Memento 系列框架思路,把智能体学习定义为记忆驱动过程,通过留存交互轨迹、典型案例与复用技能,在全新任务场景调取历史经验,优化智能体决策水平。为形式化这一思想,提出了 Stateful Reflective Decision Process(SRDP,有状态的反思决策过程)。该模型依靠动态迭代的记忆库完成读写两类核心操作:写入环节留存交互数据、完成策略评估;读取环节检索匹配经验、落地策略优化。研究依托检索增强型策略迭代,打通反思学习与强化学习的融合路径,从理论层面验证:伴随记忆样本持续扩充、覆盖更多状态空间,智能体策略可逐步收敛至最优。该成果搭建起记忆型智能体的完备理论体系,有效支撑智能体落地后动态适配复杂环境。报告同步介绍了自研 Memento-Skills 智能体系统,该工具可兼容各类商用大模型工程落地场景,具备良好工程复用价值。

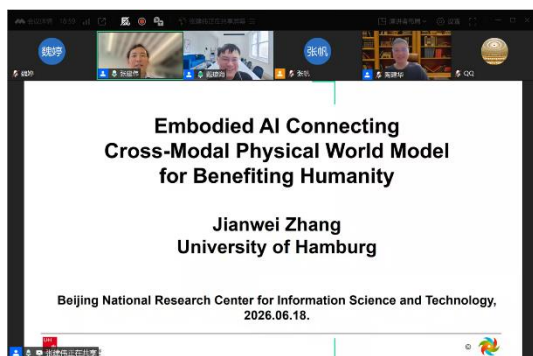


汪军作报告

问答环节,汪军就 Memento 实验对硬件、GPU 与内存需求;若反思能力依托原生 LLM 而非外部干预,如何管控反思、记忆压缩里的幻觉、错归因等问题,防止特例经验被固化为可用技能;以及模型是否原生具备主动遗忘冗余记忆、规避上下文超限的能力等问题进行充分解读。

北京信息科学与技术国家研究中心系列交叉论坛(第 117 期)举办

6 月 18 日晚,北京信息科学与技术国家研究中心系列交叉论坛(第 117 期)通过线上会议和直播的形式举办,本次论坛邀请了中国工程院外籍院士、德国工程院院士、清华大学杰出访问教授张建伟作题为“具身智能连接多模态物理世界、造福人类”的报告。论坛由清华大学信息学院院长、信息国家研究中心主任戴琼海院士和信息国家研究中心副主任陶建华教授共同主持。信息国家研究中心党政联席会成员、群体负责人以及校内外师生等 270 余人通过腾讯会议在线参加论坛,累计约 37 万人次通过上直播、新浪、百度、视频号等直播平台在线观看。



张建伟作报告

具身智能作为人工智能重要分支,聚焦智能体与物理环境交互,融合数据驱

动、感控一体、变体结构、多主体协同等技术，破解复杂物理场景自主化难题。当前，多模态大模型与具身智能机器人深度融合，为多智能体打造“智慧大脑与小脑”，实现多模态自感知、自认知、自决策、自执行与自学习，推动思考、感知、行动有机融合，可在真实环境自主完成任务。该技术将强力赋能智能机器人、工业制造、智能医疗、智慧农业等领域，为体力工作替代提供技术支撑。目前行业仍面临非结构化环境建模、灵巧操作、自然交互、高质量多模态数据生成等挑战，融合世界机理模型与大模型训练的新方法，已成为核心科研方向。未来，行业将以人类需求为导向，深化全球合作与技术交叉，完善标准体系与伦理治理，加快通用智能服务机器人在民用刚需领域落地，以技术创新服务社会发展，创造美好生活。

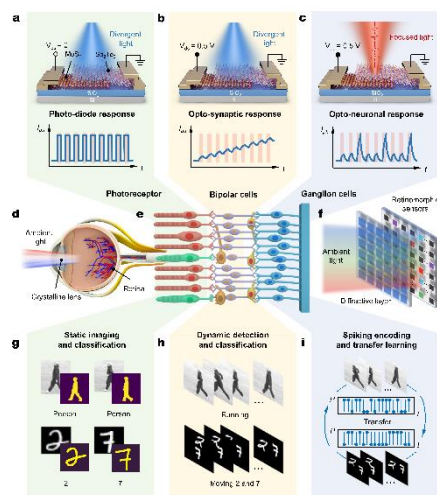
问答环节，张建伟就在大模型与具身机器人结合普遍依赖海量数据训练的现状下，如何在非结构化未知真实环境中减少海量标注数据依赖、实现高效自主学习等问题进行充分解读。

◆ 重点成果介绍

信息国家研究中心团队研制出多响应类视网膜传感器

近日，清华大学信息国家研究中心光电智能技术交叉创新群体在类脑光电感知领域取得重要突破，成功研制出一种面向可重构光电计算的多响应类视网膜传感器（MRRS），为低功耗、高通量智能视觉系统提供了全新器件与架构方案。

研究团队创新性提出并构建了基于碲化锑（ Sb_2Te_3 ）/二硫化钼（ MoS_2 ）横向异质结的两端器件，在单一器件内集成光电二极管、光电突触、光电神经元三重响应，可通过调节源漏偏置电压与入射光聚焦程度，实现三种工作模式的原位可逆切换，硬件结构极简、功能高度集成。MRRS 传感器的设计借鉴了生物视网膜中光感受器、双极细胞和神经节细胞协同处理视觉信息的机制：零偏压下，呈现快速光电二极管响应，可用于静态图像感知；低偏压和发散光照条件下，产生具备保持与累积特性的光电突触响应，可模拟短时/长时可塑性以及学习—遗忘—恢复等行为；低偏压和聚焦脉冲光刺激下，展现生物神经元典型的漏电积分—发放特性，完成动态光信号脉冲频率编码，适配复杂动态场景处理需求。



MRRS 传感器多响应机制及其光电计算任务演示



实验结果表明, MRRS 传感器在自供电状态下具备低暗电流与微秒级快速响应能力。光电突触模式可依据光脉冲宽度、功率密度和频率等光脉冲参数灵活调控突触电流, 实现从短时可塑性到长时可塑性的平滑转变。光电神经元模式基于界面陷阱介导的势垒调制机制, 当光电流达到阈值后触发神经元式发放并自动复位, 省去传统复杂电学转换电路, 单次脉冲能耗降至皮焦 (pJ) 量级, 大幅提升能效表现。

在此基础上, 研究团队进一步构建了 32×32 规模的 MRRS 类视网膜传感阵列, 并与前端空间衍射层融合, 构建“前端光场调制+传感器内并行计算”的可重构光电计算架构。输入图像与视频经光场编码、衍射层特征预处理后, 由传感阵列直接完成静态识别、动态记忆、脉冲编码等核心功能, 实现感存算一体化, 为端侧智能视觉提供创新技术路径。系统还可实现基于脉冲编码的跨任务迁移学习, 在不同视觉数据集间快速复用模型特征, 减少训练开销, 显著提升动态场景下的泛化能力与任务适配效率。

未来, 结合微型化衍射光学元件、片上集成封装和宽谱响应材料体系, 该类多响应类视网膜传感器有望应用于自动驾驶、卫星遥感、具身机器人和智能终端等高通量、低功耗视觉计算场景。

研究成果以“面向可重构光电计算的多响应类视网膜传感器”(Multi-responsive retinomorph sensor for reconfigurable optoelectronic computing) 为题, 于 6 月 5 日发表于《自然·传感》(Nature Sensors)。

清华大学信息国家研究中心为论文第一完成单位。信息国家研究中心光电智能技术交叉创新群体首席科学家、电子工程系教授方璐, 副研究员王钰言, 超感知与人机融合交叉创新群体首席科学家、集成电路学院教授任天令为论文共同通讯作者; 集成电路学院 2023 级硕士生王昱豪、电子系已出站博士后程远(现为上海交通大学副教授)、集成电路学院 2023 级博士生潘炯为论文共同第一作者。

研究得到国家科技重大专项、国家自然科学基金、北京市科技新星计划、清华大学笃实计划、科学探索奖等项目支持。

报: 清华大学党政领导、信息国家研究中心建设运行管理委员会成员、信息国家研究中心学术委员会成员、信息学院党政联席会成员、信息国家研究中心党政联席会成员

送: 相关院系、部处负责人

发: 信息国家研究中心各部门负责人

编辑: 李琳

审核: 王钰言

签发: 陈文华

联系电话: 62792099

E-mail: bnrict@tsinghua.edu.cn