

IGTA2024

2024中关村论坛系列活动——

第十九届图像图形技术与应用学术会议

2024年8月16日-18日

■ 目录

➤ 大会介绍	01
➤ 大会主席	02
➤ 组委会	03
➤ 日程	04
➤ 特邀报告	07
➤ 机器具身交互智能论坛	09
➤ 人体动作理解与生成前沿论坛	13
➤ 数字人与内容生成论坛	17
➤ 工业视觉技术与应用论坛	21
➤ 医学影像智能分析与手术导航论坛	25
➤ 多模态大模型论坛	29
➤ 技术报告	33
➤ 第五届数学建模教学与培训研讨会 探索数学世界及数学建模	34
➤ 论文列表	35

■ 大会介绍

诚挚邀请莅临2024年8月16日至18日在北京西郊宾馆盛大召开的2024中关村论坛系列活动——第十九届图像图形技术与应用国际学术会议（IGTA2024）。本次盛会由北京图象图形学学会主办，北京师范大学及中关村国际会展运营管理有限公司联合承办，北京科学技术协会支持，共同打造一场图像图形领域的学术盛宴。

本届会议以“AIGC与图像图形”为核心议题，旨在把握时代脉搏，探索人工智能生成内容（AIGC）与图像图形技术的深度融合之道。学会致力于在原有坚实基础基础上，进一步扩大会议的影响力与参与规模，通过严格的论文筛选机制，确保会议学术水平的持续提升，为本领域的科研工作者提供一个思想碰撞、灵感迸发的交流平台。

本届会议日程亮点纷呈，不仅涵盖了院士级别的特邀报告，深度剖析行业前沿趋势与最新研究成果，还精心设计了多场专题研讨会、技术展示及圆桌对话，全方位展现图像图形技术的最新进展与应用实践。此外，北京图象图形学学会将持续深化与国际知名出版机构Springer的合作关系，并携手《中国图象图形学报》、《虚拟现实与智能硬件》、《Visual Intelligence（视觉智能）》等权威期刊，共同出版高质量会议论文集，旨在将每一份创新思考与研究成果转化为推动行业发展的动力源泉，进一步提升学会的学术声誉与国际地位。

作为北京图象图形学学会的标志性活动，IGTA系列会议已历经十八载春秋，不仅荣获“北京市十佳影响力学术会议”殊荣，更见证了无数科研突破与人才成长的辉煌时刻。参会人员将在这一平台上面对面交流，共谋图像图形领域的未来发展蓝图。我们坚信，通过此次大会的深入交流与探讨，将有力推动图像图形技术的学术进步，促进产学研各界的深度融合与协同创新，为参会者拓宽学术视野、理清研究思路、提升科研能力提供宝贵契机。

让我们相聚北京，共襄盛举，携手开启图像图形技术与应用的新篇章！

■ 大会主席



王涌天教授

北京理工大学杰出教授，长江、杰青，光电学院和计算机学院博导，北京市混合现实与新型显示工程技术研究中心主任。兼任科技部有关国家重点研发专项总体专家组成员，全国信息技术标准化技术委员会委员，北京图象图形学学会理事长，中国图象图形学学会副理事长，中国光学工程学会常务理事，中国计算机学会理事，中国光学学会监事，国际标准组织（ISO）有关标准制定专家组成员，国际工程光学学会（SPIE）、美国光学学会（OSA）和英国工程技术学会（IET）Fellow等职。

长期在技术光学、虚拟现实和增强现实领域从事教学和科研工作，主要研究方向包括光学系统设计和CAD、新型三维显示、虚拟现实和增强现实、医学图像处理等方面。

发表专著4部、SCI论文320余篇，授权国内外发明专利200余项，主持制定虚拟现实和增强现实领域首批国家标准6项。有关成果得到广泛应用，获得国家技术发明奖和国家科技进步奖各1项，省部级和国家一级学会/协会科技奖励10余项。



黄华教授

北京师范大学教授，人工智能学院院长，2014年国家杰出青年基金获得者，第十四届中国青年科技奖获得者，入选第三批万人计划科技创新领军人才。

主要从事图像/视频处理、计算摄像学、计算机图形学方面的研究工作。先后主持国家自然科学基金重点项目、国家重点研发计划项目等。近年来，发表IEEE Trans论文50余篇，授权国家发明专利60余项，部分成果在国防、工业、互联网等领域得到应用。

兼任北京图象图形学学会副理事长，“新一代人工智能”科技创新2030重大项目管理专家组成员，JKW XX智能专家组副组长；中国计算机学会常务理事、教育工作委员会副主任；中国图象图形学学会常务理事、多媒体专业委员会主任；中国自动化学会理事、副秘书长；IEEE Transactions on Intelligent Transportation Systems AE；The Visual Computer AE；《计算机辅助设计与图形学学报》副主编；《电子与信息学报》副主编。

■ 组委会

组织委员会主席

- 刘 越 北京理工大学
袁晓如 北京大学
董 晶 中国科学院自动化研究所
张 曼 北京邮电大学

程序委员会主席

- 殷绪成 北京科技大学
毋立芳 北京工业大学
赫 然 中国科学院自动化研究所
马 楠 北京工业大学
杨 健 北京理工大学
马占宇 北京邮电大学
武仲科 北京师范大学
宋维涛 北京理工大学

论坛委员会主席

- 彭宇新 北京大学
李 亮 中国科学院计算技术研究所
徐婧林 北京科技大学
于 涛 清华大学

宣传与展览委员会主席

- 王醒策 北京师范大学
陆 峰 北京航空航天大学
韩向娣 中国图象图形学学报

网络推广委员会主席

- 田 运 北京师范大学
张浩鹏 北京航空航天大学



■ 日程

8月16日 北京西郊宾馆 一号楼		
	第五会议室	第六会议室
13:30-17:00	机器具身交互智能论坛 论坛主席：北京工业大学马楠教授 清华大学刘永进教授	人体动作理解与生成前沿论坛 论坛主席：北京大学彭宇新教授 北京科技大学徐婧林副教授
13:30-14:00	讲者：北京大学查红彬教授 题目：具身智能—身体、行为与环境的紧耦合	讲者：中国科学院自动化研究所徐常胜研究员 题目：不确定性感知的弱监督人体动作定位研究
14:00-14:30	讲者：国防科技大学徐凯教授 题目：融合三维感知与多模态大模型的具身智能	讲者：清华大学刘华斌教授 题目：人体无标记运动捕捉与运动生成
14:30-15:00	讲者：同济大学高炳钊教授 题目：AI大模型驱动的自动驾驶算法及车端具身实现	讲者：南京大学王利民教授 题目：多模态视频理解基础模型Intern Video
15:00-15:30	讲者：北京云迹科技股份有限公司支涛CEO 题目：AI+具身智能 未来竞争新势力	讲者：浙江大学周晓巍研究员 题目：基于神经表达的体积视频技术
15:30-16:00	讲者：清华大学高跃副教授 题目：超图计算	讲者：武汉大学涂志刚研究员 题目：视频人体动作行为识别与生成
16:00-16:30	讲者：清华大学赵思成副研究员 题目：开放环境领域自适应	讲者：北京科技大学徐婧林副教授 题目：面向行为理解的细粒度运动分析
16:30-17:00	圆桌论坛	圆桌论坛
8月17日 北京西郊宾馆 一号楼 银杏大厅		
09:00-09:15	致辞： 北京图象图形学学会理事长、北京理工大学王涌天教授 北京图象图形学学会副理事长、北京师范大学黄华教授	
09:15-09:30	表彰： BSIG优博、优硕、最美女科技工作者、第13届APMCM亚太地区大学生数学建模竞赛	
09:30-10:30	特邀报告一： 首都医科大学附属北京友谊医院王振常院士 医学影像的现状与创新	
10:30-10:45	茶歇+合影	
10:45-11:45	特邀报告二： 美国雪城大学徐瑞鸽教授 Rebecca Ruige Xu Ethical Dimensions and Creative Collaborations of AI in Art and Design	
11:45-12:00	北京超算百万机时算力基金启动仪式 技术报告： 北京超算解决方案架构师王小磊 算力服务加速数字人与内容生成技术创新发展	

12:00-13:30	午餐 五号楼 赏园餐厅	
8月17日 北京西郊宾馆 一号楼		
	第五会议室	第六会议室
13:30-17:00	数字人与内容生成论坛 论坛主席：清华大学刘焯斌教授 北京师范大学张鸿文副教授	工业视觉技术与应用论坛 论坛主席：华中科技大学周瑜副研究员 中国科学院自动化研究所陶显副研究员
13:30-14:00	讲者：北京大学彭宇新教授 题目：多维感知驱动的AIGC	讲者：清华大学丁贵广教授 题目：模型推理优化技术研究
14:00-14:30	讲者：中国科学院大学吕科教授 题目：高保真数字人建模与驱动关键技术与应用	讲者：北京邮电大学杨辉华教授 题目：工业制造表面缺陷视觉检测技术及应用
14:30-15:00	讲者：清华大学鲁继文教授 题目：多模态视觉内容生成	讲者：华南理工大学徐雪妙教授 题目：基于知识引导的场景智能理解和生成技术
15:00-15:30	讲者：哈尔滨工业大学张盛平教授 题目：超写实虚拟数字人驱动技术	讲者：东北大学宋克臣副教授 题目：工业钢管缺陷智能检测方法与应用
15:30-16:00	讲者：北京大学连宙辉副教授 题目：个性化数字人的便捷构建	讲者：中国科学院自动化研究所 赵朝阳副研究员 题目：大模型时代下的工业视觉及行业应用
16:00-16:30	讲者：中国科学院自动化研究所 朱翔昱副研究员 题目：精细三维人脸感知	讲者：华东师范大学张志忠副教授 题目：受限工业场景下的视觉智能检测技术及应用
16:30-17:00	圆桌论坛	圆桌论坛
8月18日 北京西郊宾馆 一号楼		
	第五会议室	第六会议室
09:00-12:00	医学影像智能分析与手术导航论坛 论坛主席：北京理工大学杨健教授	多模态大模型论坛 论坛主席：中国科学院大学黄庆明教授 中国科学院计算技术研究所李亮副研究员
09:00-09:30	讲者：四川大学华西医院龚启勇副院长 题目：精神影像学发展现状	讲者：中国科学院大学叶齐祥教授 题目：视觉表征模型的结构设计与物理启发
09:30-10:00	讲者：北京邮电大学马占宇教授 题目：数据受限条件下的视觉感知	讲者：北京交通大学魏云超教授 题目：从3D到4D，快速且时空一致的4D内容生成探索
10:00-10:30	讲者：湖南大学方乐缘教授 题目：眼科光学相干断层图像弱监督处理与分析	讲者：中国科学院自动化研究所刘静研究员 题目：多模态预训练模型的研究与应用

10:30-11:00	讲者：南昌大学刘巨根教授 题目：面向高维医学成像数据建模的扩散模型	讲者：哈尔滨工业大学(深圳)邵睿教授 题目：“九天”多模态大模型
11:00-11:30	讲者：北京理工大学宋红教授 题目：医学影像智能分析与处理	讲者：智象未来 (HiDream.ai) 姚霆联合创始人兼CTO 题目：多模态内容生成—掀起未来创意无限可能
11:30-12:00	圆桌论坛	圆桌论坛
北京西郊宾馆 一号楼 第三会议室		
09:00-12:00	第五届数学建模教学与培训研讨会—探索数学世界及数学建模	
09:00-10:00	讲者：南京邮电大学李雷教授 题目：数学建模的工作准备与竞赛训练	
10:00-11:00	讲者：北京林业大学李红军教授 题目：大模型与数学建模	
11:00-11:30	讲者：星推官 题目：分享科普创作心路历程，提供相关指导	
11:30-12:00	线下经验交流	

■ 特邀报告

时间

2024年8月17日 09:30-10:30

地点

一号楼 银杏大厅



王振常院士 首都医科大学附属北京友谊医院

医学影像的现状与创新

报告摘要：人民健康离不开医学影像。健康体检、亚健康检出、疾病治疗和康复评价等临床诊疗信息75%~85%来源于医学影像，评估健康和疾病的核心指标是分辨尺度和信息维度。医学影像的创新在空间分辨力和时间分辨率方面带来了很大的提升，其中空间分辨力跃升至十微米级，时间分辨率至亚毫秒级；同时，医学影像也从单模态向结构、功能、代谢、血流等多模态发展。医学影像目前的特点是数据采集的数字化、数据显示的实时化、数据传输的网络化、数据使用的共享化、数据处理的智能化、数据储存的永久化、分辨能力的显微化、影像观察的多维化、功能影像的临床化、代谢影像的定量化、分子影像的可视化、影像诊断的智能化、影像类别的融合化、影像知识的全面化。

个人简介：中国工程院院士，医学期刊知识挖掘与服务重点实验室联合主任、消化健康全国重点实验室副主任、首都医科大学医学影像学系主任，耳鸣临床诊疗与研究中心主任、北京市医学影像质量控制和改进中心主任，获“时代楷模”中国援外医疗队群体代表、全国五一劳动奖章、全国优秀科技工作者。

■ 特邀报告

时间

2024年8月17日 10:45-11:45

地点

一号楼 银杏大厅



徐瑞鸽教授 Rebecca Ruige Xu 美国雪城大学

Ethical Dimensions and Creative Collaborations of AI in Art and Design

Abstract: In the evolving landscape of AI in Generative Creativity (AIGC), ethical considerations and collaborative opportunities are transforming artistic practice and societal engagement. This talk addresses the ethical challenges of AIGC, such as algorithmic bias, cultural representation, and the democratization of creative tools. Drawing on personal experiences and philosophical reflections, we explore how AI technologies are reshaping the boundaries of artistic expression, challenging traditional notions of authorship and originality. We will examine compelling case studies where AI collaborates with artists and designers, enhancing creative processes and expanding possibilities in visual arts, design, and interactive media. By analyzing these themes through interdisciplinary lenses, this discussion aims to inspire critical discourse on the future of creativity in a digitally augmented world. Emphasis will be placed on the pivotal role of ethical awareness and collaborative innovation in shaping AI-driven art and design practices.

Biography: Rebecca Ruige Xu is a professor of computer art at the College of Visual and Performing Arts at Syracuse University. Her research interests include artistic data visualization, experimental animation, visual music, interactive installations, digital performance, and extended reality. Xu's recent work has been featured at prestigious venues such as the SIGGRAPH and SIGGRAPH Asia Art Gallery, IEEE VIS Arts Program, ISEA, Ars Electronica, Museum of Contemporary Art in Italy, Los Angeles Center for Digital Art in the USA, Colloquium Culture and Digitization in Switzerland, CYNETart in Germany, and the International Digital Art Exhibition in China.

Xu has served on the Media Arts Advisory Panel for the U.S. National Endowment for the Arts. She is the co-founder of the ChinaVIS Conference Arts Program and currently serves as the Chair of the ACM SIGGRAPH Digital Arts Committee and IEEE VIS 2024 Arts Program Co-Chair. Additionally, she is on the Executive Committee for the Association of Chinese Artists in American Academia.

<https://vpa.syr.edu/people/rebecca-ruige-xu/>

<http://rebeccaxu.com/>

机器具身交互智能论坛

第五会议室 8月16日 13:30-17:00

论坛介绍

当越来越多的智能机器走入人们的生活，协助我们的工作，如何以物理形态与用户进行自然、直观的交互，是具身交互智能所关注的核心问题。随着人工智能和机器人技术的快速发展，传感器、执行器和计算能力的持续进步，智能系统具备了能够更好地感知环境、理解用户需求，并作出相应回应的基础。本次论坛邀请知名高校及企业的优秀学者汇聚一堂，观点碰撞，探究如何利用具身智能实现自然、高效的交互技术；如何提升感知能力使智能系统能够在复杂环境和任务需求下自主决策、协同工作；如何赋予具身智能系统交互学习能力。

论坛主席



马楠教授
北京工业大学

北京工业大学教授、博士生导师，国家重点研发计划项目负责人，北京工业大学信息科学技术学院副院长，中国人工智能学会副秘书长，BSIG常务理事，北京市智能制造与机器人创新专项负责人，研究方向为交互认知、机器视觉、无人驾驶与移动机器人。以第一完成人先后获得中国图象图形学学会科技进步一等奖、中国电子学会科学技术奖【技术发明类】二等奖，主持多项国家、省部级项目，承担北汽集团、东风悦享、云迹科技等企业委托智能交互项目10项。带领团队多次在国际、国内人工智能、无人驾驶重要比赛中获得冠军，团队成果“无人驾驶云智能交互系统”获得第二届中国“AI+”创新创业大赛总决赛特等奖(2000余参赛队)。已在IEEE TIP、IEEE TNLS、ICRA、中国科学·信息科学等国内外学术期刊、会议发表论文80余篇。人才培养方面，国家级一流本科课程《智能交互技术》负责人，先后获得第六届全国教育科学研究优秀成果奖二等奖和北京市教学成果一等奖等。



刘永进教授
清华大学

清华大学计算机系长聘教授，人机交互与媒体集成研究所所长，国家杰青。主要研究方向为人机交互，情感计算，计算几何与图形学、多模态媒体智能信息处理。近五年发表论文六十余篇，包括7篇PAMI，5篇ACM TOG/SIGGRAPH，22篇TVCG/TIP/TMM/TAFFC/TASE/THMS/TCDS和21篇CVPR/ICCV/ECCV/AAAI/ICRA。一篇论文入选ESI热点论文，三篇论文入选ESI高被引论文。获得世界华人数学家联盟年度最佳论文奖(2017、2018)在内的十余项国内外知名期刊会议的最佳论文/优秀论文。六项专利进行了成果转化。培养多名博士生和硕士生获得清华大学优秀博士/硕士学位论文，以及北京市和中国图象图形学学会的优秀博士学位论文奖/提名奖，两名博士毕业生获得中国图象图形学学会石青云科学家奖，入选中国科协青年人才托举计划。两次获得中国人工智能学会吴文俊人工智能自然科学奖(2018年一等奖和2021年二等奖)，以及2021年度北京市科学技术进步奖二等奖。作为第一完成人获得2022年度中国图象图形学学会高等教育教学成果奖一等奖。

机器具身交互智能论坛

第五会议室 8月16日 13:30-17:00



查红彬教授
北京大学

具身智能：身体、行为与环境的紧耦合

报告摘要：具身智能的本质是实现心智与身体、行为、环境的多层次反馈闭环处理，强化智能体与环境的紧耦合。从生态心理学的角度来看，智能体的环境是一个随时间而变化的事件与过程所形成的系统，具有丰富的流动性。生物生存的基本要求是对变化的快速感知与实时反应，人的生存就是感知-行为循环的无限反复，这种对时变环境的强耦合形成了脑的动态处理机理。该报告将站在生态心理学的立场，说明针对具身智能的以下观点：生态环境中动态过程的存在；智能体与环境的相互依存关系（耦合）；对环境及其变化的直接感知；作为环境内身体的行为能力及其制约；智能体与环境相互作用的多层次循环。这些观点突出了与数据驱动的大模型预训练范式的对比，也对传统认知科学的研究提出了挑战。

个人简介：查红彬，北京大学智能学院教授，机器感知与智能教育部重点实验室主任。主要从事计算机视觉与智能人机交互的研究，在三维视觉几何计算、三维重建与环境几何建模、传感器即时定位与地图构建等方面取得了一系列成果。出版学术期刊及国际会议论文350多篇，其中包括IEEE T- PAMI, IJCV, IEEE T-VCG, IEEE T-RA等国际期刊以及ICCV, ECCV, CVPR等国际学术会议论文150多篇。



徐凯教授
国防科技大学

融合三维感知与多模态大模型的具身智能

报告摘要：视觉感知是机器人探索、感知和理解未知环境的最重要方式。随着三维传感和重建技术的飞速发展，三维图形正与机器人视觉深度融合，形成三维几何引导的具身感知与交互新途径，结合多模态大模型的强大常识推理和任务规划能力，最终支持机器人在三维世界中实现具身智能。本次报告围绕机器人的主动重建，主动理解，以及任务驱动的交互，汇报我们近年来的系列工作。其中，在任务驱动的交互方面，重点介绍在线语义理解驱动的物体目标导航机器人，以及多模态大模型驱动的室内物品整理机器人。最后探讨基于三维感知与交互的世界模型构建与演进，及其对具身智能发展的推动作用。

个人简介：徐凯，国防科技大学教授，国家杰出青年基金获得者。普林斯顿大学访问学者。研究方向为计算机图形学、三维视觉、具身智能、数字孪生等。在国际上较早开展了数据驱动三维感知、建模与交互工作，提出面向复杂三维数据的结构化感知、建模与交互理论方法系统。发表TOG/TPAMI/TVCG等A类论文90余篇，其中图形学顶会SIGGRAPH论文29篇（第一作者10篇）。担任图形领域顶级国际期刊ACM Transactions on Graphics的编委，以及多个领域重要会议的主席。任中国图象图形学学会三维视觉专委会副主任、中国工业与应用数学学会几何设计与计算专委会副主任。获湖南省自然科学一等奖2项（排名1和3）、中国计算机学会自然科学一等奖（排名3）、军队科技进步二等奖、军队教学成果二等奖等。

机器具身交互智能论坛

第五会议室 8月16日 13:30-17:00



高炳钊教授
同济大学

AI大模型驱动的自动驾驶算法及车端具身实现

报告摘要：随着大语言模型（Large Language Model）智能涌现能力的产生，智能基础模型（Foundation Model）技术被认为有望解决汽车自动驾驶安全长尾问题。基础模型的海量数据预训练基于自监督学习，其泛化能力优于需要标注的监督学习。本报告将从自动驾驶面临的挑战出发，讲述自动驾驶算法开发的技术演进过程，总结自监督学习的最新进展，最后给出基于自监督学习的AI大模型垂直应用于自动驾驶场景可能的技术路线。

个人简介：高炳钊，同济学长聘特聘教授，获控制理论与控制工程专业以及机械工程专业双博士学位。2015年度国家优青、入选2022年度国家重大

人才工程，自动化学会车辆控制与智能化专业委员会秘书长、IFAC T.C. Automotive Control委员。研究方向为汽车控制与智能化，发表IEEE汇刊、ASME汇刊、VSD等控制和车辆权威期刊论文60余篇，出版中英文专著2部，授权发明专利30余件，含美国发明专利2件。曾获自动化学会技术发明一等奖2项、吉林省优秀博士学位论文和多项学会优秀论文奖。



支涛CEO
北京云迹科技股份有限公司

AI+具身智能 未来竞争新势力

报告摘要：机器人从上个世纪中发展至今经历了电气时代、数字时代，如今迈入了智能时代。服务机器人从专用单一功能机器人向复合功能机器人走过了10余年的进化之路。随着大模型落地应用场景越来越广泛，服务机器人被以大模型为代表的新一代人工智能赋能，正加速具身化，通用人工智能时代即将来临。从单模态到多模态再到具身智能，作为其载体的“复合多态机器人”已成为服务机器人领域的先进代表，不仅具备高度的应用价值，更是行业新质生产力的典型产品。随着技术的不断进步和市场的逐渐成熟，复合多态机器人有望在未来发挥更加重要的作用，为服务行业带来深远的影响。

个人简介：曾创建工业机器人公司，曾任两家电子科技公司创始人兼CEO。拥有十余年传感器及自动化工作经验。2014年创立北京云迹科技股份有限公司，专注于商用智能服务机器人技术与场景落地，拥有专利上百件，获吴文俊人工智能科学技术发明奖，中国（行业）十大创新企业家，是机器人行业发展的探索者。

机器具身交互智能论坛

第五会议室 8月16日 13:30-17:00



高跃副教授
清华大学

超图计算

报告摘要: 许多生物、社会和技术系统的复杂性源于系统单元之间丰富性的相互作用。社会交流、化学反应、生命系统到生态系统中相互作用通常发生在三个或更多对象的组中，并不能简单地用二元关系来描述。有效地利用这些高阶关联结构能够增强对各类系统的建模能力，帮助理解和预测其动态行为。超图是建模此类高阶交互、即高阶关联的天然结构。实际应用中通常也面临数据不足和关联复杂难题。本报告主要介绍面向高阶关联的超图计算理论、方法及应用。具体而言，从由数据到高阶关联的映射、由高阶关联到语义表示的映射及高阶关联语义领域泛化三个维度介绍高阶关联的超图结构建模、数据和高阶关联协同的超图语义计算及面向不同领域的超图计算应用相关进展。

个人简介: 高跃，清华大学长聘副教授、博士生导师，国家青年特聘专家、DeepTech 2022年中国智能计算科技创新人物。主要研究领域为人工智能、计算机视觉及医学图像处理，提出并发展了超图计算理论及方法，实现了高阶关联的结构建模与计算，有效应对数据不足和信息关联复杂等挑战，应用于视觉目标检测、医学辅助诊断等领域。研究成果在IEEE TPAMI等发表论文百余篇，论文引用万余次，出版《Hypergraph Computation》等英文专著，多次入选科睿唯安高被引科学家和爱思唯尔中国高被引学者，担任International Journal of Computer Vision和Medical Image Analysis等国际重要期刊编委。曾获得中国电子学会自然科学一等奖、福建省科技进步一等奖等。



赵思成副研究员
清华大学

开放环境领域自适应

报告摘要: 领域自适应旨在实现分布不同的有标注源域和无标注目标域之间的自适应对齐，提升机器学习模型的泛化能力，解决人工智能的最后一公里问题。本报告将介绍课题组近年来在开放环境下面向情感计算、对象检测、语义分割等应用的领域自适应研究，包括单源领域自适应、多源领域自适应与领域泛化等三个方面，并探讨未来可能的研究方向。

个人简介: 清华大学副研究员，国家级青年人才，ACM/IEEE/CCF/CSIG高级会员，哈尔滨工业大学博士，加州大学伯克利分校和哥伦比亚大学博士后。研究方向为情感计算、多媒体、机器学习等。发表IEEE/ACM汇刊、CCF推荐

A类论文60余篇，谷歌学术引用10000余次，H指数为50。担任中国图象图形学学会CSIG情感计算与理解专委会秘书长、IEEE TIP/IEEE TAFFC/CSIG通讯编委、IEEE TFS/ACM TOMM/IEEE MM等国际期刊的首席客座编辑、NeurIPS/ACM MM/ECCV/AAAI/IJCAI/ICASSP等国际会议的领域主席或资深程序委员。主持或参与国家自然科学基金、国家重点研发计划、CCF-滴滴盖亚学者科研基金等项目。获得CSIG优秀博士论文、ACM SIGMM中国新星奖、媒体建模旗舰会议MMM 2015最佳论文Runner-up奖等奖励，入选AI 2000人工智能全球最具影响力学者榜单和全球前2%顶尖科学家榜单。

人体动作理解与生成前沿论坛

第六会议室 8月16日 13:30-17:00

论坛介绍

人体动作理解与生成是计算机视觉与多媒体领域的重要研究方向，旨在构建具备通用性、适应性的运动时空感知与生成模型，提升人体动作理解与生成的准确性和逼真度，实现对人体动作序列的识别、定位、分析、重建以及生成。在当前飞速发展的大模型驱动下，人体动作理解与生成技术在行为分析、人机交互、运动捕捉等方向的性能不断取得新突破，在个性化训练、医疗康复、智能监控、社交媒体、数字人等领域的应用不断革新。本论坛邀请了人体动作理解与生成领域的专家学者介绍最新前沿技术进展，探讨技术落地应用过程中面临的机遇、挑战与未来发展趋势。

论坛主席



彭宇新教授
北京大学

彭宇新，北京大学二级教授、博雅特聘教授、国家杰出青年科学基金获得者、国家万人计划科技创新领军人才、科技部中青年科技创新领军人才、863项目首席专家、中国人工智能产业创新联盟专家委员会主任、中国工程院“人工智能2.0”规划专家委员会专家、CAAI/CIE/CSIG Fellow、中国图象图形学学会副秘书长、提名与奖励委员会副主任、北京图象图形学学会副理事长。主要研究方向为跨媒体分析、计算机视觉、机器学习、人工智能。以第一完成人获2016年北京市科学技术奖一等奖和2020年中国电子学会科技进步奖一等奖，2008年获北京大学宝钢奖教金优秀奖，2017年获北京大学教学优秀奖。主持了863、国家自然科学基金重点等40多个项目，发表论文200多篇，包括ACM/IEEE Trans.和CCF A类论文100多篇。多次参加由美国国家标准技术局NIST举办的国际评测TRECVID视频样例搜索比赛，均获第一名。主持研发的跨媒体互联网内容分析与识别系统已经应用于公安部、工信部、国家广播电视总局等单位。担任IEEE TMM、TCSVT等期刊编委。



徐婧林副教授
北京科技大学

徐婧林，北京科技大学智能科学与技术学院副教授，北京图象图形学学会理事、副秘书长，中国图象图形学学会青托俱乐部副主席。2023年入选第九届中国科协青年人才托举工程、2022年获中国图象图形学学会优秀博士学位论文奖、2023年获中国自动化学会自然科学奖一等奖（4/5）。主要研究方向为视频动作理解、多模态细粒度分析、三维人体姿态估计与动作生成。已发表/接收20篇ACM/IEEE Trans.和CCF A类国际期刊和会议论文（一作15篇，通讯2篇）。主持国家自然科学基金面上项目、青年基金、中国博士后科学基金面上项目等。获得2023年度北京图象图形学学会“最美女科技工作者”、2022年西北工业大学优秀博士学位论文等荣誉。担任《Chinese Journal of Electronics》青年编委、《电子与信息学报》编委、《计算机科学》青年编委、《人工智能》编委等。

人体动作理解与生成前沿论坛

第六会议室 8月16日 13:30-17:00



徐常胜研究员
中国科学院
自动化研究所

不确定性感知的弱监督人体动作定位研究

报告摘要: 近年来, 基于视频数据的人体动作定位成为了动作理解领域的一大研究热点。然而, 传统的定位方法需要耗费大量的人力物力进行精细的时序标注, 这限制了其在实际应用中的可行性。为此, 研究者们提出了弱监督人体动作定位方法, 其只需要一个粗略的视频级别标签就可以进行模型训练。然而, 弱监督方法由于无法获取精细准确的监督信号, 其学习过程面临着极大的不确定性。近期, 证据深度学习作为一种高效的不确定估计方法, 在各领域展现出了优秀的性能, 其对提升弱监督人体动作定位具有很好的潜力。本报告将介绍团队近期在本领域中的研究进展, 包括双级证据学习方法、向量化证据学习方法、存-缺证据感知的音视频协同定位方法等, 相关成果发表在CVPR、IEEE TPAMI等会议和期刊上。

个人简介: 徐常胜, 中国科学院自动化研究所研究员, 国家杰出青年基金获得者, 国家万人计划领军人才, 入选国家百千万人才工程和首都科技领军人才工程, 科技部重点领域创新团队负责人, 国家重点研发计划项目首席科学家, 中国科学院王宽诚率先人才计划卢嘉锡国际团队负责人。国际电子电气工程师学会会士 (IEEE Fellow), 国际模式识别学会会士 (IAPR Fellow), 国际计算机学会杰出科学家 (ACM Distinguished Scientist)。担任国际计算机学会多媒体专委会中国区 (ACM SIGMM China Chapter) 主席。在多媒体分析, 计算机视觉, 模式识别, 图像处理等领域发表论文500多篇, 其中IEEE和ACM汇刊论文150余篇, 国际顶级会议论文150余篇。在多媒体国际顶级会议和期刊上获得最佳论文奖10余次。获得2018年中国电子学会自然科学一等奖, 2009年中国计算机学会青年科学家奖, 7次获得中国科学院优秀导师奖。



刘焯斌教授
清华大学

人体无标记运动捕捉与运动生成

报告摘要: 人体运动捕捉与运动生成在数字人、4D内容生成、具身智能有重要的应用价值。本报告首先介绍当前无标记人体运动捕捉技术的最新进展, 包括单相机和多相机条件下, 面向人体运动和人手运动等捕捉需求。基于运动捕捉结果, 介绍报告人在运动生成及人体数字化身方面的研究进展。最后, 报告面向4D内容生成及具身智能研究, 阐述人体运动生成的意义和价值。

个人简介: 刘焯斌, 清华大学长聘教授, 国家基金委杰青基金获得者。研究方向为三维视觉与影像生成。发表CVPR/ICCV/ECCV/TPAMI/TOG论文百篇。多次担任CVPR、ICCV、ECCV领域主席, 担任IEEE TVCG、CGF编委, 中国图象图形学会三维视觉专委会副主任。获2012年国家技术发明一等奖 (排名3), 2019年中国电子学会技术发明一等奖 (排名1)。

人体动作理解与生成前沿论坛

第六会议室 8月16日 13:30-17:00

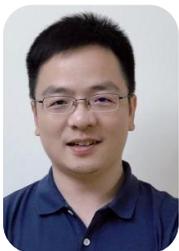


王利民教授
南京大学

多模态视频理解基础模型InternVideo

报告摘要: 构建多模态基础模型已经成为计算机视觉领域的研究热点。视频理解面临着数据维度高、信息容量大、场景变化多等核心挑战, 如何构建通用视频理解基础模型已经成为现阶段一项极具挑战的任务。本次报告将主要介绍多模态视频理解基础模型InternVideo及其背后的关键技术, 包括单模态视频自监督预训练方法VideoMAE, 多模态视频弱监督预训练方法UMT, 和多模态视频交互对话模型VideoChat。同时还将介绍多模态视频数据集InternVid和多模态视频评测基准MVBench。最后将展望多模态视频基础模型发展趋势。

个人简介: 王利民, 南京大学教授, 博士生导师, 国家海外高层次青年人才计划入选者, 科技创新2030-“新一代人工智能”重大项目青年科学家。2011年在南京大学获得学士学位, 2015年在香港中文大学获得博士学位, 2015年至2018年在苏黎世联邦理工学院 (ETH Zurich) 从事博士后研究工作。主要研究领域为计算机视觉和深度学习, 专注视频理解和动作识别, 在IJCV、T-PAMI、CVPR、ICCV、NeurIPS等学术期刊和会议发表论文100余篇。根据Google Scholar统计, 论文被引用 23000余次, 两篇一作论文取得了单篇引用接近或超过4000的学术影响力。在视频分析领域做出了一系列有重要影响力的研究工作, 例如: TSN网络架构, VideoMAE预训练方法, MixFormer跟踪器等。曾获得广东省技术发明一等奖, 世界人工智能大会青年优秀论文奖, ACM MM 2023最佳论文荣誉提名奖。入选2022年度AI 2000人工智能全球最具影响力学者榜单, 2022年度全球华人AI青年学者榜单, 2021-2023年度爱思唯尔中国高被引学者榜单。担任CVPR/ICCV/NeurIPS等重要国际会议的领域主席和计算机视觉领域旗舰期刊IJCV的编委。



周晓巍研究员
浙江大学

基于神经表达的体积视频技术

报告摘要: 沉浸式体积视频是未来媒体的重要形式, 可广泛应用于实时通信、文化艺术、体育直播等领域。传统的三维表达方法在表达能力、重建鲁棒性、渲染质量等方面存在局限。近年来兴起的基于神经表达的重建与渲染方法为体积视频提供了新的技术途径, 但在采集成本、渲染效率、存储空间等方面仍有局限。本报告将介绍我们在基于神经表达的体积视频重建与渲染等方面的工作, 并简要探讨未来的发展方向。

个人简介: 周晓巍, 浙江大学研究员, 国家级青年人才项目入选者。研究方向主要为三维视觉及其在混合现实、机器人等领域的应用。近五年在相关领域的顶级期刊与会议上发表论文70余篇, 多次入选CVPR最佳论文候选。曾获得浙江省自然科学一等奖陆增镛CAD&CG高科技奖一等奖, CCF优秀图形开源贡献奖, 入选全球前2%顶尖科学家榜单、中国高被引学者。担任国际顶级期刊IJCV编委、顶级会议CVPR/ICCV领域主席, 图形学与混合现实研讨会 (GAMES) 执行委员会主席, 视觉与学习研讨会 (VALSE) 常务委员, CSIG三维视觉专委会常务委员。

人体动作理解与生成前沿论坛

第六会议室 8月16日 13:30-17:00



涂志刚研究员
武汉大学

视频人体动作行为识别与生成

报告摘要: 视频理解作为计算机视觉领域的核心任务成为推动人工智能发展的主要力量，其中视频人体动作行为识别与生成作为视频理解的关键技术具有广泛应用而一直都备受关注。首先，报告将从人机交互的角度，介绍武汉大学行为理解与视觉感知研究组（HUVPRLab）在视频人体动作行为识别与生成方面的系列工作，主要包括人手姿态估计与重建、人体姿态估计与动作识别、人体动作生成与迁移，形成了从局部（人手）→整体（个体）→下游应用（生成）的研究范式。最后，报告将总结和展望视频人体动作行为识别与生成的发展趋势。

个人简介: 涂志刚，武汉大学测绘遥感信息工程国家重点实验室研究员，博士生导师，湖北省杰出青年基金获得者。主要研究领域为计算机视觉与视频理解，聚焦视频人体行为识别、重建与生成。发表高水平论文70余篇，其中第一/通讯作者发表中科院1区Top SCI 期刊+CCF A类顶会论文30篇，国际会议“最佳学生论文”1篇。获省部级科技奖励3项，其中主持荣获2022年获湖北省自然科学二等奖（排名1）。主持湖北省杰出青年基金、教育部联合基金（青年人才）、国家自然科学基金、腾讯犀牛鸟基金（获技术创新奖）等科研项目。开发了视频人体行为智能识别系统，成功应用“第七届世界军人运动会开闭幕式”等多个领域，被央视新闻/体育频道采访报道。入选中科院1区Top期刊IEEE TCSVT 2022年度全球三位最佳审稿人。指导研究生获ICCV2021-MMVRAC挑战赛2个赛道亚军。担任多个国际期刊、专刊、会议的编委、领域主席。



徐婧林副教授
北京科技大学

面向行为理解的细粒度运动分析

报告摘要: 细粒度运动分析旨在精细化分析视频动作序列，通过细粒度识别、时空定位、质量评价，以实现目标行为的理解，可广泛应用于智能安防、智慧医疗、智能体育、智慧传媒等领域。本报告首先介绍细粒度运动分析面临的挑战与意义，重点介绍细粒度动作质量评价、文本驱动的细粒度时空动作定位、文本驱动的三维人体姿态估计，回答如何精准辨识时空边界模糊的细粒度动作类型、如何利用文本对人体动作精细化重建等问题。上述研究工作将细粒度运动分析从二维空间延伸到三维空间，在运动分析、康复训练、体能测试等领域具有重要意义。

个人简介: 北京科技大学智能科学与技术学院副教授，北京图象图形学学会理事、副秘书长，中国图象图形学学会青托俱乐部副主席。2023年入选第九届中国科协青年人才托举工程、2022年获中国图象图形学学会优秀博士学位论文奖、2023年获中国自动化学会自然科学奖一等奖（4/5）。主要研究方向为视频动作理解、多模态细粒度分析、三维人体姿态估计与动作生成。已发表/接收20篇ACM/IEEE Trans.和CCF A类国际期刊和会议论文（一作15篇，通讯2篇）。主持国家自然科学基金面上项目、青年基金、中国博士后科学基金面上项目等。获得2023年度北京图象图形学学会“最美女科技工作者”、2022年西北工业大学优秀博士学位论文等荣誉。担任《Chinese Journal of Electronics》青年编委、《电子与信息学报》编委、《计算机科学》青年编委、《人工智能》编委等。

数字人与内容生成论坛

第五会议室 8月17日 13:30-17:00

论坛介绍

数字人和数字内容作为元宇宙中的主体，紧密连接着现实与虚拟世界。近年来，数字人和内容生成相关方向的研究受到学者的广泛关注。在智能技术的驱动下，数字人在外在形象、内在行为、智能交互水平方面取得长足进展；内容生成在动静态细节、真实感和丰富度等方面不断得到革新。本论坛聚焦数字人和内容生成领域中的前沿热点问题，邀请学者分享最新技术进展并展现前沿研究工作，探讨数字人和内容生成在多方面领域的结合应用及其未来发展趋势。

论坛主席



刘焯斌教授
清华大学

刘焯斌，清华大学长聘教授，国家杰青基金获得者。研究方向为三维视觉、数字人重建、生成与交互。发表TPAMI/SIGGRAPH/CVPR/ICCV等论文近百篇。多次担任CVPR、ICCV、ECCV领域主席，担任IEEE TVCG、CGF编委，中国图象图形学学会三维视觉专委会副主任。获2012年国家技术发明一等奖（排名3），2019年中国电子学会技术发明一等奖（排名1）。



张鸿文副教授
北京师范大学

张鸿文，北京师范大学人工智能学院副教授。主要从事以人为中心的三维视觉研究，尤其是三维数字人体的运动捕捉与生成、化身重建与驱动等课题。发表CCF-A类/领域顶刊顶会论文40余篇，其中TPAMI/TOG和CVPR/ICCV/ECCV/SIGGRAPH论文30余篇，顶会口头报告/亮点论文7篇，Google Scholar引用量2800余次。提出的动捕系列算法累计获GitHub星标量超过1200，获中国科学院优博论文/院长奖等荣誉。代表性成果详见个人主页：<https://zhanghongwen.cn>

数字人与内容生成论坛

第五会议室 8月17日 13:30-17:00



彭宇新教授
北京大学

多维感知驱动的AIGC

报告摘要: AIGC (Artificial Intelligence Generated Content) 是指人工智能自动生成并创造新内容的生产方式。随着人工智能的迅速发展, AIGC 从主要生成文本内容到如今已可以根据用户给定的文本描述, 自动生成语义一致、内容真实、符合逻辑的图像、视频等视觉内容。这是人工智能从“感知智能”迈向“认知智能”的一项重要任务, 在教育、设计、影视、创作等领域具有重要的应用价值。当前以扩散模型为代表的生成模型已经可以根据文本描述生成高质量的视觉内容, 但在精细化控制和高维信息建模上仍存在不足。其关键科学问题是如何准确地感知和建模图像视频等视觉内容中的局部细节、空间布局和时序动作等多维时空信息, 实现视觉内容的可控精细生成。本报告将介绍我们在文本到视觉内容生成的近期工作进展, 包括细节可控的图像生成、时序运动连贯的视频生成等, 并对未来发展方向及趋势进行简要讨论与展望。

个人简介: 彭宇新, 北京大学二级教授、博雅特聘教授、国家杰出青年科学基金获得者、国家万人计划科技创新领军人才、科技部中青年科技创新领军人才、863项目首席专家、中国人工智能产业创新联盟专家委员会主任、中国工程院“人工智能2.0”规划专家委员会专家、中国电子学会会士、中国人工智能学会会士、中国图象图形学学会会士、副秘书长、提名与奖励委员会副主任、北京图象图形学学会副理事长。主要研究方向为跨媒体分析、计算机视觉、机器学习、人工智能。以第一完成人获2016年北京市科学技术奖一等奖和2020年中国电子学会科技进步奖一等奖, 2008年获北京大学宝钢奖教金优秀奖, 2017年获北京大学教学优秀奖。主持了863、国家自然科学基金重点等30多个项目, 发表论文200多篇, 包括ACM/IEEE Trans和CCF A类论文100多篇。多次参加由美国国家标准技术局NIST举办的国际评测TRECVID视频样例搜索比赛, 均获第一名。主持研发的跨媒体互联网内容分析与识别系统已经应用于公安部、工信部、国家广播电视总局等单位。担任IEEE TMM、TCSVT等期刊编委。

高保真数字人建模与驱动关键技术与应用



吕科教授
中国科学院大学

报告摘要: 随着人工智能技术的不断发展, 虚拟现实、增强现实、数字孪生和元宇宙正在逐渐从概念走向现实, 数字人是其中不可或缺的角色, 是近年来产业聚焦的热点和研究的前沿。数字人是存在于非物理世界中多种计算机手段合成的具有多重人类特征的综合体, 数字人在泛娱乐、金融、文旅、教育以及医疗等领域都有应用, 本报告主要介绍数字人建模与驱动中的关键技术, 包括基于情感感知的数字人表情生成方法, 多模态协同的人体姿态估计与动作捕捉、面向多级场景的数字人三维重建与背景融合以及面向多级场景的数字人仿真测试与演示验证。

个人简介: 吕科, 中国科学院大学特聘教授、博士生导师, 国家高层次人才特殊支持“万人计划”领军人才, 科技部创新人才推进计划“中青年科技创新领军人才”, 北京市高等学校高层次人才引进与培养计划特聘教授, 鹏城国家实验室双聘教授, 国家重点研发计划“基础科研条件与重大科学仪器设备研发”专项项目负责人, 享受国务院政府特殊津贴专家。主要研究方向为图像处理、智能信息处理技术。承担国家自然科学基金、国家重点研发计划、中国科学院仪器设备、北京市教委重大专项科研项目三十余项。在国内外学术期刊和国际主流会议上发表学术论文150余篇, 出版专著两部。研究成果先后获2004年度、2009年度国家科技进步二等奖、2012年度北京市科学技术二等奖、2012年度中国电子学会电子信息科学技术二等奖, 2017年获得中国科学院(京区)成果转化奖, 2021年首届全国博士后创新创业大赛团队银奖。

数字人与内容生成论坛

第五会议室 8月17日 13:30-17:00



鲁继文教授
清华大学

多模态视觉内容生成

报告摘要: 视觉内容生成是计算机视觉的研究热点，在公共安全、文化影视、消费电子等领域有着重要的应用前景。报告将回顾视觉内容生成近年来的研究进展，主要包括生成对抗网络、高效扩散模型、多模态大模型等方法，以及在图像超分、视频修复、点云补全、文生视频等生成任务中的应用，最后对未来发展趋势进行展望。

个人简介: 鲁继文，清华大学自动化系副主任，长聘教授，博士生导师，国家杰出青年科学基金获得者，IEEE/IAPR Fellow，国家重点研发计划项目负责人，国际期刊Pattern Recognition Letters主编，中国自动化学会专家咨询工作委员会副主任，中国仿真学会视觉计算与仿真专业委员会主任。长期从事计算机视觉、模式识别、具身智能等方面研究，发表IEEE汇刊论文140余篇（其中T-PAMI论文39篇），CVPR、ICCV、ECCV论文130余篇，谷歌学术引用3万余次，获授权国家发明专利60余项，主持国家自然科学基金重点项目2项，获中国电子学会自然科学一等奖1项（第一完成人）和国家级教学成果奖二等奖1项，担任T-IP、T-CSVT、T-BIOM、PR、自动化学报等期刊编委和ACCV2026、FG2023、ICME2022、VCIP2022等会议大会主席/程序委员会主席。



张盛平教授
哈尔滨工业大学

超写实虚拟数字人驱动技术

报告摘要: 赋予情感的虚拟数字人在元宇宙，虚拟现实等多项应用中发挥着重要作用，而其中，如何通过带有情感的动作信号准确驱动虚拟数字人，是提升数字人真实感和沉浸感的关键问题。因此，为了增强虚拟数字人驱动的准确性与稳定性，我们从二维和三维的不同人体表达出发，挖掘驱动信号与人体表示在不同维度下的关联性，提出基于生成式大模型的可控人体视频生成算法和基于3D高斯的三维人体驱动算法，进而在单目拍摄的视频中学习一个可驱动的二维/三维超写实虚拟数字人表示，为情感动作的精确表达提供高效的载体。

个人简介: 张盛平，哈尔滨工业大学长聘教授/博士生导师、青年科学家工作室学术带头人，美国布朗大学博士后、香港浸会大学博士后、美国加州大学伯克利分校访问学者，入选国家级青年人才计划、山东省泰山学者青年专家、哈尔滨工业大学青年拔尖人才计划（副教授、教授）、人社部香江学者人才计划。主要研究方向为3D视觉、虚拟数字人等。主持国家自然科学基金4项、华为公司项目10余项。研究成果获2019年度黑龙江省自然科学二等奖（排名第一）。已发表学术论文80余篇，包括PNAS、IEEE T-PAMI、IJCV、ICML、CVPR、ICCV等。

数字人与内容生成论坛

第五会议室 8月17日 13:30-17:00



连宙辉副教授
北京大学

个性化数字人的便捷构建

报告摘要: 数字人制作技术在教育、宣传、社交、影视、电商等领域有着重要应用价值，在学术研究方面也面临诸多挑战，是当前学术界的一个研究热点。本报告将介绍我们在基于单相机拍摄输入的个性化数字人便捷构建方面的一些研究进展，包括前期发表的基于2D图像生成模型的人脸编辑（CVPR' 21）、肖像风格化（TOG' 22）、可控人体图像合成（TPAMI' 23），以及最近提出的基于3D生成模型的可驱动3D数字人生成（CVPR' 24）和风格化高保真3D数字分身构建（CVPR' 24）。

个人简介: 连宙辉，北京大学王选计算机研究所副教授、博士生导师，北京大学博雅青年学者，中国文字字体设计与研究中心副主任，入选国家级青年人才计划和北京市科技新星计划。研究领域为计算机图形学、计算机视觉与人工智能，主要研究方向是图形图像生成及其应用，在领域重要期刊（TOG, TPAMI, IJCV等）和会议（SIGGRAPH/SIGGRAPH Asia, CVPR, NeurIPS等）上发表论文90余篇。多次担任NeurIPS、CVPR、ICCV等国际会议领域主席，获北京市技术发明奖二等奖（排名1）、中国专利优秀奖（排名1）、吴文俊人工智能优秀青年奖等奖励。



朱翔昱副研究员
中国科学院
自动化研究所

精细三维人脸感知

报告摘要: 人脸作为个体身份和情感特征的主要载体，在日常信息传递和情感表达中起着非常重要的作用，利用三维技术建模人脸的结构和行为，是人脸分析领域中的重要课题。特别在数字人、视频生成、以及VR/AR等技术的推动下，对精确人脸感知的需求不断攀升。重建更精细的人脸结构、更生动的人脸行为，以及更准确的人脸位姿成为研究焦点。报告将结合课题组研究，从人脸结构、表情和位姿三个方面的精细重建为大家介绍三维人脸感知的最新进展。

个人简介: 朱翔昱，中国科学院自动化所副研究员，国际模式识别协会（IAPR）青年学者奖（YBIA）获得者，IEEE Senior Member，中国科学院青年创新促进会成员，长期从事三维人脸重建、人脸识别、可解释神经网络等方面的相关理论研究与应用。共发表论文80余篇，其中作为第一作者发表T-PAMI论文3篇，IJCV论文1篇。两次获得CVPR口头报告。发表文章的Google Scholar总引用次数为8000余次，最高单篇引用2000余次，14篇论文引用超过100次。获得WACV 2024、ICCV 2019人脸识别竞赛冠军、FG2017人脸微表情竞赛冠军、CCBR2017最佳学生论文奖和CCBR2022最佳海报论文奖。授权国家发明专利8项。入选北京科协青年人才托举工程、百度学术全球华人AI青年学者榜单（全球25人），受到腾讯犀牛鸟基金支持。获2021中国电子学会科技进步二等奖、中国图象图形学学会优秀博士论文提名奖。任2021及2022年ACM MM（CCF:A类）及国际生物特征识别大会（IJCB）领域主席（Area Chair），提出的人脸三维建模方法在被PyTorch官方Twitter报道，开源代码在Github上收获6000余星。相关技术已在电商、安防、金融等多个行业落地应用。研发的三维人脸及人体动捕系统在阿里数字人生、淘宝直播、天猫数字藏品等APP上线。

工业视觉技术与应用论坛

第六会议室 8月17日 13:30-17:00

论坛介绍

制造业是我国实体经济的基础，也是科技创新的主战场，我国制造业总体规模已连续14年位居全球第一。目前，所有制造类企业都面临着一个核心共性需求，即对工业制品的品质检测、质量监测。近年来，随着计算机视觉技术的快速发展，采用视觉检测技术进行产品的瑕疵检测已获得广泛应用。本论坛结合国内工业视觉技术的快速发展，旨在向学术界报告大模型时代下的工业视觉、工业制造表面缺陷检测、管道缺陷智能检测、面向罕见工业缺陷检测的小样本异常生成等重要热点技术的最新进展。

论坛主席



周瑜副研究员
华中科技大学

周瑜，华中科技大学副研究员，博导。主要研究方向为开放场景视觉分析、工业视觉。已发表高水平学术论文50余篇，包括IJCV, TIP, ICCV, ECCV, NeurIPS, 《中国科学: 信息科学》等权威期刊及会议。获首届中国图象图形学学会高等教育教学成果奖一等奖（第一完成人），中国国际“互联网+”大学生创新创业大赛国赛金奖2项、第八届吴文俊人工智能科学技术奖技术发明三等奖、VOT2019-ST冠军。主导研发的工业缺陷检测人工智能平台在我国制造业500强企业广泛应用，获湖北省委全省通报表彰，《人民日报》等广泛报道。担任CSIG视频图像通信专委会委员，CSIG武汉会员活动中心执行委员，VALSE2022计算机视觉与工业应用研讨会主席。主持或参与国家自然科学基金项目、科技部“新一代人工智能”重大项目，湖北省新型显示产业重大科技创新项目等。



陶显副研究员
中国科学院
自动化研究所

陶显，中国科学院自动化研究所副研究员，长期从事工业视觉智能检测研究，主持国自然面上、北京市面上、国自然青年、国自然地区合作、科技部2030“新一代人工智能”重大项目子课题等项目，发表IEEE trans论文十余篇，开发视觉设备和算法成功应用于中国工程物理研究院激光聚变中心，入选IEEE Senior Member，中国科学院青促会会员，曾获CVPR 2023学术竞赛亚军和最具创新奖，CAAI工业视觉比赛冠军。2019年获第十四届‘春晖杯’中国留学人员创新创业大赛高新技术组优胜奖，2022年获中科院科技促进发展奖，2023年获中国仪器仪表学会科技进步三等奖。

工业视觉技术与应用论坛

第六会议室 8月17日 13:30-17:00



丁贵广教授
清华大学

模型推理优化技术研究

报告摘要：随着大模型技术的发展，模型复杂度变得越来越高、参数越来越多，给模型的高效部署带来的巨大挑战。模型的推理优化技术成为人工智能研究的重要方向之一，如何在不损失或较少损失模型精度的前提下，减小模型的计算复杂度是人工智能开展大规模应用的重要技术，对于推动大模型在智能手机、自动驾驶等端侧场景应用至关重要。本次报告将介绍深度学习模型的推理优化技术，重点介绍卷积神经网络的“训大用小”的深度学习训练和部署方法论，并探索多模态大模型的压缩优化技术等。

个人简历：丁贵广，清华大学信息科学与技术国家研究中心党总支书记、副主任、教授，国家杰青获得者。以计算机视觉技术在国家公共安全、网络内容管理、自动驾驶、机器人和工业智能制造等领域的实际应用需求为出发点，开展视觉感知理解、视觉模型结构设计、视觉模型压缩优化等方面的研究，在网络多媒体内容处理、工业视觉缺陷检测等领域开展规模化应用，研发高效的视觉感知计算系统和平台。先后主持国家自然科学基金重点项目、重点研发项目等数十项。发表学术论文百余篇，Google Scholar引用17000余次，相关成果成功应用于快手、OPPO、京东、凌云光等单位。曾获国家科技进步二等奖、电子学会技术发明一等奖、人工智能学会吴文俊科学技术进步一等奖等奖项。



杨辉华教授
北京邮电大学

工业制造表面缺陷视觉检测技术及应用

报告摘要：表面缺陷检测利用先进的图像处理和机器学习算法，能够自动识别和分类制造过程中的产品缺陷，是实现高质量、高效率生产的关键步骤。由于缺陷的形成原因不同，表面缺陷检测具有如下突出特点：1) 缺陷形貌特征复杂，各种缺陷因素交织；2) 缺陷种类多并在样本数量上呈长尾分布；3) 因技术升级或数据标注不充分，生产过程中会出现未知类型缺陷。针对以上问题，本团队构建一系列融合多深度、多宽度的神经网络模块，以提取形貌差异较大的表面缺陷特征，并基于此模块构建了形状和尺寸感知的骨干网络。提出多种均衡特征提取的自监督学习方法，能有效提高对无标签数据和特征的利用率，从而充分捕获无偏置特征并提高长尾分布数据的检测性能。

提出基于单分支网络的无监督学习方法，能持续提取已有数据和现场增量数据的缺陷特征，从而充分提取已知缺陷和未知缺陷的综合性特征。相关研究成果在宁德时代、比亚迪等国内知名大型锂电企业，以及金川集团、佳通轮胎等国内知名企业应用。

个人简历：杨辉华，北京邮电大学教授，博导。长期从事机器学习、视觉检测及其在工业和医学领域的应用研究。主持科技部科技创新2030--“新一代人工智能”重大项目课题、国家自然科学基金面上项目、省重点研发等纵向项目，以及中石化、国家电网和比亚迪等企业项目20余项。在IEEE TCSVT、IEEE JBHI等国内外重要期刊发表学术论文40余篇，获发明专利5项，出版译著一部。曾获国家科技进步二等奖、省科技进步二等奖（排一）、中国仪器仪表学会科技进步一等奖（排二）等科技奖励。2021年任第五届图像计算与数字医学国际研讨会（ISICDM 2021）程序主席，2023年任第三届工业自动化、机器人与控制工程国际会议（IARCE 2023）执行主席。

工业视觉技术与应用论坛

第六会议室 8月17日 13:30-17:00



徐雪妙教授
华南理工大学

基于知识引导的场景智能理解和生成技术

报告摘要: 数据与知识的双轮驱动已成为人工智能发展的重要趋势之一。报告将面向工业场景,围绕视觉技术在实际应用中面临的场景复杂多变、数据缺失且质量不可控等难题,介绍一系列基于知识引导的场景智能理解和生成创新技术。

个人简介: 徐雪妙,博士,华南理工大学计算机科学与工程学院教授、博导、副院长,广州国际校区峻德书院副院长,广东省大模型与生成式人工智能工程技术研究中心主任,广东省卓越青年团队带头人,广东省特支青年拔尖人才,珠江科技新星。徐教授博士毕业于香港中文大学,研究方向为视觉智能、图形图像处理,及其在智能制造、智能交通等领域的应用。近年在国际重要期刊和会议发表论文70余篇,其中以第一或通信作者发表CCF A/Trans.论文40余篇,ESI高被引论文3篇;主持国家重点研发、国家自然科学基金、广东省及广州市重大专项等项目共15项,以第一完成人获得2022年中国图象图形学会科技进步二等奖,及2021年广东省科技进步二等奖。



宋克臣副教授
东北大学

工业钢管缺陷智能检测方法与应用

报告摘要: 工业钢管作为“现代工业的动脉”,广泛地应用于石油石化、航空航天、海洋工程等领域,缺陷智能检测系统作为保障产品质量的重要监测技术,已经成为工业钢管生产中必不可少的重要组成部分。本报告将分享课题组近年来在工业钢管缺陷检测与识别方面的研究工作:首先以无缝钢管内表面缺陷的检测需求为例,针对检测对象的属性特征,研发了一套检测设备,并分别针对缺陷“检不完”和“检不准”的难题,给出了具体解决方案;然后针对焊接钢管内部缺陷的检测需求,研究了一种疑似缺陷快速筛分方法、一种细粒度微小缺陷分割方法、以及智能化检测软件。

个人简介: 宋克臣,东北大学副教授,博士生导师,英国拉夫堡大学访问学者,沈阳市拔尖人才。入选2023全球前2%顶尖科学家榜单,担任SCI期刊《Computers, Materials & Continua》副主编、SCI期刊《Sensors》和《Coatings》客座编辑、国家自然科学基金同行通讯评审专家,捷克科学基金(GACR)同行通讯评审专家、教育部学位中心博士学位论文评审专家,IEEE TPAMI、IEEE TNNLS、IEEE TCYB等国际期刊审稿人。作为项目负责人主持国家自然科学基金与省部级项目5项,参与3项。在IEEE TII、IEEE TIE、IEEE TCSVT等国内外重要学术期刊和会议发表学术论文一百余篇,其中ESI高被引论文11篇,ESI热点论文1篇。曾获得“辽宁省自然科学学术成果奖”、多个国际期刊“杰出贡献审稿人”、“教育部博士研究生学术新人奖励”、以及“优秀硕士学位论文指导教师”等奖励。

工业视觉技术与应用论坛

第六会议室 8月17日 13:30-17:00



赵朝阳副研究员
中国科学院
自动化研究所

大模型时代下的工业视觉及行业应用

报告摘要：随着大模型技术的发展，其由语言模型逐步向融合了视觉以及音频的多模态方向进行技术演进，视觉语言大模型技术的进步使得其能够在广泛的下游场景任务中通过少样本的方式取得较好的泛化性和性能优势。本报告围绕着视觉大模型/视觉语言大模型在工业视觉缺陷检测方面的探索展开讨论，分享大模型技术在少样本、零样本缺陷检测方面的一些研究进展和探索工作，以及在实际场景中的落地应用情况。

个人简介：赵朝阳，博士，男，中科院自动化研究所副研究员，中国科学院正高级工程师，中科视语（北京）科技有限公司CTO。长期从事计算机视觉与人工智能算法相关落地应用研究，在视频内容分析、目标检测与识别、视觉预训练大模型、具身智能等方面开展相关研究工作。曾获2018年人工智能最高奖吴文俊奖，在人工智能领域CVPR, ICCV, ECCV, NeurIPS等国际会议以及TIP等国际期刊发表相关论文30余篇，获人工智能领域国内外竞赛冠军20余项，授权人工智能领域相关发明专利20余项。2018年起带领产业化团队围绕打造全国国产化人工智能重大基础设施建设的使命，建设以工业大模型为核心的开放人工智能平台，相关研究成果在智能制造、智慧能源、智慧医疗等领域进行了产品化和市场拓展，取得了良好的落地应用示范。



张志忠副教授
华东师范大学

受限工业场景下的视觉智能检测技术及应用

报告摘要：工业质量检验是电子制造、新能源、新材料品质管控的重要环节。以锂电池、3C生产为例：安全质检要求高（十亿分之一）、检测速度快（秒级）、精密部件多（数千）、组装工艺复杂（数十道工序）、产品迭代快，迫切需要质检智能化。本次报告聚焦受限工业场景下的工业视觉检测技术，结合新能源锂电池的质检需求，探讨细微缺陷成像及复原、样本受限下的工业视觉检测及模式受限下的缺陷持续学习方法，为视觉智能检测技术提供新的解决思路和关键技术支撑。

个人简介：张志忠，华东师范大学副教授，中国图象图形学学会动画与数字娱乐专委会委员，主要研究方向为机器学习以及计算机视觉，研究成果集中于多视图学习、自监督学习等。作为项目负责人，先后承担了国家自然科学基金青年项目、上海市人工智能重大专项，CAAI—Mindspore学术奖励基金，参与国家自然科学基金重大项目、国家“科技创新2030”重大项目等，在国内外权威期刊和顶级会议上共发表和接收学术论文40余篇，包括CVPR, ICCV, AAAI等。

医学影像智能分析与手术导航论坛

第五会议室 8月18日 09:00-12:00

论坛介绍

多模态医学图像分析与微创手术导航是精准诊疗发展的研究热点。本次论坛邀请全国知名的国家级人才和优秀青年学者，聚焦智能医学图像分析处理、智能医学成像、微创手术导航等重要的智能图像处理技术落地应用，向广大专业同行分享医疗领域人工智能的关键技术、创新难点及发展趋势，推动医学图像智能分析与手术导航的高质量发展。

论坛主席



杨健教授
北京理工大学

杨健，北京理工大学二级教授，博导。国家杰出青年科学基金获得者，国家新一代人工智能重大项目首席科学家、国家重点研发计划项目首席科学家，北京理工大学信息与电子学部学部委员，国家一级学科“光学工程”学科责任教授。兼任中国计算机学会数字医学分会副主任委员、中国仿真学会医疗仿真专委会副主任委员、北京医学会数字医学分会副主任委员、北京图像图形学会理事、北京光学学会理事、中国计算机学会计算机视觉专委会常务委员。

长期从事手术导航机器人、医学图像处理、计算机视觉、虚拟现实与增强现实、人工智能等方面的教学和科研工作。主持国家科技创新2030人工智能重大项目1项、国家重点研发计划项目2项、国家自然科学基金项目5项，作为核心人员参与国家973、863等项目10余项。在MIA、TMI等国际著名期刊上发表SCI论文198篇，获国家发明专利授权87项，其中11项实现转化应用。成果转化获国家三类医疗器械注册证2项、二类医疗器械注册证4项。研究成果获国家技术发明二等奖、教育部技术发明一等奖、人工智能学会科技进步一等奖、中国图象图形学学会科学技术一等奖等科研奖励10项。

医学影像智能分析与手术导航论坛

第五会议室 8月18日 09:00-12:00



龚启勇副院长
四川大学华西医院

精神影像学发展现状

报告摘要: 精神影像学 (Psychoradiology, 又名精神放射影像学) 是面向神经心理与精神疾病的放射影像 (Radiology) 学科及临床专科。该学科专业的发展得益于临床医学、心理、神经科学、生物医学工程、物理化学与计算机信息等多学科交叉融合与协同创新。其发展有助于进一步揭示精神心理疾病的脑病理机制, 并为相关临床诊疗提供客观评估方法。其亚专业方向“介入精神影像” (Interventional Psychoradiology, 又名介入精神放射影像) 的发展, 更为精神心理疾患提供精准的干预和治疗手段。本讲座就关于精神影像学基础理论的提出, 对疾病发病机制的探索、临床分型、发病风险与预后评估、影像引导的精准干预以及精神影像技术链等的发展现状与前景做了概述和总结。

个人简介: 龚启勇, 主任医师/教授, 四川大学华西医院影像学中心主任、华西厦门医院院长、华西临床医学院副院长、华西厦门医院院长、华西临床医学院副院长、国务院学位委员会学科评议组秘书长, 美国医学和生物工程院 Fellow, 长江学者特聘教授、国家杰出青年基金和国家自然科学基金委创新研究群体项目负责人、国家重点研发计划项目首席。曾任国际医学磁共振学会 (ISMRM) 精神磁共振学组主席、国际华人医学磁共振学会主席。主编北美放射医师培训教材《Psychoradiology》, 担任期刊 Psychoradiology 主编、Am J Psychiatry 副主编、Neuroradiology 编委。长期聚焦脑与精神心理疾病放射影像, 获国家自然科学基金二等奖和 4 项省部级科技奖一等奖。在 PNAS、Nature Mental Health、JAMA Psychiatry 等发表精神影像相关论著逾百篇, 被 Nature、Lancet 等引用逾万次; h 指数逾 100; 多次上榜科睿唯安“全球高被引科学家”, 并获吴阶平医药创新奖、全国创新争先奖、ISMRM Society Award、OCSMRM 杰出贡献奖等。



马占宇教授
北京邮电大学

数据受限条件下的视觉感知

报告摘要: 得益于在计算资源与硬件性能的数量级增长, 大数据和大模型成为当前人工智能发展的重要组成。然而, 在实际的生产环境中, 由于垂直专业领域的数据受限问题, 通用大模型在专业领域的能力仍然存在局限性。本报告将数据受限问题归纳为小规模、弱标注和低质量三个挑战, 并从面向小规模数据的精准特征挖掘、面向弱标注数据的鲁棒特征学习、面向低质量数据的高效型视觉感知三个方面汇报团队近期的研究进展。

个人简介: 北京邮电大学人工智能学院教授/发展规划处处长, 瑞典皇家理工大学院博士、博士后, 国家杰出青年科学基金获得者, 国务院学位委员会学科评议组成员, IEEE 高级会员, 亚太信号与信息处理协会杰出讲者, 中国图象图形学学会理事兼副秘书长、青工委副主任, 中国计算机学会计算机视觉专委会常务委员, 中国自动化学会人工智能与机器人教育专委会副主任。主要研究兴趣是人工智能、模式识别与机器学习基础理论与方法, 及其在计算机视觉、多媒体信号处理等领域的应用。在包括 IEEE TPAMI、CVPR 在内的顶级国际期刊和会议上发表论文多篇, 担任 IEEE TNNLS、IEEE TVT 等国际期刊编委 (Editor) 和 CVPR、AAAI AC 等, 授权发明专利 30 余项 (含美国发明专利一项); 先后主持国家自然科学基金委“杰青”、“优青”、联合重点等项目以及科技部“科技冬奥”重点研发计划课题、北京市自然科学基金重点项目等; 曾获中国人工智能学会“第七届吴文俊人工智能科学技术奖”一等奖, 中国图象图形学学会技术发明一等奖, “北京市科学技术奖”二等奖, 国际会议最佳论文奖等; 入选“北京市科技新星”计划; 获北京高校“优秀共产党员”、北京市教育系统“教书育人先锋”等荣誉称号。

医学影像智能分析与手术导航论坛

第五会议室 8月18日 09:00-12:00



方乐缘教授
湖南大学

眼科光学相干断层图像弱监督处理与分析

报告摘要: 深度学习已在多种图像应用领域获得突破性进展, 在医学影像领域的突出成果尤其令人振奋。目前, 现有深度学习方法仍有许多亟待改进的方面, 如需大量的数据与标签, 难以训练与收敛, 泛用性差等。本报告围绕分布式学习、弱监督学习、图像智能等新型深度学习方法, 介绍如何突破已有深度学习方法的局限。并针对人工智能在医学影像分析上的应用, 介绍深度学习如何与眼科OCT图像问题紧密结合, 基于实际问题设计深度学习模型解决多种眼科OCT图像分析中存在的临床问题。

个人简介: 方乐缘, 湖南大学岳麓学者特聘教授, 国家优青, 科睿唯安 (Clarivate Analytics) 全球“高被引科学家”, 爱思唯尔中国高被引学者, 湖南省创新领军人才。获得2019年度和2023年度国家自然科学基金二等奖2项 (排名第二和第四)、IEEE GRSS最高影响力论文奖、湖南省自然科学一等奖2项等奖项。担任SCI期刊IEEE Transactions on Image Processing、IEEE Transactions on Neural Networks and Learning System、IEEE Transactions on Geoscience and Remote Sensing、Neurocomputing等期刊编委。现主要从事深度学习、弱监督学习以及在遥感图像处理与分析等方面的研究。研究成果在国际权威期刊和会议发表论文200余篇, 其中SCI期刊发表论文100余篇 (IEEE TPAMI、IJCV、TIP等本领域顶级期刊论文70余篇), 国际权威会议论文30篇, Google scholar引用14500余次, ESI高被引22篇, ESI热点论文4篇。主持国家自然科学基金联合重点、国家重点研发课题等项目。



刘旦根教授
南昌大学

面向高维医学成像数据建模的扩散模型

报告摘要: 扩散模型是一种生成模型, 其可基于噪声初始化迭代生成高分辨率图像, 本报告主要介绍利用扩散模型进行迭代医学重建的最新进展, 包括三部分內容: 从DAE到DSM的学习, 从图像域到投影域的学习, 以及从大型数据集到小型数据集的学习, 并就以上方法在快速核磁共振成像、稀疏视图CT和光声断层扫描等方面的突出应用。

个人简介: 南昌大学教授、博导, 信息工程学院执行院长、人工智能工业研究院院长。IEEE Senior member。2012年获上海交通大学工学博士学位, 期间在中科院-德国马普学会计算生物所和中科院深圳先进技术研究院客座交流。2015年至2017年在美国UIUC和加拿大卡尔加里大学做博士后。致力于智能成像与视觉显示的系统开发与算法研究, 发表IEEE Trans 和成像与视觉显示期刊论文80余篇。为IEEE和CCF等数十个国际国内学术组织的专委会委员, CT理论与应用研究、电子与信息学报等多个期刊的编委。主持国家重点研发计划课题、自然科学基金、企业委托项目等20余项 (含国家优青项目)。参编专著5部。获江西省自然科学、吴文俊人工智能科技进步等科技奖励和江西省教学名师、省级示范研究生导师创新团队等教学荣誉近10项, 指导学生获生物医学工程学会青年优秀论文研究生奖、人工智能+和中国国际互联网+金奖等数10项。

医学影像智能分析与手术导航论坛

第五会议室 8月18日 09:00-12:00



宋红教授
北京理工大学

医学影像智能分析与处理

报告摘要：微创手术具有创伤小、患者恢复快、术后并发症少等显著优势，已广泛应用于肿瘤外科、头颈外科、骨科等疾病的诊疗。手术导航为微创手术提供“慧眼”，可大幅提升手术的精准度、安全性和效率。本报告聚焦于手术导航中的术前智能规划、术中智能感知、目标智能导引等核心问题，介绍课题组在影像分割建模、图像弹性配准融合、运动形变感知补偿等多模态影像智能分析方面的研究进展，以及研制的多模态影像融合增强现实手术导航系统与临床转化应用。

个人简介：宋红，北京理工大学计算机学院教授，教育部高等学校教学指导委员会委员、中国人工智能学会智能服务专委会常委。长期从事人工智能辅助诊断、图像分析处理、增强现实手术导航研究。作为负责人主持国家自然科学基金重点项目、面上项目、国家重点研发计划课题、“新一代人工智能”重大专项课题等10余项。在TIP、TFS、JBHI等领域权威期刊发表SCI论文90余篇，参与制定诊疗标准1项；申请/授权国家发明专利70余项，主持研制了3套具有自主知识产权的多模态图像引导手术导航系统，成果转化获国家三类/二类医疗器械注册证6项，在200余家医院和10余家企业实现创新应用。获吴文俊人工智能科技进步一等奖、中国产学研合作创新成果优秀奖。

多模态大模型论坛

第六会议室 8月18日 09:00-12:00

论坛介绍

当前，具有强大学习能力、上下文理解和多模态处理能力的AI大模型正迅猛发展，给学术界和工业界都带来了革命性的变化，基于多模态的预训练大模型将成为人工智能基础设施，以及新质生产力的代表。未来，大模型作为AI基础设施，将实现文本、图象、音频、视频统一知识表示，并朝着能推理、能问答、能总结、能创作的认知智能方向演进。本次论坛设置主旨演讲和对话环节，多形式展现当下多模态大模型的最新技术研究和行业成果，并讨论未来发展趋势。

论坛主席



黄庆明教授
中国科学院大学

中国科学院大学二级教授、讲席教授、博士生导师，国家杰出青年科学基金获得者，百千万人才工程国家级人选，享受国务院政府特殊津贴。主要研究领域为多媒体分析、计算机视觉、模式识别、机器学习等，累计发表学术论文600余篇，其中IEEE/ACM汇刊和CCF-A类会议论文300余篇，谷歌学术引用30000余次。承担新一代人工智能国家科技重大专项、国家自然科学基金重点项目和重点国际合作项目、973课题、863课题等重大项目，相关成果获得吴文俊人工智能自然科学一等奖、中国图象图形学学会自然科学一等奖、教育部科技进步一等奖等奖励，是IEEE Fellow，CCF会士，IEEE CASS北京分会主席，CCF理事，CCF多媒体技术专业委员会主任，中国图像图形学学会常务理事，北京图象图形学学会副理事长。

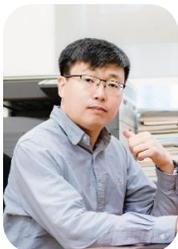


李亮副研究员
中国科学院计算
技术研究所

博导，国家自然科学基金委优秀青年基金获得者，研究方向为生成式人工智能和跨模态学习等，发表TPAMI等汇刊和CCF-A类会议论文70余篇，谷歌引用3600余次，获国际会议最佳论文奖2次，6项工作入选ESI高被引论文。主持/参与科技部科技创新2030重大项目、973课题、NSFC重点和面上项目等20余项，参与组织7次IEEE/ACM国际会议和5次国内会议，国际期刊Electronics和JCST编委，ACL、EMNLP和ACM MM等会议的领域主席，ACM SIGMM中国分会高校合作负责人，获2022年浙江省科学技术进步一等奖、2020年吴文俊人工智能自然科学一等奖。与新华社共同起草制定了的全球首个内容自动化生产标准《机器生产内容自动化分级》，相关技术在华为、淘宝、大华、华数传媒等使用，服务于2022北京冬奥会等重要任务。

多模态大模型论坛

第六会议室 8月18日 09:00-12:00



叶齐祥教授
中国科学院大学

视觉表征模型的结构设计与物理启发

报告摘要: 分析了局部卷积运算与全局注意力运算的互补性, 将局部特征与全局特征耦合形成Conformer网络结构, 显著增强视觉表征能力, 提升表征模型的性能下限。探讨了局部卷积运算造成的自监督学习信息泄露问题, 提出了Token Merging操作, 突破卷积或局部运算的局部约束, 形成高效分层Transformer表征(HiViT)与全预训练的Transformer金字塔网络(iTPN)。在ImageNet分类任务上, iTPN-Base, iTPN-Large, iTPN-Huge 分别达到了88.0%, 89.2%, 89.7%的Top-1分类精度。在模型结构设计的基础上, 探索了物理原理启发的基础模型, 从热传导视角探索表征模型, 取得了初步效果, 代码: github.com/pengzhiliang/Conformer、github.com/sunsmarterjie/iTPN、github.com/MzeroMiko/VMamba。

个人简介: 叶齐祥, 中国科学院大学特聘教授、国家杰出青年基金获得者、中科院卢嘉锡青年人才奖获得者、中国科学院优秀博士生导师、CVPR2023、NeurIPS2023、ICLR2024 Area Chairs、国际期刊IEEE TITS, IEEE TCSVT编委。主要进行视觉表征学习与目标感知方向的研究, 在CVPR, ICCV, NeurIPS等国际会议与TPAMI, TNNLS, TIP等期刊发表论文150余篇。承担了自然科学基金重点项目、开发的高精度目标感知方法支撑了华为、航天等单位的应用系统。曾获中国电子学会自然科学一等奖。培养多名博士生获中科院院长奖、中国科学院百篇优秀博士论文、博士后创新人才计划支持等。



魏云超教授
北京交通大学

从3D到4D, 快速且时空一致的4D内容生成探索

报告摘要: 近年来, 借助2D生成大模型的发展, 3D内容生成发展迅速, 然而由于缺少大量4D数据集和强有力的视频预训练模型, 动态3D(4D)内容生成却少有探索。本报告旨从3D生成视角出发, 研究快速且时空一致的4D内容生成, 将会介绍最新的4D生成工作——4DGen和Diffusion4D。通过探究如何设计时空上高保真度的监督信号, 以及如何使用4D数据集对视频生成模型进行微调, 实现从文本、图像或视频控制信号到4D内容的快速高质量生成。

个人简介: 魏云超, 北京交通大学计算机学院教授、副院长, 国家高层次人才计划获得者。曾在新加坡国立大学、美国伊利诺伊大学厄巴纳-香槟分校、悉尼科技大学从事研究工作。入选《麻省理工科技评论》中国区“35岁以下科技创新35人”(MIT TR35 China)、百度全球高潜力华人青年学者、《澳大利亚人》TOP 40 Rising Star; 获世界互联网大会领先科技奖(2023)、教育部高等学校自然科学奖一等奖(2022)、中国图象图形学学会科技技术奖一等奖(2019)、澳大利亚研究委员会青年研究奖(2019)、IBM C3SR最佳研究奖(2019)、计算机视觉世界杯ImageNet目标检测冠军(2014)及多项CVPR竞赛冠军; 发表CCF A类期刊/会议论文100多篇, Google引用超20000次。目前主要研究方向包括面向非完美数据的视觉感知、多模态数据分析、生成式人工智能等。

多模态大模型论坛

第六会议室 8月18日 09:00-12:00



刘静研究员
中国科学院
自动化研究所

多模态预训练模型的研究与应用

报告摘要: 近年来,从预训练模型到预训练大模型,从文本、音频、视觉等单模态大模型,到现在的图文、图文音等多模态预训练大模型,无论在学术界还是企业界预训练模型都得到了广泛关注与爆发式发展。多模态预训练通过联合图文音等多模态内容进行模型学习,其发展在多模态理解、搜索、推荐、问答,语音识别与合成,人机交互等应用领域中具有潜力巨大的市场价值。本报告主要包含三方面内容:分析多模态预训练模型的重要性与必要性;回顾当前多模态预训练的最新研究进展;多模态预训练模型主要应用场景与未来展望。

个人简介: 刘静,中国科学院自动化所研究员/博导,中国科学院大学岗位教授,国家优青获得者。研究方向多模态分析与理解,紫东太初大模型。曾获中国电子学会自然科学一等奖,图像图形学会科学技术二等奖,2022年世界人工智能大会“卓越人工智能引领者奖SAIL”。承担或参与多项国家自然科学基金项目、国家973课题、国家基金重大研究计划、国家重点研发等。已发表高水平学术论文150余篇,谷歌学术引用11000+次,SCI他引次数4000+次,其中有三篇被ESI列为Top1%高被引论文。在视觉计算相关领域的多项国际学术竞赛中荣获冠军10+项。



邵睿教授
哈尔滨工业大学
(深圳)

“九天”多模态大模型

报告摘要: 本报告首先将总览现有语言大模型、视觉大模型与多模态大模型的发展历程与其惊人的通用人工智能潜力,其次将重点剖析当前多模态大模型的主流模型架构与预训练任务,最后介绍我们研发的九天多模态对话大模型的主要功能与实现方法。

个人简介: 邵睿,哈尔滨工业大学(深圳)计算机科学与技术学院教授,博士生导师,入选2023年国家级青年人才计划。主要研究方向为多模态大模型、计算机视觉,以第一作者/通讯身份发表CCF-A会议与IEEE/ACM汇刊等高水平论文20余篇,包括:TPAMI, IJCV, TIFS, TNNLS, CVPR, ECCV, AAAI, ACM MM等,并多次担任CVPR, ICCV, ECCV, TPAMI, IJCV等顶级会议和期刊审稿人。GitHub代码仓库关注星数超过1000。主持/参与国家自然科学基金面上项目、青年项目、广东省自然科学基金面上项目等。担任计算机视觉旗舰会议BMVC领域主席。现负责构建哈工大深圳“九天”多模态大模型。

多模态大模型论坛

第六会议室 8月18日 09:00-12:00



姚霆联合创始人
兼CTO

HiDream.ai (智象未来)

多模态内容生成：掀起未来创意无限可能

报告摘要：生成式人工智能（AIGC）已成为技术和产业界竞相追逐的热点，全球各大科技企业都在积极拥抱 AIGC，不断快速推出相关基础模型、平台和应用。AIGC 将会从单一模态迅速迈向多模态，我们将会迎来多模态基础模型与应用的爆发期。本次分享，将回顾 AIGC 多模态基础模型进展，分析多模态内容生成技术与应用，并展望相关挑战与技术趋势。

个人简介：姚霆，生成式人工智能初创公司智象未来（HiDream.ai）联合创始人兼CTO。获2022 IEEE ICME Multimedia Star Innovator，2019 ACM SIGMM Rising Star，2019 IEEE TCMC Rising Star，2022中国智能计算科技创新人物。发表论文100余篇，谷歌引用1.5万余次，先后10余次获得国际学术竞赛冠军，获评2022 年度中国图象图形学学会科技进步奖一等奖。设计了视频分析领域标准的3D卷积神经网络Pseudo-3D Network，构建的业界首个大规模视频文本数据集MSR-VTT被全球五百余研究机构的学者下载使用，并研发了多款数百万日活用户的商业产品。曾任京东科技算法科学家和微软研究院研究员。

技术报告

银杏大厅 8月17日 11:45-12:00



王小磊解决方案架构师
北京超算

算力服务加速数字人与内容生成技术创新发展

报告摘要：在数字人与内容生成技术日新月异的今天，算力服务成为推动其创新发展的关键力量。本报告深入探讨了算力服务如何为数字人技术提供强大的支撑，推动其在交互与内容生成等方面的更进一步进展。同时，我们也探讨了如何通过算力选型和优化设计等算力服务，实现算力资源利用率的提升，助力更高效、更逼真的内容创作。展望未来，算力服务的持续升级将为数字人与内容生成技术带来无限可能，助力各行各业实现数字化转型与创新。

个人简介：王小磊，北京超算解决方案架构师，负责产品技术与解决方案，打造基于超算、智算的算力基础设施服务。历任高性能计算系统工程师、方案工程师，长期从事高性能计算机架构设计、技术研究、国产芯片与应用整合等工作，并致力于建设人工智能计算平台建设，运用超算技术助力百亿、千亿参数的大模型训练，倡导通过应用运行特征分析技术提供更加高效、稳定、可靠、高性价比的AI智算云服务，并实现技术整合、产品革新与应用落地。科技创新人物。发表论文100余篇，谷歌引用1.5万余次，先后10余次获得国

第五届数学建模教学与培训研讨会探索数学世界及数学建模

第三会议室 8月18日 09:00-12:00



李雷教授
南京邮电大学

数学建模的工作准备与竞赛训练

报告摘要: 什么叫数学模型和数学建模、数学建模的特点、数学模型及建模分类、数学建模工作的步骤、数学建模工作指导教师的基本能力、数学建模竞赛的校内筹备、数学建模竞赛的校内组织、如何进行数学建模竞赛训练、如何合理安排竞赛的时间、如何撰写竞赛论文。

个人简介: 李雷，安徽砀山人，哈尔滨工业大学数学博士，南京邮电大学数学教授，曾任高校教务处处长、校长助理等职务，省政协委员。现为南京邮电大学教学督导组组长、教学工作委员会委员、督监委委员。为国家和省多个一级学科学会常务理事、理事和会员。国家和省教学成果奖、一流专业课程

评审专家，国家和省本科教学评估专家组组长，国家重大研究计划项目、国家自然科学基金项目评审专家。十余所高校教学督导。全国大学生数学建模竞赛指导教师和阅卷专家，指导学生获得数学建模竞赛一等奖、二等奖百余项。亚太地区大学生数学建模竞赛出题、评审专家。主要学术研究方向：1、人工智能，主要是人工神经网络、智能优化算法、智能控制。2、大数据分析，主要是视觉认知、数据可视化、压缩感知。3、模式识别与智能系统，主要是深度学习、数据挖掘。4、信息网络安全检测，主要是智能信息安全检测与处理。5、应用数学，主要是分析数学在网络信息与数据分析中的应用理论与方法。主持和参与国家自然科学基金等国家级省部级科研项目十多项。发表学术论文近二百篇，SCI、EI检索50余篇。获省部级奖励十余项。出版专著多部。指导硕士博士研究生百余人。



李红军教授
北京林业大学

大模型与数学建模

报告摘要: 介绍应用大模型进行数学建模和参加数学建模竞赛时应该注意的一些问题;介绍利用大模型开展数学建模教学的实践与思考。

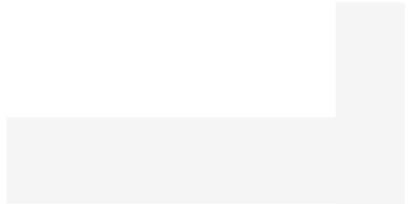
个人简介: 李红军，北京林业大学理学院教授，硕士研究生导师。2003年毕业于北京科技大学应用科学学院，获应用数学硕士学位。2012年毕业于中科院自动化研究所模式识别国家重点实验室&中法信息自动化与应用数学联合实验室，获博士学位。研究兴趣主要集中于植物几何建模、图形图像等领域。曾主持完成国家自然科学基金面上项目1项。在国内外学术刊物上发表科研论文70余篇，获中国发明专利授权十余项，登记软件著作权十余项。担任学校建模培训指导教师，曾获“北京市大学生数学建模与计算机应用竞赛优秀指导教师”称号，国家一流课程《空间解析几何》负责人。

■ 论文列表

序号	论文题目	论文作者
1	Remote Sensing Scene Classification Method Based on Multi-Scale Local Attention Network	Yi Miao,JunJie Wang, MengMeng Zhang, XiaoMing Xie,Wei Li
2	Multi-View Intention Recognition in Face-to-Face Communication	Pukun Chen, Dongdong Weng, Xiaonuo Dongye
3	Study on different apple ripeness detection based on improved YOLOv5	Yulu Cai, Fangchu Wanghan, Anping Shen, Weizheng Kong, Qianzhe Zhang
4	HSINet: A Hybrid Semantic Integration Network for Medical Image Segmentation	Ruige Zong, Tao Wang, Xinlin Zhang, Qinquan Gao, Dezhi Kang, Fuxin Lin, Tong Tong
5	Learning Features by Minimizing the Interframe Differences	Dong Zhao, Dan Zhang
6	An Iterative Method for Single Image Stripe Nonuniformity Correction	Gaojin Wen, Hongmin Wang, Yun Xu, Linpeng Li, Zhikai Wang, Ziwei Zhou, Pu Huang, Xingjian Li, Liang Long, Yuedong Zhang, Limin Wu
7	Optimized Recognition and Depth Estimation for Fruit Picking Robots	Jiancheng Liu, Tianle Jin, Jiaolai Wen
8	VT Prompt Tuning: Improving Portrait Generation With Combined Visual Prompt and Text Prompt in Stable Diffusion Model	Miao Hong, Weidong Fang, Hao Wang, Jinghe Zhang
9	VTR-former: Vision Token Rolling Transformer for Weakly Supervised Temporal Text Localization	Zeyu Xi ,Xinlang Zhou , Zilin Liu, Lifang Wu
10	CMME-YOLO: Composite Microscale Multi-Stream Enhanced Model for PCB Defect Detection	Zhijia Xie, Xiaowei Zou, Zhiwu Zhou
11	BM3D-UGanNet: A Hybrid Deep Learning Network for Low-Light Thangka Image Enhancement	Quanhong Peng, Dan Zhang, Mingquan Zhou, Jianpeng Zhang, Meihua Song,Ning Wang, Chenhao Xu

12	Research on Enhancing Garbage Recognition in Low Illumination Environments Using Attention Mechanisms	Zhenwei Lv, Yapeng Diao, Chunnian Zeng, Weiping Wang, Shufan An, Xiangyu Li
13	Stagewise Positional Encoding for Implicit Neural Representation of Image	Chengyang Yan, Linsen Chen, Qiu Shen
14	Multi-behavior Recommendation with Hypergraph Contrastive Learning	Bing Yang, Xiaolei Guo, Litao Shang, Zhigang Zhang, Yanxian Bi, Xiaopeng Gu
15	Image Enhancement Based on a Diffusion Model Guided by No-Reference Image Quality Assessment	Keqing Zhang, Lifang Wu
16	Adaptive Pixel Pair Generation Strategy for Image Matting Methods based on Pixel Pair Optimization	Jiamin Zheng, Wen Wen, Yihui Liang, Fujian Feng, Xiang Xu
17	Phase Error Correction Algorithm Based on Complementary Gray Code and Reverse Error Compensation	Pengjie Zhang, Shaoping Wang, Bin Kong
18	Semantic segmentation of remote sensing images based on U-Net	Peiquan Xie, Yanzhao Zhu, Lei Guo, Chunhong Mo, Limei Shi, Peiru Wang, Xiaotong Huang, Yaxuan Zhu, Jian Liu
19	Application of Swin_Transformer in Segmentation of Remote Sensing Objects	Ran Peng, Yu Peng, Xingchen Lan, Zitong Qiu, Bo Wang, Xianjin Chen, Shuai Gao, Tao Liu
20	EdgCNN:Thangka line drawing extraction based on CNN	Meihua Song, Dan Zhang, Jianpeng Zhang, Ning Wang, Quanhong Peng, Chenhao Xu
21	MRFFA-Net: A Multi-scale Residual Feature Fusion and Attention Mechanisms for Retinal Vessel Segmentation	Lu Cao, Guangwu Liu, Junying Gan, Jiancheng Li, Xiquan He, Min Luo
22	Long Sequences Generation for Motion Diffusion Models	Yang Zhou, Dongdong Weng
23	Fire detection based on flame enhancement for weak fires	Kuan Chen, Wen Wen, Fujian Feng, Xiang Xu, Yihui Liang

24	Microclimate Regulation in Glass Greenhouses: Simulation of Wind Velocity and Temperature Fields and Design Model of the Greenhouse Fans	Kaixin Tan, Yuhan Huang, Chenyang Jin, Jiantao Li
25	HDR-TDC: High Dynamic Range Imaging with Transformer Deformable Convolution	Shuaikang Shang, Xuejing Kang
26	Gaussian Replacement: Gaussians-Mesh Joint Rendering for Real-time VR Interaction	Xiaonuo Dongye, Hanzhi Guo, Yihua Bao, Dongdong Weng
27	Visualization of Spatial and Temporal Information of Railway Heritage in Modern China	Huaping Shen, Hanyue Wang
28	Optical flow generation method based on multi-scale feature enhancement of compressed video bitstream information	Chuang Wang, Hongwei Lin, Ziyi Tian, Haijiao Ren, Kunyu Gao, Qicai Jia
29	Contrastive Learning-Based Dual Path Fusion Network for Group Activity Recognition	Ronggui Fu, Zequan Li, Ye Xiang, Lei Lu, Ruixuan Ding, Lifang Wu
30	Enhancing Fine-Tuning Performance of Text-to-Image Diffusion Models for Few-Shot Image Generation Through Contrastive Learning	Yan Lin Zhu, Peipei Yang
31	Generalized diffusion models for non-Gaussian noise	Tianqian Zhou, Chun Fang, Xiaozhu Lin
32	Defect Detection based on Normalized Attention Mechanism and Multi-level Feature Fusion	Yiyan Chen, Zexin Zheng, Jialin Tang, Wanyu Feng, Honghui Lin, Lin Shounan
33	Exploration of New Energy Electric Vehicle Development Based on SSA-SVR Model	Xuanzhi Wang, Yajun Huang, Zhenghao Li
34	基于多头注意力增强的三阶段图像修复算法	赵阳, 丁建伟
35	生成式对抗对比学习变化检测方法	郭程旺, 张蒙蒙, 张宇翔
36	改进 YOLOv8s 的多尺度信息融合小目标检测	仲尚, 马丽, 刘文哲, 李雨豪



学会微信公众号



会议网站