



中国图像图形学会
CHINA SOCIETY OF IMAGE AND GRAPHICS



昆明理工大学
KUNMING UNIVERSITY OF SCIENCE AND TECHNOLOGY



云南师范大学
YUNNAN NORMAL UNIVERSITY



云南大学
YUNNAN UNIVERSITY

中国图像图形学会 青年科学家会议 *2026*

会议手册 Conference Handbook

云南·昆明

2026年8月3-6日



主办单位

中国图像图形学会

承办单位

昆明理工大学 云南师范大学 云南大学 CSIG青年工作委员会

协办单位

AutoDL

钻石合作

白泽智研（南京）科技有限公司 全球校园人工智能算法精英大赛

铂金合作

云南奥海信息技术有限公司 北京英博数科科技有限公司

金牌合作

北京九章云极科技股份有限公司

银牌合作

上海卓越睿新数码科技股份有限公司 上海长城汽车科技有限公司
合肥中科类脑智能技术有限公司

单项合作

趋动科技（上海）有限公司 云南天奥信息技术有限公司
昆明飞迅科技技术有限公司 云南云耕科技有限公司
昆明松影数码科技有限公司

目录 contents

○ 大会介绍·····	01
○ 大会组委会·····	02
○ 大会日程总览·····	05
○ 大会主旨报告讲者·····	17
○ 大会分论坛·····	22
 ○ 8月4日09:00-12:00	
讲习班：多模态智能体·····	23
分论坛：图像视频理解·····	26
分论坛：可视化与渲染·····	31
分论坛：开放场景下的多模态图像融合·····	36
分论坛：智能交互数字人·····	41
分论坛：自动驾驶与智能座舱·····	46
分论坛：大模型基础理论·····	51
分论坛：三维视觉与空间智能·····	55
分论坛：多模态大语言模型·····	61
青托论坛·····	66
优博论坛·····	72
 ○ 8月4日14:00-17:30	
讲习班：多模态遥感理解·····	80
分论坛：以人为中心的视觉计算·····	83
分论坛：具身智能·····	87
分论坛：视觉基础模型·····	92
分论坛：扩散模型·····	97
分论坛：大模型安全与可信人工智能·····	103
学术服务论坛：人工智能基础·····	108
学术服务论坛：计算机视觉I·····	113
学术服务论坛：交叉研究·····	118
学术服务论坛：计算机视觉II·····	123

○ 8月5日14:00-17:30

讲习班：生成式AI与智能创作·····	127
分论坛：图像复原与增强·····	130
分论坛：本地特色论坛（小语种OCR）·····	134
分论坛：医学影像处理·····	138
分论坛：AI For Science·····	144
分论坛：国产大模型的部署与优化·····	149
分论坛：视频世界模型·····	154
分论坛：大模型的高效训练与推理·····	159
博士生论坛·····	163
女科学家论坛·····	171
学术新星论坛·····	176

○ 8月6日09:00-12:00

讲习班：空间基础模型·····	184
分论坛：大模型评测·····	187
分论坛：视觉认知与情感计算·····	192
分论坛：遥感影像处理·····	198
分论坛：低空智能视觉·····	203
分论坛：AI Coding Agent·····	207
分论坛：图像视频生成与编辑·····	212
分论坛：多智能体·····	216
国自然博士生论坛·····	220
托举工程博士生专项论坛·····	226

大会介绍

中国图像图形学会青年科学家会议是由中国图像图形学会青年工作委员会发起的学术会议。会议面向国际学术前沿与国家战略需求，致力于支持图像图形领域的优秀青年学者，为青年学者们提供学术交流与研讨的平台，促进青年学者之间的交流与合作创新。会议同时邀请产业界专家与青年学者做深入交流，鼓励图像图形领域的“产学研”深度合作。

大会将于2026年8月3日-6日在昆明云安会都举办，由中国图像图形学会主办，昆明理工大学、云南师范大学、云南大学、中国图像图形学会青年工作委员会联合承办，AutoDL 协办。

本次会议是中国图像图形学领域的重要会议，已经列入中国科协的重要会议指南。大会参会人数预计将超过2000人，其中包括300多位国家级人才。会议将聚焦于图像图形领域前沿和热点，涵盖重要的研究领域和研究方向，各类高水平报告将超过200场，将是一场精彩纷呈的学术盛宴。



大会组委会

大会主席

王耀南（湖南大学）
甘健侯（云南师范大学）

余正涛（昆明理工大学）
彭宇新（北京大学）

程序委员会主席

王红斌（昆明理工大学）
杨 云（云南大学）
王瑞平（中国科学院计算技术研究所）

邵永航（云南师范大学）
董 晶（中国科学院自动化所）

组织委员会主席

李华锋（昆明理工大学）
高盛祥（昆明理工大学）
李英娜（昆明理工大学）

贾 伟（合肥工业大学）
蔡 青（合肥工业大学）
丁家满（昆明理工大学）

赞助主席

余鹏程（昆明理工大学）
钱文华（云南大学）
雷印杰（四川大学）
郭宗辉（中国海洋大学）

王林飞（云南师范大学）
谢 诚（云南大学）
任文琦（中山大学）
元 辉（山东大学）

宣传主席

朱 磊（同济大学）
马明睿（昆明理工大学）
李梦迪（云南师范大学）
谭明奎（华南理工大学）

李 凡（昆明理工大学）
金 鑫（云南大学）
崔金荣（华南农业大学）

讲习班主席

李冠彬（中山大学）

舒振球（昆明理工大学）

杨 欣（华中科技大学）

邓 欣（北京航空航天大学）

学术前沿论坛主席

赵才荣（同济大学）

张亚飞（昆明理工大学）

张鼎文（西北工业大学）

蒋 铨（北京航空航天大学）

沈 为（上海交通大学）

丛润民（山东大学）

许倩倩（中国科学院计算技术研究所）

郭裕兰（中山大学）

青托论坛主席

魏秀参（东南大学）

张 奇（同济大学）

彭春蕾（西安电子科技大学）

张勇丙（昆明理工大学）

优博论坛主席

马 超（上海交通大学）

傅慧源（北京邮电大学）

晏轶超（上海交通大学）

王青旺（昆明理工大学）

学术新星论坛主席

阚美娜（中国科学院计算技术研究所）

邹博超（北京科技大学）

刘 朋（云南大学）

胡 鹏（四川大学）

刘笑宏（上海交通大学）

博士生论坛主席

户战选（云南师范大学）

贾 旭（大连理工大学）

杨 鹏（华中科技大学）

张 杰（中国科学院计算技术研究所）

女科学家论坛主席

相 艳（昆明理工大学）

董燕妮（武汉大学）

冯 婕（西安电子科技大学）

高联丽（电子科技大学）

托举工程博士生专项论坛主席

朱翔昱（中国科学院自动化研究所）
李华光（云南大学）

徐婧林（北京科技大学）

国自然青年学生项目论坛主席

金 枝（中山大学）

梁孔明（北京邮电大学）

学术服务论坛主席

舒祥波（南京理工大学）
张 正（哈尔滨工业大学（深圳））
袁国武（云南大学）

张敬林（山东大学）
杨 勋（中国科学技术大学）

Poster主席

董 能（昆明理工大学）
徐天阳（江南大学）
江 倩（云南大学）

王金峰（昆明理工大学）
刘 羽（合肥工业大学）

网站主席

付湘琼（昆明理工大学）

赵 洋（合肥工业大学）

注册主席

谢 宁（昆明理工大学）
鄢杰斌（江西财经大学）

王路路（昆明理工大学）
张怀文（内蒙古大学）

财务主席

范林川（昆明理工大学）

段继忠（昆明理工大学）

大会日程总览

🕒 2026年8月3日14:00-20:00

国 内 容

📍 地 点

会议报到+注册

云安会堂一楼大厅

🕒 2026年8月4日09:00-12:00

国 内 容

📍 地 点

讲习班：多模态智能体

主席：李冠彬（中山大学）、徐天阳（江南大学）

讲者：林衍凯（中国人民大学）、岳翔宇（香港中文大学）、杨思蓓（中山大学）

102会议室

分论坛：图像视频理解

主席：白慧慧（北京交通大学）、刘 洋（北京大学）、李 锋（合肥工业大学）

讲者：查正军（中国科学技术大学）、施柏鑫（北京大学）
房 伟（全球校园人工智能算法精英大赛组委会秘书长）、徐 行（同济大学）
林春雨（北京交通大学）、刘 洋（北京大学）

101会议室

分论坛：可视化与渲染

主席：汪云海（中国人民大学）、胡瑞珍（深圳大学）
高 林（中国科学院计算技术研究所）

讲者：王莉莉（北京航空航天大学）、陈雪锦（中国科学技术大学）
连宙辉（北京大学王选计算机研究所）、宋维涛（北京理工大学）
龙霄潇（南京大学）

201会议室

分论坛：开放场景下的多模态图像融合

主席：吴小俊（江南大学）、刘 羽（合肥工业大学）
赵文达（大连理工大学）、刘晋源（大连理工大学）

讲者：马佳义（武汉大学）、江俊君（哈尔滨工业大学）、洪丹枫（东南大学）
邓 欣（北京航空航天大学）、李华锋（昆明理工大学）

202会议室

内 容	地 点
分论坛：智能交互数字人	
主席： 邹博超（北京科技大学）、晏轶超（上海交通大学）、马惠敏（北京科技大学） 讲者： 李 坤（天津大学）、张盛平（哈尔滨工业大学）、姚 超（北京科技大学） 朱 昊（南京大学）、晏轶超（上海交通大学）、毕秀丽（重庆邮电大学）	204会议室
分论坛：自动驾驶与智能座舱	
主席： 韩文程（山东大学）、董性平（武汉大学） 高学浩（西北工业大学）、杨思蓓（中山大学） 讲者： 杨 健（南京理工大学）、王 栋（大连理工大学） 许春景（华为技术有限公司/车BU） 马 超（上海交通大学）、丁恒辉（复旦大学）	302会议室
分论坛：大模型基础理论	
主席： 刘 勇（中国人民大学）、胡 鹏（四川大学）、李 健（北京师范大学） 讲者： 孟德宇（西安交通大学）、侯臣平（国防科技大学）、刘日升（大连理工大学） 李崇轩（中国人民大学）、李 帅（上海交通大学）	309会议室
分论坛：三维视觉与空间智能	
主席： 郭裕兰（中山大学）、雷印杰（四川大学） 胡心亭（中国科学技术大学）、姜海勇（中国科学院大学） 讲者： 杨 易（浙江大学）、张兆翔（中国科学院自动化研究所） 夏桂松（武汉大学）、戴玉超（西北工业大学） 刘家瑛（北京大学）、齐晓娟（香港大学）	401会议室
分论坛：多模态大语言模型	
主席： 常晓军（中国科学技术大学）、黄 高（清华大学）、杨 学（上海交通大学） 讲者： 沈春华（浙江大学）、冯 洋（中国科学院计算技术研究所） 高广亮（智慧树网）、刘 伟（上海交通大学） 范 琦（南京大学）、钟志航（上海交通大学）	402会议室

内 容		地 点
分论坛：青托论坛		
主席： 魏秀参（东南大学）、彭春蕾（西安电子科技大学） 张 奇（同济大学）、张勇丙（昆明理工大学）		207会议室
讲者： 梁小丹（中山大学）、余旻婧（天津大学）、鲁 国（上海交通大学） 吴 乐（合肥工业大学）、谢 诚（云南大学）、杨 旭（西安电子科技大学）		
分论坛：优博论坛		
指导专家： 马惠敏（北京科技大学）、彭宇新（北京大学） 董 晶（中国科学院自动化研究所）、贾 伟（合肥工业大学） 赵才荣（同济大学）		206会议室
主席： 马 超（上海交通大学）、晏轶超（上海交通大学） 傅慧源（北京邮电大学）、王青旺（昆明理工大学）		
讲者： 李 勇（东南大学）、晏轶超（上海交通大学）、王洪松（东南大学） 鄢杰斌（江西财经大学）、王世鹏（西安交通大学）、姜 林（电子科技大学）		
🕒 2026年8月4日12:00-14:00		
午餐&Poster		
🕒 2026年8月4日14:00-17:30		
内 容		地 点
讲习班：多模态遥感理解		
主席： 周 峻（格里菲斯大学）、舒振球（昆明理工大学）		102会议室
讲者： 肖 亮（南京理工大学）、高连如（中国科学院空天信息创新研究院） 谢卫莹（西安电子科技大学）		

内 容	地 点
分论坛：以人为中心的视觉计算	
主席： 杨 勋（中国科学技术大学）、韩军功（清华大学）、陈超奇（深圳大学） 讲者： 杨 欣（华中科技大学）、涂志刚（武汉大学）、刘安安（天津大学） 余梓彤（大湾区大学）、朱贝尔（中国科学技术大学）	101会议室
分论坛：具身智能	
主席： 丛润民（山东大学）、许华哲（清华大学）、刘航欣（北京大学） 讲者： 李坚强（深圳大学）、李 阳（北京航空航天大学）、张 伟（山东大学） 王晓波（深圳理工大学，深信服科技股份有限公司）、梁小丹（中山大学） 刘行健（大连理工大学）	302会议室
分论坛：视觉基础模型	
主席： 高广谓（南京理工大学）、齐国君（西湖大学） 胡 玮（北京大学）、舒祥波（南京理工大学） 讲者： 王利民（南京大学）、鲍秉坤（合肥工业大学）、李泽超（南京理工大学） 雷 震（中国科学院自动化研究所）、周文罡（中国科学技术大学）	309会议室
分论坛：扩散模型	
主席： 李崇轩（中国人民大学）、张 力（复旦大学&上海创智学院） 李锦兴（哈尔滨工业大学（深圳））、赵 敏（南京大学） 讲者： 张 健（北京大学）、傅朝友（南京大学）、王楠楠（西安电子科技大学） 王亚星（吉林大学）、邹征夏（北京航空航天大学）、武 宇（武汉大学）	401会议室
分论坛：大模型安全与可信人工智能	
主席： 谭 鑫（华东师范大学）、刘笑宏（上海交通大学）、蔡 青（合肥工业大学） 讲者： 向 涛（重庆大学）、高联丽（电子科技大学） 李良斌（北京英博数科科技有限公司）、王立元（清华大学） 韦星星（北京航空航天大学）、方俊峰（新加坡国立大学）	402会议室

内 容		地 点
学术服务论坛：人工智能基础		
指导专家： 鲍秉坤（合肥工业大学）、常晓军（中国科学技术大学） 蒋树强（中国科学院大学）、李厚强（中国科学技术大学） 宋井宽（同济大学）、王琼华（北京航空航天大学） 肖 俊（浙江大学）、余正涛（昆明理工大学）		201会议室
主席： 程 德（西安电子科技大学）、刘 武（中国科学技术大学） 普 楠（合肥工业大学）、杨 勋（中国科学技术大学）		
青年学者： 成 娟（合肥工业大学）、党吉圣（兰州大学）、李俊成（浙江大学） 刘 瑞（内蒙古大学）、刘佐珠（浙江大学）、齐梦实（北京邮电大学） 王浩宇（中国矿业大学）、姚涵涛（中国科学技术大学） 于灵云（中国科学技术大学）		
学术服务论坛：计算机视觉I		
指导专家： 韩军伟（重庆邮电大学）、黄庆明（中国科学院大学） 马思伟（北京大学）、史振威（北京航空航天大学） 王国胤（重庆师范大学）、徐常胜（中国科学院自动化研究所） 张勇东（中国科学技术大学）		202会议室
主席： 李 帅（山东大学）、舒祥波（南京理工大学） 谢国森（南京理工大学）、张 正（哈尔滨工业大学（深圳））		
青年学者： 邓淑敏（浙江大学）、黄步真（天津大学） 胡心亭（中国科学技术大学）、林绍辉（华东师范大学） 刘 李（香港科技大学（广州））、沈 力（中山大学） 沈 雯（同济大学）、宋霄罡（西安理工大学） 徐斌倩（新加坡国立大学）、易 冉（上海交通大学） 张宁豫（浙江大学）、周 康（中国科学院香港创新院）		

内 容		地 点	
学术服务论坛：交叉研究			
指导专家： 黄科科（中南大学）、李肯立（湖南大学） 陶建华（清华大学）、王景景（青岛科技大学） 王 亮（中国科学院自动化研究所）、吴建德（云南大学） 俞 俊（哈尔滨工业大学（深圳））、张 靖（西安交通大学） 主席： 白 琮（浙江工业大学）、张敬林（山东大学） 赵世杰（西北工业大学）、郑钰辉（青海师范大学） 青年学者： 陈 想（南京航空航天大学）、杜云涛（山东大学） 付 钰（兰州大学）、李 晴（北京理工大学） 刘雨帆（中国科学院自动化研究所）、刘政哲（岭南大学） 潘良睿（新加坡南洋理工大学）、沈 飞（新加坡国立大学） 王培松（中国科学院自动化研究所）、王乾铭（华北电力大学） 熊凤超（南京理工大学）、姚鑫杰（昆明理工大学） 祝澜耘（同济大学）		204会议室	
学术服务论坛：计算机视觉II			
指导专家： 白 翔（华中科技大学）、韩军功（清华大学）、何良华（同济大学） 纪荣嵘（厦门大学）、马惠敏（北京科技大学） 徐 凯（中国科学院工业人工智能研究所） 主席： 任文琦（中山大学）、王国庆（电子科技大学） 王文冠（浙江大学）、袁国武（云南大学） 青年学者： 陈超超（浙江大学）、甘磊磊（浙江大学） 郭海云（中国科学院自动化研究所）、刘晋源（大连理工大学） 卢宪凯（山东大学）、王 倩（广东工业大学） 杨传广（中国科学院大学）、叶育森（西安电子科技大学） 张 敏（华东师范大学）、周仕佶（北京航空航天大学） 朱一凡（北京邮电大学）、朱智勤（重庆邮电大学）			207会议室
优博俱乐部全体委员大会		206会议室	

🕒 2026年8月4日18:00-19:30

晚 餐

🕒 2026年8月4日19:00-21:00

📋 内 容

📍 地 点

学科竞赛与AI人才培养研讨会

202会议室

🕒 2026年8月4日19:30-21:30

📋 内 容

📍 地 点

CSIG青工委委员大会(闭门会议)

401会议室

🕒 2026年8月5日08:30-12:00

📋 内 容

📍 地 点

大会开幕式

- 1、领导致辞
- 2、大会筹备情况汇报
- 3、合作单位感谢牌颁发仪式

大会报告

讲者: 申恒涛（同济大学）、李 波（北京航空航天大学）、冯 丹（华中科技大学）
彭宇新（北京大学）、王大轶（北京空间飞行器总体设计部）

学会奖励宣讲

白泽智研企业宣讲

国际会议中心二楼

🕒 2026年8月5日12:00-14:00

午餐&Poster

🕒 2026年8月5日14:00-17:30

📋 内 容

📍 地 点

讲习班：生成式AI与智能创作

主席： 邓 欣（北京航空航天大学）、盛 律（北京航空航天大学）
于 茜（北京航空航天大学）

102会议室

讲者： 曹炎培（VAST）、李崇轩（中国人民大学）、于 茜（北京航空航天大学）

分论坛：图像复原与增强

主席： 潘金山（南京理工大学）、顾舒航（电子科技大学）、钱文华（云南大学）

101会议室

讲者： 白慧慧（北京交通大学）、徐 迈（北京航空航天大学）
傅雪阳（中国科学技术大学）、李 博（vivo）、金 枝（中山大学）

分论坛：本地特色论坛（小语种OCR）

主席： 毛存礼（昆明理工大学）、杨雅婷（中国科学院新疆理化技术研究所）

201会议室

讲者： 高盛祥（昆明理工大学）、肖 桐（东北大学）、飞 龙（内蒙古大学）
何向真（西北民族大学）、杨雅婷（中国科学院新疆理化技术研究所）

分论坛：医学影像处理

主席： 王 妍（华东师范大学）、王连生（厦门大学）
赵一天（中国科学院宁波材料技术与工程研究所）

207会议室

讲者： 窦 琪（香港中文大学）、刘且根（南昌大学）、雷柏英（深圳大学）
范敬凡（北京理工大学）、许言午（华南理工大学）
张鼎文（西北工业大学）、邵永航（云南师范大学）

分论坛：AI For Science

主席： 余肇飞（北京大学）、张敬林（山东大学）
欧阳万里（深圳河套学院）、张 奇（同济大学）

302会议室

讲者： 索津莉（清华大学）、杜长德（中国科学院自动化研究所）
蒋 宁（北京九章云极科技股份有限公司）、金哲依（北京大学）
黄科科（中南大学）、周 游（南京大学）

内 容		地 点
分论坛：国产大模型的部署与优化		
主席： 杨智勇（中国科学院大学）、冯骁骋（哈尔滨工业大学） 秦 兵（哈尔滨工业大学）		309会议室
讲者： 刘新旺（国防科技大学）、黄龙涛（阿里巴巴）、徐梦炜（北京邮电大学） 韩 旭（清华大学）、杨传广（中国科学院计算技术研究所）		
分论坛：视频世界模型		
主席： 王韞博（上海交通大学）、段岳圻（清华大学）、杨 啸（启元实验室）		401会议室
讲者： 徐凯（中国科学院工业人工智能研究所） 章国锋（浙江大学）、金 鑫（宁波东方理工大学） 范 略（中国科学院自动化研究所）、赵 敏（南京大学）		
分论坛：大模型的高效训练与推理		
主席： 包世龙（中国科学院计算技术研究所）、王 伟（北京交通大学） 许倩倩（中国科学院计算技术研究所）		402会议室
讲者： 马 楠（北京工业大学）、车万翔（哈尔滨工业大学）、谭明奎（华南理工大学） 张利军（南京大学）、张淼（哈尔滨工业大学（深圳））		
分论坛：博士生论坛		
主席： 贾 旭（大连理工大学）、杨 鹏（华中科技大学） 张 杰（中国科学院计算技术研究所）、户战选（云南师范大学）		202会议室
讲者： 马 跃（香港科技大学）、邓 辉（西北工业大学） 夏文科（中国人民大学）、陈子健（上海交通大学） 秦逸然（香港中文大学（深圳）） 宋子帆（同济大学）、黄思琪（复旦大学） 王中琦（中国科学院计算技术研究所） 龚天翼（香港中文大学（深圳）） 张莘芮（上海人工智能实验室）		

内 容		地 点
分论坛：女科学家论坛		
主席： 董燕妮（武汉大学）、冯 婕（西安电子科技大学） 高联丽（电子科技大学）、相 艳（昆明理工大学）		204会议室
讲者： 曹刘娟（厦门大学）、杜军平（北京邮电大学）、刘世霞（清华大学） 马惠敏（北京科技大学）、邬 霞（北京理工大学）、张 蕾（四川大学）		
分论坛：学术新星论坛		
主席： 阚美娜（中国科学院计算技术研究所）、胡 鹏（四川大学） 邹博超（北京科技大学）、刘笑宏（上海交通大学）、刘 朋（云南大学）		206会议室
讲者： 魏 坤（西安电子科技大学）、孙 元（四川大学）、汪 昱（同济大学） 包世龙（中国科学院计算技术研究所）、夏 俊（香港科技大学（广州）） 梁 科（国防科技大学）、宋启亚（湖南师范大学） 刘岱宗（武汉大学）、刁云峰（合肥工业大学）		

🕒 2026年8月5日18:00-19:30

晚 餐

🕒 2026年8月6日09:00-12:00

内 容	地 点
讲习班：空间基础模型	
主席： 杨 欣（华中科技大学）、崔兆鹏（浙江大学）	102会议室
讲者： 周晓巍（浙江大学）、姚 遥（南京大学）、穆 尧（上海交通大学）	

内 容		地 点
分论坛：大模型评测		
主席：马 珂（中国科学院大学）、王子泰（中国科学院计算技术研究所） 冯 洋（中国科学院计算技术研究所）		101会议室
讲者：纪荣嵘（厦门大学）、秦 杰（南京航空航天大学）、魏秀参（东南大学） 叶 茫（武汉大学）、刘艾杉（北京航空航天大学）		
分论坛：视觉认知与情感计算		
主席：杨 云（云南大学）、胡 斌（北京理工大学）		201会议室
讲者：胡 斌（北京理工大学）、陈恩红（中国科学技术大学） 秦 兵（哈尔滨工业大学）、马 婷（哈尔滨工业大学） 谷 雨（电子科技大学）、蔡 毅（华南理工大学） 杨 云（云南大学）		
分论坛：遥感影像处理		
主席：孙 显（中国科学院空天信息创新研究院）、邹征夏（北京航空航天大学） 段普宏（湖南大学）		
讲者：方乐缘（湖南大学）、谭 琨（华东师范大学）、唐 旭（西安电子科技大学） 殷君君（北京科技大学）、董燕妮（武汉大学）、曹相湧（西安交通大学）		207会议室
分论坛：低空智能视觉		
主席：刘 偲（北京航空航天大学）、蒋 铎（北京航空航天大学） 张 聪（山东大学）		302会议室
讲者：吴启晖（南京航空航天大学）、王 鹏（西北工业大学）、朱鹏飞（东南大学） 李冠彬（中山大学）、屠 展（北京航空航天大学）		
分论坛：AI Coding Agent		
主席：刘祥龙（北京航空航天大学）、刘艾杉（北京航空航天大学） 马宇晴（北京航空航天大学）		309会议室
讲者：杨 珉（复旦大学）、石恩升（华为）、杨 健（北京航空航天大学） 赵才荣（同济大学）、邵 婧（上海人工智能实验室）		

内 容	地 点
分论坛：图像视频生成与编辑	
主席： 王亚星（吉林大学）、武 宇（武汉大学）、易 冉（上海交通大学） 讲者： 李 玺（浙江大学）、刘 婧（天津大学）、贾 旭（大连理工大学） 张 驰（西湖大学）、全睿杰（浙江大学）	401会议室
分论坛：多智能体	
主席： 吴 飞（浙江大学）、郝建业（天津大学）、吉 炜（南京大学） 讲者： 洪晓鹏（哈尔滨工业大学）、冯 婕（西安电子科技大学） 姚权铭（清华大学）、张 岸（中国科学技术大学）、杨天培（南京大学）	402会议室
国自然博士生论坛	
主席： 金 枝（中山大学）、梁孔明（北京邮电大学） 讲者： 马云峰（湖南大学）、吴冠尧（大连理工大学）、陈思蕤（同济大学） 窦士涵（复旦大学）、郭苡辰（北京航空航天大学）、杨 震（湖南大学） 郭子晗（中山大学&上海创智学院）、李春一（上海交通大学&南洋理工大学）	204会议室
托举工程博士生专项论坛	
主席： 朱翔昱（中国科学院自动化研究所）、徐婧林（北京科技大学） 李华光（云南大学） 讲者： 宋禹成（中南大学）、岳晓辉（东北大学）、付盛豪（中山大学） 刘辰阳（北京航空航天大学）、陈 铮（上海交通大学）、许煌标（福州大学） 陈 锦（昆明理工大学）、左嘉龙（华中科技大学） 林浩坤（中国科学院自动化研究所）、李 凯（清华大学） 刘焰锋（武汉大学）、刘元晔（湖南大学）	206会议室
🕒 2026年8月6日12:00-13:30	
午 餐	

大会主旨报告讲者



申恒涛 同济大学

报告题目：世界模型

个人简介：

国家领军人才，当选欧洲科学院外籍院士、国际计算机学会会士（ACM Fellow）、国际电气与电子工程师协会会士（IEEE Fellow）、美国光学学会会士（OSA Fellow），科睿唯安全球高被引科学家。现任同济大学计算机科学与技术学院院长，长期深耕计算机科学与人工智能前沿领域，主攻世界模型、空间智能方向研究。科研成果先后荣获ACM-CCF人工智能奖、科学探索奖、教育部自然科学奖一等奖、中国电子学会自然科学奖一等奖等多项学术奖项。

报告摘要：

具身智能正面临两大根本性挑战：一是数字与物理空间的断层，传统大模型擅长处理语言与视觉信息，但难以在真实物理世界中输出高频、精准的控制指令。二是任务与设备的碎片化，不同机器人、不同任务之间模型无法迁移，效率低且成本高。世界模型被视为破解该困局的关键，通过构建对物理空间规律、物体属性与动力学关系的内在表征，驱动智能体从“感知-反应”的浅层模式，跃升至“理解-预测-行动”的深度推理。让智能体不仅能“看”能“听”，更能理解空间关系、预判物理后果，并根据环境自主行动，是通往具身智能的核心基石。本次报告将系统解读世界模型的核心要素与技术范式，梳理当前国内外研究前沿。同时分享团队在世界模型上的技术布局与最新进展，并展望未来发展趋势与产业机遇。



李 波

北京航空航天大学

报告题目：面向复杂场景的红外目标识别新方法

个人简介：

北京航空航天大学长江学者特聘教授、杰青。兼任国务院学位委员会学科评议组成员，是军委装备发展部/军事航天部队/火箭军等人工智能装备应用基础技术专家组成员，教育部教育技术与人工智能教指委副主任委员。

当前主要研究方向为计算机视觉、机器学习、知识推理、嵌入式智能系统。已主持国家、省部级课题40余项，是国家973计划项目首席科学家、国家重点研发计划项目负责人，原总装“十五”、“十二五”某边海防视频系统的型号总设计师。在本领域重要期刊和国际学术会议发表学术论文100余篇，有国内外发明专利100余件，获国家技术发明二等奖1项(排名1)、国家科技进步二等奖1项(排名1)，省部级科技成果奖11项，荣获CCF王选奖。

报告摘要：

外图像是热辐射图像，没有颜色、纹理细节弱、分辨率低，其目标识别比可见光图像更加困难，需要研究适合红外图像处理的新型表征和识别方法。本报告围绕复杂应用场景，按照表面特征、温度材质分布、纹理细节等三方面表征高动态红外图像，提出了具有物理特性的红外图像多元表征方法；针对弱小目标识别难题，发现了弱小目标与清晰目标之间存在稳定的语义差异特征，提出了基于语义特征还原的弱小目标识别方法；针对遮挡目标，构建了遮挡目标联想记忆库，提出了抗干扰的目标准确识别与定位方法。



冯 丹 华中科技大学

报告题目：面向AI加速的存算一体芯片及软硬件系统

个人简介：

华中科技大学副校长，武汉光电国家研究中心光电信息存储研究部主任，信息存储系统教育部重点实验室主任。国家杰出青年基金获得者、长江学者特聘教授、万人计划科技创新领军人才入选者，IEEE Fellow。从事信息存储系统与技术、计算机系统结构等研究，担任2个973项目首席科学家，863重大项目总体专家组组长、重点研发计划项目负责人、基金委创新研究群体项目负责人等。授权发明专利100余项，发表论文300余篇，著作3部，主持完成国家标准8项和电子行业标准3项。获国家技术发明二等奖2项、国家科学技术进步二等奖1项、国家教学成果二等奖1项。

报告摘要：

随着数据规模增长，人工智能、大数据处理中数据移动导致的请求延迟增加、系统性能下降、系统整体能效降低等问题越来越严重，存算一体技术改变了传统的将数据从存储移动至计算单元的模式，直接在存储器内实现计算，大大减少数据搬运，从而提升系统整体性能。报告将介绍我们基于新型非易失存储器的存算一体芯片探索和实践工作，近年来研制了基于国产工艺的MRAM存算芯片和ReRAM存算芯片及验证性原型，支持神经网络、图像处理等的高能效存内计算加速。



彭宇新 北京大学

报告题目：细粒度理解与物理性生成

个人简介：

北京大学二级教授、博雅特聘教授，IEEE/CCF/CAAI/CIE/CSIG Fellow，2019年国家杰青，2025年国家杰青延续资助，2019年国家万人领军人才，2018年科技部中青年科技创新领军人才，863项目首席专家，国家重点研发计划“社会治理与智慧社会科技支撑（平安中国）”重点专项总体专家组专家，中国工程院“人工智能2.0”规划专家委员会专家，中国人工智能产业创新联盟专家委员会主任，中国图像图形学会副理事长，北京图像图形学学会副理事长。主要研究方向为多媒体分析、计算机视觉、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（2项）、北京自然科学基金联合基金重点、发改委专项等40多个项目。发表TPAMI、IJCV、CVPR、NeurIPS、ICML等ACM/IEEE Trans.和CCF A类论文近200篇，获最佳论文奖2次。参加10届（10年）由美国国家标准与技术研究院（NIST）举办的国际评测TRECVID视频搜索比赛，均获第一名，参赛队伍包括斯坦福大学、卡内基梅隆大学、牛津大学等。成果应用于国家网信办、公安部、国家广播电视总局等重要单位以及华为、腾讯、快手、蔚来、美团、中国电信、中国铁塔等头部企业。担任IEEE TCSVT高级领域编委、IEEE TMM等期刊编委，培养博士生获中国计算机学会、中国电子学会等优博。

报告摘要：

多模态大模型与扩散模型作为两类代表性的基座模型，分别在视觉内容理解与生成上表现出色，但也面临重要挑战：在视觉内容理解上，多模态大模型难以识别真实世界对象的细粒度类别；在视觉内容生成上，扩散模型难以生成符合真实世界物理规律的视觉内容。针对上述挑战，本报告将首先介绍团队在细粒度识别大模型、细粒度树的分层识别、物理驱动的视频生成等方面的研究进展。然后，介绍团队在美学理解与虚拟换衣两大应用场景下的最新进展。最后，介绍大模型时代下视觉内容理解与生成技术的落地应用，并展望多模态大模型与扩散模型的未来研究方向。



王大轶

北京空间飞行器总体设计部科技委

报告题目：序列图像自主导航

个人简介：

王大轶研究员，北京空间飞行器总体设计部科技委主任，是国家杰青、国防卓青、万人领军，973项目技术首席专家、政府特殊津贴专家。长期从事观测、诊断和重构能力定量表征理论方法以及空间飞行器自主运行技术研究，入选国家自然科学基金十三五优秀成果，为我国深空探测、北斗三号等任务成功做出重要贡献。获首届国家卓越工程师奖、国家技术发明二等奖、国家科技进步特等奖、何梁何利基金科技创新奖、全国创新争先奖、钱学森杰出贡献奖等，是国家有突出贡献中青年专家。

报告摘要：

空间飞行器自主运行的两大关键核心是自主导航与自主诊断重构，自主导航是实现全自主运行的前提。报告人针对资源受限无人系统，以观测能力的定量表征为理论创新突破口，以序列图像的优化构建和信息优选为关键技术突破口，提出了基于观测能力定量表征的序列图像自主导航技术，为无人系统实现安全可靠自主运行做出贡献。

白泽智研企业宣讲



姚鑫杰

昆明理工大学，白泽智研（南京）科技有限公司

**报告题目：AI for Research新范式：从科学发现
基座模型到硅基科学家**

个人简介：

姚鑫杰，昆明理工大学信息工程与自动化学院特聘副教授，2025年毕业于天津大学，获工学博士学位。主要研究方向包括AI4R以及多智能体学习等，发表学术论文20余篇，其中在ICML、NeurIPS、AAAI、IJCV、TMM、TCYB、PR等国际会议和期刊上发表论文10余篇。申请发明专利6项，获得软件著作权10项。曾担任ICCV、ECCV等会议分论坛主席或组织委员会成员，参与世界智能大会、Valse等多项重要学术活动的组织工作。担任TPAMI、TIP、ICML、NeurIPS、ICLR、ICCV等期刊和会议审稿人。获CVPR 2023 CLVision挑战赛亚军等多项奖励，参与国家自然科学基金重点项目、新一代人工智能国家科技重大专项及国家自然科学基金青年科学基金项目（A类、B类）等科研项目10余项。

报告摘要：

人类科学研究正经历从经验、理论、计算与数据密集型范式，迈向智能驱动科学的新跃迁。面对知识规模扩张、科研流程高度手工化以及科学假设增长快于验证能力等瓶颈，AI for Research不应止于文献检索、语言生成或局部工具增强，而应重构知识、模型、实验与论文之间的协同关系。本报告围绕科学发现基座模型，阐释以领域知识与专家库为认知底座，以多智能体协同为组织机制，以科学假设生成、自动实验、可信验证和AutoPaper为核心闭环，并连接具身实验机器人，贯通数字推理与物理验证。由此，科研将从作坊式生产走向可协同、可复用、可追溯、可持续进化的智能科研体系；人类科学家也将从科研执行者转向问题架构师，与硅基智能共同拓展科学发现边界。

大会分论坛

讲习班 多模态智能体

时间：8月4日上午09:00-12:00 地点：102会议室

内 容	嘉 宾
主席：李冠彬（中山大学）、徐天阳（江南大学）	
面向通用智能的大模型自主智能体	林衍凯（中国人民大学）
从多模态理解推理到多模态智能体	岳翔宇（香港中文大学）
迈向长程多模态智能体	杨思蓓（中山大学）

论坛主席



李冠彬

李冠彬，中山大学计算机学院副院长，教授，博士生导师，国家优青。主要研究领域为图像视频内容理解与生成。迄今为止累计发表计算机学会 A 类 / 中科院一区论文 200 余篇，谷歌学术引用超过 22000 次，入选 ScholarGPS 全球前 0.05% 顶尖科学家榜单。曾获得广东省青年科技创新奖、中国图像图形学会青年科学家奖、吴文俊人工智能优秀青年奖、中国图像图形学会科学技术一等奖、ICCV2019 最佳论文提名奖、CVPR2024 最佳论文候选等荣誉。主持了包括国家自然科学基金优青、面上、重点研发课题、广东省杰青、CCF- 腾讯犀牛鸟科研基金、华为科研合作基金等 20 多项科研项目。担任广东省大数据分析与管理重点实验室副主任、广东省图像图形学会计算机视觉专委会主任等职务。担任人工智能领域顶级会议 CVPR、ECCV、AAAI 等领域主席，获得 10 余项人工智能领域国际顶级会议竞赛冠军，研究成果应用于华为智驾系统，OPPO 小布助手及智慧医疗诊断等。



徐天阳

徐天阳，江南大学副教授，博士生导师。研究方向为视频理解与多模态融合，发表 CCF-A/IEEE 汇刊 40 余篇，谷歌学术引用 8000 余次。主持国家自然科学基金面上项目、青年项目、江苏省杰青项目等。担任中国人工智能学会模式识别专委会副秘书长、江苏省人工智能学会模式识别专委会副主任。担任 IEEE TIP、Pattern Recognition 编委，CVPR、ICLR、ICPR、PRCV、VALSE 等学术会议的程序主席、领域主席、研讨会主席 / 组织者、讲习班组织者等 10 余次。相关研究工作获“中国图像图形学会优秀博士学位论文奖”，获“江苏省自然科学百篇优秀学术成果”，获 CVPR/ICCV/ECCV 等国际学术会议举办相关学术竞赛（VOT、MMVRAC、Anti-UAV、AI City Challenge、Perception Test Challenge、MARS2）冠军 10 余项，连续入选斯坦福大学全球前 2% 顶尖科学家年度榜单。

论坛讲者



林衍凯

林衍凯，中国人民大学高瓴人工智能学院预聘副教授，主要研究方向为大模型智能体，Google Scholar 统计引用达到 2.9 万次，2020-2025 年连续六年入选爱思唯尔（Elsevier）中国高被引学者和斯坦福大学全球前 2% 顶尖科学家年度榜单、入选北京市科技新星、吴文俊人工智能青年科技奖、《麻省理工科技评论》中国与 DeepTech 联合发布的“AI100 青年先锋”等。其成果获评教育部自然科学一等奖、世界互联网大会领先科技成果。现任 CCF NOI 科学委员会委员。

报告题目：虚拟现实内容高效渲染方法研究

报告摘要：近年来，大模型在理解与推理方面展现了革命性潜力。实现从被动“大脑”到主动“行动者”的跨越，即构建大模型自主智能体，是迈向通用智能的关键所在。这要求智能体不仅能被动回答，更能通过学习使用工具、与复杂动态环境交互，来自主完成现实世界中的复杂任务。本次报告将聚焦大模型自主智能体，深入探讨其面临的三大核心挑战：泛化性、长程性与自主性。报告将系统梳理智能体从模仿学习到探索学习的演进路径，探讨如何通过智能体强化学习提升其长程规划与环境泛化能力，并展望智能体如何从“被动工具”演化为真正的“主动助手”，分享我们在该领域的最新研究成果与未来发展方向。



岳翔宇

岳翔宇，香港中文大学助理教授，博士生导师。他于加州大学伯克利分校获得电子计算机博士学位，于斯坦福大学和南京大学分别获得硕士和学士学位。他在人工智能、多模态大模型、生成模型等领域有深厚的研究经验，在人工智能顶会发表论文近百篇。曾获得 ACM CCS 2025 杰出论文奖、Lotfi A. Zadeh 奖（用于表彰对于软计算领域的理论和应用做出突出贡献的 UC Berkeley 优秀博士毕业生）、全球前 2% 顶尖科学家等奖项。现 Google Scholar 引用 16000 余次，h-index 41；曾担任 CVPR, NeurIPS, ICLR, ICML, ICCV, ECCV, ACL, AAAI 等人工智能顶级会议的领域主席。他的 MiCo、LATTICE、FlashVDM 等工作已经成为字节 Seed-VL, 腾讯混元 3D 等模型的预训练和后训练范式。

报告题目：从多模态理解推理到多模态智能体

报告摘要：随着大模型从语言理解迈向多模态感知、复杂推理与行动决策，多模态智能正经历从“看懂世界”到“理解、推理并主动完成任务”的范式转变。本演讲将介绍我们在多模态智能方向的一系列研究进展，涵盖统一多模态表征学习、全模态预训练、跨模态知识迁移、视觉与视频推理、多任务多模态推理模型，以及面向生成和行动的多模态智能体。具体而言，我们将讨论如何通过多模态上下文建模提升模型对多种数据模态的统一理解能力，如何利用强化学习、思维奖励和时序奖励机制激发多模态大模型的深层推理能力，以及如何构建能够同时处理图像、视频和多样化视觉任务的通用推理模型。在此基础上，演讲还将进一步探讨多模态模型如何从被动问答走向主动搜索、工具调用、内容生成与环境交互，并结合知识密集型图像生成、三维生成、视觉到代码生成、自动驾驶世界模型和机器人操作等应用，展示多模态智能体的发展潜力。



杨思蓓

杨思蓓，中山大学计算机学院副教授，博士生导师，逸仙学者。分别在香港大学和浙江大学取得博士和学士学位。主要研究方向为跨模态视觉感知、理解、生成与交互，聚焦于多模态统一表征学习、多模态大模型、具身智能。迄今为止在 TPAMI、CVPR、ICCV 等期刊或会议发表 CCF A 类 / 中科院一区论文 60 余篇，其中以第一作者和通讯作者发表 CCF A 类论文 40 余篇，引用 3600 余次。主持国家自然科学基金面上项目、青年项目，以及浦江人才计划、上海领军人才海外计划、腾讯犀牛鸟专项等科研项目。担任 ICCV、ICLR、NeurIPS、ECCV 等顶会领域主席，入选全球前 2% 顶尖科学家榜单。

报告题目：迈向长程多模态智能体

报告摘要：多模态智能体在短时交互任务上已展现出不错的能力，而真实世界中的长程任务，跨越多轮交互、多个场景及较长时间尺度，对智能体提出了更高要求。从短程到长程的跨越，不仅是时间维度的延伸，更涉及记忆、推理与决策机制的新挑战。如何在延长时间尺度上维持感知连贯性、积累并更新任务相关记忆、进行多步推理与前瞻性规划，同时确保决策的可靠性与一致性，是构建可用长程智能体的关键。本报告将介绍相关技术进展与系统框架，讨论现有方法的适用边界与未来研究方向。

分论坛

图像视频理解

时间：8月4日上午09:00-12:00 地点：101会议室

内 容

嘉 宾

主席：白慧慧（北京交通大学）、刘 洋（北京大学）、李 锋（合肥工业大学）

关于具身智能感知与决策的探索	查正军（中国科学技术大学）
智能体驱动的电影感视频生成	施柏鑫（北京大学）
AI竞赛助力企业技术攻关与创新人才培养	房 伟（全球校园人工智能算法精英大赛组委会秘书长）
第一视角多模态视频理解	徐 行（同济大学）
从几何对齐到语义认知：全景视觉计算的技术演进与实践	林春雨（北京交通大学）
面向开放世界的人物交互感知与生成	刘 洋（北京大学）

论坛主席



白慧慧

白慧慧，北京交通大学教授、博士生导师，国家级高层次人才。主要研究方向是图像视频编码和增强等。已发表包括 IEEE 汇刊及 CCF A 类会议论文 100 余篇。获国际专利授权 3 项、国家发明专利授权 20 余项。主持国家自然科学基金重点项目 / 企业联合基金项目、北京市自然科学基金项目等。获北京市自然科学一等奖、北京市科技进步二等奖、中国电子学会青年科学家奖、中国产学研合作创新成果奖二等奖、中国人工智能学会教学成果一等奖、中国图像图形学会教学成果二等奖等。



刘 洋

刘洋，北京大学王选计算机研究所研究员，博士生导师，北大博雅青年学者。其博士毕业于英国剑桥大学计算机专业，而后曾任英国牛津大学博士后研究员。主要从事多模态感知和生成相关研究，致力于用人工智能算法解决多媒体信息分析与智能融合的相关问题。研究成果发表国际知名期刊和会议论文 90 余篇，获 7 项国际专利授权，7 项国际竞赛获奖。主持科技部国家重点研发计划课题、国家自然科学基金面上项目等项目 10 余项，部分创新性研究成果及其关键技术已在学术及产业界得到了应用。www.csyangliu.com



李 锋

李锋, 合肥工业大学人工智能创新学院副教授, 博导, 主要研究方向为计算机视觉、图像视频增强及多模态信息处理。主持国家自然科学基金青年基金, 国家自然科学基金重点项目课题负责人。在包括 IEEE TPAMI、IJCV、CVPR、ICCV、ICLR 等国际权威期刊与学术会议上发表论文 50 余篇。

论坛讲者



查正军

查正军, 中国科学技术大学讲席教授、博导, 类脑智能国家工程实验室执行主任。从事图像视频处理与分析、多模态内容理解与生成、具身智能感知与决策等领域研究。主持科技创新 2030- 新一代人工智能重大项目, 国家自然科学基金杰青、优青、重点项目, 国家创新人才引进计划青年项目等。多次获得国际会议论文奖励, 包括 CCF A 类会议 ACM Multimedia 最佳论文奖、最佳学生论文奖, AAAI 杰出论文奖等。获安徽省自然科学一等奖、安徽省教学成果特等奖、国家教学成果二等奖、中国科学院青年科学家奖、中国科学院优秀导师、霍英东教育基金会青年科学奖等。指导的研究生学位论文入选中国科学院优秀博士论文、中国人工智能学会博士论文激励计划、中国电子学会硕士论文激励计划等。担任国际会议 ICME 指导委员会成员, IEEE Trans. PAMI、IEEE Trans. MM、IEEE Trans. CSVT 等权威国际期刊编委。

报告题目: 关于具身智能感知与决策的探索

报告摘要: 具身智能感知与决策旨在赋予智能体在开放物理世界中可靠感知、持续理解与自主决策的能力, 是驱动人工智能由静态理解走向动态决策、由被动感知走向主动执行的关键。然而, 真实具身环境呈现出数据多模异构、分布动态演化、内容结构繁杂等特点, 严重制约了场景感知的可靠性和行为决策的持续性。针对上述挑战, 本报告聚焦于“具身智能的感知连续表征与策略持续推理”这一问题, 从场景表征、内容解析及行为决策三个维度, 介绍团队的近期研究探索, 并探讨相关方向的发展趋势。



施柏鑫

施柏鑫, 北京大学研究员(长聘副教授、博导), 计算机学院视频与视觉技术研究所副所长; 科技部重大专项首席科学家, 国家自然科学基金重点项目负责人, 北京智源学者, 北大-智平方具身智能联合实验室主任。日本东京大学博士, 麻省理工学院媒体实验室博士后。研究方向为计算摄像学与人工智能, 发表论文 270 余篇(包括 TPAMI 论文 37 篇, 计算机视觉三大顶级会议论文 110 篇)。论文获评 CVPR24/ICCP15 最佳论文亚军、CVPR26/3DV26/ICCV15 最佳论文候选, 获得日本大川研究助成奖(2021)、中国电子学会青年科学家奖(2024)、IJCV 杰出编委奖(2025)。担任国际顶级期刊 TPAMI/IJCV 编委, 顶级会议 CVPR 等牵头领域主席。更多信息请访问“相机智能”实验室主页: <http://camera.pku.edu.cn>

报告题目: 智能体驱动的电影感视频生成

报告摘要: 当前主流视频生成模型因缺乏对物理规律与叙事逻辑的理解, 在实现“电影感”的效果方面遇到了瓶颈。引入具备逻辑推理能力的智能体, 是跨越这一认知鸿沟的有效途径之一。本次报告将分享在视频生成的流程中, 通过引入智能体解析、结构化先验与底层扩散模型协同等技术, 支持物理光影重构、多镜头视觉连贯、音视频精准对齐、长视频叙事剪辑, 实现高维逻辑理解与底层视觉生成的高效协同, 弥补现有视频生成方法在电影感可控生成方面的能力缺失。



房伟

房伟, 高级程序员、高级工程师, 江苏省人工智能学会秘书长, 兼任江苏省工业软件软件产业链联盟、江苏省国防软件产业创新联盟秘书长, 南京市经济社会发展咨询委员会委员。目前主要从事人工智能等新一代信息技术与产业发展研究与咨询服务, 参与制定省人工智能、软件、大数据等相关产业政策, 先后主持省、市等 20 余项产业课题研究和标准制定。

报告题目: AI 竞赛助力企业技术攻关与创新人才培养

报告摘要: 全球校园人工智能算法精英大赛(AIC)由江苏省人工智能学会首任理事长、中国科学院院士、南京大学副校长周志华教授于 2019 年倡议创办, 是江苏唯一入选“全国普通高等学校学科竞赛排行榜”的国家级人工智能专项赛事, 也是国内唯一以人工智能算法为核心主题的国家级学科竞赛。赛事已连续成功举办七届, 累计吸引全球 26 个国家和地区、1000 余所高校参赛, 参赛队伍超 4.5 万支, 参与师生规模突破 10 万人。2025 第七届赛事吸引全球 900 多所高校、2 万余支团队参赛, 覆盖国内全部 211 院校, 设立 13 个省级赛区, 辐射全国 33 个省级行政区, 成为衔接高校人工智能教学、科研创新与产业需求的核心竞赛载体。2026 第八届大赛于 4 月 28 日正式启动报名, 设置 18 大赛区, 开设算法挑战赛、算法创新赛、算法应用赛、算法主题赛、产业命题赛、算法专项赛六大赛道, 覆盖基础算法研发、行业落地应用、产业真实命题等多元赛道, 全面挖掘全球高校 AI 算法创新人才。



徐 行

徐行, 现任同济大学计算机科学与技术学院教授, 青年百人计划 A 岗, 博士生导师, 国家高层次青年人才。主要研究方向为多模态信息感知与理解、跨媒体智能分析、多模态大模型应用。已发表人工智能、多媒体、计算机视觉相关领域的中科院 JCR1 区 /CCF A 类论文 80 余篇, ESI 高被引 / 热点论文 10 余篇, 授权国内外发明专利 30 余项。获得 2025 年教育部科学研究优秀成果奖自然科学奖一等奖, 2025 年中国电子学会自然科学一等奖, 2022 年吴文俊人工智能自然科学一等奖、国际多媒体大会 ACM Multimedia (CCF A) 2017 最佳论文奖, 多媒体顶级期刊 IEEE Transactions on Multimedia 2020 最佳论文奖, 人工智能顶级期刊 IEEE Transactions on Fuzzy Systems (中科院 JCR 1 区) 2024 杰出论文奖, 国际多媒体展览会 IEEE International Conference on Multimedia & Expo (CCF B) 2017 最佳会议论文铂金奖和 2022 最佳学生论文奖等科技成果及国际学术奖励。入选第十四批四川省学术和技术带头人及后备人选。主持包括国自然青年科学基金(B 类、C 类)及面上项目、四川省科技厅重大专项, 并参与科技部、国自然及省部级科研项目 20 余项。

报告题目：第一视角多模态视频理解

报告摘要：随着具身智能、混合现实以及机器人技术的快速发展, 针对由摄像头、麦克风、陀螺仪等多源感知设备共同形成的第一视角视频数据的理解需求日益凸显。然而, 现有视频理解技术多用于互联网第三视角视频, 难以有效应对第一视角视频数据存在的模态类型多样、场景变化频繁、事件语义复杂等挑战。本报告围绕视频理解的感知 - 表征 - 推理三个核心要素, 介绍课题组近期基于多感融合和表征泛化的多模态第一视角视频理解研究成果, 以及在动作识别、视频内容分析、大模型推理增强、具身 / 随身智能等场景的应用。



林春雨

林春雨, 北京交通大学教授, 博士生导师。主要研究方向: 人工智能、计算机视觉、多模态信息处理方向。在 CVPR、ICCV、ECCV、NeurIPS、AAAI、ACM-MM 等国际顶级会议及 TPAMI、TIP、TCSVT、TMM、TITS 等权威期刊发表论文 80 余篇。主持国家级、国防级、省部级等纵向课题十余项。相关研究成果已落地应用于中国中车、中国移动、大疆创新科技、国能集团等头部企业。获教育部、北京市等省部级科技奖励 2 项, 获中国电子学会、中国图像图形学会科技奖励 2 项。指导学生斩获国家级学科竞赛一等奖 2 项、二等奖 4 项, 省部级十余项。

报告题目：从几何对齐到语义认知：全景视觉计算的技术演进与实践

报告摘要：全景视觉凭借 360° 全域无死角的视场覆盖优势，已成为智能感知、沉浸式媒体、自主机器人等领域的核心感知模态。本报告以全景视觉技术发展脉络为主线，系统梳理从高精度几何对齐到高层次语义理解的完整技术体系与代表性实践成果。报告首先针对复杂大视差场景下的全景视频拼接痛点，提出全局单应性与局部薄板样条(TPS)统一参数化的柔性对齐变换，在保障重叠区域精准内容对齐的同时，实现拼接图像自然形状的保持。在此基础上，报告进一步阐释全景深度估计技术，搭建起从几何成像到空间感知的技术衔接桥梁，实现 360° 全视场深度信息的精准估计，并支撑稀疏场景下的三维重建任务。随后，报告围绕大模型技术背景，介绍端到端全景视觉基座大模型的研究进展，可统一支撑多项全景理解任务，推动技术向高层语义认知维度延伸。最后，结合自动驾驶、全景直播、无人机等典型落地场景验证技术应用成效，并对全景视觉的未来发展方向进行展望。



刘 洋

刘洋，北京大学王选计算机研究所研究员，博士生导师，北大博雅青年学者。其博士学位毕业于英国剑桥大学计算机专业，而后曾任英国牛津大学博士后研究员。主要从事多模态感知和生成相关研究，致力于用人工智能算法解决多媒体信息分析与智能融合的相关问题。研究成果发表国际知名期刊和会议论文90余篇，获7项国际专利授权，7项国际竞赛获奖。主持科技部国家重点研发计划课题、国家自然科学基金面上项目等项目10余项，部分创新性研究成果及其关键技术已在学术及产业界得到了应用。www.csyangliu.com

报告题目：面向开放世界的人物交互感知与生成

报告摘要：面向开放世界的交互感知与生成技术对推动具身智能与虚实融合应用具有关键意义。但在实际研究中仍面临诸多挑战：高质量人物交互多模态数据稀缺、细粒度交互时空标注不足，以及以人为中心的物理常识建模不足等。本次汇报将分享团队在这一领域的探索，介绍一系列基于多模态大模型和扩散模型的交互感知与生成工作，从静态图像中的人物交互关系理解出发，到解译视频中动态交互与物体状态迁移的因果关联，拓展到实现文本可控的人物交互图像和视频生成。

分论坛

可视化与渲染

时间：8月4日上午09:00-12:00 地点：201会议室

内 容

嘉 宾

主席：汪云海（中国人民大学）、胡瑞珍（深圳大学）

高 林（中国科学院计算技术研究所）

虚拟现实内容高效渲染方法研究

王莉莉（北京航空航天大学）

大规模脑微观影像中的几何建模和形态分析

陈雪锦（中国科学技术大学）

基于单目视觉的三维场景重建

连宙辉（北京大学王选计算机研究所）

虚拟环境中从多模态输入到智能生成的交互方法探索

宋维涛（北京理工大学）

三维世界建模：从重建到生成

龙霄潇（南京大学）

论坛主席



汪云海

中国人民大学信息学院教授，博导，国家级青年人才，数据工程与知识工程教育部重点实验室副主任。研究方向为大数据可视分析，近五年在 ACM SIGMOD、ACM SIGKDD、ACM SIGCHI、IEEE VIS、IEEE TVCG 等相关会议与期刊发论文 70 余篇，获得 IEEE VIS、ACM SIGCHI 多个领域顶级会议最佳论文提名奖等奖项。曾获山东省自然科学奖一等奖、中国电子学会科技进步二等奖、吴文俊人工智能自然科学二等奖，担任 IEEE TVCG、CGF、IEEE CG&A 等期刊编委。



胡瑞珍

深圳大学特聘教授，博士生导师，国家优秀青年科学基金、广东省杰出青年项目获得者。研究方向为计算机图形学，长期从事智能几何建模与处理方面的研究，发表 ACM SIGGRAPH/TOG 论文三十余篇；入选中科协青年人才托举工程；荣获亚洲图形学协会青年学者奖、全国几何设计与计算青年学者奖；荣获 ICRA 2026 最佳论文奖；担任期刊 IEEE TVCG、IEEE CG&A 和 Computers&Graphics 等国际期刊编委；担任国际会议 SGP 2024/CVM 2023/SMI 2020 Technical Paper、SIGGRAPH Asia Workshops/Technical Communications and Posters 以及 EG 2024 Short Paper 程序委员会主席，连续多年担任 SIGGRAPH 等大会程序委员会委员；担任中国图像图形学会智能图形专委会副主任、中国计算机学会计算机辅助设计与图形学专委会常委 / 副秘书长、计算机图形学与混合现实在线平台 (GAMES) 执委会主席。



高林

中国科学院计算技术研究所研究员、博士生导师、泛在计算系统研究中心副主任、中国科学院大学岗位教授。在清华大学获得博士学位。研究方向为计算机图形学、三维计算机视觉。在 SIGGRAPH、TPAMI、TVCG 等期刊会议发表论文 100 余篇，研发的人脸 AIGC 的 APP 被全球 180 余个国家或者地区的用户所使用。现任或者曾任亚洲图形学学会秘书长，CSIG 智能图形专委会秘书长，SIGGRAPH 技术论文程序委员会委员，CVPR 领域主席，IEEE TVCG 编委，作为项目负责人承担国家重点研发计划、国家自然科学基金委优青等项目，曾获得亚洲图形学会青年学者奖，吴文俊人工智能优秀青年奖，CCF 技术发明一等奖等奖励。

论坛讲者



王莉莉

北京航空航天大学计算机学院教授，国家级领军人才，虚拟现实技术与系统全国重点实验室副主任，中国电子学会虚拟现实分会副主任。主要从事虚拟现实、混合现实等方面的研究。主持了国家自然科学基金重点项目、联合项目等国家级科研项目，获国家科技进步一等奖、国家技术发明二等奖、中国计算机学会技术发明一等奖、国家高等教育教学成果一等奖。担任 IEEE TVCG 等国际期刊编委，任 IEEE VR 2026 程序委员会副主席，IEEE ISMAR 2026 程序委员会主席。

报告题目：虚拟现实内容高效渲染方法研究

报告摘要：虚拟现实要求呈现延迟低于 13 毫秒，以避免感知不一致导致的眩晕，而传统 30-60FPS 的绘制效率难以满足强实时需求。本报告聚焦于视觉感知驱动的高效渲染方法，首先介绍近眼呈现的基本理论与凝视点驱动的资源优化思路。接着，重点汇报课题组在 3D 高斯泼溅及其动态扩展领域的最新进展：结合视觉敏锐度与对比度敏感性模型，利用非均匀变换与多分辨率拟合提升绘制效率；针对 VR 交互的核心需求，介绍 Motion Hierarchical Gaussian 实现动态场景的分层控制，GaussianHand 驱动实时高保真手部化身动画，以及 GSHOI Denoiser 在复杂手物交互场景下的高质量去噪呈现。最后，对虚拟现实内容渲染的未来发展进行展望。



陈雪锦

中国科学技术大学信息学院教授，少年班学院副院长。分别于中国科学技术大学获得学士和博士学位，耶鲁大学计算机系博士后，曾于斯坦福大学、微软亚洲研究院任访问学者。主要研究方向为计算机图形学、三维视觉、脑显微影像分析。在 ACM TOG、TMI、ICCV、ICML 等期刊 / 会议上发表学术论文 100 余篇，曾获 CVM 期刊年度最佳论文提名、GDC、3DMM 会议最佳论文、ISBI 挑战赛第一名等奖励。获国家教学成果二等奖、安徽省教学成果特等奖、CSIG“石青云女科学家奖”。

报告题目：大规模脑微观影像中的几何建模和形态分析

报告摘要：实现突触级分辨率的神经元连接图谱高精度重建是解析脑功能机制的关键手段。随着连续电镜成像技术的发展，大规模神经连接组学研究已进入 PB 级数据时代。基于电镜的连接组学也被评为《自然方法》年度方法。然而，脑中数万神经元分布稠密、形态复杂，单神经元空间跨度大，从大规模电镜图像中完整重建神经元的三维精细结构面临巨大挑战。本报告将汇报我们结合神经元几何形态和电镜图像特征，在大规模电镜图像中进行神经元的精细重建、高效追踪、形态分析方面的研究工作和最新进展。



连宙辉

北京大学王选计算机研究所所长聘副教授、研究员、博士生导师，中国文字字体设计与研究中心副主任，中国图像图形学会三维视觉专委会秘书长，入选国家级青年人才计划、国家语言文字关键研究领域领航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与三维视觉，在领域重要期刊 (TOG, TPAMI, IJCV 等) 和会议 (SIGGRAPH/SIGGRAPH Asia, CVPR 等) 上发表论文 100 多篇。多次担任 CVPR、ICCV、ICML、ECCV、SIGGRAPH Asia 等国际会议领域主席 (或相当职务)，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)、ICRA 2024 服务机器人最佳论文候选、吴文俊人工智能优秀青年奖、北京大学 - 中国光谷成果转化奖、北京大学教学成果一等奖、北京大学青年教师教学基本功比赛一等奖等奖励。

报告题目：基于单目视觉的三维场景重建

报告摘要：空间智能是近年来学术界和工业界关注的热点问题，视觉三维场景重建是其中的核心技术之一，在测绘、巡检、具身智能等领域有着重要的应用价值。本报告将系统介绍近年来我们团队在“基于单目视觉的三维场景重建”方面的一些研究进展，从单图像三维重建，到单全景图室内三维重建，再到基于单目相机的三维场景主动重建，目标是构建具备鲁棒性、高效性、高精度和泛化性的单目视觉三维重建系统，迈向具备人类水准的纯视觉空间智能。



宋维涛

北京理工大学光电学院教授, 博士生导师, 国家科技重大专项项目负责人, 入选 2021 年度国家级青年人才计划, 主要研究工作包括颜色科学, 人机交互, 新型显示等方面。主持国家科技重大专项、国家重点研发计划课题等多个项目, 在 Nature Communications、IEEE TVCG、ACM CHI 等期刊会议上发表八十余篇, 授权发明专利三十余项, 牵头起草国家标准 4 项, 国际标准 2 项, 担任国际标准化组织 ISO/IEC SC24 专家组成员, 中国照明学会视觉与颜色专委会副主任委员, 中国图像图形学会虚拟现实专委会秘书长, 北京图象图形学会理事, 中国计算机学会人机交互专委会常务委员等学术兼职。

报告题目: 虚拟环境中从多模态输入到智能生成的交互方法探索

报告摘要: 随着 VR/AR/MR 等空间计算技术的发展, 实现自然高效的多模态交互与智能化三维场景构建成为重要研究方向。本文围绕多模态感知与智能生成, 开展了面向沉浸式虚拟环境的关键方法与系统研究: 在交互感知方面, 提出了基于声音识别的微手势检测、基于指环压力传感器的微手势识别、面向运动场景的磁感应面部动作识别, 以及结合手臂指向与前臂滑动的多模态交互方法, 实现从声音、手指、面部到全身的多层级自然输入; 在智能生成方面, 结合大语言模型与视觉语言模型, 实现语义理解与空间推理驱动的用户自定义 3D 场景快速构建; 在应用方面, 探索了面向认知训练的 VR 游戏设计, 将认知任务与沉浸体验结合以提升注意力、记忆和执行功能, 报告介绍了团队在面向虚拟环境中交互方法的一些探索, 希望为空间计算时代的内容创作与交互设计提供了参考。



龙霄潇

南京大学智能科学与技术学院副教授，南京大学紫金学者，国家级青年人才项目（海外）入选者。2023 年博士毕业于香港大学计算机系，导师为王文平教授（IEEE&ACM Fellow）和 Komura Taku 教授；2018 年本科毕业于浙江大学。本人长期关注三维重建与生成、空间智能、世界模型等方向的研究，谷歌学术总引用量 5000 余次，开源项目在 Github 平台累积获得超 1.5 万次星标关注。担任 CCF-A 类期刊 IEEE TVCG 副主编。带领学生获得 2025 ICRA、CVPR Maniskill-vitac 机器人视触觉操作竞赛冠军，获得 2025 年大学生挑战杯“人工智能+”挑战赛特等奖。入选 2025 年世界人工智能大会 WAIC 云帆奖明日之星，2026 年度中国三维视觉大会优秀青年学者。

报告题目：三维世界建模：从重建到生成

报告摘要：“生成式三维世界模型”是赋予机器空间认知、通向物理级具身智能的核心基石。本报告系统探讨从传统三维重建向生成式世界模型跨越的技术演进与前沿范式。报告首先回顾基于 3D 高斯（3DGS）等传统三维重建技术的进展，并深度剖析其在应对开放动态场景与复杂交互时所面临的局限性；为突破这一瓶颈，其次探讨前馈式（Feedforward）3DGS 等生成式三维建模方案，解析其如何实现从“单一观测拟合”到“具备泛化与预测能力”的物理级三维表征跃迁。接着，报告将视角扩展至视频世界模型，剖析其在捕获复杂时空规律与海量数据建模上的技术突破；最后，针对当前视频世界模型在长期物理一致性、长程漂移以及单纯依赖算力扩张（Scaling Law）等方面存在的固有缺陷进行深刻反思，探讨显式三维图形学先验与隐式生成架构相融合的破局之道，展望三维世界模型重塑自动驾驶闭环仿真与具身智能底层基础设施的未来图景。

分论坛 开放场景下的多模态图像融合

时间：8月4日上午09:00-12:00 地点：202会议室

内 容

嘉 宾

主席：吴小俊（江南大学）、刘 羽（合肥工业大学）
赵文达（大连理工大学）、刘晋源（大连理工大学）

面向应用的多模态图像融合

马佳义（武汉大学）

多模态信息感知与融合增强

江俊君（哈尔滨工业大学）

遥感大模型与对地观测应用

洪丹枫（东南大学）

开放场景下的多模态图像配准

邓 欣（北京航空航天大学）

从“看清”到“看懂”：复杂场景下的图像融合探索

李华锋（昆明理工大学）

论坛主席



吴小俊

IAPR Fellow、AAIA Fellow、AIIA Fellow、NAAI Member, 江南大学二级教授、至善教授、研究生院院长、Josef Kittler 人工智能研究院院长、教育部装备发展部联合基金创新团队负责人、科技部中英人工智能联合实验室主任、教育部 / 江苏省人工智能国际合作联合实验室主任、2006 年教育部新世纪优秀人才、江苏省 333 工程第一层次人才、国家重点研发计划项目首席科学家、国家自然科学基金重点项目负责人。从事模式识别与人工智能的研究, 在 TPAMI、IJCV、TIP、CVPR、ICCV、AAAI 等期刊和会议上发表学术论文 400 余篇, 出版学术著作 5 本, 研究论文被引用 3 万 1 千余次, h 指数为 77。研究成果获得国内外学术奖励 30 余项, 其中包括多项国际竞赛冠军、国际会议最佳论文奖、教育部科技进步一等奖、合作者 Josef Kittler 院士获中国政府友谊奖; 现任 IEEE 智慧城市指导委员会委员、中国人工智能学会模式识别专委会副主任、教育部计算机类教学指导委员会委员、中国图像图形学会理事和江苏省人工智能学会副理事长等职。



刘羽

刘羽，合肥工业大学仪器科学与光电工程学院教授、博士生导师，国家高层次青年人才计划入选者，科睿唯安全球高被引科学家(2023 至今)，爱思唯尔中国高被引学者(2020 至今)，IEEE 高级会员。主要研究方向包括多源图像融合、医学图像处理、计算机视觉等。近年来，在 IEEE TPAMI、IJCV、IEEE TIP、CVPR 等领域权威期刊和会议上发表论文 100 余篇，谷歌学术总被引 20000 余次（3 篇第一作者论文被引超 1000 次）。获中国图像图形学会(CSIG)青年科学家奖、吴文俊人工智能优秀青年奖、安徽省自然科学二等奖、IEEE TIM 等两个国际期刊年度最佳论文奖等。担任 IEEE TIP、Information Fusion、IEEE SPL、中国图像图形学报等期刊编委。



赵文达

赵文达，教授、博士生导师，国家级青年人才。研究方向为多模态遥感图像分析，在包括 CVPR, ICCV, AAAI 等本领域顶级会议以及 IEEE TPAMI, IEEE TIP 等本领域顶级期刊上发表学术论文 50 余篇。出版专著 2 部，授权专利 20 余项。获得山东省科技进步一等奖，中国指控学会技术发明二等奖，国际 IEEE 多媒体技术委员会最佳论文奖。兼任 IEEE Transactions on Image Processing 编委，中国人工智能学会智能融合专业委员会副秘书长，国际数字地球学会中国国家委员会青年科学家工作委员会副主任委员。



刘晋源

刘晋源，工学博士，软件学院、国际信息与软件学院副教授，博士生导师。研究主要研究方向为计算机视觉，聚焦端侧多模态大模型 (MLLM)、多任务统一模型 (Uni-Task Model)，相关研究成果发表于 TPAMI、IJCV、CVPR、ICCV、NeurIPS、ICML 等 CCF-A 类期刊会议 40 余篇，其中 7 篇期刊论文入选 ESI 高被引论文，6 篇受邀国际会议口头报告。授权国际 PCT/ 国内技术发明专利十余项，授权国际 PCT/ 国内技术发明专利十余项。。主持国家自然科学基金、JkW 课题、XX 部委预研等国家级纵向项目。研究成果获中国人工智能学会吴文俊自然科学一等奖、教育部自然科学二等奖、辽宁省优秀博士论文等。

论坛讲者



马佳义

马佳义，武汉大学机器人学院教授。研究方向为计算机视觉、信息融合等。以第一 / 通讯作者发表 CVPR、ICCV、IEEE TPAMI、IJCV、Cell 等中科院一区 / CCF A 类论文 200 余篇，谷歌学术引用 5 万余次，H 指数 103。曾获湖北省自然科学一等奖（序 1）、IEEE SPS 最佳论文奖、钱学森论文奖、Information Fusion 最佳论文奖。担任 Information Fusion 期刊共同主编，IEEE TIP、IEEE/CAA JAS 等期刊编委。

报告题目：面向应用的多模态图像融合

报告摘要：图像融合技术旨在将来自不同源图像中的互补信息集成于一体，生成一幅能够全面表征成像场景的融合图像，从而提升目标识别、场景感知与环境理解的能力。本报告以应用导向的多模态图像融合为核心，剖析该领域当前面临的挑战与发展机遇，并系统介绍若干代表性方法。重点围绕未配准多模态图像融合、具备退化鲁棒性的图像融合、高低层视觉任务协同下的图像融合、跨场景通用的图像融合等方向展开讨论，同时结合典型应用场景，展示其实际价值与广阔的应用前景。



江俊君

江俊君，哈尔滨工业大学人工智能学院教授、副院长，入选国家级青年人才计划。主要从事图像处理、多源信息感知融合研究，在 IEEE Transactions 汇刊和 CCF A 类会议发表论文百余篇，包括 IEEE TPAMI 12 篇，谷歌学术引用约 2.8 万次，H 指数 75，入选全球高被引科学家、全球前 0.05% 顶尖学者和中国高被引学者。主持国家自然科学基金联合重点项目和国家重点研发计划课题。曾获中国人工智能学会自然科学奖一等奖（序 1）、吴文俊人工智能优秀青年奖和 CCF 优博。

报告题目：多模态信息感知与融合增强

报告摘要：近年来，由于深度学习技术的兴起和飞速发展，利用大规模标注数据进行监督学习在封闭场景的特定计算机视觉任务上的性能取得了突破性的进展，与此同时这些方法的性能也日趋饱和，其发展面临来自真实开放世界等应用挑战。针对真实世界的视觉多模态融合与感知这一重大需求，我们围绕小样本、弱标注、多源异构和跨场景等关键核心问题开展研究工作，将为复杂环境下无人系统的智能感知与自主决策提供技术支撑。报告将主要聚焦团队近年来在开放场景下的多模态信息感知与融合增强相关方面的主要研究成果。



洪丹枫

洪丹枫, 东南大学教授, 博导, 德国洪堡学者, 日本 JSPS 学者, 科睿唯安全球“高被引科学家”, 海外优青。曾任德国宇航中心(DLR)研究员兼“光谱视觉”课题组长、法国格勒诺布尔阿尔卑斯大学-傅立叶实验室(GIPSA Lab)客座研究员、德国亥姆霍兹联合会人工智能研究院(HACUI)AI 科研顾问, 中国科学院空天信息创新研究院研究员, 博士生导师, 中国科学院大学岗位教授, 遥感与数字地球(全国)重点实验室副主任等。

主要研究方向为人工智能、多模态遥感、大模型、高光谱成像、地球科学等。在 Nature 子刊、IEEE TPAMI、IEEE TIP、Information Fusion、IEEE TCYB、IEEE TNNLS、The Innovation、CVPR、ECCV、NeurIPS、AAAI、ACM MM 等国内外著名学术期刊和会议上发表相关方向论文 260 余篇, ESI 热点论文 25 篇, ESI 高被引论文 77 篇, Google Scholar 引用 30000+。获得《麻省理工科技评论》—中国智能计算创新人物奖、《麻省理工科技评论》首批 AI100 青年先锋奖、中国指挥与控制学会青年科技奖、IEEE GRSS Highest Impact Paper Award、IEEE GRSS Early Career Award、Remote Sensing 青年科学家奖、国际人工智能促进会(AAAI)杰出程序委员奖等。担任 IEEE TPAMI、IEEE TIP、IEEE TCI、IEEE TGRS、ISPRS JP&RS、Information Fusion、The Innovation、遥感学报等期刊领域编委 / 副主编 / 编委。

报告题目：遥感大模型与对地观测应用

报告摘要：随着人工智能特别是基础模型技术的快速发展，遥感领域正从传统任务驱动范式迈向数据驱动与知识驱动相结合的大模型时代。遥感大模型通过利用海量、多模态、多时相对地观测数据进行预训练，为地物感知、环境监测、变化分析和智能决策提供了新的技术路径。本报告将介绍遥感大模型的发展背景与研究现状，重点分享团队在该方向的系列研究进展，包括光谱遥感大模型 SpectralGPT、多模态大模型 SeaMo 与 FlexiMo 等代表性工作，以及在地物分类、目标检测、变化检测等典型对地观测任务中的应用实践。最后，结合遥感智能化发展的新趋势，探讨遥感大模型面临的机遇与挑战。



邓欣

邓欣, 北京航空航天大学电子信息工程学院教授, 博士生导师, 国家级青年人才。研究领域为多模态图像处理、可解释性神经网络与多媒体内容安全, 在人工智能顶级期刊 TPAMI、图像处理顶级期刊 TIP、视频编码权威期刊 TCSVT, 计算机视觉领域顶级会议 ICCV、CVPR、ECCV 等发表论文 70 余篇, TPAMI 论文 8 篇(一作 / 通讯 6 篇)。在科研获奖方面, 获 2023 年度北京市自然科学一等奖(排 2)、第 13 届吴文俊人工智能自然科学一等奖(排 2)。在个人荣誉方面, 获第 14 届吴文俊人工智能青年科技奖、2021 年度英国 Eryl Cadwallader Davies Prize、2021 年度中国图像图形学会石青云女科学家奖, 入选 2025 年度北京市科技新星, 23-25 连续三年入选斯坦福大学全球前 2% 顶尖科学家榜单, 因成功揭榜“面向 AIGC 溯源证真的大容量数字水印技术”难题, 荣获 2025 年度华为“火花奖”。

报告题目: 开放场景下的多模态图像配准

报告摘要: 多模态图像配准旨在实现不同传感器采集图像之间的特征对齐, 是多模态图像融合的重要前期步骤。开放场景下, 不同模态图像不仅在分辨率上存在较大差异, 还常受到噪声干扰, 这给配准精度带来严峻挑战。本次报告将介绍我们近期针对开放场景下多模态图像配准的研究工作, 重点聚焦分辨率不一致与噪声干扰两类问题。我们将首先定量分析分辨率和噪声对配准精度的影响, 并基于此进行网络设计, 以实现跨分辨率且具有抗噪声能力的多模态图像配准。



李华锋

李华锋, 博士, 教授, 博士生导师, 云南省杰出青年基金获得者、云南省省级人才, 多次入选全球前 2% 顶尖科学家年度榜单。主要从事计算机视觉、图像处理与多模态智能感知等方向研究。主持国家自然科学基金项目 4 项、云南省杰出青年基金项目 1 项、云南省基础研究计划重点项目 1 项等科研项目。在 CVPR、ICCV、ICML、AAAI、ACM MM、IJCV、IEEE TPAMI、IEEE TIP、IEEE TIFS、IEEE TMM 等国际高水平期刊和会议发表论文 80 余篇, 入选 ESI 高被引论文 8 篇次、热点论文 2 篇次, 相关论文曾入选 CVPR Best Paper Candidate 和 CVPR Compute Transparency Champion。获云南省自然科学一等奖、二等奖各 1 项, 授权国家发明专利 50 余项。现任 IEEE TIP 副编辑、IEEE SPL 副编辑及 Information Fusion 期刊编委, AAAI SPC Member。

报告题目: 从“看清”到“看懂”: 复杂场景下的图像融合探索

报告摘要: 图像融合旨在整合不同图像中的互补信息, 提升图像的视觉质量与场景感知能力。随着真实应用需求的发展, 图像融合研究已从理想配准条件下的多源信息整合, 逐步拓展到面向下游任务、未配准图像、单图像信息增强以及退化条件下的鲁棒融合等复杂场景。本报告将围绕复杂场景下图像融合的若干关键问题展开, 介绍面向检测、分割等下游任务的图像融合方法, 分析未配准条件下跨模态信息对齐与融合的挑战, 探讨单张图像中的多层次信息融合机制, 并讨论退化与未配准同时存在时的鲁棒融合问题。报告最后将总结图像融合技术在复杂视觉感知中的发展趋势与应用前景。

分论坛

智能交互数字人

时间：8月4日上午09:00-12:00 地点：204会议室

目 录

嘉 宾

主席：邹博超（北京科技大学）、晏轶超（上海交通大学）、马惠敏（北京科技大学）

面向具身智能的人体重建与生成

李 坤（天津大学）

超写实虚拟数字人驱动技术

张盛平（哈尔滨工业大学）

面向人体运动状态识别的具身感知：从传感器感知到生理交互

姚 超（北京科技大学）

大模型启发的三维数字人生成与驱动

朱 昊（南京大学）

如何突破恐怖谷：从高精度数字人脸到灵动机器人脸

晏轶超（上海交通大学）

AI生成内容检测研究现状及未来挑战

毕秀丽（重庆邮电大学）

论坛主席



邹博超

北京科技大学计算机与通信工程学院副教授，北理工-哈佛大学医学院联合培养博士，北航-中电科联合培养博士后，适人协同智能重点学科与技术研究中心副主任，中国图像图形学会理事、副秘书长，社会安全风险感知与防控大数据应用国家工程研究中心特聘专家。主要从事视觉认知、情感计算、人机协同方向研究，作为负责人先后承担国家自然科学基金面上、青年项目、联合基金课题、新一代人工智能重大项目子课题，参与国家自然科学基金重大项目、入选北京市青年人才托举工程，在 IEEE Trans.、ICML、NeurIPS、AAAI、ACM MM 等发表 SCI/EI 论文四十余篇，授权发明专利十余项，曾获 2024 年北京市自然科学一等奖，2021 年中国电子科技集团科学技术一等奖。



晏轶超

上海交通大学副教授，X-Human 实验室负责人。主要研究方向为三维表征学习，发表高水平学术论文 50 余篇。主持国家自然科学基金青年 / 面上项目、上海市 2025 年度关键技术研发计划“新一代信息技术”项目、阿里、华为、字节等企业合作项目 10 余项。入选上海市海外高层次人才计划，获 2025 年中国图像图形学会科技进步二等奖（排 1），2024 年中国图像图形学会自然科学二等奖（排 2），2023 年挑战杯-揭榜挂帅专项赛特等奖。



马惠敏

北京科技大学领军学者, 二级教授, 计算机与通信工程学院院长、鼎新实验室主任、适人协同智能重点学科与技术研究中心主任, 中国图像图形学会党委副书记、会士。围绕无人系统类脑智能与人机混合智能等重大科学问题开展研究, 作为负责人承担了国家自然科学基金重点项目、国家重点专项等多项项目, 作为通讯作者在 TPAMI、CVPR 等国际高水平期刊和会议发表论文 200 余篇, 谷歌学术单篇他引 4500 余次; 获批专利 60 余项, 6 项专利实现了成果转化, 应用于教育、交通、医疗健康等领域。作为第一完成人获得北京市自然科学一等奖、吴文俊人工智能一等奖, 中国图像图形学会一等奖等奖励。

论坛讲者



李坤

天津大学计算机科学与技术学院与人工智能学院英才教授、博士生导师。主要研究方向为三维视觉, 尤其是人体三维重建与生成, 以第一作者 / 通讯作者在 IEEE TPAMI、CVPR 等国际知名期刊和会议上发表论文 80 余篇, 部分研究成果实现了产业化应用。主持了国家优秀青年科学基金、天津市杰出青年科学基金、国家重点研发计划青年科学家项目等 20 项科研项目。荣获中国图像图形学会石青云女科学家奖、ICME 最佳论文奖、FIFA Skeletal Tracking 'Light' 国际挑战赛第一名等荣誉。担任天津市人工智能学会副秘书长、ACM MM/AAAI/NeurIPS 大会领域主席、Fundamental Research 等国际期刊青年编委、VASE 2022 大会本地主席等职务。

报告题目：面向具身智能的人体重建与生成

报告摘要：具身智能旨在赋予智能体在真实物理世界中自主感知、决策与交互的能力, 而实现人体、行为及环境的数字化建模是连接现实世界与智能体学习过程的重要桥梁。然而, 当前仍面临人体行为数据获取困难、复杂交互场景建模不足以及虚实融合训练环境缺乏等挑战, 制约了具身智能系统的学习效率与泛化能力。人体重建与生成技术为解决上述问题提供了新的途径。本报告将介绍本研究团队近年来在人体三维重建、多人行为建模、数字孪生构建等方面的研究进展, 并探讨人体重建与生成技术在具身智能中的应用潜力与未来发展方向。



张盛平

哈尔滨工业大学长聘教授、如本科技首席科学家，入选国家级青年人才计划，布朗大学和香港浸会大学博士后、加州大学伯克利分校访问学者。主要研究方向为三维视觉、具身智能、文化遗产保护与传承。主持国家自然科学基金 5 项（包括联合基金重点项目 1 项）、华为公司项目 10 余项，被华为授予 2021 年“HUAWEI Ascend Expert”和 2022 年“昇腾众智星光奖”荣誉称号，荣获 2022 年度教育部 - 华为“智能基座”栋梁之师。研究成果获省部级科研奖励 5 项。已发表学术论文 100 余篇（包括 PNAS、IEEE T-PAMI、IJCV、ICML、NeurIPS、CVPR、ICCV、SIGGRAPH 等）。

报告题目：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。



姚 超

北京科技大学计算机与通信工程学院特聘教授，北科鼎新学者，入选国家高层次青年人才计划，北京市科协“青年人才托举”工程。主要研究方向包括图像处理、具身感知、人机交互等。在 TIP、TVCG、TCSVT、NeurIPS、CVPR、ICLR、UbiComp、ACM MM、AAAI、自动化学报等国内外学术期刊及会议上发表论文 80 余篇，申请国家发明专利 10 项。主持包括国家重点研发计划课题、国家重大专项课题，国家自然科学基金国际合作重点项目课题等国家级项目 10 余项，2022 年获国家级教育教学成果奖一等奖，获中国电子学会、中国人工智能学会、中国黄金学会科技进步奖等省部级科技奖励 8 项。

报告题目：面向人体运动状态识别的具身感知：从传感器感知到生理交互

报告摘要：随着智能可穿戴设备和物联网技术的迅猛发展，人体运动状态识别已从单纯的机械运动捕捉迈向深度的“具身感知”。通过开展多模态传感材料、多模态传感数据的质量增强与语义理解方法的研究，以解决复杂动态环境下运动意图理解模糊以及个体差异适应性欠佳的问题。进一步结合人体运动特性，研发了一系列面向具身数采、智慧医疗等领域的多模态传感装置。这不仅大幅提高了非受控场景下人体运动状态识别的鲁棒性，还显著提升了面向大规模疾病诊断、筛查的应用能力。该研究为下一代具身交互系统、智能康复医疗以及全天候健康监护提供了重要的理论依据和技术支撑。



朱昊

朱昊，于南京大学获得学士和博士学位，现为南京大学智能科学与技术学院准聘副教授、博导，国家级青年人才，共同指导南京大学三维视觉实验室(NJU-3DV)，同时也是成像与智能实验室的一员。长期从事三维视觉、空间智能、数字人方向的研究，发表 TPAMI、CVPR、SIGGRAPH 等高水平论文 50 余篇，研究成果被应用于国内首部真人全 CG 电影的制作和江苏省城市足球联赛(苏超)的现场拍摄与建模。入选教育部青年长江学者、中国科协青年人才托举工程、腾讯犀牛鸟专项研究计划，曾获得 IEEE CAS 杰出青年作者奖等、江苏省双创博士等荣誉。

报告题目：大模型启发的三维数字人生成与驱动

报告摘要：随着大模型成为驱动人工智能变革的新引擎，三维数字人的生成与驱动技术正加速迈入基础模型赋能的新阶段。面向高保真数字人建模的需求，课题组受大模型规模化生成与多模态对齐思想启发，构建了一系列兼顾高效率与高可控性的生成与驱动方法：一方面，借鉴基础模型的生成范式，攻克了单张图像亚秒级三维建模与文本引导的外貌自由编辑技术，并在复杂人群的单目三维重建上取得显著进展，显著降低了创作门槛；另一方面，受多模态大模型跨模态理解机制启发，实现了基于视觉、压感、按键等多模态输入的精细化动作控制。这些成果不仅赋予了三维数字人极具真实感的渲染效果与可信的动态交互能力，更为构建以三维数字人为载体的空间智能与具身智能融合系统奠定了技术基石。



晏轶超

上海交通大学副教授，X-Human 实验室负责人。主要研究方向为三维表征学习，发表高水平学术论文 50 余篇。主持国家自然科学基金青年 / 面上项目、上海市 2025 年度关键技术研发计划“新一代信息技术”项目、阿里、华为、字节等企业合作项目 10 余项。入选上海市海外高层次人才计划，获 2025 年中国图像图形学会科技进步二等奖(排 1)，2024 年中国图像图形学会自然科学二等奖(排 2)，2023 年挑战杯 - 揭榜挂帅专项赛特等奖。

报告题目：如何突破恐怖谷：从高精度数字人脸到灵动机器人脸

报告摘要：三维人脸表征是连接视觉感知、图形渲染与智能交互的核心基础，在数字人、人形机器人等领域具有关键支撑作用。报告聚焦结构化三维人脸表征这一核心命题，探索从高精度数字人脸重建到灵动机器人脸驱动的技术路径。首先剖析高精度三维人脸重建的关键技术，涵盖几何结构精准表征、纹理细节高保真还原及多源数据融合建模；进而探讨动态人脸表征的结构化建模方法，解决表情迁移、微表情生成与面部动作精准驱动等核心难题；最后延伸至机器人脸的个性化建模与灵动表达，打通数字人脸与物理机器人的表征迁移壁垒，实现高逼真度、高自然度的面部交互能力。



毕秀丽

重庆邮电大学教授、博士生导师，重庆市留学归国人员创新创业项目支持人选、重庆市“巴渝学者”特聘教授、国家级青年人才。主要研究方向为多模态数据感知与安全，先后在 TIP、TIFS、ACM SIGGRAPH、ICML、CVPR、ICCV、ACL 等国际知名期刊和会议上发表论文 100 余篇，2024、2025 年全球被引用次数最多的前 2% 科学家。主持国家自然科学基金 4 项，重庆市基金 2 项，重点参与国家重点研究计划，国家重大仪器研发项目各 1 项，成功实现 1 项重大成果转化，2 项试点应用落地，获得重庆市自然科学奖二等奖、重庆市科技进步奖一等奖、吴文俊人工智能科学进步奖自然科学奖一等奖。

报告题目：AI 生成内容检测研究现状及未来挑战

报告摘要：AI 生成式技术的快速发展正在深刻改变数字内容的生产方式，同时也带来了内容真实性验证的新挑战。未经识别的 AI 生成内容若被应用于科学研究、新闻传播或司法取证等关键场景，可能误导公众认知，甚至影响科学结论和社会信任。因此，AI 生成内容检测已成为人工智能安全领域的重要研究方向。本次报告将系统介绍 AI 生成内容检测的发展现状、关键技术及典型研究进展，重点分析当前检测方法在开放场景下面临的泛化能力不足、未知类别适应性差等核心挑战，并结合最新研究成果，探讨提升检测模型对未知 AI 生成内容识别能力的关键思路与技术方法，进一步展望 AI 生成内容检测的发展趋势与未来挑战。

分论坛 自动驾驶与智能座舱

时间：8月4日上午09:00-12:00 地点：302会议室

目 录 内 容

嘉 宾

主席：韩文程（山东大学）、董性平（武汉大学）
高学浩（西北工业大学）、杨思蓓（中山大学）

面向无人驾驶的3D场景理解与生成

杨 健（南京理工大学）

面向低速智能无人车的感知与决策方法

王 栋（大连理工大学）

通往2030完全自动驾驶之路和关键挑战

许春景（华为技术有限公司/车BU）

自动驾驶VLA世界模型构建

马 超（上海交通大学）

面向真实世界的具身智能探索：感知、规划与评测

丁恒辉（复旦大学）

论坛主席



韩文程

韩文程，山东大学控制科学与工程学院教授、博士生导师、国家级高层次青年人才（海外优青）。曾任澳门大学物联网智慧城市国家重点实验室博士后研究员。2025 年在澳门大学获得计算机科学博士学位，2022 年在北京理工大学获得计算机科学与技术硕士学位。长期深耕自动驾驶感知领域，相关算法已在极氪 9X 等热门车型落地，累计部署超过 50 万辆量产车型，在产业界具有广泛影响力。研究方向包括计算机视觉、自动驾驶感知、机器人控制、医学影像分析、AIGC 等。目前从事单目深度估计、3D 目标检测、视觉跟踪、自动驾驶视频生成、医学影像智能分析等研究工作。



董性平

董性平，博士、武汉大学计算机学院教授、博士生导师、国家级青年人才。曾任阿联酋起源人工智能研究院研究员。2019 年在北京理工大学获得博士学位，2016-2018 年曾在澳洲国立大学进行联合博士培养。已发表高水平论文 40 余篇，其中包括国际权威期刊 IEEE Trans. 汇刊 (IEEE TPAMI, IEEE TIP, IEEE TMM, IEEE TNNLS, IEEE TCYB, IEEE TCSVT) 和国际顶级视觉会议 (IEEE CVPR, ECCV)。谷歌学术 H 指数达到 27，单篇引用最高 800 余次。在 SCI 论文中，四篇论文入选 ESI 高被引，其中一篇被评为热点论文。连续入选斯坦福大学全球前 2% 顶尖科学家年度影响力榜单。曾获得中国人工智能学会优秀博士论文奖、北京市优秀毕业生等奖励与荣誉。目前从事小样本学习、目标跟踪、图像 / 视频目标分割、三维重建、自动驾驶感知与决策、多模态学习、医学人工智能等研究工作。



高学浩

高学浩，西北工业大学自动化学院副教授、硕导，博士毕业于西安交通大学人机混合增强智能全国重点实验室，后进入西北工业大学脑与人工智能实验室（韩军伟教授领衔）进行科研工作，博士期间曾前往清华大学、西湖大学、以及腾讯 AI Lab 开展合作研究。主要研究方向为面向空间智能的三维机器视觉研究、包括空间目标识别、立体场景感知、运动态势理解、多智能体交互分析等，在 IEEE TPAMI、TVCG、TIP、CVPR、NeurIPS 等国际期刊和会议上发表多篇论文，并主持国家自然科学基金青年基金项目 C 类、北京市自然科学基金、陕西省博士后资助项目等多项国家级和省部级科研项目。入选 2025 年度全国博士后创新人才支持计划 A 类（博新计划），荣获 2025 年度中国具身智能大会评选的首届优秀青年博士奖、国际人工智能科学院评选的 2026 年度亚太人工智能优秀青年学者奖等多项荣誉。



杨思蓓

杨思蓓，中山大学计算机学院副教授，博士生导师，逸仙学者。分别在香港大学和浙江大学取得博士和学士学位。主要研究方向为跨模态视觉感知、理解、生成与交互，聚焦于多模态统一表征学习、多模态大模型、具身智能。迄今为止在 TPAMI、CVPR、ICCV 等期刊或会议发表 CCF A 类 / 中科院一区论文 60 余篇，其中以第一作者和通讯作者发表 CCF A 类论文 40 余篇，引用 3600 余次。主持国家自然科学基金面上项目、青年项目，以及浦江人才计划、上海领军人才海外计划、腾讯犀牛鸟专项等科研项目。担任 ICCV、ICLR、NeurIPS、ECCV 等顶会领域主席，入选全球前 2% 顶尖科学家榜单。

论坛讲者



杨 健

杨健，南京理工大学计算机学院教授，长期从事模式识别和智能系统方面的研究，在 IEEE Transactions 及 Pattern Recognition 等国际权威期刊和顶级会议上发表论文 500 余篇，谷歌学术被引 6 万余次。曾获国家自然科学基金二等奖 1 项（第二完成人）；省部级一等奖 3 项（第一完成人）。入选 2016 年度 IAPR Fellow，获 2011 年度国家杰出青年科学基金、2014 年度教育部长江学者特聘教授；2018 年入选万人计划科技创新领军人才。

报告题目：面向无人驾驶的 3D 场景理解与生成

报告摘要：随着人工智能领域的快速发展，无人驾驶感知技术取得了长足进步，但依然无法满足复杂开放环境下无人系统安全运行的实际需求，从而导致无人驾驶汽车事故时有发生。该报告针对无人驾驶的视觉感知与生成任务，从 3D 占用预测、SLAM、4D 世界模型等方面，介绍了课题组近期的工作进展。



王 栋

王栋，大连理工大学信息与通信工程学院，教授、博导，国家级人才计划入选者。迄今在本领域顶级会议 (CVPR/ICCV) 及期刊 (TPAMI/IJCV) 发表论文 60 余篇，谷歌学术引用 2 万余次。获得国际视觉目标跟踪竞赛 VOT 冠军 (10 次)、CCF 自然科学二等奖、教育部自然科学二等奖、CVPR2020 最佳论文提名等学术奖励。研究工作获得国家自然科学基金优秀青年科学基金、区域联合重点项目等资助。

报告题目：面向低速智能无人车的感知与决策方法

报告摘要：本报告主要汇报团队近期在面向低速智能无人车的感知与决策方法上的阶段性进展，主要包括单机器人目标跟踪与多机器人协同探索两方面。在单机器人场景中，通过融合多源感知信息提升目标跟踪的稳定性与鲁棒性；在多机器人场景中，基于分布式协同机制实现高效环境探索与信息共享，从而提高整体探索效率。实验与仿真结果表明，所提出方法在目标跟踪精度与协同探索性能方面具有良好效果。



许春景

许春景, 1999 年于武汉大学获数学学士学位, 2002 年北京大学数学硕士学位, 2009 年获香港中文大学工程学博士学位。2009 年至 2012 年就职于中国科学院深圳先进技术研究院, 历任助理研究院和副研究员。2012 年加入华为, 历任中央研究院媒体实验室高级工程师、主任工程师、以及诺亚方舟实验室计算视觉实验室主任。2021 年转入智能驾驶产品线, 任 AI 首席专家, 参与设计和开发 ADS1.0 到 5.0 历代产品。主要的研究兴趣在机器学习和计算机视觉。

报告题目: 通往 2030 完全自动驾驶之路和关键挑战

报告摘要: 2030 年左右人类是否能够真正的实现第一个广泛使用的具身智能系统 --- 真正高度场景泛化能力的自动驾驶? 其主要的发展技术路径, 关键性的挑战性的问题是什么? 是从业的研究和工程人员一直在探寻的问题。我们将简要的回顾技术的演进脉络, 介绍当前所处的状态以及目前技术框架中的挑战问题, 以及推演未来几年的研究和技术演进趋势。



马 超

马超, 上海交通大学计算机学院、人工智能研究院教授, 博士生导师。上海市浦江人才、中国图象图形学学会优博。上海交通大学与加州大学默塞德分校联合培养博士。澳大利亚机器人视觉研究中心(阿德莱德大学)博士后研究员。主要研究多模态大模型与自动驾驶问题。谷歌学术引用 1 万 7 千余次, 连续入选爱思唯尔中国高被引学者(2020-2025)。任中国图象图形学学会优博俱乐部主席、青年工作委员会副秘书长。担任计算机视觉顶级会议 CVPR、ICCV、AAAI、ICLR、NeurIPS 领域主席, IEEE Trans. on Multimedia (TMM)、Journal of Artificial Intelligence Research (JAIR) 编委。主持国家自然科学基金优秀青年基金项目。获中国图象图形学学会青年科学家奖、MMM 2024 唯一最佳论文奖、华为技术合作领域 2021 年度优秀技术成果奖。

报告题目: 自动驾驶 VLA 世界模型构建

报告摘要: 视觉 - 语言 - 动作 (VLA) 和世界模型是端到端自动驾驶领域的两大主流技术路线。本报告主要汇报团队在统一视觉 - 语言 - 动作模型和世界模型 (VLA-World) 以进一步提升自动驾驶性能方面的工作。首先, 通过预训练实现轨迹控制未来帧生成的世界模型, 再设计 VLA 的感知 - 生成 - 思考 - 规划的监督微调方式, 最后利用强化学习优化 VLA 模块。VLA 世界模型框架简洁高效, 大量实验表明, 显著优于当前领先的 VLA 模型以及世界模型自动驾驶方法。



丁恒辉

丁恒辉, 复旦大学青年研究员, 博导, 入选国家海外高层次人才、上海市海外高层次人才、小米青年学者, 任上海市计算机学会副秘书长。2016 年于西安交通大学获学士学位, 2020 年于新加坡南洋理工大学获博士学位。曾在 TikTok AI Lab、MMLab@NTU、ETH Zurich 担任研究员 / 博士后。主要从事计算机视觉、多模态、场景理解、AIGC、世界模型等研究。过去 5 年内共发表 100 多篇 CCF-A 类论文。担任 IEEE TIP 期刊编委、Pattern Recognition 期刊编委, 担任多个国际顶级会议的 Area Chair 或 Senior Area Chair, 如 CVPR、NeurIPS、ICML、ICLR、ECCV、AAAI、ACM MM 等。

报告题目: 面向真实世界的具身智能探索: 感知、规划与评测

报告摘要: 本报告聚焦真实世界具身智能在自动驾驶等开放动态场景中的关键问题, 围绕复杂感知、自主规划与交互式评测展开。首先介绍 MOSEv2 在复杂视频感知与场景理解中的数据构建和评测价值, 支撑系统应对遮挡、快速运动、多目标干扰等挑战。随后, 介绍面向开放环境的无人机自主导航系统 AutoFly, 探讨其在无人系统感知、环境建模与规划决策中的探索, 分析智能体如何基于连续观测理解场景变化并形成行动策略。最后, 针对当前世界模型与具身智能系统在闭环交互评测方面的不足, 报告介绍 WBench 交互式视频世界模型评测体系, 从任务执行、环境反馈与动态交互等维度评估模型在自动驾驶等真实应用场景中的可用性与泛化能力。

分论坛

大模型基础理论

时间：8月4日上午09:00-12:00 地点：309会议室

目 录

嘉 宾

主席：刘 勇（中国人民大学）、胡 鹏（四川大学）、李 健（北京师范大学）

机器学习的“变”与“不变”

孟德宇（西安交通大学）

大模型架构泛化与幻觉抑制理论

侯臣平（国防科技大学）

受限感知场景下的优化定义生成式学习理论与方法

刘日升（大连理工大学）

推敲之间、渐成其文：扩散大语言模型

李崇轩（中国人民大学）

Toward a Reliable Theory of In-Context Learning:
From Mechanism to Robustness

李 帅（上海交通大学）

论坛主席



刘 勇

刘勇，中国人民大学高瓴人工智能学院院长聘副教授、研究员，国家级青年人才。长期致力于机器学习基础理论与算法的研究，在机器学习相关领域期刊会议发表论文百余篇。



胡 鹏

胡鹏，四川大学计算机学院教授，博士生导师，国家级青年人才，CSIG 青工委副秘书长。主要研究方向为可信机器学习和多媒体分析。2019-2020 年在新加坡科技研究局（A*STAR）担任研究员。目前已在 Nature 子刊、TPAMI、IJCV、CVPR 等 JCR 一区期刊及 CCF A 类会议上发表长文 80 余篇，包括多篇 ESI 热点论文（前 1%）和高被引论文（前 1%）。研究成果获得中国图像图形学会（CSIG）青年科学家奖、中国图像图形学会（CSIG）自然科学一等奖、四川省自然科学二等奖、中国人工智能学会教学成果激励计划一类成果（序 1）。担任 TPAMI、TIP、CVPR、ICCV、NeurIPS、ICML、ICLR 等期刊和会议审稿人 /（高级）领域主席 / 高级程序委员。主持国家自然科学基金（联合重点、面上、青年）项目、国家重点研发计划课题、JKW GF 科技创新特区项目、四川省重点研发基金项目等。



李健

李健，北京师范大学人工智能学院副教授、博士生导师。主要研究大模型基础理论、高效智能体，关注资源约束下大模型的信息压缩、能力保持与泛化机制。已发表学术论文40余篇，Google Scholar 引用1000余次；其中以第一作者在 IEEE TIT、JMLR、Artificial Intelligence、ICML、NeurIPS 等人工智能领域顶级期刊和会议发表论文十余篇。主持国家自然科学基金面上项目、青年项目、国家重点研发计划子课题、中国博士后科学基金特别资助等项目，累计项目经费300余万元。曾入选微软亚洲研究院“铸星计划”、中国科学院特别研究助理资助、中科院信工所优才计划 A 类与 B 类、百度研究院 AIDU 人才计划等，并获 PRICAI 2021 最佳学生论文奖。

论坛讲者



孟德宇

孟德宇，西安交通大学数学与统计学院教授。长期致力于机器学习基础理论与算法的研究，在机器学习相关领域期刊会议发表论文百余篇。现任 TPAMI 等七个期刊编委。

报告题目：机器学习的“变”与“不变”

报告摘要：在深度学习快速迭代的浪潮中，前沿研究聚焦于从“变”的角度构建机器学习方法，如增加数据 / 标记规模、设计创新网络架构、构建多样学习模式等。然而，从机器学习的基础研究视角，我们发现机器学习的各个环节中存在更为本质的“不变性”规律与内涵，如数据高维标记空间的低维特征模式、网络基础模块的不变 / 等变性结构本质、学习模式设计的内在统一性规律等。把握这些不变性内涵，有利于我们更深刻理解提升机器学习泛化性、鲁棒性、可解释性机理的理论方法途径，更合理利用这些基础原理设计更加简洁合理的学习模式，构建更加具有深刻内涵的机理 - 数据双驱动、知识 - 网络相融合的有效机器学习方法。基于此，本报告将介绍针对高维标记空间低维不变性隐空间提炼的“标记分布建模”理论与方法、针对网络基础卷积模块旋转 - 尺度 - 仿射等变性结构刻画的“参数化卷积”理论与方法、针对机器学习方法超参设置不变性规律提炼的“模拟学习方法论”理论与方法，从而尝试探讨对机器学习方法如何从“变”中提炼其“不变”内涵的方法论思想，为机器学习的基础研究与工程应用提供一种可参考的视角。



侯臣平

侯臣平，国防科技大学教授，博士生导师，国自科创新群体核心成员。主要从事人工智能基础方面的研究工作，在自适应学习理论与应用等方面取得了系列研究成果，并成功应用于国防领域。近年来，在 IEEE TPAMI、JMLR 等国内外刊物和会议上以第一 / 通讯作者发表学术论文 150 余篇（包括 IEEE/ACM 汇刊长文 60 余篇），担任 ICML、IJCAI、AAAI 等会议的 AC/SPC/PC 等，是 PR、FCS 等 SCI 期刊的编委，多篇论文进入 ESI 各层次高被引论文。主持军科委基础加强重点项目、科技部 2030 重大项目课题、国家自然科学基金重点项目课题等 15 项。获中国图像图形学会自然科学一等奖 1 项、省部级一等奖 1 项，获国家杰青、优青、湖南省杰青等。

报告题目：大模型架构泛化与幻觉抑制理论

报告摘要：大模型时代的智能应用场景日益复杂，泛化能力提升与推理幻觉抑制已成为人工智能研究中的重要挑战。本报告围绕“大模型架构泛化与幻觉抑制理论”展开研究，从基础功能单元、架构组件单元以及组件单元级联方式三个层面出发，探讨不同结构设计对大模型泛化能力与推理稳定性的影响，并进一步研究如何通过新型架构设计提升大模型可信泛化与推理能力。最后，对该研究方向的发展趋势与未来挑战进行总结和展望。



刘日升

刘日升，大连理工大学教授、软件学院副院长、教育部重点实验室副主任，国家优秀青年科学基金获得者，IET Fellow，CCF、CSIG 杰出会员。主要从事人工智能基础理论、优化方法与无人系统智能技术研究，发表 CCF 推荐 A 类及 IEEE 汇刊论文 100 余篇。以第一完成人获吴文俊人工智能自然科学奖一等奖，并获教育部自然科学奖二等奖、CCF-IEEE CS 青年科技奖、CAA 青年科技奖等。担任 ICML、NeurIPS、ICLR、CVPR 等会议领域主席，TPAMI、TIP、TCSVT 等期刊编委。

报告题目：受限感知场景下的优化定义生成式学习理论与方法

报告摘要：极端受限感知场景中，少劣观测、设备受限、机理不清与任务耦合使传统生成式学习面临“缺根少魂”的瓶颈。本报告将重点汇报团队针对上述挑战开展的优化定义生成式学习理论与方法研究工作，主要涉及面向生成式学习的双层优化理论，任务协同学习、生成约束学习与机理嵌入的学习方法，以及其在弱光增强、红外成像、水下视觉、RAW 图像复原和弱小目标探测等受限感知任务中的应用。



李崇轩

李崇轩，中国人民大学高瓴人工智能学院副教授、博士生导师。致力于生成模型基础理论、建模范式、大规模训练策略和高效采样算法的研究，带领团队研制了扩散语言模型 LLaDA，开源模型下载量破 600 万次，多项成果部署于 DALL·E 2、Stable Diffusion、Vidu 等行业领先大模型。成果发表于 NeurIPS、ICML、ICLR、TPAMI、Nature Communications 等，谷歌学术引用 1.4 万余次。获机器学习领域顶级国际会议 ICLR 2022 杰出论文奖；北京市自然科学基金一等奖（排名 2）；吴文俊人工智能自然科学一等奖（排名 5）。主持国家自然科学基金青年基金 B 类、重大研究计划培育项目等，入选智源学者、吴文俊优秀青年奖、北京市科技新星。担任 IEEE TPAMI 编委和 ICLR/ICML/NeurIPS 等国际会议领域主席；作为主编出版《大模型十讲》教材，指导博士生入选国家自然科学基金青年学生基础研究项目、字节跳动奖学金计划等。

报告题目：推敲之间、渐成其文：扩散大语言模型

报告摘要：大模型时代的智能应用场景日益复杂，泛化能力提升与推理幻觉抑制已成为人工智能研究中的重要挑战。本报告围绕“大模型架构泛化与幻觉抑制理论”展开研究，从基础功能单元、架构组件单元以及组件单元级联方式三个层面出发，探讨不同结构设计对大模型泛化能力与推理稳定性的影响，并进一步研究如何通过新型架构设计提升大模型可信泛化与推理能力。最后，对该研究方向的发展趋势与未来挑战进行总结和展望。



李 帅

李帅, 上海交通大学人工智能学院特聘副教授、约翰·霍普克罗夫特计算机科学中心副主任、上海创智学院全时导师, 研究可自主决策适应动态环境的强化学习理论与方法、扩散模型与大模型的分析, 迄今共发表学术论文 100+ 篇, 包含上海交通大学首篇机器学习理论顶会 COLT 论文等, 其中第一 / 通讯作者发表 CCF-A 类论文 50+ 篇, 10 余项理论提升成果仍保持理论最优。担任 TMLR 的执行编委 (Action Editor), NeurIPS、AAMAS 的高级领域主席 (Senior Area Chair), 机器学习顶会 ICML、ICLR、ACL、UAI 等的领域主席 (Area Chair), 受邀于群体智能顶级会议 AAMAS 上给出多智能体在线学习 (2024) 与马尔可夫博弈理论基础 (2025) 的教程、IJCAI 上给出多智能体在线学习 (2025) 的教程和 DAI 上给出多智能体在线学习的教程 (2025), 主持国自然重大培育项目、面上基金、青年基金, 参与国自然重大研究计划、科技部 2030 新一代人工智能重大项目。曾获得 AAAI-IAAI Deployed Application Award (2025)、上海市东方英才拔尖项目、上海市扬帆人才计划、上海徐汇光启人才、华为火花奖、香港政府外展合作奖、谷歌博士奖学金、国际 SAT 竞赛并行求解赛道铜牌、腾讯优秀导师奖等。首位博士生已入职南方科大担任助理教授 (2024), 指导学生获得 CCF- 多智能体优博 (2025)、百度奖学金 (全球仅 10 位)、商汤奖学金、腾讯犀牛鸟优秀学生奖、上海市优秀毕业生、国家奖学金、美团北斗计划、腾讯青云计划等。

报告题目: Toward a Reliable Theory of In-Context Learning: From Mechanism to Robustness

报告摘要: In-context learning (ICL) enables language and sequence models to infer tasks from a few demonstrations without updating their parameters. Despite its empirical success, a central theoretical question remains: what learning algorithm does a model execute during inference, and when is its behavior reliable?

In this talk, we present a unified optimization-based view of ICL across Transformers and Mamba state-space models. We first explain how changing contextual representations can induce a controlled shift in the effective prediction target, closely paralleling a weight update. We then show how Transformer hidden states can carry an implicit predictor and how attention and MLP layers approximate gradient descent on deep neural-network parameters. This mechanism extends theoretical guarantees for ICL beyond linear regression to classification and smooth function classes. Moving beyond attention-only models, we analyze the duality between Mamba-2's structured state-space duality layers and linear self-attention. A single SSD layer can realize one gradient step, whereas hybrid SSD-LSA architectures preserve the information required for multi-step approximate stochastic gradient descent. Expressive power alone, however, does not guarantee reliable ICL. We therefore study prompt-order sensitivity introduced by positional encoding and find that larger reorderings cause greater prediction changes, while longer contexts reduce the effect of a fixed-size perturbation. Together, these results show that a complete theory of ICL must explain both what models learn from a prompt and how reliably they learn it.

分论坛

三维视觉与空间智能

时间：8月4日上午09:00-12:00 地点：401会议室

内 容

嘉 宾

主席：郭裕兰（中山大学）、雷印杰（四川大学）

胡心亭（中国科学技术大学）、姜海勇（中国科学院大学）

从三维生成到可交互世界：面向具身智能的
空间资产编辑、场景重构与世界模型

杨易（浙江大学）

4D世界模型：基座、应用与评测

张兆翔（中国科学院自动化研究所）

Street-Level 3D Scene Generation from
Single Satellite Image

夏桂松（武汉大学）

4D动态世界模型构建

戴玉超（西北工业大学）

语义对齐的多模态空间感知与生成

刘家瑛（北京大学）

面向开放三维世界的世界建模、生成和交互

齐晓娟（香港大学）

论坛主席



郭裕兰

中山大学教授，博士生导师，国家高层次青年人才。主要研究领域为三维视觉与空间智能，包括三维重建、理解、生成及智能系统应用。主持国家自然科学基金联合重点项目等纵向项目 10 余项，发表学术论文 200 余篇，谷歌学术总被引 2.7 万余次，入选 Clarivate 全球高被引科学家、Elsevier 中国高被引学者，获中国计算机学会自然科学一等奖、吴文俊人工智能优秀青年奖等奖励。担任中国图像图形学会三维视觉专委会副主任，IEEE Transactions on Image Processing 高级领域编辑，曾担任首届中国空间智能大会主席，历届中国三维视觉大会组委会主席，CVPR、ICCV、ECCV、NeurIPS、ICML 等国际会议领域主席。



雷印杰

雷印杰, 教授, 国家级青年人才、四川省杰青、四川省千人计划, 四川大学网络空间安全学院党委书记。主要研究方向为“多模态认知计算”。近年来, 主持了包括国家自然科学基金重点项目、国家重点研发计划课题、国防基础科研重点项目、四川省杰出青年科学基金等三十余项国家级、省部级及企业委托科研项目。在人工智能等领域的国内外顶级学术会议和期刊上发表学术论文 100 余篇, 包括 IEEE TPAMI、TIP、CVPR、ICCV、ECCV 等, 并申请了三十余项国家发明专利。先后担任 CVPR、ICCV、ECCV、ACM MM、ECCV、AAAI、WACV 等人工智能领域国际知名学术会议的领域主席(AC)、高级程序委员会委员(SPC)。



胡心亭

胡心亭, 中国科学技术大学人工智能与数据科学学院教授, 博士生导师, 国家级青年人才。主要研究方向为视觉智能、多模态感知与生成、鲁棒高效视觉学习、可控生成与交互式多模态智能。相关成果发表于 CVPR、ICCV、NeurIPS、ICLR、ECCV 等国际重要会议 20 余篇。多次担任 NeurIPS、ICLR 等国际会议领域主席。



姜海勇

姜海勇, 中国科学院大学人工智能学院副教授。主要研究方向为三维深度学习、几何建模与分析、鲁棒三维场景感知。近年来先后主持了北京市科技新星项目、国家自然科学基金项目以及国家重点研发子课题。相关成果先后发表于 CVPR、ICCV、NeurIPS、IJCV、ACM ToG、IEEE TVCG 等国际重要会议与期刊。

论坛讲者



杨 易

IEEE Fellow、IAPR Fellow, 浙江大学求是讲席教授、国家特聘专家。现任浙江大学人工智能研究所所长、微软-教育部视觉感知重点实验室主任、人工智能省部共建协同创新中心副主任。相关论文在谷歌学术上引用超过 9 万次, H-index 达 152, 近 8 年连续入选科睿唯安全球“高被引学者”。他的科研成果获得国内外广泛认可, 荣获教育部全国优秀博士论文(2010)、浙江省自然科学一等奖(2015)、澳大利亚计算机学会颠覆创新金奖(2016)、谷歌学者研究奖(2016)、澳大利亚科研终身成就奖(2019)、亚马逊机器学习科研奖(2020)、AAAI 最具影响力论文(2021)、ACM MM 唯一最佳论文奖(2023)、中国图像图形学会自然科学奖一等奖(2023)、浙江省科学技术进步一等奖(2024)、首届祖冲之奖-人工智能前沿创新奖(2024)、吴文俊人工智能科学技术奖自然科学一等奖(2025)等多项奖项, 并在国际科研竞赛中获得 30 余次世界冠军。

报告题目: 从三维生成到可交互世界: 面向具身智能的空间资产编辑、场景重构与世界模型

报告摘要: 具身智能的发展依赖于能够承载感知、行动、交互与试错的高质量三维环境。传统仿真环境通常依赖人工建模和规则驱动, 存在资产生产成本高、场景多样性不足、交互扩展困难等问题。近年来, 三维生成模型与三维编辑技术的快速进展, 为构建可扩展、可控、可演化的具身仿真环境提供了新的技术路径。本报告将围绕“三维生成—三维编辑—世界模型”这一主线展开。三维生成解决虚拟环境中的资产供给问题, 使文本、图像或多模态条件能够快速转化为三维对象、材质与结构; 三维编辑进一步解决环境可控修改问题, 使已有资产和场景能够根据任务需求完成语义编辑、局部替换、结构变形与上下文融合。在此基础上, 具身仿真环境不再只是静态三维场景的集合, 而逐渐演化为具备状态表示、物理约束、交互反馈与长期一致性的生成式世界。



张兆翔

中国科学院自动化研究所研究员、博士生导师，模式识别北京市重点实验室常务副主任，中国科学院大学人工智能学院副院长，国际模式识别学会会士、亚太人工智能学会会士，“教育部长江学者特聘教授”，研究方向是人工智能、具身智能、智能体学习。先后主持了国家人工智能 2030 重大专项、中国科学院先导、国家自然科学基金重点项目、联合基金重点、重点国际（地区）合作等多项国家级重点项目，获北京市科技进步奖一等奖、北京市科技奖杰出青年中关村奖、中国电子学会科技进步一等奖等。

报告题目：4D 世界模型：基座、应用与评测

报告摘要：本报告首先对主流世界模型进行了梳理，从数据维度的角度分析了当前世界模型领域的发展现状，重点探讨了重建生成一体化世界模型的技术特点和现有局限，提出了一种可拓展的通用 4D 世界模型 NeoVerse 及其训练方法。报告进一步分享了该通用 4D 世界模型在自动驾驶和具身智能两个重点下游领域的垂直应用。最后介绍课题组围绕视频世界模型的全方位评测所进行的最新工作，对下一代通用世界模型的预期特性进行展望。



夏桂松

武汉大学弘毅特聘教授，现任人工智能学院执行院长、国家多媒体软件工程技术研究中心副主任。长期从事人工智能、计算机视觉、空天智能信息处理等研究，主持国家自然科学基金杰青、优青、重点等纵向项目 30 余项，系列成果发表论文 150 余篇，谷歌学术引用 4 万余次，并在国家重要工程中应用。获湖北省自然科学一等奖等省部级一等奖 7 项。兼任 2 个 SCI 一区期刊编委、中国图像图形学会遥感图像专业委员会副主任。

报告题目：Street-Level 3D Scene Generation from Single Satellite Image

报告摘要：随着城市数字孪生、自动驾驶导航和增强 / 虚拟现实 (AR/VR) 技术的快速发展，从易于大规模获取的卫星图像直接构建逼真的地面街道级 3D 场景，成为了计算机视觉与摄影测量领域一项极具潜力和挑战性的前沿任务。然而，卫星俯视视角与街道平视视角之间存在巨大的几何形变与信息不对称，这为视角的转换与 3D 结构的推断带来了极大的困难。本报告将围绕跨视角场景理解与 3D 生成的演进路线，系统性地探讨如何克服这一难题，实现从“单幅卫星图像”到“全面街道级 3D 场景”的跨越式生成。



戴玉超

戴玉超, 西北工业大学长聘教授、博士生导师, 校学术委员会委员, 陕西省信息获取与处理重点实验室主任。研究方向为机器视觉、空间智能与具身智能, 主持国家自然科学基金青年科学基金 A 类项目(原国家杰青项目)、联合基金重点项目、国家级纵向重点项目等项目。近年来在 TPAMI、IJCV、ICCV、CVPR、NeurIPS 等国际著名期刊和会议上发表论文 70 余篇, 谷歌学术引用超过 17000 次, H 因子 63。获 CVPR 最佳论文奖、陕西省自然科学奖一等奖、CSIG 青年科学家奖等奖项。担任国际顶刊 TPAMI 编委、IJCV 编委、APSIPA 杰出讲者、China3DV 大会主席、中国空间智能大会程序主席和 CVPR、ICCV、ECCV、ICML、NeurIPS 等国际会议领域主席。

报告题目: 4D 动态世界模型构建

报告摘要: 具身智能致力于实现能够像人类一样, 在物理世界中具有自主感知、理解、推理并执行复杂任务能力的通用智能体, 构建具备时空预测能力的“世界模型”是其核心底座。本报告汇报课题组在动态场景重建与理解方面的系列工作, 探讨如何从海量视觉数据中构建 4D 动态具身世界模型, 并展望其在深空探测、空天协同等高动态与极端环境下应用与发展趋势。



刘家瑛

刘家瑛, 北京大学王选计算机研究所博雅特聘教授, IEEE Fellow, 教育部“长江学者奖励计划”特聘教授, 电子出版新技术国家工程研究中心副主任。研究领域为智能媒体计算, 发表 IEEE/ACM 汇刊与 CCF A 类论文 100 余篇, 谷歌学术引用 2.9 万余次。担任 IEEE Trans.on Image Processing, IEEE Trans.on Circuit System for Video Technology, IEEE Multimedia Magazine 等期刊编委, IEEE ICME/ACM ICMR 指导委员会委员, ACM MM Asia 2024 大会主席, ACM MM Asia 2025/2023, ACM ICMR 2021/IEEE ICME 2021 大会程序主席, CVPR/ICCV/ECCV/ACM MM 会议领域主席等。高校计算机专业优秀教师奖励计划获得者, 获省部级奖励三项、CSIG 石青云青年女科学家奖、北京大学教学卓越奖、王选青年学者奖等。

报告题目: 语义对齐的多模态空间感知与生成

报告摘要: 基于语义对齐的多模态空间感知与生成技术, 通过融合视觉、语言、姿态等多维语义线索, 赋予模型对空间结构与内容关系的深层理解能力。然而, 真实场景中的空间语义复杂多变, 不同模态间的对齐误差易导致生成内容失真或行为误判, 如何在有限监督下实现鲁棒的跨模态语义对齐与空间生成, 仍是亟待攻关的难题。本报告将围绕该方向开展研究, 从图像、视频背景空间和视频前景人体三个层次递进, 首先提出基于语义循环一致性的图像编辑方法, 利用循环理解对齐后训练提升编辑保真度; 进而面向视频背景空间, 设计去噪轨迹对齐策略以保护前景并实现自然换背景与重光照; 最后在视频前景的三维人体理解任务中, 引入姿态模态对齐机制, 借助多模态大模型增强对动态行为的精准感知, 并缓解以人为中心的自我视角的空间感知幻觉。



齐晓娟

齐晓娟现任香港大学电气与计算机工程系副教授。她于香港中文大学获得博士学位，并曾在加拿大多伦多大学、英国牛津大学以及英特尔视觉计算研究组开展科研合作与学术交流。她的研究致力于赋予机器在开放环境中感知、理解与重建视觉世界的能力，重点关注具身智能系统中的真实场景应用。迄今为止，她已在计算机视觉、计算机图形学与机器学习等领域的顶级国际会议（包括 SIGGRAPH Asia、CVPR、ICCV、NeurIPS 等）发表论文 100 余篇，其中多篇入选口头报告。齐晓娟博士曾获得 SIGGRAPH Asia 最佳论文荣誉提名奖，并入选 2024 年 IEEE AI “10 to Watch” 以及 MIT TR35 China。她还积极服务学术界，长期担任 ICCV、CVPR、NeurIPS、ICML 和 AAAI 等国际顶级会议的领域主席 (Area Chair)。

报告题目：面向开放三维世界的世界建模、生成和交互

报告摘要： Real-world AI systems perceive and act through continuous streams rather than isolated images or completed videos. This makes streaming understanding and generation essential for embodied agents, robotics, AR, and interactive multimodal systems, where models must process incoming observations causally, maintain compact memory, and produce stable outputs in real time. In this talk, I will present three recent works toward this goal. First, I will discuss streaming depth perception, where single-image geometry foundation models suffer from temporal scale-shift drift when applied frame by frame. Our Dynamic Feature Normalization method stabilizes latent feature statistics over time, enabling consistent streaming 3D reconstruction while preserving the accuracy of the frozen pretrained backbone. Second, I will introduce streaming vision-language understanding, focusing on spatial reasoning under partial observations, long-term memory, and active exploration. Third, I will discuss streaming generation, where efficient long-video generation requires memory-efficient training, quantized inference, and real-time decoding. Together, these works point toward a new paradigm of streaming AI: systems that continuously perceive, remember, reason, and generate under real-world temporal and computational constraints.

分论坛 多模态大语言模型

时间：8月4日上午09:00-12:00 地点：402会议室

内 容

嘉 宾

主席：常晓军（中国科学技术大学）、黄 高（清华大学）、杨 学（上海交通大学）

Multi-Modal Models for Embodied Intelligence with Geometry-Aware Generation and Reasoning

沈春华（浙江大学）

端到端高效构建多模态交互大模型

冯 洋（中国科学院计算技术研究所）

学科大模型建设思路分享——参数凝聚智慧，训练铸造应用

高广亮（智慧树网）

基于多模态大模型的视觉感知理解

刘 伟（上海交通大学）

从视频生成到长时稳定的世界模型

范 琦（南京大学）

从原始视频到多模态空间智能：大模型时代的3D数据自动化构建

钟志航（上海交通大学）

论坛主席



常晓军

常晓军，中国科学技术大学讲席教授，国家级高层次人才，中国科学院海外引进杰出人才，民盟盟员，澳大利亚研究理事会青年研究奖获得者。主要从事具身智能、多模态大模型及类脑多模态大模型等方面的研究。主持包括国家自然科学基金区域联合基金，科技部新一代人工智能重大专项等在内的国家级科研项目十余项。相关科研成果发表在 IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI)、International Journal of Computer Vision (IJCV) 等国际顶级期刊以及 CCF A 类会议上共计 150 多篇，Google Scholar 引用次数超过 2 万次，其中 21 篇论文入选 ESI 高被引 / 热点论文。2019 年至 2024 年连续入选科睿唯安高被引科学家，入选爱思唯尔 2024“中国高被引学者”榜单。担任 IEEE TCSVT、IEEE TNNLS、ACM TOMM 等国际顶级期刊的副主编，并在多个 CCF A 类国际会议中担任领域主席 (Area Chair)，积极参与国际学术社区建设与服务。



黄 高

黄高，清华大学长聘副教授，博导。研究领域为人工智能基础模型及算法，在 Nature Machine Intelligence、IEEE TPAMI 等期刊、会议发表论文 200 余篇，总引用 10.5 万余次，代表作 DenseNet 单篇引用逾 6 万次。获教育部、CAAI、CSIG 自然科学一等奖、CVPR Best Paper 和 NeurIPS Best Paper Runner-Up、MIT TR35、AI 2000 人工智能最具影响力学者；担任 IEEE TPAMI 等国际期刊编委和多个人工智能顶级会议领域主席。



杨 学

杨学, 上海交通大学自动化与感知学院特聘教轨助理教授, 主持国家自然科学基金项目, 曾入选第十届中国科协青年人才托举工程、中国计算机学会优博、上海市启源青年学者, 并连续两年入选爱思唯尔中国高被引学者名单和连续三年入选全球前 2% 顶尖科学家榜单。研究方向为多模态大模型、空间智能和遥感影像解译。在 CCF-A 类顶级期刊和会议上发表 70 余篇论文, 包含一篇最具影响力 AAAI21 论文榜首、CVPR2026 最佳论文候选、遥感期刊年度最佳论文, 谷歌学术引用累计超过 1.2 万次, 并担任 ICLR/AAAI/NeurIPS 顶级会议领域主席。

论坛讲者



沈春华

Chunhua Shen has been a University Chair Professor at Zhejiang University since 2022, where he serves as Deputy Director of the State Key Laboratory of CAD&CG and leads research in artificial intelligence, machine learning, and computer vision. Prior to joining Zhejiang University, he spent nearly two decades in Australia, holding key positions including: Principal Applied Scientist at Amazon Australia, Full Professor at the University of Adelaide and the Australian Institute for Machine Learning, and Chief Investigator at the Australian Centre for Robotic Vision, National ICT Australia (NICTA), and the Australian National University.

He earned his BSc and MSc from Nanjing University, an MPhil from the Australian National University, and a PhD from the University of Adelaide, Australia. His research centers on computer vision, statistical machine learning, and multi-modal models, with influential contributions to anchor-free object detection (e.g., FCOS) and deep learning fundamentals. As a highly cited scholar, his Google Scholar citations exceed 100,000 as of May 2026, with an H-index of 152. Several of his research outputs, including image parsing and object detection technologies, have been successfully deployed on tens of millions of mobile devices and applied in autonomous driving scenarios. His honors include the Australian Research Council Future Fellowship (2012) and the Chang Jiang Scholars Distinguished Professorship (2022).

报告题目: Multi-Modal Models for Embodied Intelligence with Geometry-Aware Generation and Reasoning

报告摘要: This talk presents our advances in bridging multi-modal models with embodied intelligence by focusing on 3D/4D scene construction and generation, reinforcement learning-based visual reasoning, and geometry-aligned video synthesis. We first introduce methods for accurate 3D representation and 4D dynamic scene reconstruction from multi-modal inputs. We then leverage reinforcement learning to enable robust reasoning over images and videos for sequential decision-making. To improve physical consistency, we explicitly inject camera pose information into video generation frameworks, producing geometrically aligned and spatially coherent outputs. These contributions enhance the perception, reasoning, and generation capabilities of embodied systems, supporting more reliable and realistic interactions in dynamic environments.



冯 洋

冯洋, 中国科学院计算技术研究所研究员、博士生导师, 研究方向为自然语言处理和大模型, 担任国际计算语言学学会 (ACL) 秘书长以及 ACL/EMNLP/IJCAI 等国际会议高级领域主席。主导研发了百聆系列大模型, 先后获得 ACL2019 唯一最佳论文奖、中国人工智能学会第七届理事会“杰出贡献奖”、CCF 自然语言处理专委会“青年新锐奖”、中文信息学会“钱伟长中文信息处理科学技术奖 - 汉王青年创新奖”一等奖等, 主持多项国家重点研发计划项目、国家自然科学基金委项目等。

报告题目: 端到端高效构建多模态交互大模型

报告摘要: 由于人类通过多模态信息与世界发生交互, 因此多模态大模型成为通往通用人工智能的必经之路。然而多模态大模型面临着模态异构、数据短缺、模态对齐困难、实时交互等挑战, 比训练文本大模型难度更大。针对这些挑战, 本报告将介绍我们的一些探索, 包括构建支持全双工交互语音大模型、高效视觉大模型和端到端三模态大模型。



高广亮

智慧树 AI 事业部区域总监，在教育信息化领域深耕多年，专注于创新性 AI 技术的研究和应用，AI 技术的多领域包括语音技术、图像技术、云平台服务、大数据服务、以及 AI 课程研发等方面。目前负责京沪云贵川地区高校 AI 整体解决方案相关业务。

报告题目：学科大模型建设思路分享——参数凝聚智慧，训练铸造应用

报告摘要：1、学科知识体系智能构建；2、AI 赋能全流程教学工具；3、个性化自适应学习成效；4、跨模态复杂任务处理能力。



刘 伟

刘伟现为上海交通大学自动化与感知学院特聘副教授、博士生导师，入选国家高层次青年人才、上海市海外高层次人才和上海市浦江人才。刘伟博士近年来主要研究多模态大模型和基于深度学习的计算机三维视觉任务包括三维重建、深度估计、三维检测等。他已主持多项国家级和上海市科委项目，并与华为等企业存有广泛的项目合作，相关成果发表学术论文八十余篇，现担任国际期刊 Transactions on Image Processing 期刊编委（Associate Editor）。

报告题目：基于多模态大模型的视觉感知理解

报告摘要：本次报告将从三项多模态视觉相关研究展开介绍。第一，针对 CLIP 受文本编码器限制、仅适配短文本对齐的缺陷，我们提出 Fix-CLIP 模型，将词元上限由 77 拓展至 248，实现精细化图文匹配，在图文检索任务中性能超越同期同类算法。第二，面向多模态智能体复杂视觉推理存在观测偏差、忽略执行反馈的痛点，本文提出 V-ABS 行动观测驱动束搜索框架，搭配熵自适应加权算法，并依托 80k+ 自建微调数据集优化推理路径，该方案适配开源与闭源模型，在主流基准测试中取得最优结果。第三，为解决现有深度估计算法缺失对象级信息、鲁棒性薄弱的问题，我们创新性结合 SAM 模型提出 RoSAMDepth 框架，通过特征增强与区域约束策略优化自监督深度估计效果，综合性能位居行业前列。



范 琦

范琦，南京大学智能科学与技术学院副教授，紫金学者，博士生导师，国家高层次青年人才。范琦博士毕业于香港科技大学，主要研究方向为计算机视觉和世界模型，近年以通讯作者和第一作者身份在计算机视觉顶级会议和期刊（CVPR、ECCV、ICCV、ICLR、IJCV 等）发表学术论文近三十篇。近两年主持国家人才项目、部委人才项目、国自然、省自然青基、CCF 基金、企业横向等项目，主持项目经费累计超千万。

报告题目：从视频生成到长时稳定的世界模型

报告摘要：近年来，多模态生成模型推动视频生成从短视频合成走向长时序、可交互的世界模拟。随着生成时长增加，模型不仅要保证局部画面质量，还需在持续演化过程中维持稳定一致的场景结构与内容状态。本报告围绕长视频生成与视觉世界建模这一主线，介绍如何从整体上规划长视频的演化过程，在生成过程中及时修正逐步累积的偏差，并将历史场景中的空间结构压缩为可持续利用的隐式记忆，从而支持超出有限上下文范围的长期一致生成。这些工作旨在推动视频生成模型从视觉内容生成进一步走向能够长期维持空间与内容一致性与稳定性的视觉世界模型。



钟志航

钟志航，副教授、博士生导师，上海市海外人才引进计划青年项目获得者。曾任职于浦江国家实验室，作为核心成员参与研发的“书生·浦动”系统已服务于巴黎奥运会、汤尤杯、中网等国际重大赛事，并在 CCTV5 等主流媒体平台多次播出与报道，相关成果获 2025 年度“上海设计 100+”全球竞赛 AI 创新设计奖（行业应用）。

钟志航于 2023 年毕业于东京大学计算机科学系，获博士学位；2018 年毕业于浙江大学机械电子工程系（竺可桢学院混合班），获荣誉学士学位。求学期间曾获东京大学系长学术奖（全系每年 1 名博士）、国家优秀自费留学生奖学金 B 类（全球 50 名、日本领区 2 名）、日本学术振兴会青年科学家 Fellowship (JSPS DC) 等多项荣誉。

研究方向涵盖空间智能体、神经渲染与计算摄影等领域，在国际顶级计算机视觉和机器学习会议 (CVPR、ECCV、ICCV、ICML) 及期刊 (TPAMI、IJCV) 发表论文近 30 篇，文章多次获得 Oral、Spotlight，以及 CVPR 2026 Award Candidate 的认可。

报告题目：从原始视频到多模态空间智能：大模型时代的 3D 数据自动化构建

报告摘要：空间智能正在成为多模态大模型与三维视觉交叉中的重要方向。围绕“大模型时代 3D 视觉还能做什么”这一问题，本报告将介绍我们在空间数据自动化构建方面的探索。我们提出 Holi-Spatial，首次尝试从原始视频流出发，构建一条端到端、全自动的空间数据生成管线，将高保真 3D Gaussian Splatting 重建、开放词汇感知、2D-to-3D 实例 Lifting 与场景级 refinement 有机结合，在无需 3D 传感器和人工逐场景标注的条件下，自动生成兼具几何结构与多层语义监督的空间数据。基于该管线，我们进一步构建了 Holi-Spatial-4M 数据集，为空间智能的数据规模化提供了一条新路径，并验证了其对 3D grounding 与 spatial reasoning 等任务的显著增益。进一步地，报告还将延伸至 SpaceDG，探讨空间智能在模糊、低照度等非理想视觉观测下的鲁棒性问题。通过这两项工作，我们希望说明：在大模型时代，3D 视觉的重要价值并未减弱，而是在向空间智能的多模态数据引擎与真实世界能力支撑进一步演化。

分论坛

青托论坛

时间：8月4日上午09:00-12:00 地点：207会议室

内 容

嘉 宾

主席：魏秀参（东南大学）、彭春蕾（西安电子科技大学）
张 奇（同济大学）、张勇丙（昆明理工大学）

物理具身基础模型：表征、架构及真机评测

梁小丹（中山大学）

基于一致性约束的多模态感知与生成

余旻婧（天津大学）

生成式图像视频压缩

鲁 国（上海交通大学）

面向用户建模的记忆机制：挑战、技术与应用

吴 乐（合肥工业大学）

图机器学习异常检测基础模型

谢 诚（云南大学）

资源受限条件下的多模态感知与推理

杨 旭（西安电子科技大学）

论坛主席



魏秀参

魏秀参, 东南大学青年首席教授、博士生导师, 国家基金委青年科学基金项目 B 类(原优青)获得者、国家重点研发计划首席青年科学家、江苏省杰青。主要从事计算机视觉、机器学习和机器人方面研究, Google Scholar Citations 逾 10000 次, 目前担任 IEEE TPAMI/TIP/TMM/TCSVT 编委, NeurIPS/CVPR/ICCV/AAAI/IJCAI 等领域主席, 并任中国图像图形学会(CSIG)理事、Fundamental Research/ 电子学报(英文版)青年编委、CSIG 青托俱乐部主席、中国计算机学会计算机视觉专委会副秘书长、CSIG 青工委副秘书长等。曾入选 CSIG 青年科学家奖、吴文俊人工智能优秀青年奖等。主编教材《解析深度学习》已重印 8 次并入选“十四五”国家重点出版物出版规划、中国计算机学会推荐教材和省部级重点教材等。



彭春蕾

彭春蕾，西安电子科技大学教授、博导、党支部书记。长期从事可控可解释可靠的视觉身份智能分析与识别技术研究，围绕视觉身份的智能生成、伪造检测和鲁棒识别核心问题发表论文90余篇，主持国家自然科学基金面上项目，以排名第二参与2项国家自然科学基金重点项目。入选中国科协青年人才托举工程、中国人工智能学会优秀青年奖、陕西省特支计划青拔、陕西省青年科技新星、段宝岩科教创新奖，获湖北省技术发明一等奖、中国图像图形学会自然科学二等奖、陕西省优博、中国图像图形学会优博，指导研究生入选电子学会硕士学位论文激励计划。担任CSIG青工委委员、CCF CV专委会委员、Valse执行领域主席委员会委员等。



张 奇

张奇，同济大学计算机科学与技术学院特聘副教授，先后获得北京理工大学计算机科学与技术工学博士学位和悉尼科技大学分析学哲学博士学位。入选国家级人才（海外）、上海白玉兰海外高层次人才、上海市浦江人才计划，主持和参与国家自然科学基金项目、科技部中重大专项青年科学家项目、科技部重点研发。主要研究方向为多模态信息处理，具体包括脑视觉编解码、时间序列分析、多模态模型压缩等，在人工智能和数据科学领域的顶级国际期刊和会议上发表论文100余篇。近年来担任AAAI、IJCAI、ICLR、NeurIPS、ICML、ACM MM、CVPR等会议程序委员或领域主席，并在多个国际期刊担任客座编辑和审稿人包括TPAMI、TKDE、TSMC、TOIS、Machine Learning、TIST等，并在The Web Conference、KDD等多个高水平会议上进行受邀讲座和专题讲座（Tutorial），担任CSIG青工委副秘书长、上海计算机学会计算机视觉专委会副秘书长、YOCSEF上海学术AC委员。



张勇丙

张勇丙，博士，硕士生导师。2024年12月毕业于昆明理工大学获计算机科学与技术博士学位。主要从事自然语言处理、信息检索和文本摘要等方向的研究。近年来，以第一作者或通讯作者在PR、EAAI、ESWA、NAACL、BIBM等国际期刊和会议发表学术论文10余篇，授权国家发明专利8项。主持中央引导地方科技发展资金项目1项（财政经费400万元），参与国家自然科学基金、云南省重大科技专项计划项目、云南省自然科学基金重点项目及企事业合作开发项目等5项。2022年获云南省科技进步特等奖1项。

论坛讲者



梁小丹

梁小丹，中山大学教授，博导，获得国家级万人计划青年拔尖人才，教育部 U40 人才项目，科技部人工智能重大专项项目负责人等支持，Google Scholar 引用超过 40,000 余次。获 IEEE CVPR 2024 Best Paper Finalist, 广东省科技进步一等奖，阿里巴巴达摩院青橙奖（全国计算机领域 35 岁以下十名）、CSIG 石青云青年女科学家奖、世界人工智能大学优秀论文奖，吴文俊人工智能优秀青年奖、中国图像图形学会科技技术一等奖、连续 5 年入选爱思唯尔中国高被引学者榜、ACM MM 中国新星奖、福布斯中国 30 Under 30 科学榜、斯坦福大学“全球前 2% 顶尖科学家榜单”，终生科学影响力榜、国际 ScholarGPS2024 年全球前 0.05% 顶尖科学家榜单等。并于 2025 年荣获深圳光明区五一劳动奖章、劳模称号、工匠人才创新工作室优秀工作室奖，多次获得华为火花奖。十余次担任 CCF-A 类顶级学术会议 ICCV、CVPR、NeurIPS 等的领域主席，并为 CVPR 监察委员会。科研成果转化为 70 余项授权发明专利，并落地于华为、字节、腾讯、联想、DeepSeek 等国内科技领军企业的产品。

报告题目：物理具身基础模型：表征、架构及真机评测

报告摘要：本报告围绕具身智能尚未迎来“GPT 时刻”的核心问题，系统梳理视觉 - 语言 - 动作 (VLA) 基础模型的技术趋势、核心挑战与未来方向，明确具身智能与大模型、机器人范式的路径差异。当前具身智能形成 VLA、世界模型、智能体三大主流范式，正逐步走向融合，分别支撑动作对齐、环境预测、高层决策能力。其核心瓶颈在于互联网语义知识与物理世界执行的深层鸿沟，具体表现为物理先验缺失、跨模态对齐精度不足、数据采集成本高、实时性与安全性难以平衡、评测体系不统一等问题。重点介绍 A1、AtomicVLA、eWAM 等创新模型，通过自适应早退、原子动作分解、VLA 与世界模型原生融合等技术，显著提升模型效率、泛化性与持续学习能力。同时提出 PhyAgentOS 标准化框架与 ManipArena 真机评测基准，解决硬件差异、评测不可复现等行业痛点。



余旻婧

余旻婧，天津大学英才副教授，主要开展计算机图形、人机交互与认知科学的交叉研究。入选中国科协青年人才托举工程，主持国家自然科学基金、天津市自然科学基金等多项科研项目，并获中国图像图形学会石青云女科学家奖、吴文俊人工智能自然科学奖二等奖等奖励。已在人工智能、计算机图形学、计算机视觉、认知计算、多媒体、人机交互等领域的顶级期刊与会议发表论文四十余篇，包括 IEEE TPAMI、ACM TOG、IEEE TVCG、CVPR、ICCV、IEEE TAFFC、ACM MM、ACM CHI、IJHCS 等；相关工作曾获世界华人数学家联盟年度最佳论文奖、全国几何设计与计算学术会议 (GDC) 最佳论文奖、IEEE 机器人与仿生学国际会议 (ROBIO) 最佳学生论文提名奖、北京青年优秀科技论文等荣誉。

报告题目：基于一致性约束的多模态感知与生成

报告摘要：多模态感知与生成的核心挑战之一，在于如何保持不同模态之间的一致性：模态间的细微错位，往往会导致识别偏差、生成失真乃至体验破坏。要实现稳健、可信的多模态感知与生成，需跨越三个层面的一致性鸿沟——感知层面，各模态可靠性随样本与个体动态漂移，固定融合规则难以泛化；生成层面，跨模态信号在时序、几何与语义上难以对齐，易产生不自然现象；应用层面，感官通道与虚拟场景缺乏对齐，削弱交互的临场感。本报告以“一致性约束”为主线，介绍课题组近期围绕这三个层面的系列研究：在感知层面，提出按可靠性动态组织与路由的多模态情绪识别方法，提升跨个体场景下的稳健性；在生成层面，面向语音、音乐等信号驱动的表情与动作生成，引入几何约束与循环一致性监督，使生成结果与输入在时序、几何与语义上忠实对齐；在此基础上，将上述能力落地于沉浸式交互，通过感官反馈与虚拟场景的动态对齐提升临场感，并在文化遗产体验等真实场景中得到验证。



鲁 国

鲁国，上海交通大学电子工程系副教授、博士生导师。主要研究方向是视频编码与多媒体处理。在 IEEE T-PAMI、CVPR 等上发表论文 50 余篇。作为项目负责人先后承担自然科学基金委青年、面上、重点子课题、重点研发子课题以及华为等企业项目等。担任权威期刊 IEEE T-CSVT 编委，IJCV 等期刊客座编辑，ICLR /NeurIPS 等会议领域主席，在 CVPR/ACMMM 上组织多次智能编码讲习班。入选中国科协青年人才托举工程，获首届 IEEE 视觉信号处理与通信新星奖、中国图像图形学会优秀博士论文奖，上海交通大学优秀博士论文奖、IEEE WCSP/VCIP 等会议论文奖等。

报告题目：生成式图像视频压缩

报告摘要：生成式图像视频压缩编码利用大规模视觉先验，在极低码率下实现内容重建与细节补全，为传统压缩框架带来了新的发展空间。本报告将简要梳理生成式图像视频压缩编码的发展脉络，介绍其在感知质量优化、生成先验利用等方面的基本思路与关键问题。随后，报告将重点汇报本团队围绕生成式图像与视频压缩开展的相关工作，包括视频先验驱动的图像压缩方法，以及面向实际部署的轻量化实时视频编码技术。最后，报告将结合相关 benchmark 和近期研究进展，讨论该方向在评价标准、模型效率、时空一致性和应用落地方面的机遇与挑战。



吴 乐

吴乐, 合肥工业大学人工智能创新学院副院长。主要研究领域包括数据挖掘、数据价值、个性化推荐、因果推断应用研究等。近五年在知名国际期刊及国际会议发表论文七十余篇。主持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、国家级青年人才项目等国家级、省部级及应用合作项目数十项。研究成果获得了华为火花奖, 入选了安徽省新时代青年先锋、全球华人 AI 学者榜单、中国科协青年人才托举工程, 获得了吴文俊人工智能优秀青年奖、中国人工智能学会优秀博士论文奖等, 指导了 4 名学生获得了省级优秀博士 / 硕士论文。担任了 IEEE Trans. on Big Data, AI OPEN, Frontiers of Computer Science 期刊编委, 及多个顶级国际会议组委会成员或者高级程序委员会委员。

报告题目: 面向用户建模的记忆机制: 挑战、技术与应用

报告摘要: 大语言模型的快速发展正在深刻重塑个性化推荐与用户建模的研究范式: 用户偏好表示从传统的隐式嵌入向基于语义的持续记忆演进。在此背景下, 记忆机制作为实现长时泛在个性化推荐的关键基础设施, 面临多重核心挑战: 如何将用户复杂、多维的兴趣偏好组织为结构化的语义记忆, 并随用户行为的演变实现动态更新? 如何在长期持续交互的场景下高效管理记忆的写入与查询, 在有限资源约束下保障记忆质量? 如何在跨场景、跨域的用户建模任务中实现记忆的精准检索与选择性迁移, 规避无关记忆引入的噪声与负迁移风险? 本报告将围绕上述问题, 分别介绍课题组在结构化演化记忆、面向长期运行智能体的高效记忆构建算法以及基于记忆检索的跨域序列推荐三项工作上的近期思考与研究进展, 并探讨以记忆为中心的智能体个性化范式为用户建模研究带来的新机遇。



谢 诚

谢诚博士毕业于上海交通大学软件学院。从事图机器学习和工业数字化相关研究。入选中国科学技术协会“青年托举工程”人才; 入选“云南省高层次人才引进计划青年项目”和“东陆青年学者项目”。以第一作者(通讯), 在 NIPS、ICLR、AAAI、IJCAI、KDD、TII、TITS、TAI、TIM、TETCI、TCSS、IOTJ 等工业信息化、人工智能等方向的顶级期刊和会议发表学术论文 60 余篇, 学术引用 2000 余次。主持图机器学习和工业信息化相关国家自然科学基金 2 项, 省部级项目 10 余项, 获上海市科技进步奖二等奖, 云南省自然科学二等奖。指导学生获“国创赛(互联网+)”国家级金奖 2 项。主持国家一流课程《软件设计与体系结构》。其研究成果转化服务地方经济建设产生经济价值超过 2700 余万元。

报告题目：图机器学习异常检测基础模型

报告摘要：本报告系统梳理图异常检测（Graph Anomaly Detection, GAD）从专用模型向基础模型演进的技术脉络。针对传统 GAD 方法依赖特定数据环境、泛化能力受限的痛点，首先阐释跨域图异常检测（Generalist GAD, GGAD）的解决思路，即通过最小化域间分布差异与知识迁移机制，实现模型从标注充足的源域向稀疏目标域的高效适配。进而，聚焦于文本属性图（Text-Attributed Graphs, TAG）这一关键载体，探讨如何利用大语言模型的语义推理能力与图神经网络的结构建模优势，构建统一的“文本 - 拓扑”双模态表征空间。在此基础上，深入剖析图异常检测基础模型的核心架构，包括面向异构图数据的通用编码器设计、基于 SKILL 和 Agent 的灵活推理范式以及对未知场景的零样本泛化能力。通过整合上述技术，该基础模型旨在突破“一域一模型”的开发壁垒，为复杂图数据环境下的异常识别提供通用、鲁棒的解决方案，并展望未来在开放世界场景中的落地挑战与研究前景。

**杨 旭**

杨旭，西安电子科技大学副教授，陕西省青年拔尖人才支持计划入选者，主要从事多模态内容理解、持续学习等相关领域的研究，在 IEEE Trans 一区期刊和 CCF-A 顶级会议上发表论文 50 余篇。主持国家自然科学基金面上、青年项目、陕西省重点研发计划，作为骨干成员参与国家重点研发计划重点专项、国家自然科学基金重点基金等项目多项。入选陕西省科协青年人才托举工程，获 2023 陕西省自然科学奖一等奖、2021 年吴文俊人工智能科学技术奖优秀博士论文奖以及 2021 ACM XI' AN 优博奖。担任 VASLE 执行领域委员会委员、CSIG 青工委委员等。

报告题目：资源受限条件下的多模态感知与推理

报告摘要：多模态大模型的爆发式发展，极大推动了知识推理、语义理解等任务的性能上限，但模型规模与计算开销的快速增长，使其难以直接部署在边缘终端等资源受限环境中。在保障语义完整性与任务精度的前提下，探索体系化的轻量化技术路径，实现多模态模型的高效推理与落地部署，已成为当前人工智能领域的核心研究方向之一。本次报告围绕“资源受限条件下的多模态感知与推理”核心主题，结合系列前沿研究成果，从知识压缩、模型量化、部署优化三个层面系统展开。主要内容包括面向多模态模型的 Token 剪枝方法，并聚焦基础模型的轻量化实践，探讨低资源约束下高性能与效率平衡策略，最后延伸至边缘部署场景，介绍架构搜索与跨架构云边适配技术。

分论坛

优博论坛

时间：8月4日上午09:00-12:00 地点：206会议室

内 容

嘉 宾

指导专家：马惠敏（北京科技大学）、彭宇新（北京大学）

董 晶（中国科学院自动化研究所）、贾 伟（合肥工业大学）、赵才荣（同济大学）

分论坛主席：马 超（上海交通大学）、晏轶超（上海交通大学）

傅慧源（北京邮电大学）、王青旺（昆明理工大学）

共情智能：从面部原语到多模态协同

李 勇（东南大学）

人体视觉的协同表征学习

晏轶超（上海交通大学）

多模态人体动作对齐研究及其应用

王洪松（东南大学）

面向虚拟现实人机交互的全景图像质量评价与优化

鄢杰斌（江西财经大学）

持续学习与大语言模型知识编辑

王世鹏（西安交通大学）

基于大脑时空网络解码的精神疾病智能诊疗研究

姜 林（电子科技大学）

指导专家



马惠敏

马惠敏，北京科技大学领军学者，二级教授，计算机与通信工程学院院长、鼎新实验室主任、适人协同智能重点学科与技术研究中心主任，中国图像图形学会党委副书记、会士。围绕无人系统类脑智能与人机混合智能等重大科学问题开展研究，作为负责人承担了国家自然科学基金重点项目、国家重点专项等多项项目，作为通讯作者在 TPAMI、CVPR 等国际高水平期刊和会议发表论文 200 余篇，谷歌学术单篇他引 4500 余次；获批专利 60 余项，6 项专利实现了成果转化，应用于教育、交通、医疗健康等领域。作为第一完成人获得北京市自然科学一等奖、吴文俊人工智能一等奖，中国图像图形学会一等奖等奖励。



彭宇新

彭宇新, 北京大学二级教授、博雅特聘教授, IEEE/CCF/CAAI/CIE/ CSIG Fellow, 2019 年国家杰青, 2025 年国家杰青延续资助, 2019 年国家万人领军人才, 863 项目首席专家, 国家重点研发计划“社会治理与智慧社会科技支撑(平安中国)”重点专项总体专家组专家, 中国工程院“人工智能 2.0”规划专家委员会专家, 中国人工智能产业创新联盟专家委员会主任, CSIG 副理事长。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。主持了 863、国家自然科学基金重点(2 项)、北京自然科学基金联合基金重点、发改委专项等 40 多个项目。成果应用于国家网信办、公安部等重要单位以及华为、腾讯、快手、蔚来、美团、中国电信、中国铁塔等头部企业。担任 IEEE TCSVT 高级领域编委、IEEE TMM 等期刊编委, 培养博士生获中国计算机学会、中国电子学会等优博。



董晶

董晶, 中国科学院自动化所研究员、博士生导师, CSIG/CCF/IEEE 高级会员, CAAI 杰出会员, 中国科学院青年创新促进会会员, 国家高层次青年人才计划入选者, 目前担任中国图像图形学会(CSIG)理事、秘书长、女工委副主任, IEEE 亚太区执委、奖励委员会主席、IEEE 信号处理协会全球成员发展主席。主要从事人工智能安全与对抗、计算机视觉、多媒体内容取证等前沿方向的技术研究, 她先后以课题(子课题)负责人承担了 4 项国家自然科学基金、3 项国家重点研发课题、1 项工信部国家重大专项子项及 20 余项省部级科研项目(课题)。已在国际权威期刊及学术会议上发表学术论文 100 余篇, 已授权 30 项发明专利含 3 项美国专利。她曾获 4 次最佳(优秀)论文奖、2016 年度 IBM 学院奖、2019 年度中国人工智能学会杰出贡献奖、2020 年度 CSIG 石青云女科学家奖、2021 年度北京青年优秀科技论文奖、2021 年度 CSIG 科技奖二等奖、2021 年度吴文俊人工智能科学技术奖(技术进步科普类)、2022 年度中国发明协会创业创新奖一等奖、2022 年度国家广电总局 MediaAIAC 大赛深度合成技术应用类一等奖、2023 年度吴文俊人工智能科学技术奖技术发明一等奖、2024 全国大数据与计算智能挑战赛一等奖以及 2025(第二届)“金灵光杯”中国互联网创新大赛总决赛金奖。



贾伟

贾伟, 博士, 合肥工业大学计算机与信息学院教授, 博士生导师, 智能科学与技术系系主任。中国图像图形学会青年工作委员会副主任。中国图像图形学会生物特征识别专委会(筹)副主任。中国自动化学会模式识别与机器智能专业委员会常务委员。30 余次担任国际国内顶级会议的程序主席、组织主席等。已在 TPAMI、IJCV、TIP、TIFS、CVPR、ICCV 等权威期刊和顶级会议上发表论文 CCF A 类及中科院 1 区论文 70 多篇。H 因子 41。2020-2025 连续 6 年入选前 2% 顶尖科学家。主要研究兴趣为人工智能、生物特征识别、计算机视觉、模式识别等。



赵才荣

赵才荣, 工学博士, 同济大学计算机科学与技术学院教授, 博士生导师, 上海市计算机学会计算机视觉专委会主任, 中国图像图形学会青工委秘书长。入选东方英才, 担任 Pattern Recognition、中国图象图形学报等期刊编委。主要研究领域: 计算机视觉。已在 TPAMI、CVPR 等发表学术论文 60 余篇。主持了国家自然科学基金重点、面上等项目 5 项。研究成果获 2022 年上海市科技进步一等奖, 2023 年上海市自然科学二等奖, 2025 年吴文俊人工智能科技进步二等奖。获 2024 年全国人工智能应用场景创新挑战赛总决赛特等奖, 获 2024 年上海市计算机学会教学成果一等奖。指导学生获中国电子学会优秀硕士论文, 获上海市计算机学会优秀博士论文、优秀硕士论文。

论坛主席



马 超

马超, 上海交通大学计算机学院、人工智能研究院教授, 博士生导师。上海市浦江人才、中国图像图形学会优博。上海交通大学与加州大学默塞德分校联合培养博士。澳大利亚机器人视觉研究中心(阿德莱德大学)博士后研究员。主要研究多模态大模型与自动驾驶问题。谷歌学术引用 1 万 7 千余次, 连续入选爱思唯尔中国高被引学者(2020-2025)。任中国图像图形学会优博俱乐部主席、青年工作委员会副秘书长。担任计算机视觉顶级会议 CVPR、ICCV、AAAI、ICLR、NeurIPS 领域主席, IEEE Trans. on Multimedia (TMM)、Journal of Artificial Intelligence Research (JAIR) 编委。主持国家自然科学基金优秀青年基金项目。获中国图像图形学会青年科学家奖、MMM 2024 唯一最佳论文奖、华为技术合作领域 2021 年度优秀技术成果奖。



晏轶超

晏轶超, 上海交通大学计算机学院、人工智能研究院副教授, 博士生导师, X-Human 实验室负责人。主要研究方向为人体视觉表征学习, 发表高水平学术论文 50 余篇。主持国家自然科学基金青年/面上项目、上海市 2025 年度关键技术研发计划“新一代信息技术”项目、阿里、华为、字节等企业合作项目 10 余项。入选上海市海外高层次人才计划, 获 2025 年中国图像图形学会科技进步二等奖(排 1), 2024 年中国图像图形学会自然科学二等奖(排 2), 2023 年挑战杯-揭榜挂帅专项赛特等奖。



傅慧源

傅慧源，北京邮电大学计算机学院教授、博士生导师，国家级青年人才、北京市杰青。担任北邮国内合作与成果转化处副处长、物联网技术中心主任。主要研究方向包括图像视频处理、视频物联网、多媒体系统等。主持国家重点研发计划课题、国家自然科学基金(3项)、国防科技创新特区专项、北京市杰出青年科学基金专项、北京市科技新星计划专项等科研项目。发表高水平学术论文80余篇，已授权40余项国际/国家发明专利，相关成果应用于中国移动、中国电信、华为、小米、浙江大华等公司。研究成果获教育部科技成果一等奖、中国电子学会科技进步一等奖、中国通信学会科技进步二等奖等科研奖励，获CVPR 2026图像智能去噪挑战赛冠军、CVPR 2026智能驾驶挑战赛冠军、CVPR 2024 AI生成图像质量评估挑战赛冠军、MM 2020通用异常事件识别挑战赛冠军、国防科技创新特区人工智能挑战赛冠军等竞赛奖励。



王青旺

王青旺，昆明理工大学教授，高校GIS新锐，主要研究方向为空天多源遥感协同解译和AI for Science。承担国家自然科学基金项目3项(面上、青年C、地区)，省部级项目4项(包括云南省优青)，企业开发项目4项，在人工智能、遥感领域期刊和会议上发表论文100余篇，其中ESI高被引论文7篇，会议最佳论文4篇，授权国家发明专利30余项，获批软件著作权13项，相关科研成果已应用于安录国际、昆明滇和等企业，获中国电子学会自然科学一等奖(序4)，中国有色金属工业协会科技进步二等奖(序3)，中国指挥与控制学会科技进步二等奖(序8)，中国产学研合作促进会科技创新二等奖(序4)。指导学生参加中国大学生计算机设计大赛、全国大学生数学建模等学科竞赛，获国家级奖项18项，省部级24项；获批教育部产学研合作协同育人项目2项，发表教改论文2篇。

论坛讲者



李 勇

李勇, 东南大学副教授, 博士毕业于中国科学院计算技术研究所, 先后在香港城市大学, 新加坡南洋理工大学留学和工作, 入选 2021 年 CSIG 优博及 2025 年教育部海外引才计划。长期聚焦单 / 多模态情感计算方面的研究工作, 在领域内国际期刊 / 会议发表论文 50 余篇 (含第一作者 IEEE T-PAMI 3 篇), 作为 AI 方向负责人执行香港特别行政区“早期阿尔兹海默症患者康复训练范式效果精准评估”研究课题, 主持国家自然科学基金面上项目、蚂蚁集团校企合作项目、智慧医疗及情感陪伴机器人项目等, 现任中国图像图形学会优博俱乐部副秘书长。

报告题目：共情智能：从面部原语到多模态协同

报告摘要：共情智能 (Empathic AI) 是实现人工智能“情智兼备”的核心愿景, 而提升模型在复杂场景下的细微情感理解能力仍是当前研究的重难点。本报告围绕复杂场景稳健情感计算, 系统介绍申请人在细粒度情感表征、多模态协同理解与可信决策等方面的研究进展: 首先, 针对细微情感样本稀缺与跨域泛化难题, 率先构建基于大规模无标注视频的解耦情感自监督学习框架, 实现关键情感线索的自动发现与跨场景迁移, 被国际同行评价为细微表情分析领域的突破性进展; 进一步, 提出解耦式多模态蒸馏、非完备模态重建与生成学习框架, 揭示多模态知识动态蒸馏机理, 实现多模态情感语义的统一建模与协同增强, 并被国内外多家高校和企业广泛采用与拓展; 最后, 面向复杂自然场景中的遮挡、场景变化和个体差异问题, 提出多粒度情感自动编码与个性化可信决策机制, 大幅度提升开放场景下情感感知的鲁棒性, 其中代表性成果 (第一作者) 入选 IEEE T-IP 年度 Top-25 最受关注论文, 单篇他引超过 1080 次, 被多名国际知名学者评价为该领域代表性工作。本报告尝试构建“细粒度情感原语建模—跨模态关联对齐—复杂场景可信决策”的技术范式, 并分享相关技术在情感陪护机器人, 智慧医疗, 用户体验大模型等领域的应用落地体会。



晏轶超

晏轶超, 上海交通大学计算机学院、人工智能研究院副教授, 博士生导师, X-Human 实验室负责人。主要研究方向为人体视觉表征学习, 发表高水平学术论文 50 余篇。主持国家自然科学基金青年 / 面上项目、上海市 2025 年度关键技术研发计划“新一代信息技术”项目、阿里、华为、字节等企业合作项目 10 余项。入选上海市海外高层次人才计划, 获 2025 年中国图像图形学会科技进步二等奖(排 1), 2024 年中国图像图形学会自然科学二等奖(排 2), 2023 年挑战杯 - 揭榜挂帅专项赛特等奖。

报告题目: 人体视觉的协同表征学习

报告摘要: “人”是图像与视频中的关键目标, 在复杂场景下实现精准的人体视觉表征学习, 不仅是图像视频理解与生成的核心问题, 也是三维角色建模、人形机器人等前沿应用的共性关键技术。要得到强泛化、高精度、高动态的自动化人体视觉表征, 需要跨越三条鸿沟: 一是环境干扰与表观特征之间的泛化鸿沟, 二是三维视觉与三维图形之间的精度鸿沟, 三是局部运动与整体变化之间的语义鸿沟。本报告将介绍晏轶超在人体视觉的协同表征学习方面的研究: 在表观感知层面, 提出上下文感知人体表观特征协同学习方法, 跨域泛化能力提升 25%; 在三维建模层面, 提出拓扑一致约束三维人体建模理论, 突破亚毫米重建精度; 在运动推理层面, 提出语义渐进人体运动表征与推理方法, 攻克人体运动的高效理解、数字人与机器人运动的可控生成难题。



王洪松

王洪松, 东南大学计算机科学与工程学院博士生导师。2018 年博士毕业于中国科学院自动化研究所, 同年获中国图像图形学会优秀博士学位论文提名奖。主要研究以人为中心的视觉感知和 AI for Science 等方向。以第一或通讯作者发表 / 录用 CCF-A 会议 / 期刊 20 余篇; 获得国际 / 国内顶级会议级别竞赛冠军、亚军 8 项; 主持自然科学基金青年基金(C 类)、江苏省青年基金、国家自然科学基金专项项目课题等项目。

报告题目: 多模态人体动作对齐研究及其应用

报告摘要: 随着多模态大模型、基础模型和数字人技术的快速发展, 人体动作与文本、语音、图像等多模态信息之间的语义对齐问题受到广泛关注。多模态人体动作对齐旨在建立不同模态与人体运动表示之间的统一语义空间, 实现跨模态理解、检索、生成与编辑, 为数字人、虚拟现实、人机交互及具身智能等应用提供基础支撑。本报告首先系统介绍多模态人体动作的研究进展, 然后主要介绍我们在文本生成动作、音频生成动作、统一动作表示和零样本行为理解等典型任务的研究成果。最后, 报告总结多模态人体动作对齐研究面临的数据稀缺、语义抽象、模态异构、动作表示统一性不足等问题, 并结合大模型、世界模型和具身智能的展望发展趋势。



鄢杰斌

鄢杰斌, 江西财经大学计算机与人工智能学院副教授、人工智能系副主任、硕士生导师, 江西省人工智能检验检测中心技术负责人, 江西省 C 类人才, 入选中国科协青年人才托举工程项目、CSIG 博士学位论文激励计划、江西省“赣鄱俊才”支持计划青苗人才, 曾获江西省自然科学奖一等奖(排二)。

报告题目: 面向虚拟现实人机交互的全景图像质量评价与优化

报告摘要: 本次报告首先介绍我们在视觉内容理解与显著区域检测、内容感知的视觉信号质量评价和基于质量反馈的感知优化三方面的研究成果。然后, 介绍我们在虚拟现实人机交互的全景图像质量评价与优化开展的工作, 旨在解决当前该领域面临的三大关键问题: 视口预测依赖导致质量评价可迁移性弱、单模态模型泛化性能不足, 以及视觉增强优化目标与实际应用需求脱节。报告主要讲解对应方案, 包括: 基于几何结构感知的统一全景图像质量评价模型、多模态大语言模型先验知识增强的模型和面向虚拟现实人机交互的视觉增强模型。



王世鹏

王世鹏, 西安交通大学助理教授, 入选 CSIG 博士学位论文激励计划、西安交通大学青秀人才、国家资助博士后研究人员计划(B 档)。研究方向为持续自主化学习, 核心成果发表在 IEEE TPAMI、CVPR、AAAI 等人工智能和计算机视觉领域的国际顶级期刊 / 会议, 并在国际机器视觉顶级会议 CVPR 上进行口头报告, 被中国图像图形学会遴选为青年科学家优秀论文。主持科技部重点研发计划子课题 1 项、国自然青年科学基金 1 项、博士后面上项目 1 项。担任 Nature Communications、IEEE TPAMI、IEEE TNNLS、IEEE TIP、IEEE TMI、CVPR、ICCV 等顶级期刊 / 会议审稿人。

报告题目: 持续学习与大语言模型知识编辑

报告摘要: 随着大语言模型的广泛应用, 如何对其内部庞大的知识库进行精准、高效且安全的更新与编辑, 成为了当前研究的热点。本报告将围绕“持续学习与大语言模型的知识编辑”这一主题, 系统介绍本课题组在该方向的系列研究进展: 首先, 探讨基础持续学习机制, 介绍基于特征零空间缓解灾难性遗忘的方法, 阐述如何利用参数空间的几何特性在不干扰已有能力的前提下融合新知识; 其次, 针对大语言模型的精准更新需求, 报告基于子空间分析的大语言模型知识编辑方法, 探索如何通过定位和微调来实现知识的高效局部干预; 最后, 介绍安全感知的大语言模型知识编辑, 讨论如何在注入新知识的同时防范潜在风险, 维持模型的价值对齐与输出安全。



姜 林

姜林，围绕多模融合时空网络成像及疾病应用研究，以第一 / 通讯作者在 Nature Communications、Science Advances、Nature Mental Health、Research 等发表论文 20 余篇，授权 / 申请发明专利、软著 7 项，出版专著 1 部。主持国自然青年科学基金项目(C 类)、博士后面等项目，入选 2025 中国图像图形学会博士学位论文激励计划。成果入选 2023 中国脑机接口十大进展，被中国生物医学工程学会、中国人工智能学会等评选为青年优秀论文一等奖。任中国中文信息学会情感计算专业委员会委员, Exploration 期刊青年编委。

报告题目：基于大脑时空网络解码的精神疾病智能诊疗研究

报告摘要：精神疾病临床诊疗长期面临诊断主观性强、疾病异质程度高、症状重叠与共病现象普遍等难题，在实现个体化、精准化诊疗方面仍存在显著挑战。通过从单模态到多模态、从宏观到微观的跨尺度整合分析，我们开发了系列具有临床转化潜力的客观诊疗标记物，并在此基础上构建了面向精神疾病的脑机接口闭环调控系统，为精神疾病的客观评估和个体化治疗提供新思路与方法支持。

讲习班

多模态遥感理解

时间：8月4日下午14:00-17:30 地点：102会议室

内 容

嘉 宾

主席：周 峻（格里菲斯大学）、舒振球（昆明理工大学）

物理模型驱动AI及其应用

肖 亮（南京理工大学）

高光谱遥感技术与应用

高连如（中国科学院空天信息创新研究院）

多模态基础模型及边缘具身智能研究进展

谢卫莹（西安电子科技大学）

论坛主席



周 峻

周峻现为澳大利亚格里菲斯大学信息与通信技术学院教授、副院长，IEEE Fellow。曾任澳大利亚国立大学和澳大利亚国家信息通信技术研究院研究员。研究方向包括模式识别、高光谱成像和计算机视觉，及其在遥感、环境、农业和医学中的应用。在 IEEE TIP、IEEE TGRS、PR、AAAI、IJCAI、CVPR 和 ICCV 等图像处理、遥感、计算机视觉等领域的顶级期刊和会议上发表论文 300 余篇。担任 Springer Nature Computer Science 主编，IEEE TGRS、IEEE TMM、Pattern Recognition、Journal of Soils and Sediments 等国际期刊副主编，IEEE IGARSS 2025 等多个国际会议主席和程序委员会主席。2024 至 2025 年任澳大利亚模式识别学会主席。



舒振球

舒振球，昆明理工大学信息工程与自动化学院教授、博士生导师，云南省人工智能学会副理事长，入选云南省“兴滇英才”支持计划青年人才。长期致力于多模态学习、高光谱图像处理领域的理论与应用研究。以第一作者及通讯作者在 AAAI、IJCAI、TGRS、TNNLS、TIM、TCBB、TBD、TALLIP、PR 等国内外知名期刊会议上发表学术论文 70 余篇，曾任多个 SCI 期刊客座编委。主持国家级、省部级科研项目 12 项，其中国家自然科学基金 3 项、云南省重大科技专项课题 1 项、云南省基础研究重点项目 1 项、中国博士后科学基金 1 项。研究成果获云南省自然科学一等奖、中国电子学会科技进步三等奖各 1 项，授权国家发明专利 20 余项。

论坛讲者



肖亮

肖亮, 南京理工大学紫金卓越教授, 国家级领军人才, 江苏省 333 工程领军科技人才, 计算机学院 / 人工智能学院副院长, 高维信息智能感知与系统教育部重点实验室主任、视觉传感与智能计算江苏省重点实验室副主任, 江苏省人工智能学会智能遥感专委会主任, CCF 理事。主持和承担科技部重点研发计划课题、科技部重大仪器专项任务等国家级任务十余项, 研发低空多模态遥感大模型, 物理 AI 计算引擎, 智能光谱仪与解译系统等, 获得行业应用, 获省部级科技奖 7 项, 发表 IEEE 汇刊、CVPR 等期刊和会议论文 100 余篇, 授权发明专利 40 余项, 多项进行成果转化。

报告题目：物理模型驱动 AI 及其应用

报告摘要：近年来, 深度学习的成功发展呈现出两条相互交织的技术路径: 数据驱动与物理模型驱动。数据驱动方法以大规模数据标注与端到端学习为核心, 依托算力提升, 在计算机视觉、自然语言处理等领域取得了突破性进展; 而物理模型驱动方法则强调对问题结构的先验建模, 如稀疏性驱动网络、算法迭代展开网络、物理信息神经网络、及结构化概率模型, 在数据稀疏或可解释性要求高的场景中展现出独特优势。本报告首先回顾两类方法的发展历程、典型代表及其在计算成像、科学计算、智能遥感等任务中的应用成效; 进而分析当前“数据 + 模型”融合的新范式, 如隐式神经表示、神经符号学习及基于扩散模型的生成方法; 报告旨在为深度学习研究者提供一条兼顾性能与可解释性的方法学思考路径。



高连如

高连如, 中国科学院空天信息创新研究院研究员、博导, 研究方向为高光谱遥感图像处理与信息提取, 是 IET Fellow。主持国家自然科学基金、国家重点研发计划课题、国家高分专项项目等 10 余项; 已发表学术论文 300 余篇, 其中 SCI 收录 190 余篇, ESI 高被引论文 37 篇; 授权国家发明专利 30 项; 出版《高光谱图像信息提取》等学术著作 3 部。获得国家科技进步二等奖、中国科学院杰出科技成就奖、北京市自然科学一等奖、军队科技进步一等奖、湖北省科技进步一等奖各 1 项; 获得 IEEE GRSS 最高影响力论文奖和国际高光谱顶会 WHISPERS 的杰出论文奖各 1 项。现担任国际数字地球学会中国国家委员会成像光谱对地观测专业委员会主任委员, 是 IEEE TGRS 期刊的副主编以及 Chinese Geographical Science、《遥感学报》的编委, 担任 IGARSS 国际会议评奖委员会委员和 WHISPERS 国际会议技术委员会委员。

报告题目：高光谱遥感技术与应用

报告摘要：高光谱遥感图像具有图谱合一的特点和优势，在资源、环境、生态等领域具有巨大的应用价值，与此同时其独特的立方体式数据结构也给地物信息提取带来了挑战，如波段细分引发噪声问题、混合像元和同物异谱影响地物分类精度、背景干扰和高时效性信息提取是难点。围绕上述问题，过去十余年间以传统机器学习为基础在高光谱遥感图像处理与信息提取方面提出了许多方法，并且随着大数据和人工智能技术的飞速发展，面向高光谱遥感图像信息智能提取的深度学习模型近年来发展迅速。本报告将介绍高光谱遥感技术和应用的发展历程和现状，并在高光谱遥感图像信息提取技术领域已有研究工作基础上，重点讲述近年来该方向发展的新方法和新进展。



谢卫莹

谢卫莹，华山学者特聘教授、博士生导师，国家级青年人才，IEEE Senior Member，担任 IEEE TIP、PR、TGRS、NN 国际期刊编委。研究方向为遥感在轨实时处理及分布式边缘具身智能，发表多篇中国科学院一区 TOP 期刊及 CCF A 类会议论文，其中 ESI 高被引论文 13 篇，热点论文 2 篇。授权国家发明专利 20 余项并在高分对地观测和深空探测等领域完成转化应用。主持国家自然科学基金优秀青年基金项目、联合重点项目、深空探测学科先导项目子课题、国家型号任务、国家重大项目子任务等项目。主持省部级教改项目 2 项。获吴文俊人工智能自然科学一等奖、中国电子学会科技进步二等奖、石青云女科学家奖，入选全球前 2% 顶尖科学家榜单（连续三年）、全国优秀创新创业导师、中国科协青年人才托举工程。

报告题目：多模态基础模型及边缘具身智能研究进展

报告摘要：多模态基础模型在复杂场景精准感知和智能应用赋能等方面具有重要研究价值。本次报告将围绕多模态基础模型及具身智能面临的关键问题，介绍非结构化场景思维链决策数据集及具身智能控制数据集构建、基础模型设计和轻量化部署实现等关键技术，探索复杂场景下精准高效的边缘具身智能体系。

分论坛 以人为中心的视觉计算

时间：8月4日下午14:00-17:30 地点：101会议室

内 容

嘉 宾

主席：杨 勋（中国科学技术大学）、韩军功（清华大学）、陈超奇（深圳大学）

开放场景视觉空间感知

杨 欣（华中科技大学）

以人为中心视频行为识别与生成

涂志刚（武汉大学）

群体智能驱动的多模态认知计算

刘安安（天津大学）

微视觉计算

余梓彤（大湾区大学）

多模态情感推理大模型：从稳健感知到忠实决策

朱贝尔（中国科学技术大学）

论坛主席



杨 勋

杨勋，中国科学技术大学信息科学技术学院特任教授，博士生导师，国家高层次青年人才项目获得者；曾在新加坡国立大学计算机学院 NExT 研究中心担任博士后研究员；2021年回国加入中国科学技术大学信息学院。近年来主要关注跨媒体分析与推理以及人工智能模型可靠性等方向的研究。在 CCF A 类刊物以及 IEEE/ACM Transactions 国际期刊上发表一作 / 通信作者论文 50 余篇。主持国家自然科学基金联合基金重点项目、面上项目、长三角联合攻关基础研究项目。现担任 CCF 多媒体技术专委会执行委员、CSIG 青工委副秘书长、CSIG/CCF 高级会员；担任 IEEE Transactions on Big Data、IEEE Transactions on Fuzzy Systems 和 Multimedia Systems 期刊的编委，多次担任教育部、国家自然科学基金、北京市等人才 / 重点类项目的函评专家。



韩军功

韩军功，清华大学自动化系，兴华讲席教授（长聘），国家海外高层次人才，国家信息研究中心‘类脑智能技术’交叉创新群体负责人。2024 年 9 月回国，回国前任英国谢菲尔德大学（Top 100）计算机学院讲席教授。研究方向涵盖多模态感知、智能医疗、类脑机器学习与具身智能，推动 3 项人工智能技术实现产业化，音视频检索成果获 CSI 创新奖并助力企业斩获艾美奖。论文被引 4.2 万次，一作单篇最高被引超 2100 次，连续入选全球前 2% 科学家及 0.05% 高被引学者，CSRankings 人工智能方向英国学者（曾）排名第六（谢菲尔德大学第一）。现任多本 IEEE Transactions 副主编，主持 2 项国家自然科学基金重点项目，BMVC 2025 大会主席，国际模式识别学会（IAPR）会士。



陈超奇

陈超奇，深圳大学特聘教授、博士生导师，入选国家级青年人才、广东省青年拔尖人才、深圳大学“百人计划”。博士毕业于香港大学，卡内基梅隆大学访问学者，研究方向为可信机器学习、多模态大模型、空间智能等。在 IEEE TPAMI、NeurIPS、CVPR、ICCV、ICLR 等国际顶级期刊和会议发表论文 40 余篇，其中第一 / 通讯作者发表 CCF-A 类论文 20 篇，4 篇受邀国际会议口头报告，2 篇期刊论文入选 ESI 高被引论文，相关工作已在智慧医疗、智能交通、工业诊断等实际场景中部署应用。

论坛讲者



杨欣

杨欣，华中科技大学教授、博导、国家级青年人才；主持国家重点研发计划、国基金联合重点、国际合作等国家级项目，以一作或通讯作者发表 TPAMI、IJCV、CVPR 类期刊会议论文 100 余篇，多次入选斯坦福全球前 2% 顶尖科学家榜单；获湖北省技术发明一等奖、湖北省教学成果特等奖、中国图像图形学会石青云女科学家奖、中国医学科技奖青年科学技术奖、ISMAR 最佳论文提名奖等奖励；指导学生获得挑战杯全国特等奖、中国国际互联网 + 金奖、CSIG 挑战赛年度总冠军等国内外权威竞赛奖项 10 余次；担任 IEEE-TMI、IEEE-TVCG 等多个期刊编委、IEEE BISP 技术委员会委员、CSIG 青工委副秘书长、CCF 多媒体技术专委会副秘书长等职务。

报告题目：开放场景视觉空间感知

报告摘要：开放场景下的视觉空间感知在泛化性、实时性和鲁棒性方面仍面临严峻挑战。本报告将介绍视觉空间感知模型与方法的近期科研进展，重点涵盖单目、双目及多视角深度估计，以及视觉几何基础模型等关键技术方向，并探讨其在无人机等低功耗移动平台上的应用。



涂志刚

涂志刚, 武汉大学教授、博士生导师, 国家级青年人才(岗位: 新一代人工智能), 全球前 2% 顶尖科学家入选者。研究领域: 计算机视觉、视频理解, 聚焦“以人为中心”视频行为识别与生成。主持国家自然科学基金重点基金, 国家重点研发课题(国拨 > 300 万)、湖北省杰出青年基金、教育部联合基金(青年人才类)、腾讯犀牛鸟基金(技术创新奖)等项目。发表高水平论文 100 余篇, 第一 / 通讯作者中科院 1 区 Top SCI 期刊 +CCF A 类顶会论文近 60 篇。获 2022 年湖北省自然科学二等奖(排名 1) 等省部级科技奖励 3 项。指导学生获 2024 国家自然科学基金“青年学生项目”、2024 中国国际大学生创新大赛(互联网+)“全国金奖”。担任 CAAI 模式识别专委会 / 中国仿真学会 - 视觉计算与仿真专委会副秘书长、10 余个国际期刊 / 会议的编委、领域主席等职务。开发了视频人行为智能识别系统, 成功应用“第七届世界军人运动会开闭幕式”等领域, 被央视采访报道。

报告题目: 以人为中心视频行为识别与生成

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重要力量, 其中“以人为中心”的视频行为识别与生成作为视频理解的关键技术具有广泛应用, 一直是学术界与工业界的关注焦点。本报告, 首先将介绍武汉大学行为理解与视觉感知研究组(HUVPRLab)在“以人为中心”视频行为识别方面的系列工作, 包括低照度行为识别、细腻度行为识别等。其次, 将介绍视觉大模型驱动的视频行为生成方面的工作, 包括数字人运动生成、语音动作协同生成等。形成了以人体“视觉特征为载体、运动捕捉为牵引、多模态大模型为基石”的行为理解研究范式。最后, 报告将总结和展望“以人为中心”视频行为识别与生成的发展趋势。



刘安安

刘安安, 天津大学讲席教授, 人工智能学院副院长, 国家杰青。从事多模态信息感知、生成与决策, 社会模拟器构建研究, 服务于网络媒体安全审核与引导等。作为负责人承担国家自然科学基金联合重点、国家重点研发计划等; 以第一完成人获天津市科技进步特等奖; 发表高水平论文百余篇; 担任 CSIG 图像智能边缘计算专委会副主任等职务, 担任 4 个 IEEE/ACM 期刊编委, 连续 8 年担任 CCF-A 类国际会议 ACM Multimedia 领域主席。

报告题目: 群体智能驱动的多模态认知计算

报告摘要: 社会模拟器作为数字时代人工智能与社会科学交叉创新研究的方法论革命, 其本质是构建人工社会实验室, 对复杂社会系统进行“计算解构”与“涌现重构”, 实现从静态观察到动态干预的转变, 并通过多智能体的认知建模与多尺度社会网络的耦合演化, 在虚拟空间中创造具有涌现特性的社会“数字孪生”。本报告将汇报团队基于群体智能驱动的社会模拟器在群体认知领域的创新工作, 探索群体智能在媒体内容流行度预测、虚假信息认知、个性化信息传播等重大社会议题中验证其决策支持价值, 为数字社会治理提供了新的科学范式。



余梓彤

余梓彤, 大湾区大学长聘副教授, 国家高层次青年人才, CCF 东莞秘书长, CSIG 生物特征识别专委会副秘书长, GDSIG 情感计算专委会主任, 东莞市智能信息技术重点实验室副主任, 研究方向为微视觉计算。发表 Cell 子刊、TPAMI 等高水平文章 50 篇, 谷歌总引用 1 万余次, 授权国家发明专利 23 项。主持国自然优青(海外)/面上/青年项目、广东省联合重点项目、CCF-腾讯犀牛鸟项目等。获国家优秀留学生奖、IEEE 生物识别理事会新星奖(全球每年 1 人)、广东省图像图形学会科技进步一等奖(R1)/自然科学一等奖(R1)。

报告题目: 微视觉计算

报告摘要: 微视觉计算专注于捕捉、解析与增强人眼难以察觉的微弱视觉信号, 在医疗诊断、工业质检、情感计算及安全监测等领域展现出重要潜力。例如在情感计算领域, 理解人机交互场景下的生命体征变化、微表情和微姿态有助于欺骗检测。本报告将介绍以人为中心的微视觉计算基础模型和方法。



朱贝尔

朱贝尔, 中国科学技术大学信息科学技术学院教授、国家级青年人才、博士生导师。博士毕业于新加坡南洋理工大学; 本科和硕士毕业于清华大学, 曾在南洋理工大学从事博士后研究。主要研究方向包括多模态大模型推理、图像与视频可控生成、鲁棒机器学习等, 并长期围绕情感计算、心理健康、人机交互智能等交叉应用开展研究。迄今已在 NeurIPS、ICLR、ICML、CVPR、ICCV、AAAI、ACL、ACM MM、IEEE TPAMI/TIP 等国际顶级会议和重要期刊发表论文 30 余篇, 多篇论文入选 Oral、Spotlight、Highlight 展示, 并长期担任多个国际顶级会议与期刊审稿人。

报告题目: 多模态情感推理大模型: 从稳健感知到忠实决策

报告摘要: 本报告将围绕多模态情感推理大模型的关键挑战与最新研究进展展开, 重点介绍“稳健感知—忠实推理—协同决策”的研究链条。首先, 针对真实场景中常见的情感冲突问题, 即不同模态表达的情感信号不一致, 报告将介绍冲突感知多模态情感推理基准 CA-MER 以及 MoSEAR 方法, 通过模态专属专家与注意力重分配机制缓解模型对单一模态的偏置。其次, 针对现有情感大模型在推理过程中存在的证据利用不足和跨模态幻觉问题, 报告将介绍 OPPO 方法, 通过多模态证据覆盖奖励与模态忠实性约束, 提升模型推理链对视觉、音频和情感线索的充分利用与真实依赖。最后, 报告将讨论情感推理中的快慢思考协同问题, 介绍 MER-R1 如何结合快思考的直觉召回能力与慢思考的精确筛选能力, 使推理不仅提升可解释性, 也真正改善最终情感识别性能。通过以上研究, 本报告展示多模态情感推理大模型从“看得准、听得全”, 到“想得真、说得清”, 再到“判得稳、用得好”的发展路径, 为可信、可解释、可应用的情感智能系统提供新的研究思路。

分论坛 具身智能

时间：8月4日下午14:00-17:30 地点：302会议室

内 容

嘉 宾

主席：丛润民（山东大学）、许华哲（清华大学）、刘航欣（北京大学）

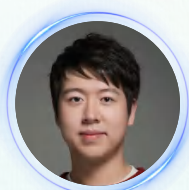
大模型时代的具身智能机器人感知与决策探索	李坚强（深圳大学）
脑机接口赋能脑疾病诊疗关键挑战与研究进展	李 阳（北京航空航天大学）
基于强化学习的腿足机器人运动控制研究与进展	张 伟（山东大学）
面向空间清洁的具身操作与导航实践	王晓波（深圳理工大学，深信服科技股份有限公司）
物理具身基础模型：表征、架构及真机评测	梁小丹（中山大学）
智能检测机器人：医工应用探索	刘行健（大连理工大学）

论坛主席



丛润民

丛润民，山东大学控制科学与工程学院教授、博士生导师，国家高层次青年人才、全球高被引科学家，主要研究方向包括模式识别与机器学习、计算机视觉、具身智能等。主持了包括国家自然科学基金（面上、青年）、科技创新 2030 子课题在内的 10 余项科研项目。在 TPAMI、IJCV、TIP、NeurIPS、CVPR、ICCV、ICML 等 CCF-A、IEEE/ACM Trans 上发表论文 140 余篇，ESI 热点论文 2 篇、ESI 高被引论文 20 篇，谷歌引用 19000 余次；授权国家发明专利 50 余项。担任 IEEE TNNLS 等 4 个 SCI 期刊编委，获省部级 / 学会科学技术奖 4 项、教学成果奖 3 项，IEEE SPS Best Paper Award、IEEE Chester W.Sall Memorial Award 等论文奖励 10 项、CVPR/ECCV 等国际竞赛冠军 3 项、亚军 2 项。



许华哲

许华哲，破壳机器人创始人，清华大学交叉信息研究院助理教授、博士生导师。博士后就读于斯坦福大学，博士毕业于加州大学伯克利分校，本科毕业于清华大学电子工程系。长期从事通用具身智能研究，致力于让机器人在真实家庭环境中学习和工作，推动通用机器人走进千家万户。作为顶级智能机器人会议 CORL'23 最佳系统论文得主，在 IJRR、RSS、NeurIPS 等顶级期刊和会议发表论文一百余篇，曾担任 IJCAI、ICRA、ICLR、CORL 等会议的领域主席或副主编。



刘航欣

刘航欣，本科毕业于美国弗吉尼亚理工大学，后于加州大学洛杉矶分校获得机械工程硕士和计算机科学博士学位。主要研究方向为智能机器人感知、认知与学习方法，机器人本体设计与人机交互。在 Science Robotics 和 Nature Machine Intelligence 发表论文 3 篇，并在 TPAMI/TRO/IJCV/RAL/ICRA/IROS 等国际期刊和会议上发表论文 60 余篇，曾获 2019 年 ACM 中国图灵大会最佳论文奖和 2023 年 IROS 移动操作方向最佳论文提名。

论坛讲者



李坚强

李坚强，深圳大学讲席教授，IEEE/IET Fellow，国家杰出青年基金获得者，国家重点研发计划项目首席科学家，入选国家级青年人才，广东省杰出青年基金获得者，深圳市高层次人才，深圳大学人工智能学院院长，大数据系统计算国家工程实验室执行主任，广东省大数据分析工程中心主任，广东省人工智能教学团队负责人。主持国家杰出青年基金项目，国家级青年人才项目，国家自然科学基金重大仪器、重点项目，广东省杰出青年基金等省部级项目。荣获吴文俊人工智能科技进步一等奖（第一完成人），中国自动化学会科技进步一等奖（第一完成人），广东省科技进步二等奖（第一完成人），深圳市自然科学科学二等奖（第一完成人），连续 5 年入选斯坦福大学发布的全球 Top 2% 顶尖科学家榜单。

报告题目：大模型时代的具身智能机器人感知与决策探索

报告摘要：基于数据的深度学习、大模型可有效提高智能系统感知决策能力，而机器人与智能系统网络协同、多智能体协同、云端协同来有效提高机器人的感知与执行效率，报告介绍了围绕机器人与智能系统网络协同感知、协同网络构建以及优化分析决策展开的研究，相关研究成果应用于特种监测机器人以及医疗健康智能监测系统，并介绍了结合大模型的机器人感知与决策探索。



李 阳

李阳, 北京航空航天大学生物与医学工程学院院长、医学科学与工程学院院长、教授、博士生导师, 国家级领军人才。2011 年毕业于英国谢菲尔德大学自动控制与系统工程系, 获哲学博士学位。2012 年 12 月入选北京航空航天大学“卓越百人”海外人才计划。2013 年 2 月起在北京航空航天大学模式识别与智能系统实验室工作。主要从事医学影像数据分析与处理、脑机接口与神经康复工程等方面的相关研究工作。主要研究手段包括脑电图 (EEG), 结构及功能性磁共振成像 (DTI/fMRI), 以及深度学习等计算分析方法。研究成果以第一 / 通讯作者在国际学术刊物上发表 SCI 论文 60 多篇, 包括 IEEE Transactions on Medical Imaging、Medical Image Analysis、Annals of Neurology 及自然语言处理领域 CCF-A 会议 ACL 等。近五年主持国家级、省部级等科研项目 10 余项, 其中包括国家自然科学基金重点、国家重点研发计划专项项目、北京市自然科学基金专题重点项目等。曾获英国谢菲尔德大学“哈里沃辛顿”学术奖、英国“优秀自费留学生奖”, 研究成果获吴文俊人工智能自然奖、中国体视学学会自然科学一等奖及“青年科技奖”等。担任国际期刊 international Journal of Biomedical Engineering and Science 编委等。

报告题目: 脑机接口赋能脑疾病诊疗关键挑战与研究进展

报告摘要: 本报告介绍与临床深度合作, 介绍当前脑机接口与神经调控研究进展及挑战, 采用多模态个体化脑网络与小样本深度学习融合的研究方法, 结合临床知识驱动, 探讨基于电生理数据、多模态结构 / 功能影像数据驱动模型假设, 如何用于揭示脑疾病的脑活动机制, 解决神经慢性脑疾病发病机理不明且误诊率高、临床难以有效检测大脑异常活动机制等问题。揭示了临床神经慢性脑疾病的潜在脑功能异常活动机制, 为探索神经慢性脑疾病的个体化脑功能网络异常连接机理及神经调控治疗提供了理论基础和技术支持, 为神经系统个体化神经调控与康复治疗干预提供了量化评估分析工具。



张 伟

张伟, 山东大学教授、博士生导师, 主要从事机器学习研究。入选国家级领军人才计划, 近年来承担国家自然科学基金重大 / 重点项目、国家重点研发计划、山东省重大专项等。代表论文以第一 / 通讯作者发表在 TPAMI、TRO、IJCV、TIP、NeurIPS、CVPR、ICCV、IJCAI、AAAI、CoRL、ICRA 等人工智能与机器人领域权威期刊和会议, 获中国、美国、日本等发明专利授权 80 余项, 获省部及国际学术奖励 10 余项。

报告题目：基于强化学习的腿足机器人运动控制研究与进展

报告摘要：腿足机器人运动控制是当前机器人领域的核心研究方向之一。相较于轮式、履带式等其他机器人平台，腿足机器人运动控制面临复杂度高、稳定性差等痛点，对控制算法提出了极高要求。传统控制方法依赖于精确的模型与环境先验，在面对未知扰动与复杂地形时适应性有限。近年来，强化学习技术发展迅猛，为腿足机器人运动控制提供了全新的范式，使机器人能够通过与环境的持续交互自主学习控制策略，从而逐步掌握动态平衡、步态生成与复杂地形穿越等高级技能。本报告将重点介绍课题组在基于强化学习的腿足机器人运动控制方面开展的前沿探索和相关研究进展。



王晓波

王晓波，深圳理工大学助理教授 / 特聘副教授，博导，深信服具身智能首席科学家。研究方向为具身智能机器人、计算机视觉，长期从事具身操作与导航的研究实践，致力于推动机器人在真实场景（如空间清洁、家庭服务）的落地。已发表相关学术论文 40 余篇，第一作者获 ICME Best Student Paper Runner-Up 奖，担任 Pattern Recognition 期刊副主编。

报告题目：面向空间清洁的具身操作与导航实践

报告摘要：传统的“2D 地面”清洁机器人已实现规模化应用，但真实的物理世界本质上是三维、高度动态且非结构化的。突破二维平面的维度限制，研发具备复杂环境认知、人机共融导航、自主决策与精细交互能力的“3D 空间”清洁机器人，是具身智能迈向通用化的必然趋势。本报告将重点介绍课题组在空间清洁机器人领域的前沿探索与研究实践，探讨具身智能在真实物理场景中的落地路径。



梁小丹

梁小丹，中山大学教授，博导，获得国家级万人计划青年拔尖人才，教育部 U40 人才项目，科技部人工智能重大专项项目负责人等支持，Google Scholar 引用超过 40,000 余次。获 IEEE CVPR 2024 Best Paper Finalist，广东省科技进步一等奖，阿里巴巴达摩院青橙奖（全国计算机领域 35 岁以下十名）、CSIG 石青云青年女科学家奖、世界人工智能大学优秀论文奖，吴文俊人工智能优秀青年奖、中国图象图形学会科技技术一等奖、连续 5 年入选爱思唯尔中国高被引学者榜、ACM MM 中国新星奖、福布斯中国 30 Under 30 科学榜、斯坦福大学“全球前 2% 顶尖科学家榜单”，终生科学影响力榜、国际 ScholarGPS2024 年全球前 0.05% 顶尖科学家榜单等。并于 2025 年荣获深圳光明区五一劳动奖章、劳模称号、工匠人才创新工作室优秀工作室奖，多次获得华为火花奖。十余次担任 CCF-A 类顶级学术会议 ICCV、CVPR、NeurIPS 等领域主席，并为 CVPR 监察委员会。科研成果转化为 70 余项授权发明专利，并落地于华为、字节、腾讯、联想、DeepSeek 等国内科技领军企业的产品。

报告题目：物理具身基础模型：表征、架构及真机评测

报告摘要：本报告围绕具身智能尚未迎来“GPT 时刻”的核心问题，系统梳理视觉 - 语言 - 动作 (VLA) 基础模型的技术趋势、核心挑战与未来方向，明确具身智能与大模型、机器人范式的路径差异。当前具身智能形成 VLA、世界模型、智能体三大主流范式，正逐步走向融合，分别支撑动作对齐、环境预测、高层决策能力。其核心瓶颈在于互联网语义知识与物理世界执行的深层鸿沟，具体表现为物理先验缺失、跨模态对齐精度不足、数据采集成本高、实时性与安全性难以平衡、评测体系不统一等问题。重点介绍 A1、AtomicVLA、eWAM 等创新模型，通过自适应早退、原子动作分解、VLA 与世界模型原生融合等技术，显著提升模型效率、泛化性与持续学习能力。同时提出 PhyAgentOS 标准化框架与 ManipArena 真机评测基准，解决硬件差异、评测不可复现等行业痛点。



刘行健

刘行健，大连理工大学科技创新研究院院长，人工智能学院（机器人与智能系统研究院）、高性能精密制造全国重点实验室教授，博士生导师，国家级青年人才。研究方向为精密机器人测量与作业。近五年，主持国家自然科学基金、国家重点研发计划 - 青年科学家项目、国家重大专项课题等科研任务；发表高水平论文 50 余篇，授权中国发明专利 10 余项，研究成果在国内外企业医院落地应用；获教育部青年科学奖、中国科协青年托举、大连市优青等荣誉；担任 IEEE/ASME TMECH、IEEE RAL、控制理论与应用等期刊编委，IEEE IROS 2025 政府论坛主席、IEEE CASE 2023/2024/2025 副主编等学术职务。

报告题目：智能检测机器人：医工应用探索

报告摘要：智能检测机器人是体现国家高端制造和智慧医疗水平的重要标志，对推动制造强国、健康中国等战略具有重要意义。目前检测机器人技术性能仍然有限，如：难以恢复弱纹理强反光劣化目标表面信息、难以深入受限深腔内部实现精准重建等，制约了其在非侵入式医学监测和工业检测等领域的应用。本报告围绕劣化目标信息恢复与受限内部空间重建的科学问题，重点汇报开胸术中心脏高质量视觉监测、非朗伯工业零件高精度测量和受限深腔结构内窥式扫描仪器三个方面的智能检测机器人医工应用探索，以及其在医疗临床和生产制造等领域的应用情况，最后报告未来研究计划及展望。

分论坛

视觉基础模型

时间：8月4日下午14:00-17:30 地点：309会议室

目 录

嘉 宾

主席：高广谓（南京理工大学）、齐国君（西湖大学）
胡 玮（北京大学）、舒祥波（南京理工大学）

InternVideo：通用视频理解大模型体系	王利民（南京大学）
特征解耦驱动的跨模态视觉生成	鲍秉坤（合肥工业大学）
开放环境下视觉表面缺陷检测与生成	李泽超（南京理工大学）
Recent Advances in Medical Foundation Models	雷 震（中国科学院自动化研究所）
手语视频理解与生成	周文罡（中国科学技术大学）

论坛主席



高广谓

高广谓，南京理工大学计算机科学与工程学院（人工智能学院，软件学院）教授，博导。研究方向为端侧智能视觉感知与理解。主持和参与国家自然科学基金、江苏省自然科学基金、科技创新 2030-“新一代人工智能”重大专项等多项课题。在国际权威期刊 IEEE TPAMI/TIP/TMM、IJCV 以及权威国际会议 CVPR、ACMMM 等发表论文 70 余篇，ESI 高被引论文 6 篇。江苏省“六大人才高峰”高层次人才，获江苏省科学技术奖一等奖(7/11)1 项，指导研究生获 2024 年江苏省优秀专业学位硕士学位论文奖 1 项，入选斯坦福大学发布的 2025 年度全球前 2% 顶尖科学家年度榜单。IEEE/CCF/CSIG/CAAI/CAA 高级会员，国际期刊 Pattern Recognition(CCF-B) 和 IEEE Transactions on Image Processing(CCF-A) 编委，中国图像图形学会青工委委员，Valse 执行 AC。



齐国君

齐国君, 西湖大学机器感知与学习 (MAPLE) 实验室负责人, 国际电气和电子工程师协会会士 (IEEE Fellow), 国际模式识别联合会会士 (IAPR Fellow), 国际计算机协会杰出科学家 (ACM Distinguished Scientist), 中国科学技术大学郭沫若奖得主。长期开展人工智能, 特别是深度学习及其在生成式人工智能 (AIGC)、多模态理解与生成式大模型、2D/3D 智慧创作、虚拟现实等方面的研究。现任 IEEE Fellow 评选委员会 (IEEE Fellow Committee) 评委, IEEE T-IP/T-MM/T-CSVT、ACM T-KDD、Pattern Recognition 等期刊的编委, 并担任 IEEE 信息处理学会会议理事会的理事。



胡 玮

胡玮, 北京大学王选计算机研究所所长聘副教授、研究员, 国家级青年人才, 北京大学博雅青年学者。研究方向为三维视觉, 包括三维生成和理解。在国际权威期刊 IEEE TPAMI/TIP/T-MM、IJCV 以及国际知名会议 CVPR、ICCV、ICLR、NeurIPS 等发表论文 100 余篇, 获得计算机视觉国际顶级会议 CVPR 2021 年度最佳论文提名、2020-2021 年 IEEE Multimedia Rising Star Award (Honorable Mention Award)、多媒体领域国际旗舰会议 ICME 2020 年度 Best Student Paper Runner Up Award 和 2021 年北京市科学技术进步二等奖。担任国际顶级期刊 IEEE Signal Processing Magazine 编委、国际旗舰会议 ICME 2026 程序委员会主席 (TPC Co-Chair)。



舒祥波

舒祥波, 南京理工大学计算机学院 / 人工智能学院 / 软件学院副院长、教授, 社会安全信息感知与系统工信部重点实验室副主任、国家优青获得者、江苏省杰青获得者。研究方向为视觉具行为计算, 在 TPAMI、CVPR、ICCV、ICLR 等期刊 / 会议上发表论文 100 余篇, 其中 ESI 高被引论文 8 篇; 获江苏省自然科学一等奖、中国电子学会自然科学一等奖、ACM MM 2015 最佳论文提名、MMM 2016 最佳学生论文、CCBR 2026 最佳论文、江苏省优博、中国人工智能学会优博、江苏省行业领域优秀科技进展、全球前 2% 顶尖科学家 (2021-2025 年); 承担国家自然科学基金联合重点项目、国家重点研发课题、国家重大科研仪器项目课题、国家自然科学基金面上项目、江苏省杰出青年基金等项目等。担任 CSIG 青工委副秘书长, 以及 TNNLS、TCSVT、TIFS 等期刊编委。

论坛讲者



王利民

王利民, 南京大学计算机学院教授, 博士生导师, 国家级青年人才计划入选者。研究领域为计算机视觉和多模态大模型, 在 IJCV、T-PAMI、CVPR、ICCV、NeurIPS 等期刊和会议发表论文 100 余篇, 谷歌学术引用 4.9 万余次, 两篇一作论文引用超过 4000 次。主持科技创新 2030-" 新一代人工智能 " 重大项目、江苏省自然科学基金攀登项目等。带领团队研发了首个通用视频理解大模型体系 InternVideo, 全球用户下载量超过 600 万, 被 Google、Meta、NVIDIA 等知名企业使用, 产生了重要国际影响力。曾获得广东省技术发明一等奖, 蚂蚁 InTech 科技奖, ACM MM 2023 唯一最佳论文提名奖、首届世界人工智能大会青年优秀论文奖。担任 CVPR/ICCV/NeurIPS 等会议的领域主席和 TPAMI/IJCV 等期刊的编委。

报告题目: InternVideo: 通用视频理解大模型体系

报告摘要: 视频理解已经成为人工智能领域现阶段核心难题, 面临着数据维度高、信息容量大、场景变化多等诸多技术挑战, 如何构建通用视频理解大模型已经成为一项极具挑战的任务。本报告将介绍通用书生视频大模型体系 InternVideo 的发展历程: 2022 年发布首个通用视频理解基础模型 InternVideo1.0; 2024 年发布多模态视频理解大模型 InternVideo2.0; 2025 年发布全新升级版本 InternVideo2.5。除此之外, 还将介绍 InternVideo 视频理解大模型最新研究进展: 1. 多任务视频理解模型强化学习框架 VideoChat-R1&TempR1; 2. 长视频压缩架构 VideoChat-Flash; 3. 长视频多轮感知与推理框架 Video-o3; 4. 流式视频处理与分析架构 VideoChat-Online&StreamForest。最后将展望视频理解大模型技术发展趋势。



鲍秉坤

鲍秉坤，合肥工业大学计算机与信息学院院长，教授，博士生导师，国家杰出青年科学基金获得者。研究方向为多媒体计算、社交多媒体、计算机视觉、人工智能等。主持国家科技创新 2030-“新一代人工智能”重大项目、国家自然科学基金重点项目、江苏省重点研发计划等。荣获第十九届江苏省青年科技奖、江苏省青年女科学家奖、2018 年度电子学会科学技术（自然科学类）一等奖。获 ACM TOMM 年度最佳论文奖。担任 IEEE TMM、TCSVT、ACM TOMM、MMSJ 等期刊编委。

报告题目：特征解耦驱动的跨模态视觉生成

报告摘要：当前，跨模态视觉生成方法普遍面临特征耦合的共性挑战，导致生成过程可控性弱、内容精准度低。特征解耦将原始数据中混杂的特征分解为相关性低且具有明确语义的潜在特征，为实现精准可控的视觉生成提供了有效途径。本报告重点介绍针对跨模态生成过程构建的“模态对齐解耦 - 语义增补解耦 - 时空表征解耦”方法体系。首先，将跨模态输入中的耦合语义拆解至不同粒度的语义子空间，实现对象、属性、关系等多层次独立解耦生成；其次，在潜在表征空间中引入外部知识，增强表征解耦能力；最后，通过时 - 空表征解耦，实现运动过程与空间结构的独立控制与连贯生成。该体系为细节可控、高效灵活的跨模态视觉生成提供了系统化支撑。



李泽超

李泽超，南京理工大学计算机科学与工程学院 / 人工智能学院 / 软件学院教授、院长，2008 年本科毕业于中国科学技术大学，2013 年博士毕业于中国科学院自动化研究所。研究兴趣主要是多模态智能分析、计算机视觉等。主持国家杰出青年科学基金、新一代人工智能国家科技重大专项课题、国家自然科学基金联合基金重点项目等；入选国家“万人计划”青年拔尖人才，在 IEEE TPAMI、IJCV 和 CCF A 类会议发表论文 60 余篇；第一完成人获得 2024 年度江苏省科技进步一等奖，第二完成人获得 2022 年度中国电子学会自然科学一等奖和 2020 年度江苏省科学技术一等奖，第三完成人获得 2017 年度江苏省科学技术一等奖等；获得 ACM MM Asia 2020 和 2024 两年度最佳论文奖等；担任 IEEE TPAMI、TCSVT、TMM 等期刊编委，曾担任 IEEE TNNLS、Information Science 等期刊编委。

报告题目：开放环境下视觉表面缺陷检测与生成

报告摘要：图像视频大数据智能分析与识别在多种实际应用中具有至关重要的作用，比如工业生产、瑕疵检测、无人驾驶等。为此，我们研究了开放环境下视觉异常检测问题，主要介绍面向视觉异常检测的特征学习、语义分割以及异常检测等，提出了基于块对角结构的特征学习方法与特征选择方法，协同分析通道上下文和空间上下文的语义分割模型，基于特征蒸馏学习、分区记忆、自监督分割引导和非对称蒸馏分割的异常检测方法，进而介绍可控文生图的增强 - 控制 - 蒸馏模型、反演适配器、风格迁移模型等，最后介绍一些相关的实际应用情况。



雷震

雷震, 男, 博士, IEEE Fellow, IAPR Fellow, AAIA Fellow。中国科学院自动化研究所多模态人工智能系统全国重点实验室研究员, 中国科学院大学人工智能学院岗位教授, 中国科学院香港创新院人工智能与机器人研究中心教授, 博士生导师, IEEE 生物识别委员会 (IEEE Biometric Council) 杰出讲者。主要研究方向为视频分析与理解, 多模态大模型, 生物特征识别, 发表论文 200 余篇, Google Scholar 文章引用次数 40000 余次, H-index:92, 爱思唯尔中国高被引学者 (2020-2024), 入选全球前 2% 顶尖科学家榜单。

报告题目: Recent Advances in Medical Foundation Models

报告摘要: In this talk, I will introduce two recent advances in medical foundation models in CAIR. One is the surgical LLM for surgical video understanding. By integrating the spatial focus and temporal awareness, a versatile large multimodal model is trained, which supports multiple tasks simultaneously. In the second part, I will introduce the EchoCare ultrasound large foundation model which are trained from more than 4,500,000 images. It achieves the state-of-the-art performance on 10+ medical downstream tasks with an average improvement of 3% to 5% over existing methods.



周文罡

周文罡, 中国科学技术大学教授, 人工智能学院副院长, 国家“优秀青年基金”获得者。主要研究方向为计算机视觉、具身人工智能和机器博弈。发表 CCF A 类论文 100 余篇, Google 学术总引用 2.8 万次。担任国际期刊 IEEE TMM 编委。主持科技部重大项目课题、国家基金委联合基金重点项目等。曾获吴文俊人工智能科技进步一等奖、中科院优秀导师奖等, 连续 6 年 (2020-2025) 入选斯坦福全球前 2% 科学家, 入选中科院青促会优秀会员、中国科协青年人才托举工程。培养博士生获中科院优秀博士学位论文奖 (2 人次)、中国图像图形学会优秀博士学位论文奖。

报告题目: 手语视频理解与生成

报告摘要: 手语作为一种视觉语言, 具备词汇、语法及语用层面的丰富结构, 是聋人群体日常沟通与信息获取的核心工具。手语视频理解与生成的目标是建立手语视频与其对应的语义文本之间的双向映射关系, 从而有效打通听人与聋人之间的沟通壁垒, 推动信息无障碍环境的构建。本报告将系统分析当前手语视频理解与生成所面临的主要技术挑战, 并重点介绍我们近期在手语视频翻译 (如从手语视频到口语文本的端到端识别与翻译模型) 和手语视频合成 (如基于文本或语义标签生成逼真、连贯手语动作的生成模型) 两个方向上的研究探索。此外, 结合真实应用场景, 展示我们在智能手语系统方面的落地尝试。

分论坛 扩散模型

时间：8月4日下午14:00-17:30 地点：401会议室

目 录

嘉 宾

主席：李崇轩（中国人民大学）、张 力（复旦大学&上海创智学院）
李锦兴（哈尔滨工业大学（深圳））、赵 敏（南京大学）

可控视觉生成及优化视角探析	张 健（北京大学）
基于离散扩散模型的理解与生成	傅朝友（南京大学）
面向边端高效推理的视觉内容生成	王楠楠（西安电子科技大学）
面向图像生成基座模型的高效生成研究	王亚星（吉林大学）
生成式遥感大模型研究进展	邹征夏（北京航空航天大学）
高效可控视频生成	武 宇（武汉大学）

论坛主席



李崇轩

李崇轩，中国人民大学高瓴人工智能学院副教授、博士生导师。致力于生成模型基础理论、建模范式、大规模训练策略和高效采样算法的研究，带领团队研制了扩散语言模型 LLaDA，开源模型下载量破 600 万次，多项成果部署于 DALL·E 2、Stable Diffusion、Vidu 等行业领先大模型。成果发表于 NeurIPS、ICML、ICLR、TPAMI、Nature Communications 等，谷歌学术引用 1.4 万余次。获机器学习领域顶级国际会议 ICLR 2022 杰出论文奖；北京市自然科学一等奖（排名 2）；吴文俊人工智能自然科学一等奖（排名 5）。主持国家自然科学基金青年基金 B 类、重大研究计划培育项目等，入选智源学者、吴文俊优秀青年奖、北京市科技新星。担任 IEEE TPAMI 编委和 ICLR/ICML/NeurIPS 等国际会议领域主席；作为主编出版《大模型十讲》教材，指导博士生入选国家自然科学基金青年学生基础研究项目、字节跳动奖学金计划等。



张 力

张力, 复旦大学大数据学院教授, 上海创智学院全时导师, 获国家级青年人才计划。博士毕业于伦敦玛丽女王大学电子工程与计算机科学系, 曾任职于牛津大学工程科学系博士后, 剑桥三星人工智能中心研究科学家。获上海海外高层次人才计划、上海科技青年 35 人引领计划 (35U35)、世界人工智能大会青年优秀论文奖; 发表人工智能期刊与会议论文 100 余篇, 论文总被引 2.8 万次。担任人工智能国际会议 NeurIPS、CVPR、ECCV 领域主席。



李锦兴

李锦兴, 现为哈尔滨工业大学(深圳)计算机科学与技术学院教授、博士生导师, 荣获深圳海外高层次人才“孔雀计划”、广东省博士后人才引进计划。已发表 70 余篇学术论文, 其中中科院一区期刊和 CCF A 类会议共 60 余篇, 包括 TIP、TNNLS、TCYB、TCSVT、TSMC、CVPR、NeurIPS、ACM MM、AAAI 等。主持多项科研项目, 包括国自然面上项目、国自然青年项目、深圳市科技重大专项(联合单位负责人)、深圳市优秀科技创新人才项目、深圳市高端人才科研启动经费; 重视产学研结合, 与河钢集团、航天科工等机构深入合作, 实现钢铁缺陷自动检测 / 裁剪等, 获河北省电子学会科学进步一等奖等。



赵 敏

赵敏, 南京大学人工智能学院准聘助理教授、博士生导师, 清华大学“水木学者”。研究方向为视频生成与视频世界模型, 是国产视频生成大模型 Vidu 的核心作者之一。以第一作者或共同第一作者身份在 NeurIPS、ICML、ICLR 等国际顶级会议发表论文十余篇, 累计引用 1000 余次。代表性工作包括视频长度外推方法 RIFLEx、实时交互视频生成方法 Causal Forcing 系列, 以及全链路开源视频世界模型框架 minWM。相关成果已被 HunyuanVideo、CogVideoX 等主流开源视频生成框架集成应用。目前主要开展视频生成、实时交互视频世界模型、长时程视频生成等方向研究。

论坛讲者



张 健

张健博士, 北京大学长聘副教授、博士生导师, 信息工程学院院长助理, 视觉信息智能学习实验室(VILLA)负责人。长期专注于“视觉重建生成与安全”研究, 已在 TPAMI、IJCV、TIP、CVPR、NeurIPS、ICLR 等高水平国际期刊和会议上发表论文 140 余篇, 谷歌学术引用逾 1.9 万次, h-index 68, 其中单篇一作最高引用超 1800 次; 相关开源项目 Github 星标量破 10k; 个人连续入选斯坦福全球前 2% 顶尖科学家“终身科学影响力”以及“年度科学影响力”榜单; 6 次获国际期刊 / 会议最佳论文奖、1 次 CVPR NTIRE 全球挑战赛冠军、1 次全球 AI 攻防挑战赛季军; 荣获 2023 年度华为 MindSpore 学术奖励基金项目优秀奖、2024 年度 OPPO 产学研“优秀合作伙伴”; 2025 年华为难题揭榜“火花奖”等; 主持国家科技重大专项课题、国自然、深圳市重点等科研项目 10 余项; 相关技术成功应用于创维、字节跳动、OPPO、兔展、Stability AI 等国内外知名企业的核心产品; 11 项授权专利通过增资入股方式完成科技成果转化; 担任深圳市人工智能学会青工委主任、JVCIR 等国际期刊编委等。个人主页: <https://jianzhang.tech/cn/>。

报告题目: 可控视觉生成及优化视角探析

报告摘要: 本报告从最优化理论视角出发, 系统剖析传统信号重建与前沿可控视觉生成之间的内在联系与本质。报告人基于近年来在视觉重建与生成领域的工作梳理, 包括结构组稀疏表示(GSR,TIP14)、迭代收缩阈值启发式网络(ISTA-Net,CVPR18)、去噪扩散零空间模型(DDNM,ICLR23)、免训练扩散模型操控(FreeDoM,ICCV23)、空间精准可控生成(T2I-Adapter,AAAI24)、拖拽式图像编辑(DragonDiffusion,ICLR24)、运动内容解耦式视频编辑(ReVideo,NeurIPS24)、拖拽式图生视频(OmniDrag,IJCV25)以及生成式视频编辑(GenCompositor,ICLR26)等为范例, 阐释如何通过最优化理论将这些研究有机串联和巧妙统一, 进一步揭示该理论在推动可控视觉生成发展中的指导价值。



傅朝友

傅朝友, 南京大学智能科学与技术学院研究员、助理教授、博导, 入选中国科协“青年人才托举工程”。2022 年博士毕业于中科院自动化所模式识别实验室。研究方向为多模态内容分析, 谷歌学术引用 1 万余次, 2 篇一作单篇引用过千次, 开源项目累计获得 2 万余次 GitHub Stars。代表性工作包括 VITA 多模态大模型系列 (VITA-1.0/-1.5、VITA-Audio), MME 多模态评测基准系列 (MME、Video-MME/-v2) 和 Awesome-MLLM 等。担任 Pattern Recognition/IEEE T-BIOM 期刊编委、ICLR/ICML 会议领域主席、CSIG 青工委委员、CCF-AI/-CV 专委会执行委员。曾获小米青年学者 - 科技创新奖、华为紫金学者、世界人工智能大会云帆奖、中科院院长特别奖、IEEE Biometrics Council Best Doctoral Dissertation Award、北京市优博、中科院优博。

报告题目：基于离散扩散模型的理解与生成

报告摘要：长期以来, 自回归架构主导着大语言模型的发展, 而扩散模型的发展为大语言模型的构建提供了新的思路。本报告聚焦扩散模型在理解与生成领域的前沿进展。报告首先介绍以扩散模型为骨架的大语言模型相关技术, 探讨其与自回归架构的性能对比, 分析其在高效生成方面的独特优势。随后, 报告扩展至多模态领域, 分享基于扩散模型实现多模态统一理解与生成的最新方案。最后, 本报告将展望上述技术的演进趋势, 探讨扩散模型在未来大模型发展中的方向。



王楠楠

王楠楠, 教授, 博士生导师, 西安电子科技大学空天地一体化综合业务网全国重点实验室副主任。在 IEEE TPAMI、IJCV、CVPR、ICCV 等期刊和会议上发表论文 200 余篇, 授权专利 40 余项, 实现专利转化 7 项, 牵头制定行业标准 2 项、参与制定国际国家标准 4 项, 获中国图像图形学会自然科学奖一等奖 (R1)、教育部自然科学奖一等奖 (R2)、陕西省科学技术奖一等奖 (R5)、陕西青年科技奖、中国人工智能学会优秀博士学位论文、陕西省优秀博士学位论文等奖项; 主持国家自然科学基金优秀青年基金项目、联合基金重点项目等项目。相关研究成果已应用于航天、公共安全、华为、腾讯、阿里等。目前是陕西省重点领域科技创新团队负责人、三秦特支计划创新团队负责人, 担任国际期刊 Visual Computer 的 Co-Editor-in-Chief 以及 Neural Networks 编委等, 同时担任中国电子学会青年科学家俱乐部会员, VALSE 常务领域主席, 中国图像图形学会制度建设工作委员会秘书长、学术工委委员、青年工作委员会委员, 中国计算机学会计算机视觉专委会委员、人工智能专委会委员、多媒体专委会委员, 中国人工智能学会模式识别专委会委员, 中国图像图形学会多媒体专委会委员、遥感图像专委会委员等。

报告题目：面向边端高效推理的视觉内容生成

报告摘要：针对视觉内容生成模型高计算复杂度引起的高能耗和高碳排放问题，本报告从架构设计、推理步数和参数量化三个方面，探索图像生成模型中神经网络规模压缩和稳定训练方法，实现绿色低碳的大模型轻量化技术。具体包括：(1) 优化网络架构，以降低模型参数规模或降低参数激活量；(2) 生成模型推理阶段的单步算法设计，以降低模型推理时延；(3) 模型的低比特量化，以实现模型的进一步缩小以及稳定训练。本报告的三项研究内容相辅相成，有望构建面向边端需求的大模型轻量化完备解决方案，实现资源消耗与模型性能之间的极佳平衡。



王亚星

王亚星，吉林大学人工智能学院教授，博士生导师，入选海外高层次人才引进计划青年项目。西班牙巴塞罗那自治大学博士。研究方向为视觉生成模型、迁移学习、无监督学习等。在 TPAMI, ICLR 等期刊会议发表论文 50 余篇。担任 ICLR/NeurIPS/ICML 等领域主席，在国际顶级期刊和会议 TPAMI、ICLR 等多次担任期刊和会议审稿人，ECCV/ICCV 组织多场 workshop。多模态语言翻译国际竞赛中荣获第一名、2022 年粤港澳大湾区(黄埔)国际算法算例大赛(遥感目标检测赛道)亚军(2/116 队伍)。主持 / 参与国家自然科学基金青年项目、三星人脸特定人脸合成、宁波市文旅个性化大模型、重庆长安 AI 智能体等项目。

报告题目：面向图像生成基座模型的高效生成研究

报告摘要：随着生成式人工智能在图像与视频领域的爆发式增长，提升模型推理效率已成为加速技术落地与工业应用的关键。本文针对当前主流生成模型存在的效率瓶颈，提出了三项创新性的改进方案：其一，针对扩散模型串行迭代导致的推理慢问题，研发了关键表征动态缓存与复用技术；其二，通过构建时间步无关的编码器架构，有效解决了模型蒸馏在追求“一步生成”时难以兼顾质量与多样性的挑战；其三，针对视觉自回归模型的表征冗余，提出特征自适应降维及单分支推理技术，为高效生成式 AI 的发展与应用提供了系统性的技术支撑。



邹征夏

邹征夏, 北京航空航天大学教授 / 博导, 国家级青年人才, 研究方向为遥感图像处理。在 Proceedings of the IEEE (最受欢迎论文)、Nature 子刊(封面论文)、TPAMI(国家自然科学基金委报道)、TIP、TGRS、CVPR、ICCV 等重要期刊和会议发表论文 100 余篇, 谷歌学术论文引用 1.4 万次, 担任遥感权威期刊 TGRS 和图像处理权威期刊 TIP 编委; 主持国家级和省部级项目十余项, 研究成果在航天国防、灾害应急与深空探测领域实现工程化部署与应用, 获麻省理工科技评论 AI100 青年先锋奖。

报告题目: 生成式遥感大模型研究进展

报告摘要: 结合领域知识生成逼真且具有场景理解能力的图像与视频数据, 已成为人工智能发展的重要方向。遥感作为对地观测与空间信息获取的重要技术手段, 发展具有遥感领域特色的生成式大模型, 对于数字地球构建、遥感场景模拟、数据增强以及下游智能解译等任务具有重要意义。在本次报告中, 讲者将介绍课题组近期在遥感生成式大模型方面的相关工作, 包括全球范围遥感图像生成模型 MetaEarth、多模态遥感图文生成模型 Text2Earth、三维遥感生成式大模型 MetaEarth3D 等。最后, 将对生成式遥感大模型的潜在应用场景和未来研究方向进行介绍。



武 宇

武宇, 武汉大学人工智能学院副院长、教授、博导, 国家重点研发计划实施方案专家组专家, 国家高层次青年人才。主持国家新一代人工智能重大项目课题等多项国家级项目。主要从事多模态大模型、大模型推理决策相关研究, 发表顶会论文 60 余篇, 谷歌学术引用 8000 余次, 单篇引用超 1000 次, 获得行业广泛关注。长期担任 NeurIPS、CVPR、ICML 等人工智能国际顶会领域主席, 担任 CVPR 2023 大会主席、主要组织者。获 2020 年谷歌博士奖(全球共 10 位获奖者)、2024 年 AAAI(国际人工智能促进会)学术明星奖, 获 6 项 CVPR/ICCV 国际学术竞赛的冠军。

报告题目: 高效可控视频生成

报告摘要: 本次演讲将从扩散模型的基本原理出发, 系统阐释视觉生成任务中可控输入的来源, 即条件信息如何有效引导生成过程。在此基础上, 进一步探讨构建高精度、高质量视觉生成方案的技术路径, 以满足实际应用对细节保真度与生成一致性的严格要求。此外, 还将介绍当前主流的视频生成方法, 分析其在长时序建模、运动连续性、运动物理性等关键方面的技术特点与局限。同时, 讨论多种高效视频生成方案, 包括模型结构优化与采样过程加速, 旨在实现兼具高效率与高保真度的流式视频生成。

分论坛 大模型安全与可信人工智能

时间：8月4日下午14:00-17:30 地点：402会议室

目 录

主席：谭 鑫（华东师范大学）、刘笑宏（上海交通大学）、蔡 青（合肥工业大学）

大模型安全分析与防御	向 涛（重庆大学）
Multimedia Intelligence toward AI Safety: From Safety Assessment to Safety Alignment	高联丽（电子科技大学）
EBFlex：企业级AI算力管理平台	李良斌（北京英博数科科技有限公司）
脑启发的持续学习与自主适应	王立元（清华大学）
大模型时代的视觉物理对抗攻击	韦星星（北京航空航天大学）
基于准则的智能体安全的思考与展望	方俊峰（新加坡国立大学）

论坛主席



谭 鑫

谭鑫，华东师范大学计算机科学与技术学院青年研究员（紫江青年学者），院长助理；上海人工智能实验室双聘副研究员；博士生导师。入选第十届中国科协青年人才托举工程、中国图像图形学会博士学位论文激励计划、上海市东方英才青年人才项目、上海市扬帆计划、晨光计划。他的主要研究兴趣是研究面向具身的生成式数据生成与智能体训练。他主持纵向项目包括新一代人工智能国家科技重大专项课题、国家自然科学基金项目（青年、国际合作交流）及上海市新一代信息技术专项等省部级项目 6 项。他还主持 CCF- 腾讯犀牛鸟基金（优秀结题）、华为校企合作等，也是华东师范大学 - 智元机器人联合实验室负责人。他以一作 / 通讯在 TPAMI、TIP、CVPR、ICCV、ICLR 等国际知名期刊和会议发表论文 50 余篇。他还担任多个知名期刊（Pattern Recognition, The Visual Computer 等）的编委（Associate editor）、ICLR 等会议的领域主席和国际多个顶级期刊和会议的审稿人等。



刘笑宏

刘笑宏, 上海交通大学计算机学院副教授、博士生导师、上海创智学院全时导师。研究方向为计算机视觉与多媒体信息处理, 现已在顶级国际会议与期刊上发表学术论文 100 余篇, 其中 CCF-A 类或中科院一区论文 80 余篇, 引用 8000 余次, h 指数 43。获中国人工智能学会吴文俊人工智能青年科技奖、华为“难题揭榜”火花奖、美团科研合作卓越创新奖, 入选全球前 2% 顶尖科学家榜单入选、上海市领军人才(海外)、上海市浦江人才。主持国家自然科学基金面上、青年项目, 参与重点项目, 承担包括华为、腾讯、美团、OPPO 等多家公司的产学研项目。担任多媒体领域知名期刊 ACM TOMM 副编辑, 多个顶级会议的领域主席, 中国图像图形学会青工委副秘书长、上海市计算机学会计算机视觉专委会副秘书长。



蔡 青

蔡青, 合肥工业大学“黄山学者”特聘教授、博士生导师, 先后荣获山东省“泰山学者”青年专家、山东省优青、山东省人工智能科学技术奖自然奖(1/3)和优秀青年奖等。CSIG 青工委副秘书长、CCF-CV 专委会委员、CCF-MM 专委会委员、CCF-AI 专委会委员、CSIG 3DV 专委会秘书处成员、安徽省人工智能学会脑机接口专委会副主任。主要研究领域为人工智能与医工交叉方向, 具体包括医学图像处理、脑机接口、具身智能等。相关工作已以第一或通讯作者身份发表在国内外重要学术会议、期刊: CVPR、AAAI、ACM MM、IEEE TIP 等。担任 IEEE TIP 等期刊副主编。主持国家自然科学基金(面上 & 青年)、山东省自然科学基金(优青 & 青年)等项目。

论坛讲者



向 涛

向涛, 二级教授, 重庆大学计算机学院院长, 教育部长江学者特聘教授, 国家重点研发计划首席科学家, 新一代人工智能国家科技重大专项、移动信息网络国家科技重大专项责任专家, 重庆市英才计划创新领军人才, 中国图像图形学会理事, 重庆计算机学会理事长。主要研究方向为人工智能安全、数据安全、区块链、隐私保护等。主持国家重点研发计划项目、国家自然科学基金联合重点项目等, 获重庆市自然科学一等奖、科技进步一等奖, 中国电子学会自然科学二等奖、科技进步二等奖等。

报告题目: 大模型安全分析与防御

报告摘要: 随着大模型技术迅猛发展及在各行各业的广泛应用, 其强大的能力背后所伴随的安全与伦理风险日益凸显, 已成为制约其健康发展的关键挑战。本报告旨在系统地剖析当前大模型所面临的威胁, 并汇报团队在大模型安全分析及防护的相关研究成果。报告将首先对大模型的发展历史、关键技术和应用背景进行简要概述, 深入分析模型的特点与风险, 详细阐述大模型在训练数据和推理应用等环节中的脆弱性。随后, 报告将介绍团队在大模型安全分析与防御方面的最新研究成果, 并对大模型安全领域的未来发展趋势进行了展望, 保障人工智能系统的安全应用。



高联丽

高联丽，电子科技大学，教授博导，国家领军人才，国家自然科学基金优青项目获得者。专注于多媒体理解、AI 安全、生成式 AI 与具身智能方面的研究。迄今为止，在人工智能领域重要国际期刊和国际会议发表高水平论文 100+, Google Scholar 引用 1.2 万 +, H 指数 60。主持多项国家级和省部级项目。担任多媒体领域国际期刊 IEEE TMM 编委、多个 CCF-A 类国际会议领域主席（如 CVPR, ICCV, ECCV, NeurIPS, ACM MM, AAAI）、2024 年中国科协青年科学家沙龙 - 通用人工智能青年科学家沙龙执行主席。获 2019 年阿里巴巴达摩院第二届青橙奖（全国仅 10 人）、2020 年国际 IEEE Technical Community on Multimedia Computing Rising Star Award（年度国际唯一）、2025 年 CSIG 青年女科学家奖、2025 年教育部自然科学一等奖、2025 年电子学会一等奖等。

报告题目：Multimedia Intelligence toward AI Safety: From Safety Assessment to Safety Alignment

报告摘要：本报告聚焦大模型安全议题，围绕人工智能安全领域两大核心研究方向：AI 安全评估与 AI 安全对齐展开汇报，系统梳理本团队近年来的前沿研究工作与突破性创新成果。



李良斌

李良斌，博士毕业于中国科学院大学。具有 10 年以上 AI 平台及应用开发经验，研究方向包括深度学习训练框架，虚拟机化及调度系统，大语言模型优化及应用。现任英博数科技术专家，云研发负责人，从 0 到 1 搭建英博云服务的技术体系，支撑以容器服务为基础的算力商业应用。加入英博数科之前，曾分别担任阿里云 PAI 团队核心研发，声智科技技术负责人。在阿里云期间主导开发的 PAI-TF 通信系统及 Runtime 引擎，广泛应用于淘系搜索、推荐、广告等多条业务线，获得 5~7 倍的性能提升；在声智科技期间，推进自研大语言模型 AzeroGPT 的技术迭代，成功落地行业内首批基于 LLM 的政务、医疗、税务应用系统。

报告题目：EBFlex：企业级 AI 算力管理平台

报告摘要：EBFlex 是英博数科推出的企业级 AI 算力管理平台，定位为一套可管、可调度、可运维、可私有化交付的一体化基础设施底座，而非单纯的调度工具。它面向 GPU 集群碎片化、多集群协同难、运维复杂度指数级上升等实际挑战，提供从硬件到 AI 应用的全栈能力：统一资源池整合多集群 GPU 实现全局算力池化，智能调度基于多维信息进行最优任务匹配，统一存储管理加速数据流转，网络优化与故障自愈保障分布式训练稳定性，并创新引入 Agentic 自主排障系统，将分散的监控证据组织成可审计的排障结论，大幅降低运维门槛。在实际落地中，EBFlex 帮助客户将 GPU 利用率从 30% 提升至 80% 以上，高优先级任务排队时间缩短 90%，运维效率提升 4 倍，真正让企业 AI 团队聚焦算法创新，而非底层基础设施的琐碎维护。



王立元

王立元，清华大学心理与认知科学系助理教授、博士生导师。研究方向为脑启发的持续学习理论与方法，并推动其在科学智能、智慧医疗、具身智能等关键场景中的应用，研究成果以第一/通讯作者多次发表在Nature Machine Intelligence、Nature Communications、TPAMI、Patterns、NeurIPS、ICML、ICLR等人工智能相关领域的顶级期刊和会议。入选北京市科技新星、中国人工智能学会首届清源学者、世界人工智能大会云帆奖和青年优秀论文奖。

报告题目：大模型驱动的自主智能体和安全风险

报告摘要：持续学习是连接人工智能与生物智能的重要研究方向，旨在使智能系统能够在动态开放环境中持续获取、整合、更新和利用知识。本报告将从人工神经网络的灾难性遗忘问题出发，介绍脑启发的持续学习计算模型，面向预训练大模型的高效持续学习方法、以及脑启发的具身智能体空间记忆，推动人工智能从即时反应走向记忆支撑的空间认知，从一次性学习走向长期能力演化。



韦星星

韦星星现为北京航空航天大学人工智能学院教授，博士生导师，国家级青年人才，入选年度全球前 2% 顶尖科学家榜单。主要研究方向为计算机视觉、安全可信人工智能。在TPAMI,IJCV 和 CVPR,ICCV 等人工智能领域顶级期刊和会议发表学术论文 100 余篇（其中Nature Communications 论文 1 篇,TPAMI/IJCV 论文 14 篇,ESI 高被引论文 5 篇）,Google 引用 6000 余次,授权发明专利 20 余项（其中 2 项已经成果转化）。作为负责人主持科技部“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（2 项）、国防基础科研计划项目、华为 / 腾讯校企合作项目。多次担任 NeurIPS, CVPR, AAAI, IJCAI 等国际顶会的领域主席和高级程序委员会委员。荣获国防技术发明一等奖。

报告题目：大模型时代的视觉物理对抗攻击

报告摘要：随着人工智能深度介入安全攸关场景，智能模型固有的脆弱性为智能感知带来了严峻挑战。攻击者可通过构造对抗样本诱导模型误判，威胁物理空间安全。因此，深入挖掘大模型时代的感知漏洞，是构建可信智能感知的核心。本报告重点探讨大模型时代视觉物理对抗攻击的三大转变，并基于此提出“几何对抗攻击（Geometric Adversarial Attack）”新框架。该框架通过优化目标的几何属性来应对上述演变：在姿态维度，通过脆弱姿态挖掘实现“零扰动”攻击，利用目标拓扑缺陷诱导大模型误判；在形状维度，基于形状建模确保攻击在不同波谱成像下具备物理一致性，突破多模态防御；在位置维度，利用敏感位置引导技术锁定几何关键区域，显著增强黑盒迁移效能。本研究揭示了多模态大模型在物理感知层面的安全软肋，为构建高可信智能系统提供了系统性的风险评估依据。



方俊峰

方俊峰, 新加坡国立大学 NExT Lab 博士后研究员, 博士毕业于中国科学技术大学, 研究方向为人工智能可信与安全, 以第一 / 通信作者身份发表论文 30 余篇, 其中包括多篇顶级会议如 ICLR、NeurIPS、ICML Oral 论文(平均录用率 < 2%)与顶级期刊如 TPAMI, 所获荣誉与提名包括: 博士国家奖学金、ICLR 2025 最佳论文、安徽人工智能学会优秀博士(唯二)、2025 世界人工智能大会优秀论文、中国科学技术大学优秀博士等。

报告题目: 基于准则的智能体安全的思考与展望

报告摘要: 随着 AI 系统从大语言模型走向具备推理、工具调用和交互决策能力的智能体, 安全问题正从静态问答场景升级为动态、多源、实时的系统性挑战。本报告将围绕智能体安全中的数据、训练、部署和后防护问题展开, 讨论 On-Policy Distillation 在高效政策内化中的作用, Policy-as-Skills 在安全奖励解耦中的意义, Noisy-injected RL 在真实环境鲁棒性中的价值, 以及 Streaming Safety 与 Latent Reasoning 对实时智能体防护的重要性, 讨论基于 Rubric 或 Policy 的训练与部署对齐对于 agent 安全的重要性。

学术服务论坛 / 人工智能基础

时间：8月4日下午14:00-17:30 地点：201会议室

内 容

主席：程 德（西安电子科技大学）、刘 武（中国科学技术大学）
普 楠（合肥工业大学）、杨 勋（中国科学技术大学）

指导专家：鲍秉坤（合肥工业大学）	常晓军（中国科学技术大学）
蒋树强（中国科学院大学）	李厚强（中国科学技术大学）
宋井宽（同济大学）	王琼华（北京航空航天大学）
肖 俊（浙江大学）	余正涛（昆明理工大学）

青年学者：成 娟（合肥工业大学）	党吉圣（兰州大学）
李俊成（浙江大学）	刘 瑞（内蒙古大学）
刘佐珠（浙江大学）	齐梦实（北京邮电大学）
王浩宇（中国矿业大学）	姚涵涛（中国科学技术大学）
于灵云（中国科学技术大学）	

指导专家



鲍秉坤

鲍秉坤，合肥工业大学计算机与信息学院院长，教授，博士生导师，国家杰出青年科学基金获得者、中组部 WR 计划青年拔尖人才、江苏省杰青，江苏省双创人才。研究方向为多媒体计算、社交多媒体、计算机视觉、人工智能等。主持多项国家和省部级项目，包括国家科技创新 2030-“新一代人工智能”重大专项、国家自然科学基金重点项目 2 项、江苏省重点研发计划等。在相关领域发表论文百余篇、申请专利五十余项。荣获第十九届江苏省青年科技奖、江苏省青年女科学家奖、2018 年度电子学会科学技术（自然科学类）一等奖。荣获多媒体领域的 ACM 汇刊 TOMM 2016 年度最佳论文奖、IEEE MM 2017 年度最佳论文奖、Multimedia Modeling 2019 年度最佳论文 Runner Up 奖。荣获 ICME 2020 Outstanding Areas Chair。担任 IEEE Trans. on Multimedia、IEEE Trans. on Circuits and Systems for Video Technology、ACM Trans. on Multimedia Computing, Communications and Applications、Multimedia Systems 编委、中国图像图形学会女工委副主任。



常晓军

常晓军，中国科学技术大学讲席教授，国家级高层次人才，中国科学院海外引进杰出人才，民盟盟员，澳大利亚研究理事会青年研究奖获得者。主要从事具身智能多模态大模型及类脑多模态大模型等方面的研究。主持包括国家自然科学基金区域联合基金，科技部新一代人工智能重大专项等在内的国家级科研项目十余项。相关科研成果发表在 T-PAMI、IJCV 等国际顶级期刊以及 CCF A 类会议上共计 150 多篇，GoogleScholar 引用次数超过 2 万次，其中 21 篇论文入选 ESI 高被引 / 热点论文。2019 年至 2024 年连续入选科睿唯安高被引科学家，入选爱思唯尔 2024“中国高被引学者”榜单。担任 IEEE TCSVT、IEEE TNNLS、ACM TOMM 等国际顶级期刊的副主编，并在多个 CCFA 类国际会议中担任领域主席 (AreaChair)，积极参与国际学术社区建设与服务。



蒋树强

蒋树强，中国科学院大学特聘教授，中国科学院计算技术研究所客座研究员，博士生导师，具身智能计算北京市重点实验室主任，先后担任期刊《IEEE TMM》、《ACM ToMM》、《IEEE Multimedia》、《计算机研究与发展》、《JCST》、《CAD 学报》编委，中国人工智能学会具身智能专委会主任，中国计算机学会多媒体专委会副主任、中国自动化学会网络计算专委会副主任、ACM SIGMM 中国分会副主席。主要研究方向是多媒体内容分析、多模态具身智能技术和食品计算。主持承担科技创新 2030-“新一代人工智能”重大项目、国家自然科学基金青年基金 A 类(杰青)、B 类(优青)、重点等项目 20 余项，发表论文 200 余篇，获授权专利 20 余项，多项技术应用到实际系统中，先后获省部级或学会奖励 5 项。



李厚强

李厚强，中国科学技术大学信息科学技术学院教授，多媒体计算与通信教育部 - 微软重点实验室主任，IEEE Fellow，长江学者特聘教授，国家“杰出青年基金”获得者，国务院学位委员会“智能科学与技术”学科评议组成员。主要研究方向包括：图像处理与计算机视觉、视频编码与通信、强化学习与机器博弈等。主持国家科技创新 2030-“新一代人工智能(2030)”重大项目、国防科技创新特区智能博弈重大项目、国家基金委重点项目、973 项目课题、863 项目等国家级科研项目；发表论文 300 余篇；获授权发明专利 80 余项，有 45 项技术提案被 H.264、H.265 和 MPEG-3D 等视频编码国际标准接受；担任国际著名期刊 IEEE TCSVT 和 IEEE TMM 编委，担任多媒体领域国际旗舰会议 ICME 2021 大会共同主席；获 2019 年国家技术发明二等奖、2015 年国家自然科学二等奖、2012 年安徽省科学技术一等奖，2013-2016 连续 4 年获得中科院优秀研究生导师奖。



宋井宽

宋井宽，同济大学计算机科学与技术学院教授，国家高层次青年人才(海外)，国家杰出青年科学基金获得者。主要研究方向为多模态、具身智能。在多媒体、计算机视觉、人工智能等领域的重要会议和期刊发表论文180余篇，谷歌学术引用18000余次。担任国际SCI期刊IEEE TMM、ACM TOMM等编委，担任多个期刊的评审和多个国际顶级会议(MM'18-'24, IJCAI'18-'24)的领域主席。主持自然科学基金委重点、科技部重点研发课题等多项国家级项目。



王琼华

王琼华是教育部长江学者特聘教授、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、国际信息显示学会、美国光学学会、国际光学工程学会、中国光学学会、中国光学工程学会会士、Photonix等期刊编委和国家重点研发计划项目负责人，入选爱思唯尔与斯坦福大学“全球前2%顶尖科学家和中国高被引学者榜单”。研究方向为显示与成像技术。研制了裸眼集成光场3D显示器、全息3D显示系统、液体透镜和连续光学变焦显微镜等。发表SCI收录论文约400篇，出版科学出版社书籍3部，获准美国和中国发明专利近200件，以第一完成人荣获2023年度国家技术发明二等奖和3项省部级/学会科技一等奖，荣获“全国三八红旗手”称号。1988-2001年在电子科技大学获得学士、硕士和博士学位并任职到副教授，2001-2004年在美国中佛罗里达大学光学学院从事博士后，2004-2018年在四川大学任教授，2018年起在北京航空航天大学任教授。



肖俊

浙江大学求是特聘教授，人工智能研究所副所长，教育部“长江学者奖励计划”特聘教授。多次入选全球前2%顶尖科学家（人工智能与图像处理领域）及爱思唯尔中国高被引学者；国家科技创新2030-新一代人工智能重大项目“新一代人工智能科教创新开放平台”负责人/首席科学家、国家重点研发计划“工业软件”重点专项“大模型驱动下产业信息智能分析与服务关键技术研究及示范应用”负责人/首席科学家。曾获吴文俊人工智能科学技术奖自然科学奖一等奖（2024）、浙江省科技进步一等奖（2025、2021）、浙江省自然科学奖一等奖（2015）及浙江省科技进步二等奖（2022）。



余正涛

余正涛，昆明理工大学教授，博导，云南省人工智能重点实验室主任，入选国家高层次人才特殊支持计划科技领军人才，云南省科技领军人才等。主要从事大模型、机器翻译、语音识别及合成等人工智能方向研究，主持完成国家重点研发计划项目、国家自然科学基金重点项目、云南省科技重大专项项目等 60 余项。在 TIP、Artificial Intelligence、ACL、AAAI 等国际顶级期刊和会议发表学术论文 300 余篇，授权国家发明专利 100 余项。获得云南省科技进步特等奖、云南省自然科学一等奖等省部级科技奖励 10 余项。

论坛主席



程 德

程德，西安电子科技大学教授、博士生导师。分别于2011和2017年在西安交通大学获得本科和博士学位，美国卡内基梅隆大学联合培养博士，曾任华为公司主任工程师。主要研究方向为计算机视觉、机器学习等，近年来已在该领域权威期刊/会议发表论文100余篇，第一/通讯作者发表中国科学院一区Top权威期刊/CCF-A 类国际会议80篇（其中包含T-PAMI/IJCV 7篇，CCF-A类会议论文35篇），谷歌学术总引用4000余次，一作单篇论文最高被引1600余次，ESI高被引论文2篇，授权发明专利30余项，相关技术应用于研究所和企业。目前主持国家自然科学基金面上项目2项、陕西省青年科学基金B类、国家重点研发计划子课题、陕西省重点研发计划、国家自然科学基金区域联合基金重点项目一课题等项目。



刘 武

刘武，中国科学技术大学特任教授，入选国家级青年人才计划。在重要国际会议和期刊上发表论文100余篇，曾获得IEEE T-MM、IEEE MM和IEEE ICME等最佳论文奖，以及CAAI吴文俊人工智能科技进步特等奖、天津市科技进步特等奖、ACM中国新星奖、中科院优秀博士论文奖、中科院院长特别奖等，入选了《麻省理工科技评论》亚太区“35岁以下科技创新35人”，中关村论坛-AI100青年先锋，并担任了IEEE T-MM和Pattern Recognition Associate Editor，IEEE ICME 2022和ACM MM Asia 2021技术委员会主席，2026年中国多媒体大会技术委员会主席，IET Fellow评审委员会委员等。



普楠

普楠，合肥工业大学计算机与信息学院教授、博士生导师，国家级青年人才。长期致力于大模型驱动的可信多媒体计算研究，聚焦多模态感知与理解中的持续学习、联邦学习和新类发现等方向，相关成果在行人重识别、多模态问答、AI4Science实际应用中广受认可。目前主持国家级青年人才项目1项。已在国内外重要期刊与会议上发表学术论文50余篇，其中以第一/共同第一作者在CCF-A类会议/期刊及计算机视觉三大顶会（TPAMI、NeurIPS、CVPR、ICCV、ECCV等）发表论文10篇。兼任国际人工智能顶级会议ICLR 2025及多媒体旗舰会议ACM MM 2024&2025领域主席，并获评ACM MM 2024杰出领域主席。



杨勋

杨勋，中国科学技术大学信息科学技术学院特任教授，博士生导师，国家高层次青年人才项目获得者；曾在新加坡国立大学计算机学院NExT研究中心担任博士后研究员；2021年回国加入中国科学技术大学信息学院。近年来主要关注跨媒体分析与推理以及人工智能模型可靠性等方向的研究。在CCF A类刊物以及IEEE/ACM Transactions国际期刊上发表一作/通信作者论文50余篇。主持国家自然科学基金联合基金重点项目、面上项目、长三角联合攻关基础研究项目。现担任CCF多媒体技术专委会执行委员、CSIG青工委副秘书长、CSIG/CCF高级会员；担任IEEE Transactions on Big Data、IEEE Transactions on Fuzzy Systems 和 Multimedia Systems 期刊的编委，多次担任教育部、国家自然科学基金、北京市等人才/重点类项目的函评专家。

学术服务论坛

计算机视觉I

时间：8月4日下午14:00-17:30 地点：202会议室

内 容

主席：李 帅（山东大学）、舒祥波（南京理工大学）
谢国森（南京理工大学）、张 正（哈尔滨工业大学（深圳））

指导专家：韩军伟（重庆邮电大学）
马思伟（北京大学）
王国胤（重庆师范大学）
张勇东（中国科学技术大学）

黄庆明（中国科学院大学）
史振威（北京航空航天大学）
徐常胜（中国科学院自动化研究所）

青年学者：邓淑敏（浙江大学）
胡心亭（中国科学技术大学）
刘 李（香港科技大学（广州））
沈 雯（同济大学）
徐斌倩（新加坡国立大学）
张宁豫（浙江大学）

黄步真（天津大学）
林绍辉（华东师范大学）
沈 力（中山大学）
宋霄罡（西安理工大学）
易 冉（上海交通大学）
周 康（中国科学院香港创新院）

指导专家



韩军伟

韩军伟，重庆邮电大学副校长，曾任西北工业大学自动化学院院长，IEEE Fellow，国家级领军人才，国家重点研发计划项目首席科学家。主要研究方向是人工智能、模式识别、类脑计算、医学影像处理等。在领域顶级期刊/会议如：Nature子刊，Proceedings of the IEEE，IEEE TPAMI，CVPR，MICCAI等发表学术论文150余篇，论文被引用近5万次。获2021年度和2023年度IEEE地球科学与遥感学会最有影响力论文奖、国际期刊IEEE TCSVT 2021最佳论文奖。获陕西省科学技术一等奖、吴文俊人工智能技术发明一等奖等10项省部级科技奖。担任IEEE TPAMI、《中国科学：信息科学》等多个国内外期刊编委。



黄庆明

黄庆明，中国科学院大学讲席教授，国家杰青，百千万人才工程国家级人选，享受国务院政府特殊津贴。主要研究方向为多媒体计算、图像视频分析、机器学习、计算机视觉等，主持承担了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和重点国际合作项目、863课题、973课题等国家级项目的研究工作，已在国内外权威期刊和重要国际会议上发表学术论文700余篇（Google学术引用44000余次，H指数95），其中IEEE/ACM汇刊和CCF-A类论文400余篇（包括T-PAMI论文40余篇），出版专著3部，申请国内外发明专利70余项（40余项已授权）。研究成果获得吴文俊人工智能自然科学一等奖、中国图像图形学会自然科学一等奖、教育部科技进步一等奖等10余项国家学会和省部级奖励。任IEEE Fellow，IEEE CASS北京分会主席，CCF理事，CCF会士，CCF多媒体技术专业委员会主任，中国图像图形学会常务理事，北京图像图形学会副理事长。



马思伟

马思伟，北京大学博雅特聘教授，北京大学计算机学院党委副书记，视频与视觉技术国家工程研究中心副主任，IEEE Fellow，国家杰出青年科学基金获得者。主要从事高效视频编码技术与标准研制、沉浸式媒体处理与压缩、多媒体信号处理与系统等研究，已在视频编码领域国际权威期刊和会议上发表论文300余篇，谷歌学术引用超17000次，获授权发明专利50余项。曾获国家技术发明奖一等奖、国家技术发明奖二等奖、国家科学技术进步奖二等奖、腾讯科学探索奖、IEEE 1857国际标准贡献奖等多项奖励。目前担任国家重点研发项目“沉浸式媒体技术在文旅场景的集成创新与应用”首席科学家，持续探索沉浸式媒体技术在文旅场景中的集成创新与应用。近年来，他将研究方向拓展至沉浸式文旅与智能艺术的交叉领域，致力于探索计算美学、计算艺术学等新兴方向，推动技术与艺术在文化场景中的深度融合。



史振威

史振威，北京航空航天大学二级教授，国家杰出青年科学基金获得者，国家级创新团队负责人，北航校学术委员会委员。主要从事遥感图像处理、遥感大模型、机器学习和时空智能方面的研究，发表科研论文300余篇（其中IEEE Trans期刊论文100余篇），论文引用27000余次，获2025年度国际地球科学与遥感协会（GRSS）最高影响力论文奖。担任学术期刊《IEEE Transactions on Geoscience and Remote Sensing》、《Pattern Recognition》、《中国图象图形学报》和《数据采集与处理》等编委；担任《Infrared Physics & Technology》和《中国空间科学技术》的副主编。现任中国图像图形学会监事。



王国胤

王国胤, 教授, 工学博士, 博士生导师, 重庆师范大学校长, 国家自然科学基金委员会创新研究群体负责人, 类脑认知计算与特殊儿童教育康复重庆市重点实验室主任、网络空间大数据智能安全教育部重点实验室主任、大数据智能计算示范型国家国际科技合作基地主任、旅游多源数据感知与决策技术文化和旅游部重点实验室主任。2009 年入选人社部“新世纪百千万人才工程”国家级人选, 2019 年入选首批“重庆英才·优秀科学家”, 国家级人才, 是国务院特殊津贴专家、国家重点研发计划项目首席科学家(2016、2021)、全国优秀教师。曾任国际粗糙集学会(IRSS) 理事长(2014-2017)、中国人工智能学会(CAAI) 副理事长(2014-2025) 和中国计算机学会(CCF) 理事(2008-2023), 现任中国认知科学学会(CSCS) 副理事长、CAAI 监事、重庆市人工智能学会(CQAAI) 理事长, 是教育部高等学校教育技术与人工智能教学指导分委员会委员, 是 IRSS 会士、CAAI 会士和 CCF 会士, 是《Int. J. of Approximate Reasoning》、《计算机学报》等 10 余种期刊编委。组织召开国际学术会议 30 余次, 应邀在国际学术会议作大会特邀报告 30 多次。主要研究领域包括: 大数据、人工智能、认知计算、机器学习、粒计算、粗糙集等, 主持国家重点研发计划项目、国家自然科学基金重点项目、863 计划项目等国家级 / 省部级科研项目数十项。带领的教学科研团队分别评为“国家级教学团队”和“国家自然科学基金委员会创新研究群体”。获国家级高等教育教学成果二等奖、吴文俊人工智能科技进步奖一等奖、重庆市自然科学一等奖、重庆市科技进步一等奖、重庆市教学成果一等奖等教学科研成果奖励 10 余项, 是 2024 年度吴文俊人工智能科技进步奖科技贡献奖获得者。



徐常胜

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、科技部重点领域创新团队负责人、国家万人计划领军人才、国家百千万人才工程入选者、首都科技创新领军人才、中科院王宽诚率先人才计划卢嘉锡国际团队负责人, IEEE/IAPR Fellow和ACM 杰出科学家。担任ACM多媒体专委会中国区(ACM SIGMM China Chapter) 主席。发表 IEEE/ACM会刊论文200余篇, CCF-A类会议论文150余篇, 在多媒体领域国际顶级会议和期刊上获得最佳论文奖10余次, 以第一完成人获中国电子学会自然科学一等奖和技术发明一等奖、中国人工智能学会技术发明一等奖, 7次获得中国科学院优秀导师奖。



张勇东

张勇东，教授，博士生导师，中国科学技术大学信息科学技术学院执行院长，传播内容认知全国重点实验室首席科学家，IEEE Fellow。国家自然科学基金委创新研究群体项目负责人，国家杰出青年科学基金获得者，“万人计划”科技创新领军人才。主要研究方向为网络空间多媒体内容安全，曾获国家自然科学奖二等奖1项，国家技术发明奖二等奖1项和国家科技进步奖二等奖1项。所发表论文谷歌引用3万余次，连续多年入选科睿唯安全球高被引学者。获授权国家发明专利100余项，相关技术在著名国际评测中多次取得优异成绩。研究成果有效应用于国家网络空间内容安全领域。担任《中国通信》副主编，国家重点研发计划-“变革性技术关键科学问题”重点专项总体专家组成员，国家重点研发计划-“社会治理与智慧社会科技支撑”重点专项总体专家组成员。

论坛主席



李帅

李帅，山东大学，教授，齐鲁青年学者，山东省优青。致力于高效视频编码及分析、计算机视觉及人工智能的研究，已发表国际期刊及会议论文100余篇。荣获3次国际学术会议最佳论文奖（IEEE ISCAS、IEEE BMSB、IIH-MSP），并荣获IEEE Transactions on Broadcasting最佳论文奖提名。代表性工作独立循环神经网络（IndRNN）获得国际多位知名专家引文的肯定，包括图灵奖共同得主之一Yoshua Bengio，单篇引用1300+次，并被集成到多个深度学习软件库中。



舒祥波

舒祥波，南京理工大学计算机学院/人工智能学院/软件学院副院长、教授，社会安全信息感知与系统工信部重点实验室副主任、国家优青获得者、江苏省杰青获得者。研究方向为视觉具行为计算，在TPAMI、CVPR、ICCV、ICLR等期刊/会议上发表论文100余篇，其中ESI高被引论文8篇；获江苏省自然科学一等奖、中国电子学会自然科学一等奖、ACM MM 2015最佳论文提名、MMM 2016最佳学生论文、CCBR 2026最佳论文、江苏省优博、中国人工智能学会优博、江苏省行业领域优秀科技进展、全球前2%顶尖科学家（2021-2025年）；承担国家自然科学基金联合重点项目、国家重点研发课题、国家重大科研仪器项目课题、国家自然科学基金面上项目、江苏省杰出青年基金等项目等。担任CSIG青工委副秘书长，以及TNNLS、TCSVT、TIFS等期刊编委。



谢国森

谢国森，南京理工大学计算机科学与工程学院教授、博士生导师，国家级青年人才，江苏特聘教授。2016年于中国科学院自动化研究所获工学博士学位，先后在新加坡、阿联酋留学和工作。研究方向为计算机视觉、多模态图像理解、零/少样本学习、工业缺陷检测等，在领域内国际期刊/会议发表论文100余篇，包括TPAMI、IJCV、TIP、NeurIPS、CVPR等。多次入选全球前2%顶尖科学家榜单，曾获国际会议MMM 2016最佳学生论文奖。主持多项国家级和省部级科研项目，包括国家自然科学基金、江苏省人才项目等。受邀担任 IEEE TIP、Pattern Recognition等权威期刊编委和ICLR等顶级会议的领域主席。



张正

张正，哈尔滨工业大学（深圳）长聘教授，教育部青年长江学者，广东省珠江学者。长期从事高效能多模态机器学习的研究，专注于高效与可信大模型，出版中英文学术专著/教材5部，发表IEEE/ACM 汇刊和CCF A类论文100余篇，谷歌引用1.5万次，入选全球前0.05%顶尖科学家榜单。主持国家级和省部级自然科学基金以及阿里巴巴创新研究计划、华为合作基金等科研项目10余项。受邀担任IEEE TIFS高级领域编辑、Information Sciences高级编辑以及IEEE TIP、IEEE TAFRC等期刊编委，常年担任ICML、NeurIPS、ICLR、CVPR、AAAI、ACM MM等顶级会议的领域主席。

学术服务论坛

交叉研究

时间：8月4日下午14:00-17:30 地点：204会议室

国内 容

主席：白 琮（浙江工业大学）、张敬林（山东大学）
赵世杰（西北工业大学）、郑钰辉（青海师范大学）

指导专家：黄科科（中南大学）
陶建华（清华大学）
王 亮（中国科学院自动化研究所）
俞 俊（哈尔滨工业大学（深圳））

李肯立（湖南大学）
王景景（青岛科技大学）
吴建德（云南大学）
张 靖（西安交通大学）

青年学者：陈 想（南京航空航天大学）
付 钰（兰州大学）
刘雨帆（中国科学院自动化研究所）
潘良睿（新加坡南洋理工大学）
王培松（中国科学院自动化研究所）
熊凤超（南京理工大学）
祝澜耘（同济大学）

杜云涛（山东大学）
李 晴（北京理工大学）
刘政哲（岭南大学）
沈 飞（新加坡国立大学）
王乾铭（华北电力大学）
姚鑫杰（昆明理工大学）

指导专家



黄科科

黄科科，中南大学教授、博导、自动化学院副院长。入选长江学者特聘教授、国家万人计划青年拔尖人才、湖南省科技创新领军人才、省杰青等。长期从事工业智能与系统相关研究工作，在IEEE Trans.、IFAC旗舰期刊、CCF A类会议等发表高水平论文100余篇，授权中国/美国等发明专利30余项。先后主持国家重点研发计划项目、国家重大科技专项、基金委重大研究计划/面上等10余项。曾获国家科技进步二等奖（排2）、湖南省科技进步一等奖（排2）、中国有色金属学会优秀青年科技奖、IEEE TCSDM Young Professional Award等。兼任IEEE TASE、IEEE SMCM、IET CPS等期刊编委和中国有色金属学会自动化学术委员会副秘书长、中国人工智能学会智能融合专委会常务委员、中国自动化学会控制理论专委会/过程控制专委会委员、中国计算机学会计算机应用专委会执行委员等。



李肯立

李肯立, 湖南大学副校长, 国家超级计算长沙中心主任, 全国创新争先奖获得者, 入选国家高层次人才计划, CCF 会士, 宝钢优秀教师, 享受国务院政府特殊津贴。教育部高效能计算学科创新引智基地负责人、高性能计算应用软件技术教育部工程研究中心主任。国家超级计算创新联盟副理事长、新一代人工智能产业技术创新联盟专家委员会委员、高性能计算国家重点研发计划总体专家组成员、IEEE 高级会员、CCF 长沙主席、湖南省计算机学会副理事长、IEEE-TC/TSUSC/TII 副编辑、《计算机研究与发展》编委。主持国家重点研发计划、国家重大科研仪器研制项目、基金委重点项目、JWKJW 重点项目等国家和省部级项目 40 余项。以排名第一获国家科技进步二等奖 2 项、中国专利金奖 1 项、省部科技一等奖 4 项, 获国家科技进步一等奖创新团队奖 (排名 13)。主要研究领域为高性能计算调度与应用, 发表 CCF 推荐 A 类会议 / 期刊、IEEE/ACM 汇刊论文 200 余篇。获授权专利和软件著作权共 150 余项。



陶建华

陶建华, 清华大学自动化系长聘教授, 北京信息科学与技术国家研究中心副主任, 国家杰出青年基金获得者, 国家级领军人才, 享受国务院政府特殊津贴, 连续多年被爱思唯尔评为全球 Top2% 顶尖科学家。主要从事多源信息融合、模式识别、大数据处理等方向, 在 IEEE TPAMI、IEEE TAFRC、IEEE TASLP、AAAI、IJCAI、ICASSP 等国内外主要期刊或会议上发表论文 300 余篇, 先后负责国家重点研发计划项目、国家自然科学基金重点项目、国家 863 重点项目、国家发改委项目等 20 余项重点科研任务。研究成果大量应用于百度、阿里、腾讯、华为等企业, 并在公安部、网信办等发挥重要作用, 成果获 2022 年中国人工智能学会吴文俊技术发明特等奖、2021 年和 2024 年中国电子学会技术发明一等奖、2018 年中国电子学会技术进步一等奖, 并多次在国内外学术会议上获奖。连目前担任中国人工智能学会会士和常务理事、中国计算机学会会士等, 并担任 TAC、TASLP、Speech Communication、计算机研究与发展等多个主要国内外期刊编委, 同时担任 Interspeech、ACII、IEEE ICSP、IEEE MLSP 等会议大会主席或程序委员会主席。



王景景

教育部长江学者特聘教授, 全国巾帼建功标兵, 国务院特殊津贴专家, 山东省泰山学者, 山东省先进工作者。现任青岛科技大学校务委员、信息科学技术学院院长, 教授/博导。主持国家自然科学基金项目 5 项 (其中重点类项目 2 项) 以及国防项目等, 在国内外高水平学术刊物和重要学术会议发表论文六十多篇, 其中以第一作者和通信作者发表 SCI 论文 27 篇; 申请国家发明专利 59 件, 授权 49 件, 转化 5 件。成果获山东省技术发明一等奖 (排 1)、山东省教学成果特等奖 (排 1)。担任中国人工智能学会智能融合专委会常务委员, 国际计算机协会 (ACM) 青岛分会副主席, 山东省深海装备物联网实验室主任。



王亮

王亮，中国科学院自动化研究所研究员，博导，国家杰青，IEEE/IAPR/CIE/CCF/-CAAI/CSIG Fellow。目前是多模态人工智能系统全国重点实验室副主任，中国计算机学会计算机视觉专委会副主任，中国图像图形学会副理事长。担任IEEE TPAMI、IEEE TIP等国际期刊编委、MIR的副主编，曾是IEEE TIFS、IEEE TSMC-B、PR等国内外学术期刊的编委，以及10余次国内外重要会议的大会、程序或组织主席。主要从事计算机视觉、模式识别、机器学习等相关领域的研究。截至目前，已申请或授权发明专利100余项，出版编专著10余部，已发表或接收论文400余篇，获得5项国内外会议最佳论文奖。目前，谷歌学术引用6.5万余次，H-index是115（2014-2025连续入选ESI中国高被引科学家名单）。



吴建德

吴建德，云南大学教授，入选国家级高层次领军人才、云岭学者，兼任中国有色金属工业智能系统与先进控制工程技术研究中心主任等职务。主要从事复杂工业过程故障检测与智能控制、工业大数据分析建模等方面的研究。先后主持国家自然科学基金区域联合基金重点项目、云南省重大科技专项等科研项目20余项；在IEEE汇刊等国内外重要学术期刊发表论文70余篇；获授权发明专利40余项。研究成果先后荣获国家科学技术进步二等奖、云南省科学技术进步一等奖、云南省自然科学一等奖、中国自动化学会自然科学一等奖等多项奖励。



俞俊

国家杰出青年基金获得者，现任哈尔滨工业大学（深圳）智能科学与工程学院院长、二级教授、博导。浙江大学计算机学院本科、博士，新加坡南洋理工大学博后。曾分别入选国家优青与青年长江等人才计划。主持基金委重点项目、科技部重点研发项目、省杰青人才项目等。致力于图像处理与分析、多模态内容理解的研究。发表IEEE/ACM Trans及CCF A类论文百余篇，授权国家发明型专利30余项，谷歌引用2.3万余次。任IEEE TMM、TCSVT、Pattern Recognition等汇刊编委。曾以第一作者获IEEE TMM, TIP, TCYB等期刊的最佳论文奖，获省自然科学一等奖2021（排名第一）。成果在人民日报、阿里等落地。



张 靖

张靖，国家级人才计划获得者，量子科学与技术专家。2001年与2006年分别于清华大学数学系与自动化系获得学士与博士学位，之后在清华大学做博士后并留校，2022年加入西安交通大学电信学部，现任西安交通大学电信学部自动化科学与工程学院教授，博士生导师。2016年得到国家自然科学基金优秀青年基金支持。2017年入选教育部青年长江学者。2021年入选教育部长江学者特聘教授。主要从事量子与微纳系统控制、光学微腔传感器等研究。在IEEE Transactions on Automatic Control, Nature, Nature Photonics, Nature Physics Reviews, Physics Reports, Physical Review Letters等物理与控制领域权威期刊发表论文60余篇。关于光机械系统中混沌传递与随机共振的工作被Nature Photonics选为2016年6月刊的封面热点论文，并入选“2016中国光学重要成果”。作为第一作者工作获国际自动控制联合会(IFAC)青年作者奖，该奖项是自动化领域最重要国际奖项之一，这是我国高校系统第一次获此奖项。

论坛主席



白 琮

白琮，博士，浙江工业大学计算机科学与技术学院教授，博士生导师。浙江省杰出青年基金及延续资助项目获得者。全球前 2% 科学家入选者。兼任中国气象学会气象人工智能专业委员会副秘书长。研究方向为多媒体信息处理与智慧气象。主持国自然联合重点、面上及科技部重点研发计划课题等省部级以上项目 10 项；在国内外知名学术期刊或会议上发表论文 100 余篇，包括 Nature Communications、JGR:Atmospheres、GRL、IJCV、IEEE TPAMI、IEEE TIP、IEEE TGRS、ICML、NeurIPS、KDD 和 CVPR 等领域内的权威期刊或会议上。曾获 CSIG 科技进步二等奖、CCF 技术发明二等奖等，并 5 次获得国际 / 国内知名学术会议的最佳 / 优秀论文奖。担任《Displays》的编委会委员，IEEE《Journal of Selected Topics in Applied Earth Observations and Remote Sensing》副编辑。



张敬林

张敬林，博士毕业于法国雷恩国立应用科学院，现任山东大学控制科学与工程学院教授，主要从事人工智能、海洋遥感以及工业大数据分析等研究，曾获得山东省杰青、泰山学者特聘教授以及全球前2%顶尖科学家等荣誉。担任教育部数字化教学工程中心专家委员、数字家庭网络国家工程研究中心技术委员会委员、中国人工智能学会组织工作委员会委员；发表Science Bulletin、Nature Communication、IJCV、TPAMI、TNNLS、TMM、CVPR、ICML、AAAI、KDD等高水平会议和期刊论文100余篇，授权发明专利30余项，主持科技部重点研发计划项目、基金委联合基金重点项目在内的国家级课题7项，获得CSIG科技进步二等奖、CCF技术发明二等奖以及广东省技术发明一等奖等多个省部学会奖励。



赵世杰

赵世杰，西北工业大学长聘教授，博士生导师，青年科学基金（B类）获得者。主要研究方向是脑机接口、医学影像处理、人工智能算法等，以第一作者/通讯作者（含共同）在Nature Communications、IEEE TMI、Medical Image Analysis等领域重要期刊和会议发表论文40余篇。主持国家自然科学基金联合基金重点支持项目、国家重大科技专项课题、军口国家级纵向项目(300万)等项目10余项。担任医学图像计算青年研讨会(MICS)副主席、中国图像图形学会视觉认知与计算专业委员会委员、中国人工智能学会生物信息学与人工生命专业委员会委员、中国生物医学工程学会医学图像信息与控制分会委员、MICCAI 2022 communication chair等多个学术兼职；担任Cyborg and Bionic Systems和iRadiology期刊青年编委。



郑钰辉

郑钰辉，青海师范大学教授、博导，国家级青年人才。现为计算机学院院长、藏语智能全国重点实验室副主任、藏文信息处理教育部重点实验室主任；中国图像图形学会理事、中国人工智能学会青工委常务委员。围绕社会与空间感知计算，主持国家自然科学基金重点项目、重大研究计划课题、国际交流与合作项目等多项国家级课题。已在计算机视觉领域国际顶刊\顶会上发表学术论文百余篇，授权国家发明专利70余项。曾获2025年吴文俊人工智能自然科学奖一等奖、2023年度江苏省高校科学技术奖（自然科学类）一等奖、2022年度江苏省科学技术奖二等奖、2024年度CCF科技成果奖二等奖、2025年度江苏省自然科学百篇优秀学术成果论文等奖励。

学术服务论坛

计算机视觉II

时间：8月4日下午14:00-17:30 地点：207会议室

内 容

主席：任文琦（中山大学）、王国庆（电子科技大学）
王文冠（浙江大学）、袁国武（云南大学）

指导专家：白 翔（华中科技大学）	韩军功（清华大学）
何良华（同济大学）	纪荣嵘（厦门大学）
马惠敏（北京科技大学）	徐 凯（中国科学院工业人工智能研究所）
青年学者：陈超超（浙江大学）	甘磊磊（浙江大学）
郭海云（中国科学院自动化研究所）	刘晋源（大连理工大学）
卢宪凯（山东大学）	王 倩（广东工业大学）
杨传广（中国科学院大学）	叶育森（西安电子科技大学）
张 敏（华东师范大学）	周仕佶（北京航空航天大学）
朱一凡（北京邮电大学）	朱智勤（重庆邮电大学）

○ 指导专家



白 翔

白翔，华中科技大学软件学院院长、教授、博士生导师。IEEE Fellow、IAPR Fellow。国家杰出青年基金获得者、优秀青年基金获得者、“万人计划”青年拔尖人才。兼任机器视觉与智能系统湖北省工程研究中心主任、国家防伪工程中心副主任，国际期刊 Pattern Recognition 副主编(A-EIC)。长期从事人工智能、计算机视觉、模式识别等方面研究，在 Nature Machine Intell.、IEEE TPAMI、CVPR 等国际一流期刊和国际会议发表论文 150 余篇，谷歌学术总引用 5 万余次，H 指数 113，连续 11 年入选 Elsevier 中国高被引学者。担任国际顶级期刊 IEEE TPAMI 编委，顶级会议 CVPR、ICCV、ECCV、AAAI、IJCAI、NeurIPS 的领域主席，国际文档分析与识别会议 ICDAR 2025 大会主席。曾获 ACL 2024 最佳论文奖（Best Paper Award）、ICCV 2025 最佳论文候选（Finalist for Best Paper Award）2019 年国际模式识别协会青年学者奖（IAPR/ICDAR Young Investigator Award）、2019 年美国人工智能协会杰出资深委员奖（AAAI Outstanding SPC Award）、2024 年湖北省青年科技创新奖、2023 年湖北省自然科学一等奖（第一完成人）、2021 年中国图像图形学会自然科学一等奖（第一完成人）和 2021 年全国科技系统抗击新冠肺炎疫情先进个人。



韩军功

韩军功，清华大学自动化系，兴华讲席教授（长聘），国家海外高层次人才，国家信息中心‘类脑智能技术’交叉创新群体负责人。2024年9月回国，回国前任英国谢菲尔德大学（Top 100）计算机学院讲席教授。研究方向涵盖多模态感知、智能医疗、类脑机器学习与具身智能，推动3项人工智能技术实现产业化，音视频检索成果获CSI创新奖并助力企业斩获艾美奖。论文被引4.2万次，一作单篇最高被引超2100次，连续入选全球前2%科学家及0.05%高被引学者，CSRankings人工智能方向英国学者（曾）排名第六（谢菲尔德大学第一）。现任多本IEEE Transactions副主编，主持2项国家自然科学基金重点项目，BMVC 2025大会主席，国际模式识别学会（IAPR）会士。



何良华

何良华，同济大学计算机科学与技术学院副院长，教育部“长江学者奖励计划”特岗学者，教授，博士生导师。主要研究方向有人工智能与模式识别、认知计算，包括医学图像分析、基于核磁共振成像的静息态脑功能网络等。在IEEE TNNLS、IEEE TSMCS、IEEE TIP、CVPR、ICCV、AAAI等知名学术期刊及会议累计发表论文100多篇，授权/申请发明专利26项，软件著作权7项、申请行业标准3项。主持多项国家级及省部级基金项目，其中主要包括国家自然科学基金、教育部重点项目、上海市基础研究重点项目。以第一完成人获得上海市科技进步一等奖1项、教育部科技进步二等奖1项，以主要完成人获得教育部科技进步一等奖2项、二等奖1项、自然科学二等奖1项、上海市技术发明一等奖1项。



纪荣嵘

纪荣嵘，厦门大学教授、博士生导师，厦门大学校长助理、人工智能研究院负责人、多媒体可信感知与高效计算教育部重点实验室主任，主要研究方向为计算机视觉、多媒体技术和机器学习。研究侧重点是开发大数据驱动的计算机智能处理系统，进行视觉场景理解、视觉行为分析、多媒体内容理解、及视觉内容检索。当前的研究兴趣还包括面向视觉搜索的紧凑描述子、社会多媒体大数据分析、场景深度推理、人群行为理解、以及遥感图像分析等。百余篇科研论文发表于国际期刊 PAMI、IJCV、TIP 和国际会议 CVPR、ICCV、IJCAI、AAAI、ACM Multimedia 上，曾获得 2007 年微软学者奖、2011 年 ACM Multimedia 最佳论文奖、2015 年黑龙江省自然科学二等奖，2016 年教育部技术发明一等奖，2018 年福建省科技进步一等奖，2020 年福建省科技进步一等奖，2022 年霍英东青年科学奖。连续两次担任国家科技创新 2030 “新一代人工智能”重大项目首席科学家（2022、2026），CVPR 2021-2022、ACM Multimedia 2018-2022 领域主席，《中国科学：技术科学》副编辑、Valse（视觉与学习青年学者研讨会）2017 大会主席、IEEE MMSP 2015、ACM ICMR 2014、IEEE VCIP 2014、ACM MMM 2015、IEEE ISM 2015 等国际会议本地专题领域主席、AAAI 2015、CVPR 2013、ICCV 2013、ACM Multimedia 2010-2015 等四十余个国际会议的程序委员会委员、IEEE Multimedia Communication 技术委员会 Visual Analysis and Content Management for Communications 专家组主席（2014 至今）。曾任中国图像图形学会青工委副主任、任中国图像图形学学会学术工委副主任、中国自动化学会模式识别与机器智能专业委员会副主任、中国自动化学会粒计算与多尺度分析专委会副主任、中国计算机学会计算机视觉专委会常务委员、中国计算机学会学术工委委员、教育部电子信息类教指委人工智能专业建设咨询委员会委员。2005 年于哈尔滨工程大学获得本科学位，2011 年于哈尔滨工业大学获得博士学位，2011 到 2013 年于美国哥伦比亚大学任博士后研究员。2013 年加盟厦门大学信息学院人工智能系，2014 年获国家优秀青年科学基金，2017 年国家级青年人才获得者，2020 年获国家杰出青年科学基金，2024 年入选享受国务院政府特殊津贴专家。



马惠敏

北京科技大学领军学者，二级教授，博导，中国图像图形学会会士，北京市“三八红旗奖章”获得者。担任计算机与通信工程学院副院长、人工智能研究院副院长、鼎新实验室主任。兼任中国图像图形学会副理事长/秘书长、中国女科技工作者协会理事、北京图像图形学学会副理事长。长期从事复杂环境无人系统视觉认知、人机混合智能研究。作为第一完成人获得北京市自然科学一等奖、吴文俊人工智能科技创新一等奖、中国图像图形学会技术发明一等奖、教育部技术发明奖二等奖等奖励，多次在国际最大的自动驾驶数据集（KITTI）评测中获得第一名，在驾驶员状态预测国际数据集（Brain4Cars）上获得最好的成绩。作为负责人承担了国家自然科学基金重点项目、国家重点专项等项目。作为通讯作者在TPAMI、IJCV、TIP、CVPR、NIPS等高水平期刊和会议上发表论文100余篇。



徐凯

中国科学院工业人工智能研究所研究员，曾任国防科学技术大学教授，2011年获国防科技大学计算机学院博士学位，曾为西蒙弗雷泽大学、普林斯顿大学访问学者。主要研究方向包括数据驱动的几何处理与建模、三维视觉、具身智能及数字孪生技术，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。主持国家自然科学基金青年科学基金A类项目、重点项目，发表ACM SIGGRAPH/Transactions on Graphics等CCF A类论文100余篇，其中计算机图形学顶级会议SIGGRAPH论文29篇。担任ACM Transactions on Graphics、IEEE Transactions on Visualization and Computer Graphics等国际期刊编委，兼任中国图像图形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。曾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2项（排名1和3）、军队科技进步二等奖，获国家优秀青年科学基金及湖南省杰出青年基金资助。

论坛主席



任文琦

任文琦，中山大网络空间安全学院教授/博导，信息技术教育部重点实验室副主任，从事图像处理、计算机视觉、人工智能安全领域的研究。在CCF-A类和中科院一区发表学术论文100余篇，Google学术引用26000余次，入选爱思唯尔中国高被引学者，ScholarGPS全球前0.05%顶尖科学家榜单。多次担任CVPR、ICCV、ICML等AI和CV领域重要学术会议的领域主席，担任IEEE Transactions on Image Processing编委(Associate Editor)。主持国家自然科学基金优秀青年基金、联合重点项目及华为、腾讯、美团等多项企业合作项目。获中国计算机学会优博论文奖、吴文俊人工智能优秀青年奖、中国电子学会自然科学一等奖（序2）、浙江省自然科学一等奖（序2）。



王国庆

王国庆, 国家级青年人才、全国新时代青年先锋、国家重点研发计划青年首席科学家, 电子科技大学计算机学院教授、博导。任四川省人工智能学院副院长、首批全国高校“双带头人”教师党支部工作室书记、四川省 AI 青年突击队队长。入选四川省和成都市人才计划、江苏省 U35 计划。研究领域包括场景智能体、具身智能, 目前在 IEEE TPAMI、IJCV、ICML、CVPR、UIST 等发表论文 100 余篇, 出版学术专著 2 部。获 ACM MM 最佳论文提名奖、中国图像图形学会青年科学家奖、一级学会科技进步一等奖、全球遥感类竞赛一等奖等, 指导学生获 CSIG 优硕。主持国家自然科学基金联合基金重点项目 (2 项)、国家重点研发计划项目 1 项、国防重大项目及课题等项目 10 余项。担任 ACM MM、AAAI、ICME 等会议领域主席和主题主席、Pattern Recognition 等 SCI 期刊客座编委。



王文冠

王文冠, 浙江大学人工智能学院百人计划研究员, 海外优青。主要研究方向为具身智能、神经-符号计算、AI4Science。在顶级期刊和会议 (如 Nat. Comm.、TPAMI、IJCV、ICLR、ICML、NeurIPS、CVPR、ICCV) 发表学术论文 120 多篇。谷歌学术引用 29, 000 余次。曾获青橙奖、ACM 中国新星奖、澳大利亚研究理事会 (Australian Research Council) 优秀青年基金 (Discovery Early Career Researcher Award, DECRA) 等奖励。



袁国武

袁国武, 云南大学信息学院教授/博导, 学科竞赛中心主任, 从事计算机视觉、图像处理、天文信息处理方面的研究。入选云南省“兴滇英才支持计划”产业创新人才、云南省“万人计划”青年拔尖人才、云南省高等学校骨干教师、云南省高层次人才研修访学计划等。获得云南省自然科学二等奖、全国宝钢优秀教师奖、谷歌 (Google) 青年教师奖教金。在中科院一区 and CCF 列表期刊会议等发表学术论文 40 余篇。主持国家自然科学基金项目 2 项, 云南省应用基础研究计划重点项目等 20 余项。

讲习班

生成式AI与智能创作

时间：8月5日下午14:00-17:30 地点：102会议室

内 容

嘉 宾

主席：邓 欣（北京航空航天大学）、盛 律（北京航空航天大学）
于 茜（北京航空航天大学）

生成式3D中的原生表征、结构化建模与动态推演

曹炎培（VAST）

视觉世界模型

李崇轩（中国人民大学）

智能设计生成

于 茜（北京航空航天大学）

论坛主席



邓 欣

北京航空航天大学电子信息工程学院教授，博士生导师，国家级青年人才。研究领域为多模态图像处理、可解释性神经网络与多媒体内容安全，在人工智能顶级期刊 TPAMI、图像处理顶级期刊 TIP、视频编码权威期刊 TCSVT，计算机视觉领域顶级会议 ICCV、CVPR、ECCV 等发表论文 70 余篇，TPAMI 论文 8 篇（一作 / 通讯 6 篇）。在科研获奖方面，获 2023 年度北京市自然科学一等奖（排 2）、第 13 届吴文俊人工智能自然科学一等奖（排 2）。在个人荣誉方面，获第 14 届吴文俊人工智能青年科技奖、2021 年度英国 Eryl Cadwallader Davies Prize、2021 年度中国图像图形学会石青云女科学家奖，入选 2025 年度北京市科技新星，23-25 连续三年入选斯坦福大学全球前 2% 顶尖科学家榜单，因成功揭榜“面向 AIGC 溯源证真的大容量数字水印技术”难题，荣获 2025 年度华为“火花奖”。



盛 律

盛律，北京航空航天大学教授，博导，入选国家级青年人才、智源学者、小米青年学者、斯坦福全球前 2% 顶尖科学家排行榜。主要研究方向为三维视觉、具身智能和多模态大模型。在 IEEE TPAMI/ACM TOG/IJCV 以及 CVPR/ICCV/ICLR/ECCV 等重要国际期刊和会议发表论文 80 余篇，Google Scholar 显示被引用数 9500 余次。现任 ACM Computing Surveys 编委，CVPR/ICLR/ECCV/ACM Multimedia/AAAI 等领域主席，以及多个领域顶会顶刊审稿人和程序委员。CCF 高级会员，任 CCF 和 CSIG 多个专委会执行委员，Valse 执行领域主席。主持或参与多项国家自然科学基金、科技部重点研发计划和省部级重点研发计划项目。



于茜

于茜, 北京航空航天大学“卓越百人”副教授、博士生导师, 兼任软件学院教学实验中心副主任, 入选第九届中国科协青年人才托举工程。研究方向为计算机视觉与生成式 AI, 当前聚焦 CAD、矢量图及草图的理解与可控生成。主持国家自然科学基金面上项目和青年项目、CCF- 百度松果科研基金项目等, 获评 CCF 产学合作基金优秀项目案例; 同时作为课题骨干参与两项新一代人工智能重大项目。已发表高水平学术论文 50 余篇, 谷歌学术引用逾 3400 次。研究成果曾荣获英国机器视觉大会 (BMVC) 最佳论文奖, 并获得 BBC 等国内外媒体广泛报道。现任 Pattern Recognition 副主编, 并担任 CVPR、ECCV、ACM Multimedia、NeurIPS 等计算机视觉与多媒体领域顶会的领域主席及审稿人。指导的学生曾入选北航博士研究生卓越学术基金项目及腾讯青云计划等。

论坛讲者



曹炎培

曹炎培, VAST 联合创始人兼首席科学家。清华大学计算机科学与技术系博士 (2009-2018), 师从胡事民院士, 专注计算机图形学与生成式 AI 研究。在学术与开源生态领域, 曾获 Pacific Graphics 2014 最佳论文奖、SIGGRAPH Asia 2024 及 CVMJ 最佳论文荣誉提名。由其主导的 threestudio、TripoSR 等知名 3D AI 开源项目累计 GitHub Star 数超 20000, 并获 2023 年度 CCF 优秀图形开源项目奖。曾任 4D 重建创业公司 OwlII 创始团队成员及 CTO (2019 年被快手收购)、快手 Y-tech 高级研究员、腾讯 AI Lab 与 PCG ARC Lab 专家研究员及三维方向负责人。在实时 4D 重建、人体捕捉及生成式 3D 大模型等领域, 完成了从底层算法架构创新到千万级产品全链路落地的多次实践。

报告题目: 生成式 3D 中的原生表征、结构化建模与动态推演

报告摘要: 近年来, 随着生成式人工智能的爆发, 三维几何表达与建模技术正经历一场底层范式的重构。传统的离散几何表征 (如多边形网格) 与现代神经网络隐式表达在生成式架构中发生了深度的碰撞与融合。本报告将梳理 3D 生成技术的发展脉络, 探讨如何打破生成速度、质量与工业管线可用性之间的“不可能三角”。课程将从底层三维原生表征出发, 进而介绍结构化重拓扑、语义解耦与自动化绑定等关键三维生成问题的解决路径, 并拓展至场景生成与交互式世界模型, 为下一代三维生成架构的设计提供参考。



李崇轩

李崇轩，中国人民大学高瓴人工智能学院副教授、博士生导师。致力于生成模型基础理论、建模范式、大规模训练策略和高效采样算法的研究，带领团队研制了扩散语言模型 LLaDA，开源模型下载量破 600 万次，多项成果部署于 DALL·E 2、Stable Diffusion、Vidu 等行业领先大模型。成果发表于 NeurIPS、ICML、ICLR、TPAMI、Nature Communications 等，谷歌学术引用 1.4 万余次。获机器学习领域顶级国际会议 ICLR 2022 杰出论文奖；北京市自然科学一等奖（排名 2）；吴文俊人工智能自然科学一等奖（排名 5）。主持国家自然科学基金青年基金 B 类、重大研究计划培育项目等，入选智源学者、吴文俊优秀青年奖、北京市科技新星。担任 IEEE TPAMI/TMLR/MLJ 编委和 ICLR/ICML/NeurIPS 等国际会议领域主席；作为主编出版《大模型十讲》教材，指导博士生入选国家自然科学基金青年学生基础研究项目、字节跳动奖学金计划等。

报告题目：视觉世界模型

报告摘要：本次讲习班将梳理视觉世界模型主流技术路径，从生成式建模的角度探讨相关方法的主要技术瓶颈，如实时交互、几何一致性等，以此为线索回顾流式生成与蒸馏方法，基于动作控制的视频生成、三维物体生成、三维场景生成和四维场景生成的近期进展。



于茜

于茜，北京航空航天大学“卓越百人”副教授、博士生导师，兼任软件学院教学实验中心副主任，入选第九届中国科协青年人才托举工程。研究方向为计算机视觉与生成式 AI，当前聚焦 CAD、矢量图及草图的理解与可控生成。主持国家自然科学基金面上项目和青年项目、CCF- 百度松果科研基金项目等，获评 CCF 产学合作基金优秀项目案例；同时作为课题骨干参与两项新一代人工智能重大项目。已发表高水平学术论文 50 余篇，谷歌学术引用逾 3400 次。研究成果曾荣获英国机器视觉大会 (BMVC) 最佳论文奖，并获得 BBC 等国内外媒体广泛报道。现任 Pattern Recognition 副主编，并担任 CVPR、ECCV、ACM Multimedia、NeurIPS 等计算机视觉与多媒体领域顶会的领域主席及审稿人。指导的学生曾入选北航博士研究生卓越学术基金项目及腾讯青云计划等。

报告题目：智能设计生成

报告摘要：随着数字设计范式不断向智能化演进，具备精确编辑能力与无损缩放特性的结构化设计资产（如矢量图形与参数化 CAD 模型）的应用需求日益增长。然而，此类资产的传统创建过程高度依赖人工操作与专业软件，存在技术门槛高、制作周期长等瓶颈，难以满足自动化与规模化的生产需求。本报告将立足于该类数据固有的“视觉 - 代码”双模态特性，系统梳理矢量图与 CAD 可控生成领域的最新研究进展，并探讨其未来的发展趋势。

分论坛 图像复原与增强

时间：8月5日下午14:00-17:30 地点：101会议室

内 容

嘉 宾

主席：潘金山（南京理工大学）、顾舒航（电子科技大学）
钱文华（云南大学）

多退化图像复原技术探索

白慧慧（北京交通大学）

面向人机混合的生成式压缩方法

徐 迈（北京航空航天大学）

低质量视觉内容高效增强与分析

傅雪阳（中国科学技术大学）

从生成式影像到 Agent 影像：移动摄影的新范式

李 博（vivo）

面向真实场景的高效图像复原

金 枝（中山大学）

论坛主席



潘金山

潘金山，南京理工大学计算机科学与工程学院教授、博士生导师；主要从事图像视频复原与增强等相关底层视觉问题的研究；发表CCF A类论文110余篇，谷歌学术引用2万余次；研究工作曾获国家优秀青年科学基金、联合基金重点项目等的资助；目前担任IEEE TPAMI等期刊的编委以及CVPR等国际会议的领域主席。



顾舒航

顾舒航，电子科技大学教授，国家级青年人才。长期围绕视觉先验建模方法开展视觉信号增强、压缩与生成应用研究。在国际顶级会议和期刊(T-PAMI, IJCV, CVPR, ICCV, NIPS等)发表论文50余篇，谷歌学术的总引用次数20000余次。近年来，主持国家自然科学基金优秀青年基金(海外)项目、国家自然科学基金面上项目、四川省创新群体项目以及多项横向课题。曾获得中国人工智能学会吴文俊人工智能奖自然科学一等奖，北京市科学技术奖自然科学一等奖。



钱文华

钱文华，云南大学信息学院教授、博士生导师，本科生院副院长；主要从事计算机视觉、图形图像处理、文化计算等方面的研究；云南省“兴滇英才支持计划”产业创新人才，云南省云岭系列“万人计划”青年拔尖人才，云南省“视觉与文化计算”创新团队带头人，云南省中青年学术技术带头人；目前担任中国图象图形学会理事、《中国图象图形学报》等期刊编委。

论坛讲者



白慧慧

白慧慧，北京交通大学教授、博士生导师，国家级高层次人才。主要研究方向是图像视频编码和增强等。已发表包括IEEE汇刊及CCF A类会议论文100余篇。获国际专利授权3项、国家发明专利授权20余项。主持国家自然科学基金重点项目/企业联合基金项目、北京市自然科学基金项目等。获北京市自然科学一等奖、北京市科技进步二等奖、中国电子学会青年科学家奖、中国产学研合作创新成果奖二等奖、中国人工智能学会教学成果一等奖、中国图像图形学会教学成果二等奖等。

报告题目：多退化图像复原技术探索

报告摘要：图像复原技术广泛落地于手机成像、图像编辑、航空航拍、户外监控等多元现实场景。本报告围绕多退化图像复原展开系列研究，核心成果包含四类针对性算法：互补图像盲超分方案，处理实拍场景未知模糊退化；轻量化超分网络，平衡重建精度与边缘终端部署需求；基于渐进重建思路设计的图像修复模型，修复大面积孔洞遮挡失真；创新循环提示学习框架，仅用单一模型同步消除雨、雾、积雪等多种天气复合退化。在此基础上，进一步拓展事件视频超分、无人机遥感去雾等细分行业复原方案。大量公开基准数据集与实景测试结果表明，所提各类算法在像素指标与人眼感知层面均优于同期主流方法，构建起适配静态图像、动态视频、航拍影像、户外实拍的多退化复原技术体系，为各类复杂失真图像清晰化任务提供多元化可行方案。



徐 迈

徐迈，国家级领军人才特聘教授，北航电子信息工程学院副院长、中国图像图形学会青工委副主任。以第一/通讯作者在IJCV、IEEE TPAMI、等国际期刊及C会议上发表论文200余篇，多篇论文入选ESI高被引论文/热点论文。获IEEE/ACM国际会议/期刊最佳论文奖6项、提名奖2项。担任IEEE TIP编委、IEEE TMM编委、IEEE J-STSP的首席客座编委，连续两年获IEEE TMM杰出编委奖。作为负责人承担科研项目20余项，包括国家自然科学基金首批原创探索、重点、优青、面上、青年等项目以及京市杰青等项目。研究成果获北京市自然科学一等奖（序1）、教育部技术发明一等奖（序1）、中国人工智能学会自然科学一等奖（序1）、24届中国科协求是杰出青年奖。

报告题目：面向人机混合的生成式压缩方法

报告摘要：在大数据时代，网络视频与图像数据规模呈爆炸式增长，给存储、传输与智能分析带来了严峻挑战。人类能够借助视觉语义理解、上下文推理以及跨模态信息补全，实现远超像素级重建的高效感知压缩。近年来，生成式模型的快速发展为面向语义的压缩方法提供了新的技术路径，使压缩目标从“信号保真”逐步拓展到“语义保持”与“任务适配”。本报告首先介绍生成式模型与语义压缩的最新发展趋势，重点分析语义表征能力提升及其在压缩任务中的作用。在此基础上，进一步探讨面向人类感知与机器视觉任务协同优化的混合编码模型，通过语义表征建模、任务相关特征压缩以及特定熵编码模块设计，探索大模型下图像压缩失真对理解的影响，实现对视觉语义信息的高效表达与压缩，为新一代智能图像与视频压缩技术提供思路。



傅雪阳

傅雪阳，中国科学技术大学特任教授、博导，国家优青获得者，IEEE Senior Member，厦门大学与美国哥伦比亚大学联合培养博士，从事图像视频处理与分析、计算机视觉、人工智能等方面的研究。发表 IEEE 汇刊和 CCF-A 类国际会议长文 100 余篇，10 篇入选 ESI 高被引论文。论文谷歌学术引用 1 万 6 千余次，4 篇一作 / 共一论文单引超过 1 千次。近四年连续入选爱思唯尔中国高被引学者和全球前 2% 科学家榜单。获安徽省自然科学一等奖、CSIG 青年科学家奖、《Signal Processing》期刊唯一最佳论文奖等学术奖励。担任 IEEE TNNLS 副编辑、MIR 青年编委、《中国图象图形学报》专题编委等，并多次担任 ICLR、NeurIPS、AAAI、PRCV 等国内外学术会议的领域主席。主持基金委优青、面上、青年基金等多个科研项目。担任 CSIG 青年工作委员会和多媒体专委会委员、CAAI 青年工作委员会委员、CSF 视觉计算与仿真专委会委员等。获 CVPR、ICCV 等国际会议竞赛冠军 7 项。

报告题目：低质量视觉内容高效增强与分析

报告摘要：在真实应用场景下，视觉数据在采集与传输过程中常受多种环境因素干扰，导致复杂的质量退化现象，严重制约了后续视觉系统的性能表现。本报告将介绍团队近期在低质量视觉内容增强与分析方面的探索，主要包括鲁棒视觉表征、轻量架构设计及跨域泛化学习等。通过构建高效增强与分析模型，旨在提升视觉系统在复杂降质环境下的感知精度和适应能力。



李 博

李博，博士毕业于南京大学，为南京大学拔尖计划首批成员、南京大学及江苏省优秀博士，入选上海明珠人才计划。长期深耕人工智能领域，研究方向覆盖计算机视觉、多模态大模型、智能体系统与 AIGC 等前沿方向，兼具扎实的基础研究能力与突出的工程转化能力。学术成果丰硕，累计发表高水平论文 80 余篇，其中 CCF A 类会议与 SCI 一区顶刊论文 60 余篇，多篇成果发表于 CVPR、ICCV、ECCV、NeurIPS、ICLR、ICML、AAAI、TPAMI、TIP、TMM 等国际顶会及期刊；曾获 CVPR 2026 Best Paper Finalist；学术服务方面，担任 NeurIPS SAC、ICLR AC，长期在国际顶级学术会议承担领域主席工作。产业实践经验丰富，曾任职于腾讯 YouTu Lab 并获评技术大咖；现任 vivo 影像算法总监，全面负责 vivo Blueimage Lab 蓝图实验室，同时担任 vivo 影像产学研及高校合作负责人，统筹校企联合科研、学术交流、联合人才培养等协同创新工作；主导大模型与智能体方向核心技术研发与系统落地，牵头负责 vivo 端侧垂类 Agent 研发及端侧 AIGC 画质大模型研发落地，全面支撑 Blueimage 高端影像品牌构建。累计申请发明专利 40 余项，其中授权专利超过 20 项。

报告题目：从生成式影像到 Agent 影像：移动摄影的新范式

报告摘要：移动摄影正在进入一个新的阶段。过去，计算摄影主要围绕成像链路中的增强、复原、融合与渲染展开；今天，生成式模型开始直接参与影像内容的理解、重建、编辑和创作。围绕这一变化，本报告将结合我在生成式影像与移动计算摄影方向的近期研究，讨论移动摄影如何从单点算法能力走向可交互、可控、可优化的智能影像系统。



金 枝

金枝，中山大学教授，博士生导师，广东省珠江青年学者，深圳市优青，中山大学“逸仙学者”，IEEE Senior Member,中国图像图形学会CSIG 交通视频专委会副秘书长、青工委副秘书长、多媒体专委会委员，主要从事恶劣环境视觉感知增强、空间智能感知等方面的研究和应用。在图像处理及计算机视觉领域顶级期刊及会议发表论文94篇。作为第一发明人已授权国家发明专利17项，已授权软件著作权10项。主持国家自然科学基金面上、青年、中德/中英国际合作等各类项目18项。带领学生连续三年（2024-2026）在CVPR NTIRE挑战赛上获全球冠军。

报告题目：面向真实场景的高效图像复原

报告摘要：图像复原技术在智能终端和自动驾驶等领域至关重要。然而，传统方法依赖理想合成数据，面对退化复杂的真实场景时泛化性差，且深度模型计算量大，难以在边缘端高效部署。针对上述挑战，本报告聚焦于“面向真实场景的高效图像复原”，从两个维度展开：一是真实退化建模，通过物理仿真或非监督学习提升模型在未知环境下的鲁棒性；二是高效架构设计，通过轻量化网络与算子优化，在保证复原质量的同时大幅降低计算延迟。本报告为推动图像复原技术从理论走向工业级落地提供了高效的解决方案。

分论坛

本地特色论坛（小语种OCR）

时间：8月5日下午14:00-17:30 地点：201会议室

内 容

嘉 宾

主席：毛存礼（昆明理工大学）、杨雅婷（中国科学院新疆理化技术研究所）

复杂场景东南亚小语种文字识别技术

高盛祥（昆明理工大学）

让大模型说好每一种语言：大模型多语言泛化与迁移探索

肖 桐（东北大学）

蒙古文古籍识别关键技术研究与应用推广

飞 龙（内蒙古大学）

语言、图像与艺术的交汇：
多模态智能驱动的民族语言与文化计算

何向真（西北民族大学）

多语言智能：低资源场景的落地实践

杨雅婷（中国科学院新疆理化技术研究所）

论坛主席



毛存礼

毛存礼，教授，博导，云南省“兴滇英才支持计划”云岭学者，云南省中青年学术和技术带头人，南亚东南亚语言语音信息处理教育部工程研究中心主任。CCF 自然语言处理专委会委员，CIPS 机器翻译与多语言信息处理专委会委员，CSIG 文档图像分析与识别专委会。获得云南省科技进步特等奖，云南省自然科学一等奖，云南省科技进步一等奖，云南省有突出贡献优秀专业技术人员三等奖等省部级奖励各一项。主持国家重点研发计划、国家自然科学基金联合基金重点、地区，云南省重大科技专项计划等国家及省部级项目，以及企业委托研发项目等 10 余项。以第一作者 / 通讯在 EMNLP、TALLIP、软件学报等国内外权威会议及期刊发表论文 60 余篇，以第一发明人获得国家授权发明专利 20 余项、登记软件著作权 20 余项。主要研究方向为自然语言处理与机器翻译、语音识别与合成、OCR。



杨雅婷

杨雅婷，中国科学院大学博士研究生导师、研究员，中国科学院新疆理化技术研究所多语种信息技术研究室副主任，新疆民族语音语言信息处理重点实验室副主任。主持国家自然科学基金、国家重点研发计划、新疆维吾尔自治区重大科技专项等国家及省部级项目 22 项，入选新疆高层次引进人才、“天山英才”青年科技拔尖人才等 9 项人才计划。在国内外重要学术刊物和会议发表学术论文 110 余篇，组建并带领了一支扎根边疆长期从事多语言智能信息处理的科研团队，科研成果获得省部级科技进步一等奖 2 项，数字中国创新大赛全国总决赛二等奖 1 项。

论坛讲者



高盛祥

高盛祥，博士，昆明理工大学教授、博导，云南省人工智能重点实验室常务副主任。国家网信创新人才，云南省“兴滇英才支持计划”产业技术领军人才、云南省中青年学术和技术带头人、昆明理工大学“明德卓越”学者、云南省委办公厅优秀智库专家。中国中文信息学会信息检索专委会委员、中国人工智能学会语言智能专委会委员、中国人工智能学会机器学习专委会委员、中国计算机学会昆明分部执行委员会委员。

长期从事自然语言处理、机器翻译、信息检索、语音识别与合成等人工智能方向研究，主持国家自然科学基金-联合基金重点项目、面上项目、地区基金项目 4 项，国家重点研发计划子课题 1 项，云南省重大科技专项、云南省基础研究重大项目等省部级科研项目 11 项，及其他企事业横向合作开发项目 15 项；发表论文 160 篇，SCI/EI 收录 79 篇；授权发明专利 68 项；获得吴文俊人工智能科技特等奖、云南省科技进步特等奖、钱伟长中文信息处理科技一等奖、云南省自然科学一等奖、三等奖各 1 项。

报告题目：复杂场景东南亚小语种文字识别技术

报告摘要：东南亚文字体系庞杂、书写形态多样，叠加低分辨率、复杂版面与多字体混排等现实干扰，主流 OCR 模型在该区域的识别性能严重受限。这一困境既源于高质量标注数据的匮乏，更深层地受制于模型对东南亚文字独特字形结构（如上标下标叠加、非连续笔画组合及语际字形差异）的有效表征能力不足。本报告首先介绍我们构建的涵盖越南语、缅甸语、老挝语、泰语等多语言真实场景图像数据集，为低资源 OCR 研究奠定数据基础；继而针对模型表征瓶颈，阐述我们提出的图像特征增强与语言知识融合的识别方法；最后展示该技术在多语言 OCR 平台中的落地成效与未来展望。



肖桐

肖桐，博士，东北大学教授、博士生导师，东北大学计算机学院人工智能系主任，东北大学自然语言处理实验室主任，小牛翻译(NiuTrans)联合创始人。2006—2009 年赴日本富士施乐、微软亚洲研究院访问学习，并于 2013—2014 年赴英国剑桥大学开展博士后研究。中国中文信息学会理事，入选全球前 2% 顶尖科学家榜单，获得中国中文信息学会“钱伟长中文信息处理科学技术奖”一等奖，中国人工智能学会“吴文俊人工智能科学技术奖”特等奖，教育部“高校计算机专业优秀教师奖励计划”等奖励。主要研究领域包括自然语言处理、机器翻译、语言模型等。撰写专著《机器翻译：基础与模型》、《Foundations of Large Language Models》。作为项目技术负责人，成功研发了 NiuTrans、NiuTensor 等开源系统，在 WMT、CCMT/CW-MT、NTCIR 等国内外评测中 30 余次获得冠军。

报告题目：让大模型说好每一种语言：大模型多语言泛化与迁移探索

报告摘要：大模型在英语任务上展现出强大的推理与理解能力，但这种能力高度以英语为中心，在多语言尤其是低资源场景下显著退化。这也引出一个值得思考的问题，多语言能力的鸿沟会随基座模型增强而自动弥合吗？它究竟是数据规模问题，还是能力如何在语言间泛化与迁移的问题？本报告将围绕我们近期的系列工作展开。我们发现，大模型的多语言能力并非随规模自然涌现，而是“局部化”地集中在特定层中，需被有针对性地激活与迁移。我们据此精准微调底层的“语言特异层”，以极低代价将英语的强能力迁移到数十种语言，并进一步延伸至视觉语言模型与强化学习场景。最后，我们将结合这些实践，展望面向更多语言、更多模态的统一多语言大模型。



飞龙

飞龙，现任内蒙古大学计算机学院、人工智能学院教授、博士生导师、蒙古文智能信息处理技术国家地方联合工程研究中心副主任、内蒙古自治区多语种人工智能技术重点实验室主任、内蒙古大学北疆民族语言文化交融与传播研究中心副主任。担任内蒙古蒙古语言文字信息化协会会长、中国电子学会计算机分会副主任等社会职务。入选内蒙古突出贡献专家、内蒙古“草原英才”、内蒙古“英才兴蒙”工程高层次四类人才，内蒙古杰出青年基金项目获得者，获内蒙古“五一劳动奖章”、内蒙古“最美职工”、内蒙古“青年五四奖章”等荣誉。获教育部高等学校科学研究优秀成果奖(科学技术)——科技进步二等奖(排1)、内蒙古科技进步一等奖(排1)、中国计算机学会 CCF 科学技术奖——科技进步杰出奖(排1)、第十二届中国技术市场协会金桥奖项目二等奖(排1)、王选新闻科技二等奖(排1)、中国计算机学会 CCF 科技成果奖——科技进步二等奖(排2)等；发表学术论文 140 余篇，其中在 TASLP、SIGIR、ACL、ICASSP 等国际顶级期刊和会议发表论文 40 多篇；授权发明专利 11 项；主持国家社科基金专项重大项目 1 项、内蒙古政府重大基础建设项目 1 项、国家自然科学基金项目 3 项、国家重点研发计划项目课题 1 项、内蒙古自然科学基金杰出青年项目 1 项等 30 余项科研项目。

报告题目：蒙古文古籍识别关键技术研究与应用推广

报告摘要：蒙古文古籍是中华民族珍贵的非物质文化遗产，完整记录了蒙古族的历史脉络、文化习俗与科技成就，是中华各民族交往交流交融的重要实物见证。当前，大量蒙古文古籍面临纸张老化、字迹褪色、虫蛀破损等自然损毁风险，而传统人工转录、数字化的方式存在效率低下、成本高昂、标准不统一且易对古籍造成二次损伤等问题，同时蒙古文手写古籍特有的连写特性、异体字繁多、书写风格差异大等特点，也给自动化识别带来了极大的技术挑战，严重制约了蒙古文古籍的保护、传承与活化利用。团队在蒙古文古籍图像预处理与破损修复、粘连字符精准分割、多风格手写蒙古文端到端识别、古籍文本语义纠错与结构化提取等方面进行了系统性研究，探索蒙古文古籍识别技术与大语言模型、知识图谱等前沿技术的融合发展路径，为推动中华优秀传统文化的创造性转化与创新性发展提供技术支撑。



何向真

何向真, 教授, 博士生导师, 现任中国民族信息技术研究院副院长, 甘肃省民族语言文化智能信息处理重点实验室主任, 语言与文化计算教育部重点实验室副主任, 中国计算机学会自然语言处理专委会执行委员、语音对话与听觉专委会执行委员, 中国中文信息学会自然语言生成与智能写作专委会委员、机器翻译与多语言信息处理专委会委员, 甘肃省电子商务专家。“全国专业技术人才先进集体”核心成员, 甘肃省“语言智能计算”优秀导师团队核心成员。主要研究领域为多语种自然语言处理、多模态人机交互等。

报告题目: 语言、图像与艺术的交汇: 多模态智能驱动的民族语言与文化计算

报告摘要: 本报告聚焦“语言、视觉与艺术”的跨模态交汇, 探讨多模态智能技术在民族文化计算中的前沿应用与发展范式。报告将围绕自然语言深度处理、高保真视觉建模以及传统艺术形态的数字重构展开, 致力于推动多模态文化内容的智能生成与沉浸式交互。通过精准的语义建模与虚拟交互载体构建, 全面提升文化传播的体验感; 借助生成式 AI 与跨模态交互技术, 实现传统文化艺术的数字化再现与创新表达。



杨雅婷

杨雅婷, 中国科学院大学博士研究生导师、研究员, 中国科学院新疆理化技术研究所多语种信息技术研究室副主任, 新疆民族语音语言信息处理重点实验室副主任。主持国家自然科学基金、国家重点研发计划、新疆维吾尔自治区重大科技专项等国家及省部级项目 22 项, 入选新疆高层次引进人才、“天山英才”青年科技拔尖人才等 9 项人才计划。在国内外重要学术刊物和会议发表学术论文 110 余篇, 组建并带领了一支扎根边疆长期从事多语言智能信息处理的科研团队, 科研成果获得省部级科技进步一等奖 2 项, 数字中国创新大赛全国总决赛二等奖 1 项。

报告题目: 多语言智能: 低资源场景的落地实践

报告摘要: 立足复杂场景下的多语言多模态自然语言理解与生成需求, 以丝绸之路经济带核心区及周边国家低资源语言为切入点, 在不同的技术发展阶段探索多语言智能的低资源场景落地实践。以多语言多模态联合表征优化机制、文化相似性的迁移学习策略以及低资源场景下参数高效微调方法为理论依据, 形成“融疆融路”特色多语言大模型, 并支撑细分领域的应用。从原子能力到产品矩阵, 探索多语言智能在低资源场景下的落地实践。

分论坛

医学影像处理

时间：8月5日下午14:00-17:30 地点：207会议室

内 容

嘉 宾

主席：王 妍（华东师范大学）、王连生（厦门大学）

赵一天（中国科学院宁波材料技术与工程研究所）

具身智能驱动的手术机器人软组织自主操作

窦 琪（香港中文大学）

面向原始数据域生成式表示的PET/CT成像重建

刘且根（南昌大学）

生成式人工智能赋能眼科疾病诊断

雷柏英（深圳大学）

微创手术导航场景感知关键技术与应用

范敬凡（北京理工大学）

Beyond Generalization: Learning Under Distribution Shift in the Foundation Model Era
从泛化到适应：基础模型走向真实世界的关键挑战

许言午（华南理工大学）

AI交叉医学的初步探索：多模态智能内窥镜系统

张鼎文（西北工业大学）

混合现实技术在外科手术中的应用研究

邵永航（云南师范大学）

论坛主席



王 妍

王妍，华东师范大学研究员，博士生导师，上海市多维度信息处理重点实验室副主任。国家级青年人才获得者，入选中国图像图形学会(CSIG)石青云女科学家、上海市海外高层次人才、上海科技青年35人引领计划、上海市启明星(A类)等。致力于研究多维成像技术与医学信号分析。在Nature Medicine, TPAMI等顶级期刊和会议上发表论文70余篇。在国内较早开展高光谱病理成像与医学分析研究，提出了系列多维度识别分析算法并将其应用于生物组织样本的检测分析。谷歌学术引用18,000余次。研究成果被美国国家公共广播电台(NPR)报道，应用在JHMI重大医学项目；光谱成像等分析技术的成果已获得中华人民共和国医疗器械注册证；获得2023中国产学研合作创新成果一等奖、2024上海市优秀发明创新金奖、2025 MICCAI最佳数字病理理论奖；指导研究生获2025挑战杯国家二等奖。担任国际会议CVPR、MICCAI领域主席、ICME、PRCV资深领域主席、CISP-BMEI24大会副主席和多个SCI期刊的编委，担任第八届全国成像光谱对地观测学术研讨会青年论坛主席。



王连生

王连生，现为厦门大学信息学院教授，医学院双聘教授，博士生导师，数字福建健康医疗大数据研究所副所长，福建省医学会放射学分会AI学组副组长，海医会智能医学影像与信息化专委会副主任，厦门大学医学院医学人工智能研究院负责人，IEEE Transactions on Medical Imaging的Associate Editor, MICS主席。长期从事医学影像处理研究，主持和参与多项科研项目，包括国家自然科学基金仪器专项、科技部科技创新2030重大项目、国家卫健委临床研究项目、国家重点研发项目、国家自然科学基金面上和青年项目等，发表包括Nature Machine Intelligence、Nature Communications、Cell Reports Methods、Cell Patterns、人工智能顶会CVPR/AAAI等相关研究论文100余篇，获得腾讯犀牛鸟科研奖、CSPE Young Investigator、福建省科技进步一等奖（排1）、福建省科技进步二等奖（排2）、2023年厦门大学田昭武交叉学科一等奖，带领团队先后11次在国际医学影像比赛中获得冠军。



赵一天

赵一天，任中国科学院宁波材料所研究员，博导，所属先进诊疗实验室副主任。主要从事人工智能 + 眼科医学影像分析的研究，近几年聚焦于眼科多模态图像，开展了眼、脑、心脏等相关疾病的智能诊断算法研究及设备研发。先后获国家优秀青年基金、科技部重点研发计划青年科学家、浙江省杰出青年基金、中国科协青年人才托举工程的支持。在Nature Machine Intelligence、npj Digital Medicine、IEEE PAMI、IJCV、IEEE-TMI、MedIA、CVPR、MICCAI等行业顶级国际期刊和会议发表论文近200篇，学术总引超12000次，单篇最高引用次数超2700次。现兼任IEEE-TIP、IEEE-TMI、Medical Physics等期刊编委/副编。成果获浙江省科技进步二等奖、宁波市自然科学一等奖及2025年“中国眼科十大原创进展”，并有2项成果获批医疗器械注册证。

论坛讲者



竇琪

竇琪，香港中文大学计算机科学与工程系副教授，博导，国家级青年人才。研究方向聚焦于人工智能和机器人前沿技术在医疗健康领域应用的交叉学科研究，涵盖智能手术机器人，智慧医疗，安全具身智能等。获得香港中文大学青年学者研究成就奖（2025）和模范教学奖（2022）。竇教授现担任港中文智能医学与扩展现实研究所副所长，天石机器人研究所研究员，InnoHK医疗机械人创新技术中心研究员，InnoHK香港物流机械人研究中心副总监。曾获国家教育部自然科学二等奖，IEEE EMBS Early Career Achievement Award（全球每年1人），何享健青年科学家，香港科学会青年科学家。在人工智能和机器人领域国际期刊和会议上发表论文百余篇，谷歌引用4万余次，H-index 80。代表性论文曾获MICCAI 2022青年科学家最佳论文奖，IJCARs-MICCAI 2021和2022最佳论文奖，IEEE ICRA 2021医疗机器人最佳论文奖，IEEE TBME 2021 Prize Paper Award, MedIA-MICCAI 2017最佳论文奖等。担任多个国际期刊副编，国际会议NeurIPS, ICML, ICRA, CVPR领域主席，MICCAI 2026大会联席主席。

报告题目：具身智能驱动的手术机器人软组织自主操作

报告摘要：随着我国人口老龄化问题日益凸显，社会对就医的需求激增，传统服务模式面临挑战。具身智能技术的快速发展有望为医疗健康领域带来新的解决方案，尽管目前处于很早期阶段，但应用前景广阔。本报告将介绍团队围绕具身智能创新技术，在手术机器人应用场景下的研究进展。面向腔镜手术机器人自主操作，介绍团队全栈自研的具身智能模拟器 SurRoL，以及使用该模拟器实现的感知决策 VPPV 框架，完成可泛化的软组织自主操作，例如：组织牵引，血管夹闭。此外，本报告还将尝试探讨医疗具身领域当前面临的技术挑战与未来趋势，分享我们在科研中的体会与思考。



刘旦根

刘旦根，南昌大学教授、信息工程学院执行院长。主要研究集中于智能成像系统与视觉显示的系统搭建、信号处理和算法实现，合作发表 IEEE Trans 和成像与视觉显示权威期刊论文 100 余篇。为 IEEE 和 CCF 等数十个国际国内学术组织的专委会委员，IEEE TMI、电子与信息学报等多个期刊的编委。主持国家自然科学基金优青和联合重点项目、企业委托项目等 30 余项。获江西省自然科学奖一等奖、吴文俊人工智能技术发明一等奖等科技奖励 6 项。指导学生获会议最佳论文奖和中国国际互联网 + 竞赛金奖等 30 余项。

报告题目：面向原始数据域生成式表示的PET/CT成像重建

报告摘要：医学成像的观测域数据（如PET/CT弦图等）承载了完整的物理采样信息，其分布特性与图像域数据存在本质差异，往往呈现非均匀和非规则性。面向原始观测域数据开展建模与生成，提供了一条更符合“第一性原理”的路径：直接在观测空间刻画采样信号与噪声共同塑造的分布规律，使生成模型能够从源头学习“信号如何被获取”。本报告聚焦原始域数据的结构特性与生成式学习范式，分享两大研究前沿：一是挖掘数据特性并协同生成模型，将复杂的不规则观测分布映射为多个可学习可泛化的表征空间，促进快速高分辨成像；二是引入生成式基础模型构建通用原始域先验学习范式，实现跨任务一致的生成与推断能力。共同为高可信、可迁移的智能成像提供新的理论框架与技术路径。



雷柏英

雷柏英, 深圳大学特聘教授、博导、国家级青年人才, 主要研究方向为医学图像处理和人工智能。在 IEEE TPAMI 等以第一 / 通讯作者发表 SCI 论文 100 余篇 (含 10 余篇 ESI 高被引 / 热点论文)。谷歌学术总引用 1.6 万余次, H 指数 67。主持重点研发计划课题、国家自然科学基金重点等项目。现任 IEEE TCYB、TNNLS 等 SCI 期刊编委。现为 IEEE BISP 等技术委员会委员、IEEE EMBS 杰出讲师、IEEE 广州分部 WIE 主席, MICCAI 领域主席 (2021-2026) 等。连续入选“全球前 2% 顶尖科学家”终身影响力榜单, 获 CSIG 石青云女科学家奖、广东省图像图形学会一等奖、中国发明协会二等奖和中国人工智能学会二等奖。

报告题目: 生成式人工智能赋能眼科疾病诊断

报告摘要: 视网膜作为人体唯一可以直接无创观测微血管循环与神经组织的部位, 可为多种眼部及全身疾病的早期发现提供重要线索。尽管具有这一观测优势, 眼底疾病仍是全球不可逆性致盲的主要原因之一, 其发病往往较为隐匿, 若错失早期干预窗口, 可能造成严重视功能损伤甚至致盲。近年来, 人工智能技术, 特别是深度学习在医学影像分析领域展现出较大的应用潜力, 为眼底疾病的自动化分析提供了新的思路。然而, 现有的眼底智能分析方法大多基于判别式人工智能范式, 其本质上是建立从图像到标签的映射, 这种机制决定了其性能高度依赖于理想化的标准数据, 即要求数据满足独立同分布且具备充分的高质量标注。这种对标准高质量数据的依赖与临床非标准化数据之间的内生矛盾, 使人工智能方法在应对数据标注稀缺、长尾分布、图像异质性以及模态缺失等真实临床挑战时面临明显瓶颈。针对上述临床与技术问题, 本文将生成式人工智能引入眼底影像分析领域。与判别式人工智能侧重于寻找决策边界不同, 生成式人工智能更关注学习眼底影像的底层数据分布, 有望为应对临床非标准化数据带来的挑战提供新的思路。



范敬凡

范敬凡, 北京理工大学光电学院教授、博士生导师, 入选国家优青、北京市高聚工程、北京市科技新星等人才计划。长期从事计算机视觉、增强现实手术导航相关领域研究, 主持国家重点研发计划项目 1 项、课题 1 项, 及国家自然科学基金项目 3 项, 在 IEEE TIP、IEEE TMI、MedIA 等期刊和会议上发表论文 100 余篇, 已参与完成多项医学图像辅助诊疗与手术导航系统的研发, 相关知识产权已申报 / 授权 60 余项国家发明专利, 成果转化获国家三类 / 二类医疗器械注册证 6 项。作为主要完成人获吴文俊人工智能科技进步一等奖、电子学会科技进步一等奖、中国专利优秀奖等奖项与荣誉。

报告题目: 微创手术导航场景感知关键技术与应用

报告摘要: 微创手术决策依赖术前三维医学图像、患者体位、手术器械位姿、术中二维内镜影像等多源信息因采集方式不同, 空间坐标不统一, 导致手术过程面临术前多模医学图像关联分析难、手术器械与病灶位置定不准、皮下组织器官在术中实时影像中不可视的临床难题。对此, 研究团队提出了由多模医学图像配准融合、高维空间位姿配准融合、术中影像透视融合组成的多源空间配准融合新方法, 为微创手术导航信息共融提供理论支撑, 取得了系列学术成果, 显著扩充了手术规划信息量、提高了空间定位精准度、增强了重要组织器官可视性。



许言午

许言午博士，华南理工大学长聘教授，前百度智慧医疗首席科学家，人工智能与数字经济广东省实验室（琶洲实验室）研究员，WHO 数字健康咨询委员会专家，新加坡眼科研究所客聘教授，IEEE 高级会员，中国医药教育协会数字影像与智能医疗专委会副主任委员，中国计算机学会(CCF)数字医学分会常委，中国医药教育协会智能眼科分会常委。发表了 160 余篇国际期刊及会议论文(含 30 多篇十分以上顶刊论文和 4 篇高被引论文)，谷歌引用 15000 余次，H 指数 52，累计主持包括国家“四大慢病”科技重大项目课题、国家自然科学基金面上、广州市人才团队等 4500 万纵向经费和 8400 万横向经费，并与两家广东企业建立联合创新中心赋能企业高质量发展。他从 0 到 1 负责获批全国首张多病种 AI 三类证（同时也是国内首张青光眼 AI 三类证），入选 2021 年“中国眼科学十大进展”，入围 2022 年工信部和国家药监局合办的第一届“AI 医疗器械创新任务揭榜”。他先后获聘公安部引智计划特聘专家、浙江省特聘专家、北京市特聘专家，连续 4 年入选斯坦福大学发布的“全球前 2% 顶尖科学家”榜单。

报告题目：Beyond Generalization: Learning Under Distribution Shift in the Foundation Model Era

从泛化到适应：基础模型走向真实世界的挑战

报告摘要：近年来，以视觉基础模型(Foundation Models)为代表的新一代人工智能技术推动了计算机视觉领域的快速发展，在图像分类、目标检测、语义分割以及多模态理解等任务上取得了突破性进展。然而，大量研究表明，在真实应用场景中，即使性能优异的基础模型也常常面临显著的性能退化问题。数据采集设备、成像条件、场景环境以及用户群体的变化都会导致训练数据与测试数据之间出现分布漂移(Distribution Shift)，从而严重影响模型的泛化能力和部署可靠性。围绕这一挑战，本报告将结合近年来基础模型与医学人工智能领域的最新研究进展，系统介绍从“泛化(Generalization)”到“适应(Adaptation)”的发展脉络，探讨真实世界环境下模型持续学习与持续演化的关键技术。报告将重点分享团队在测试时适应(Test-Time Adaptation)、持续测试时适应(Continual Test-Time Adaptation)以及参数高效在线适应(Parameter-Efficient Online Adaptation)等方向的研究成果，包括基于空间缓存机制的测试时适应方法、基于扩散模型的持续适应框架，以及面向基础模型的在线适配器设计等工作。这些研究突破了传统“训练一次、永久部署”的范式，使模型能够在无需源数据或大规模重新训练的条件下持续适应不断变化的真实环境。进一步地，报告将讨论基础模型时代分布漂移问题的新特征与新挑战，分析提示学习(Prompt Learning)、适配器学习(Adapter Learning)、生成模型辅助适应等技术路线的发展趋势，并展望面向医疗、工业视觉、自动驾驶以及机器人等复杂场景的持续适应智能系统。未来人工智能系统的核心能力不再只是泛化能力，而是在开放环境中持续感知变化、持续学习新知识并持续适应新任务的能力。



张鼎文

张鼎文，西北工业大学自动化学院教授、博导，国家优秀青年科学基金获得者、科睿唯安“全球高被引科学家”，2015赴美国卡耐基梅隆大学进行为期2年的访问研究，致力于建立面向开放环境下、具备动态学习能力的新一代计算机视觉学习框架。迄今为止，作为第一作者/通讯作者在领域内国际重要期刊及会议发表学术论文60余篇，其中包含T-PAMI,I-JCV,IEEE SPM,T-IP,CVPR,ICCV,Science China:Information Science等，曾入选中国博士后创新人才计划、AI华人青年学者榜单,获吴文俊人工智能优秀青年奖、2021 IEEE TCSVT最佳论文奖、中国图像图形学会优秀博士论文奖等奖励。担任中国图像图形学会青年工作委员会副秘书长、中国图像图形学会优博俱乐部副主席，任IEEE TMM、TCSVT、PR等刊物（客座）编辑。

报告题目：AI 交叉医学的初步探索：多模态智能内窥镜系统

报告摘要：新时代下医学影像数据的不断丰富与增长形成了医学影像大数据，对医学影像大数据的智能解译成为了提升医疗水平，科技惠医惠民的关键。本报告从多个层次与角度解读人工智能技术对医学影像及辅助诊断的赋能方式，并探讨如何突破传统定位，结合 AI 优势,推出新的医学成像设备,使 AI 下沉至成相前端,从而深刻推动医学技术的进步。



邵永航

邵永航，教授，博导，云南师范大学科技处处长，博士毕业于澳大利亚Deakin University智能信息系统研究所（IISRI），主持国家自然科学基金3项；主持云南省重点研发计划1项，课题2项、云南省军民融合等省部级项目8项；发表学术论文90余篇，其中SCI检索60余篇。荣获“兴滇英才支持计划”产业创新人才专项、青年人才专项、“云南省优秀青年基金”、“云南省青年五四奖章”、云师大“联大青年”、“南方教育奖”等称号。现任“云南省光电信息技术重点实验室”主任、云南省高校“光电探测与感知技术”工程研究中心、“云南省高校颜色与光电成像技术创新团队”、以及云南师范大学“中-澳视觉与AI联合实验室”负责人，现任中国光学学会颜色光学专业委员会委员、中国光学学会光电技术专委会委员、中国照明学会视觉与颜色专业委员会秘书等。申请人代表性成果已在IEEE Communications Surveys&Tutorials,IEEE TII,IEEE TMM,OE,AFM IEEE JBHI以及IEEE SMC等期刊发表。

报告题目：混合现实技术在外科手术中的应用研究

报告摘要：为解决传统二维医学影像存在的三维空间关系缺失、术中动态导航脱节等问题，混合现实技术通过将患者个体化的医学影像数据(CT/MRI)转化为高精度三维立体模型，并与真实手术场景、实时腔镜影像流动态融合，为外科医生提供直观的“透视”视野，显著提升手术精准度与安全性。本报告聚焦混合现实(MR)技术在外科手术中的应用和发展，从深度学习与空间配准算法出发，解析高保真性的手术场景重建和在动态手术环境中3D解剖模型实时配准导航的挑战与难点。最后，报告将分析讨论基于混合现实技术下手术导航的优势、远程沉浸式会诊系统的搭建、软组织实时形变建模的跨学科攻关路径和研究方向。

分论坛

AI For Science

时间：8月5日下午14:00-17:30 地点：302会议室

内 容

嘉 宾

主席：余肇飞（北京大学）、张敬林（山东大学）
欧阳万里（深圳河套学院）、张 奇（同济大学）

可解释医学信息智能计算研究进展	索津莉（清华大学）
大模型类脑认知机理解析研究	杜长德（中国科学院自动化研究所）
智能科研基础设施：释放青年科学家的AI for Science创新力	蒋 宁（北京九章云极科技股份有限公司）
知识与数据双驱动的植被4D建模	金哲依（北京大学）
AI4Materials：高端钎合金智能设计与制备过程控制思考与探索	黄科科（中南大学）
物理驱动的计算光学显微成像	周 游（南京大学）

论坛主席



余肇飞

余肇飞，北京大学人工智能研究院研究员、博士生导师，北京大学博雅青年学者，国家优秀青年基金获得者。主要研究方向为类脑计算、神经形态计算，担任类脑脉冲大模型北京市重点实验室副主任、中国图象图形学会类脑视觉专委会秘书长。在 Nature Biomedical Engineering、Nature Communications、Science Advance、IEEE Transaction 汇刊和 NeurIPS、ICML、CVPR 等顶级会议上发表论文 100 余篇，主持“脑科学与类脑研究”国家科技重大专项青年科学家项目（A 类）、国家自然科学基金联合重点、北京市科技新星等项目，担任 ICML、NeurIPS、ICLR 等会议领域主席，曾获中国自动化学会自然科学奖一等奖。



张敬林

张敬林，博士毕业于法国雷恩国立应用科学院，现任山东大学控制科学与工程学院教授，主要从事人工智能、海洋遥感以及工业大数据分析等研究，曾获得山东省杰青、泰山学者特聘教授以及全球前 2% 顶尖科学家等荣誉。担任教育部数字化教学工程中心专家委员、数字家庭网络国家工程研究中心技术委员会委员、中国人工智能学会组织工作委员会委员；发表 Science Bulletin、Nature Communication、IJCVC、TPAMI、TNNLS、TMM、CVPR、ICML、AAAI、KDD 等高水平会议和期刊论文 100 余篇，授权发明专利 30 余项，主持科技部重点研发计划项目、基金委联合基金重点项目在内的国家级课题 7 项，获得 CSIG 科技进步二等奖、CCF 技术发明二等奖以及广东省技术发明一等奖等多个省部学会奖励。



欧阳万里

欧阳万里，香港中文大学教授，深圳河套学院副院长，曾任悉尼大学电子信息工程学院研究主任。其团队在ImageNet和COCO竞赛多次获得第一。获IEEE TCSVT期刊最佳论文，ACL杰出论文，两篇文章入选paperdigest CVPR/ICCV最有影响力的文章。入选「人工智能全球2000位最具影响力学者榜」前100名学者。担任人工智能领域顶级期刊TPAMI, IJCV副编，曾任CCF-A类会议CVPR、NeurIPS、ICCV资深领域主席。其团队现已推出OWL·灵鉴，书生·风乌（气象）、书生·瞳真（裸眼3D）、书生·丰登（育种）、书生·翼飞（翼型设计）、书生·鉴原（化学）、书生·言普（谱学）等科学领域模型。



张 奇

张奇，同济大学计算机科学与技术学院教授，先后获得北京理工大学计算机科学与技术工学博士学位和悉尼科技大学分析学哲学博士学位。入选国家级人才（海外）、上海白玉兰海外高层次人才、上海市浦江人才计划，主持和参与国自然青年项目、科技部中重大专项青年科学家项目、科技部重点研发。主要研究方向为多模态信息处理，具体包括脑视觉编解码、时间序列分析、多模态模型压缩等，在人工智能和数据科学领域的顶级国际期刊和会议上发表论文100余篇。近年来担任AAAI、IJCAI、ICLR、NeurIPS、ICML、ACM MM、CVPR等会议程序委员或领域主席，并在多个国际期刊担任客座编辑和审稿人包括TPAMI、TKDE、TSMC、TOIS、Machine Learning、TIST等，并在The Web Conference、KDD等多个高水平会议上进行受邀讲座和专题讲座（Tutorial），担任CSIG青工委副秘书长、上海计算机学会计算机视觉专委会副秘书长、YOCSEF上海学术AC委员。

论坛讲者



索津莉

索津莉，清华大学自动化系，长聘副教授、博士生导师。近年来主要从事计算影像学理论与关键技术研究，包括高通量计算摄像与计算显微、单像素压缩成像、生物医学数据编码与传输等，在领域国际期刊与会议上发表高质量学术论文 100 余篇，包括顶级期刊 Nat. Photonics、Nat. Commun、Nat. Biomed. Eng、PNAS、IEEE TPAMI、IJCV、OSA Optica 等。先后承担国家自然科学基金项目（面上项目、重大仪器研制专项子课题、优青项目、重点项目）、科技部重点研发计划子课题、北京市自然科学基金重点项目等。关键技术获授权国家发明专利 30 余项、国际发明专利 3 项。相关成果获 2025 年度中国电子学会青年科学家奖、2023 年中国图像图形学会技术发明一等奖（第 1 完成人）、2016 年国家科技进步二等奖（第 3 完成人）、2015 年度电子学会科学技术奖一等奖（第 2 完成人）。

报告题目：可解释医学信息智能计算研究进展

报告摘要：医学决策高度依赖影像判读、时序监测与多源证据融合，但当前深度学习模型在临床落地中面临“黑箱”困境——医生难以信任缺乏依据的预测，患者亦无法理解诊断逻辑，尤其在像素级病灶边界判定、ICU 病情陡变前的因果归因以及跨指南知识检索与推理等场景中，可解释性的缺失直接制约了智能系统从实验室走向床旁。本报告将汇报课题组的三项工作：针对影像报告生成中线索不明确的问题，提出可解释的像素级病灶分割方法，为后续诊断及跨模态融合提供直接依据；引入因果推断构建 ICU 病情恶化早期预警模型，以区分相关性与因果性，进而输出具备可干预性的预警信号；此外，还面向医疗健康领域设计科研 Agent，构建从文献证据到数据验证的全链条可回溯逻辑链路。诚挚期待与与会专家展开深入交流，共同推动可解释医学信息计算技术发展。



杜长德

杜长德，中科院自动化所副研究员，围绕生物智能与人工智能的交叉前沿，通过认知神经科学与人工智能的深度融合，揭示大脑信息处理原理，并推动新一代人工智能的发展。在神经编解码、大模型机理解析、脑机融合智能等方面发表了 50 多篇高水平论文，包括 Nature Machine Intelligence、IEEE TPAMI、ICLR、ICML 等。主持国家自然科学基金面上项目。曾获得“小米青年学者”、IEEE ICME 最佳论文奖(亚军)，担任《The Innovation Informatics》青年编委，研究成果被《MIT Technology Review》报道。个人主页：<https://changdedu.github.io/>。

报告题目：大模型类脑认知机理解析研究

报告摘要：随着多模态大语言模型 (MLLM) 的兴起，一个关键问题随之而来：这些先进的 AI 模型是否发展出与人类相似的物体概念表征？我们借鉴认知心理学的经典“三选一”范式，利用大语言模型 (LLM) 和 MLLM 收集了数百万次关于上千种自然物体的相似性判断。通过对这些海量行为数据进行表征学习，我们解析出模型内部用于概念判断的 66 个核心维度。研究发现，这些维度不仅稳定、可泛化，并且具有高度的可解释性，涌现出了与人类极为相似的语义结构。将这些大模型的概念表征与人类大脑数据 (fMRI) 进行对比，发现在负责物体、场景和人脸识别的关键脑区 (如 EBA, PPA, FFA) 中存在高度一致性。



蒋宁

九章云极产品生态总监，毕业于中科院计算所。首个面向商用人工智能开发平台的标准——《人工智能开发平台系统功能要求》核心编写专家。具有人工智能与机器学习、企业级大规模分布式系统、移动嵌入式软件、云计算、大数据产品、智能体的架构、开发、技术服务和管理实施、培训、售前、产品生态、算力平台建设和运营经验。主导多门云计算课程开发，线上线下学员逾万人。

报告题目：智能科研基础设施：释放青年科学家的 AI for Science 创新力

报告摘要：AI for Science 要真正进入高效的科研生产环节，关键已不只是提出更好的算法，而是将模型智能、工程系统与学科实践组织成能够持续运行的科研体系。当前，大量科研时间仍消耗在环境搭建、算力配置、流程衔接和重复验证上，许多优秀成果也难以跨团队复制与推广。针对这些问题，我们尝试建设面向科研 Agent 的智能科研基础设施，以高阶 Token、面向 Agent 的智能算力为底座，以多智能体协作与工作流编排为中枢，协助 AI for Science 各领域的青年科学家，围绕物理、化学、生物、医学、材料、CV 等重点方向开展共创，帮助他们减少重复性工程投入，提升科研效率，聚焦真正重要的科学问题与领域方法，将科研探索沉淀为经得起验证、能够重复使用、便于传播并持续完善的开放科研能力。



金哲依

金哲依，北京大学城市与环境学院研究员、长聘副教授。曾先后在美国斯坦福大学、Atlas AI 公司、明尼苏达大学从事科研、教学与产品研发工作，担任美国国家人工智能中心 - 气候智慧农林系统研究所(AI-CLIMATE)新技术研发负责人，并获得美国国家科学基金会职业生涯发展奖(NSF CAREER Award)。主要研究方向为植被 3D 建模、数字孪生、知识引导机器学习，及其面向农业生产和复杂生态系统管理的落地应用，相关论文发表在 Science, 多个 Nature 子刊和 Remote Sensing of Environment 等杂志。担任 Earth's Future 和 Journal of Plant Ecology 副主编，《中国科学：生命科学》编委等。

报告题目：知识与数据双驱动的植被 4D 建模

报告摘要：植被 4D 建模是指重建植物踢的三维结构和时间动态，是构建植被数字孪生系统，支撑智慧农业和复杂生态系统管理的基础。针对现有植被重建方法难以兼顾 3D 结构视觉高保真与生态学准确性的瓶颈，本研究提出了知识与数据双驱动的植被 4D 建模框架。主要探索包括：基于 NeRF 和 3D GS 技术的作物重建和关键生理参数提取；引入生物物理模型，衔接点源测量与大尺度监测；基于幼苗形态，融合植物学知识推断中短期演化趋势，实现单株和场景时序生长的前向预测。本研究旨在提升大模型的植被生成能力，最终研发能够兼具时空一致性与生理生态合理性的植被 4D 建模智能体。



黄科科

黄科科，中南大学教授、博导、自动化学院副院长。入选长江学者特聘教授、国家万人计划青年拔尖人才、湖南省科技创新领军人才、省杰青等。长期从事工业智能与系统相关研究工作，在IEEE Trans.、IFAC旗舰期刊、CCF A类会议等发表高水平论文100余篇，授权中国/美国等发明专利30余项。先后主持国家重点研发计划项目、国家重大科技专项、基金委重大研究计划/面上等10余项。曾获国家科技进步二等奖（排2）、湖南省科技进步一等奖（排2）、中国有色金属学会优秀青年科技奖、IEEE TCSDM Young Professional Award等。兼任IEEE TASE、IEEE SMCM、IET CPS等期刊编委和中国有色金属学会自动化学术委员会副秘书长、中国人工智能学会智能融合专委会常务委员、中国自动化学会控制理论专委会/过程控制专委会委员、中国计算机学会计算机应用专委会执行委员等。

报告题目：AI4Materials：高端钛合金智能设计与制备过程控制思考与探索

报告摘要：钛合金是航空航天、导弹、核反应堆等战略装备的关键材料，被誉为“尖端金属”。然而，传统依赖经验试错的材料研发与制备过程控制范式难以满足多牌号、高性能等高端钛合金的需求。本报告立足高端钛合金高性能制备需求，分析其在材料设计、微观组织监测和制备过程精准调控方面面临的挑战性难题；然后，汇报 AI 在新材料研发、组织性能预测和制造过程优化等方面的研究进展，探讨 AI 提升材料研发效率、组织预测能力和过程调控水平的应用价值；最后，汇报研究团队面向高端钛合金智能设计、微观组织演化监测、物理场精准调控软件系统等研究探索，阐述 AI 赋能高端钛合金的发展思考。



周游

周游，南京大学医学院助理教授、博士生导师。2019 年获清华大学自动化系博士学位；博士期间前往美国康涅狄格大学生物医学工程系交流访学。主要聚焦计算光学显微与智能成像算法研究。主持国家自然科学基金、江苏省青年基金、江苏省青年科技人才托举工程等项目。一作 / 通信在 Laser Photonics Rev., Photonics Res., IEEE T Med. Imaging, ECCV, AAAI 等期刊和国际会议上发表论文 20 余篇。授权国家发明专利 7 项、美国专利 3 项。担任中国图像图形学会三维成像与显示专委会委员，以及 Chinese Optics Letters、BME Frontiers 等期刊青年编委。

报告题目：物理驱动的计算光学显微成像

报告摘要：人工智能技术已深度赋能生物医学成像，涵盖图像去噪、三维重构、超分辨率及多模态分割等关键任务。然而，实际成像场景常受限于高质量数据集稀缺和泛化能力不足的难题。针对上述瓶颈，本研究引入物理驱动的自监督学习范式。创新性地将光学衍射传播模型融入隐式神经表示（INR），构建衍射神经网络（DNF），实现光学显微场景下的高精度无透镜叠层成像；进一步提出基于光学传递函数（OTF）约束的椭球坐标编码INR方法，显著提升宽场荧光显微的三维体积重建质量，并将该框架拓展至心脏三维超声成像领域。此外，针对傅里叶光场成像（FLFM）模态，提出基于3D高斯泼溅技术的3D高斯自适应层析方法（3DGAT）；进一步融合时空约束与数字自适应光学（DAO）技术，发展出3D+t的4DGAT方法，最终实现动态场景下的荧光高保真四维成像。

分论坛 国产大模型的部署与优化

时间：8月5日下午14:00-17:30 地点：309会议室

目 录

嘉 宾

主席：杨智勇（中国科学院大学）、冯骁骋（哈尔滨工业大学）
秦 兵（哈尔滨工业大学）

大小模型协同知识推理

刘新旺（国防科技大学）

大模型压缩蒸馏技术在企业级落地的探索与实践

黄龙涛（阿里巴巴）

端侧具身智能系统

徐梦炜（北京邮电大学）

面向大模型训练推理的国产软硬件广谱适配

韩 旭（清华大学）

面向视频生成与3D理解重建基座模型的高效优化：
量化、稀疏化与缓存加速

杨传广（中国科学院计算技术研究所）

论坛主席



杨智勇

中国科学院大学副教授，博士生导师，研究方向聚焦可信、高效机器学习，在 CCF-A 类期刊及会议共发表论文 70 余篇，其中 TPAMI 20 余篇（一作 6 篇），ICML/NeurIPS 20 余篇（spotlight/oral 8 篇），担任 ICML/NeurIPS/ICLR 领域主席，JCR 一区期刊 Information Fusion/Pattern Recognition/Neural Networks 编委，获吴文俊青年科技奖、CCF 优秀博士学位论文激励计划（原 CCF 优博）、中国图像图形学会自然科学奖一等奖、吴文俊科技进步二等奖小米学者、首届百度 AI 全球华人新星百强、首届亚洲可信机器学习奖励金 6 项 CCF-A 类会议国际竞赛冠军。



冯骁骋

冯骁骋，哈尔滨工业大学计算学部社会计算与信息检索研究中心教授(入选校青年人才拔尖计划)、博导，国家高层次青年人才。人工智能学院副院长，中文信息处理黑龙江省重点实验室主任助理。研究兴趣包括自然语言处理、文本生成、机器翻译等。在 ACL、AAAI、NeurIPS、TKDE、TOIS 等 CCF A/B 类国际会议及期刊发表论文 70 余篇。据 Google 学术统计，论文引用数量超 17000 次，两篇论文入选 Google 学术 2020 年高被引榜单，一篇论文入选 Paper Digest EMNLP 2020 Top Ten 高引论文。担任 NIPS、ICML、AAAI、IJCAI、ACL 等国际会议程序委员会高级 / 普通成员；兼任鹏城实验室双聘学者、中国图像图形学会青工委委员、中国中文信息学会青年工作委员会副主任、2023-2024 中国计算机学会 YOCSEF 哈尔滨分论坛主席等。入选中国科协第六届青年人才托举工程，曾获一级学会中国中文信息学会优秀博士学位论文奖、教育部科技创新一等奖、吴文俊人工智能一等奖、黑龙江省科技进步一等奖、二等奖，2023 中国计算机学会自然语言处理专委会青年新锐奖和 2022 年世界人工智能大会 (WAIC)- 云帆奖。主持国家自然科学基金 B 类(原国家优秀青年基金)、科技创新 2030—新一代人工智能重大项目子课题一项、国家实验室重点项目课题一项、国家自然科学基金面上和青年项目各一项；黑龙江省重点研发项目一项、黑龙江省优秀青年基金一项，MSRA Collaborative Research Program；与华为、腾讯、科大讯飞、微软等国际一流互联网公司保持长期科研合作关系。



秦 兵

秦兵，哈尔滨工业大学长聘教授，社会计算与交互机器人研究中心主任。中国中文信息学会常务理事，中国中文信息学会情感计算专委会主任，科技部科技创新 2030—“新一代人工智能”重大项目管理专家组专家。研究方向：自然语言处理及大模型技术等。发表论文近 200 篇，累计被引 3 万余次。连续多年入选爱思唯尔高被引学者榜单，入选 2020 年福布斯中国科技女性榜及 2025 年全球顶尖 2% 科学家。获教育部科学研究优秀成果奖工程技术一等奖，黑龙江省科技进步一等奖、CCF 科技成果奖科技进步一等奖、钱伟长中文信息处理科学技术奖一等奖等奖项。

论坛讲者



刘新旺

刘新旺, 国防科技大学计算机学院教授、博导。主要研究方向为机器学习、数据挖掘等。国家杰青 (2023)、国家优青 (2019) 获得者。国家自然科学基金委重点项目、科技创新 2030 重大项目负责人, 基金委创新群体 A 类核心成员。共发表 CCF A 类期刊 / 会议论文 150 余篇, ESI 高被引论文 20 篇, 谷歌学术引用共 2.7 万余次, 连续 4 年入选全球 2% 顶尖科学奖榜单 (2022-2025)。担任 IEEE T-KDE, IEEE T-NNLS, IEEE T-CYB 等国际顶刊 AE, 以及 ICML, NeurIPS 等国际顶会领域主席。研究成果获中国计算机学会自然科学一等奖 (2025)、中国人工智能学会吴文俊自然科学一等奖 (2024)、北京市科技进步一等奖 (2024)、湖南省自然科学一等奖 (2014, 2021)、中国图像图形学会自然科学二等奖 (2024, 2025) 等。

报告题目: 大小模型协同知识推理

报告摘要: 基于既有背景知识、环境信息及历史行为推演出全新的、高价值知识, 是搜索推荐、金融风控、医疗诊断以及军事决策等领域的核心能力。随着大型语言模型 (Large Language Model, LLM) 的快速发展, 借助其强大的推理能力从复杂信息中挖掘潜在知识, 已成为当前研究的重点。然而, LLM 庞大的参数规模导致其在资源受限场景中难以直接部署与应用。针对这一挑战, 课题组提出了一套基于大小模型协同的知识推理算法体系, 包括面向轻量模型的自适应思考推理、结合跳跃思考与潜层思维链的高效推理机制以及以思维链为引导的结构化知识推理策略等。该系列方法在显著降低计算成本的同时, 全面提升了推理的精度与鲁棒性。



黄龙涛

黄龙涛, 阿里巴巴安全 AGI 实验室负责人, 研究领域涉及安全垂域大模型的构建与应用、大模型安全等。在 ACL、AAAI、IJCAI、WWW、SIGIR 等人工智能领域顶级会议及期刊发表论文 50 余篇, Google Scholar 累计引用 2000 余次, 曾获得 2022 年吴文俊人工智能科技进步二等奖、入选第二届中国科协优秀科技论文遴选计划、中科院青促会。

报告题目: 大模型压缩蒸馏技术在企业级落地的探索与实践

报告摘要: 大模型参数激增带来推理延迟和算力成本高企, 成为业务规模化落地的核心障碍。本报告从企业实战出发, 系统介绍我们在压缩蒸馏方向上的算法创新与应用实践。针对传统蒸馏在知识迁移中学习难度失配及自回归误差累积问题, 我们提出了自适应精细粒度度量与状态分布约束新范式, 显著提升蒸馏效率与训练收敛速度。在落地应用方面, 构建覆盖通用蒸馏到垂直领域适配的全链路生产流水线, 产出多个轻量级高效小模型, 已在高流量业务中规模化上线, 在业务效果持平前提下, 大幅节省了推理成本。



徐梦炜

徐梦炜，北京邮电大学计算机学院副教授，博士生导师。主要研究领域为端侧具身智能算法和系统软件，相关成果发表于 MobiCom/ATC/ASPLOS/ 软件学报等国内外重要会议和期刊，获 USENIX ATC 2024 最佳论文奖。主持国家自然科学基金 B/C 类、2030 重大项目课题等多个项目，入选中国科协青年人才托举工程，北京市科技新星，微软亚洲研究院“铸星计划”访问学者。

报告题目：端侧具身智能系统

报告摘要：具身智能是人工智能下一个十年的重要范式，是抵达 AGI 的必经之路。具身智能强调智能体在复杂物理环境中的持续感知、理解、决策与行动闭环，为人工智能从信息处理走向真实世界任务执行提供了新的技术路径。随着模型能力的提升及端侧算力平台的持续演进，将具身智能能力部署到终端设备之上，正成为推动人工智能走向实时化、个体化与普适化应用的重要趋势。本报告将围绕端侧具身智能的关键技术展开讨论，重点介绍在端侧环境感知与行为建模、真实交互数据支撑、模型高效部署与推理加速等方面的研究进展。



韩旭

韩旭，清华大学计算机系助理研究员，研究方向为人工智能、自然语言处理、智能计算系统，在国际顶级学术会议及期刊发表论文百余篇，Google Scholar 他引 1.9 万余次，创建大模型开源社区 OpenBMB，相关开源项目在全球最大开源社区 Github 上累计获得数万星标。曾获得教育部自然科学一等奖、世界互联网大会领先科技奖、钱伟长中文信息处理科学技术奖，曾入选中国计算机学会(CCF)优博激励计划、博士后创新人才支持计划、清华大学水木学者计划、清华大学优秀博士后、《麻省理工科技评论》“中国区 35 岁以下科技创新 35 人”。

报告题目：面向大模型训练推理的国产软硬件广谱适配

报告摘要：规模法则驱动下的大模型训练与推理，离不开大规模分布式算力的高效支撑。相较于英伟达算力生态，国产算力及其配套软件栈呈现出类型多样、接口不统一、特性差异显著等特点，为基于国产算力开展大模型训练与推理带来了诸多挑战。本报告将重点介绍如何构建广谱适配国内主流芯片的大模型训练推理框架，如何结合国产芯片的计算与通信特性开展分布式通信计算协同优化，并展示基于国产软硬件体系训练大模型所取得的效果。



杨传广

杨传广，中国科学院计算所智能算法安全全国重点实验室副研究员、硕导，近期研究方向为视频生成与 3D 理解重建的高效性问题，目前已在 CCF-A 类期刊会议上发表论文 30 余篇，曾获博士后创新人才支持计划、CCF 博士学位论文激励计划提名、中国科学院院长特别奖、中国指挥与控制学会科技进步一等奖，现任 CAAI 青工委委员和 CICC 具身智能专委会委员。主持国自然青 C、中国科学院青年重点项目和腾讯 IEG 校企合作项目。

报告题目：面向视频生成与 3D 理解重建基座模型的高效优化：量化、稀疏化与缓存加速

报告摘要：扩散模型与 Transformer 架构推动了视频生成和 3D 理解重建的快速发展，但模型规模扩大、时空 token 序列增长和三维表示复杂化，也带来了显存占用高、推理延迟长和部署成本大的问题。本报告围绕生成式视觉模型的压缩与加速，介绍视频扩散 Transformer 中的低比特量化、时序蒸馏、微调样本选择、稀疏 token 优化以及稀疏注意力方法；进一步讨论长视频生成与视频世界模型中的异质 token 缓存、场景记忆管理和历史 KV 状态调控机制；最后介绍面向 VGGT、SAM3D 等 3D 理解重建基座模型的后训练量化、异质计算分配和几何复杂度自适应加速方法。报告旨在探讨高效生成式视觉智能的发展：从简单地静态压缩模型参数，走向动态的面向时空结构、token 重要性、几何异质性与记忆机制的任务感知型高效推理。

分论坛 视频世界模型

时间：8月5日下午14:00-17:30 地点：401会议室

内 容

嘉 宾

主席：王韞博（上海交通大学）、段岳圻（清华大学）、杨 啸（启元实验室）

基于重建与生成结合的3D世界模型

章国锋（浙江大学）

融合机理与数据的具身世界模型

徐 凯（中国科学院工业人工智能研究所）

从表征角度思考构建和利用世界模型技术

金 鑫（宁波东方理工大学）

世界模型：物理与数字空间中的双向探索

范 略（中国科学院自动化研究所）

视频生成：从内容生成到视频世界模型

赵 敏（南京大学）

论坛主席



王韞博

王韞博，上海交通大学计算机学院副教授，长期从事机器学习、强化学习研究，CCF 优博，国家自然科学基金原创探索项目负责人，NeurIPS、ICML、AAAI等会议领域主席。主要学术成果覆盖生成、仿真、决策的世界模型完整技术栈：一、视频预测基础架构 PredRNN等成果应用于国家气象预报系统，提升我国短临气象预报精度619%，被500余篇论文作为基线对比；二、提出NeuroFluid、OrbiSim等可微物理仿真方法，显著提升世界模型的Real2Sim物理建模能力；三、提出IsoDream、LS-Imagine、VeoRL等隐空间世界模型方法，完善了有模型强化学习算法体系。研究成果在中央气象台、中国南方电网、太平洋保险集团成功应用。



段岳圻

段岳圻，清华大学电子工程系副教授，博士生导师，研究方向为三维计算机视觉。在相关方向以第一 / 通讯作者发表计算机视觉与模式识别领域 CCF-A 类国际期刊会议论文 40 余篇，主持国家重点研发计划课题、国家自然科学基金面上项目、北京市自然科学基金重点项目、中国科协青年人才托举工程项目等，获公安部和中国电子学会科学技术奖一等奖 3 项。担任 IEEE TCSVT 期刊编委，CVPR、ICCV、ECCV、ICML、ICLR 等国际会议领域主席。



杨 啸

杨啸，启元实验室世界模型研究中心负责人，博士毕业于清华大学计算机系，师从朱军教授，发表国际顶会或顶刊TPAMI、NeurIPS、ICML、CVPR、等40余篇，第一作者10余篇，谷歌学术引用9800余次，带领团队获得重要国际安全竞赛冠军5次，包括ICRA 国际竞赛冠军，国家级无人试验竞赛一等奖；作为项目负责人或技术负责人执行经费合计超2.4亿元，相关成果已在多个领域得到广泛应用。此外多次担当TPAMI，NeurIPS，CVPR等审稿人，并参与举办ICML 及AAAIWorkshop，曾获清华大学优秀博士论文，清华大学优秀毕业生等。

论坛讲者



章国锋

章国锋，浙江大学求是特聘教授，博士生导师，国家杰出青年科学基金获得者。主要从事三维视觉、增强现实与空间智能方面的研究，尤其在 SLAM、三维重建和生成方面取得了一系列重要成果，开源了一系列相关系统和算法的源代码，是 OpenXRLab 扩展现实开源平台的主要发起人。曾获 2011 年全国优秀博士学位论文奖、2020 年浙江省技术发明奖一等奖（排名第 4）、2021 年浙江省自然科学奖一等奖（排名第 2）以及国际顶级会议 ISMAR 2020 唯一最佳论文奖。担任国际顶级期刊 IJCV 编委，以及《Virtual Reality & Intelligent Hardware》、《计算机辅助设计与图形学学报》和《中国图象图形学报》等期刊编委，中国图像图形学会虚拟现实专委会副主任、增强现实核心技术产业联盟副理事长、浙江省人工智能学会增强现实分会副会长。

报告题目：基于重建与生成结合的3D世界模型

报告摘要：近年来，随着大模型的快速发展，三维重建与生成技术都取得了显著进展，而且两者技术的结合既可以提升重建的鲁棒性和完整度也能提升生成的质量和时空一致性，已经成为一个重要的发展趋势。本次报告主要介绍我们团队近年来在三维场景的高效重建与生成方面的研究工作，并展示相关的应用。例如，根据输入的一张图像，可以生成一个静态的3D场景并在其中实时漫游，甚至能够走进一个单目视频，实时地在动态场景中自由视角漫游。



徐 凯

徐凯，中科院工业人工智能研究所研究员。曾任国防科技大学教授、普林斯顿大学访问学者、加拿大西蒙弗雷泽大学客座教授。研究方向为计算机图形学、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，提出面向复杂场景的结构感知、机理增强的具身交互理论方法系统，并规模化应用于重工、船舶、汽车等智能制造领域。主持国家自然科学基金青年科学基金 A 类（原杰青）、B 类（原优青）、重点项目等。发表 TOG/TPAMI/TVCG/TIP 等 A 类论文 100 余篇。连续多年入选全球高被引学者榜单和全球前 2% 顶尖科学家榜单。担任图形领域顶级国际期刊 ACM Transactions on Graphics、IEEE Transactions on Visualization and Computer Graphics 的编委，Computational Visual Media 的领域执行编委。担任中国图像图形学会智能图形专委会副主任、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 2 项（排名 1 和 3）、军队科技进步二等奖、军队教学成果二等奖、中国电子学会青年科学家奖。

报告题目：融合机理与数据的具身世界模型

报告摘要：真实物理环境下的具身智能训练成本极高，使得基于仿真平台的 Sim2Real 迁移成为主流方案。然而，构建兼具高保真度与强泛化性的仿真环境仍面临巨大挑战。虽然生成式 AI 为世界模型提供了新范式，但单纯依赖像素级生成的世界基础模型往往需要海量算力与数据支撑。为此，我们依托多模态大模型，构建融合机理表征与物理约束的世界模型，以实现真实世界的深层理解与精准推演，从而赋能具身智能训练。本报告将汇报我们在机理与数据融合的具身世界模型方面的探索，及其在工业场景中的应用。



金 鑫

金鑫，宁波东方理工大学助理教授、博导，浙江省青年拔尖人才，近期研究聚焦在世界模型技术在具身智能中的应用，在 TPAMI、CVPR、ICCV、ECCV、NeurIPS 等期刊会议发表高水平论文 40 余篇，谷歌引用 9000 余次，多项成果被应用在微软、阿里巴巴、吉利汽车等产品中带来经济价值。曾获中国科学院院长特别奖 ACM SIGAI 优博、IEEE 电路与系统学会第二届视觉信号处理与通信新星奖、2026 IEEE MSA TC 最佳论文奖，斯坦福全球前 2% 顶尖科学家 24 25 榜单。IEEE VSPC 委员、首届 CAAI 具身智能专委会委员、Valse 执委、CSIG 青工委 & 多媒体专委会委员，在 CVPR24、ICCV25、NeurIPS25、ECCV24 等国际顶级计算机视觉及人工智能会议多次组织 AI 表征学习、具身智能与生成技术相关的 Tutorial 和 Workshop。

报告题目：从表征角度思考构建和利用世界模型技术

报告摘要：世界模型正成为连接数字世界与物理世界的核心技术范式，其价值在于赋予智能体空间理解、推理与交互的能力，使机器真正“走进”物理世界。本报告将系统梳理空间智能从3D感知、世界模型构建到具身任务执行的完整技术链路，并重点介绍团队在该领域的代表性工作：面向自动驾驶与机器人场景的UniScene统一生成框架；将世界模型引入VLA架构的DreamVLA模型，通过“世界表征嵌入”增强长程推理与动作预测能力；Inter-VLA&SceneScribe-1M等基准数据集，填补第一人称视角下具身交互数据及空间几何视频场景数据的空白；DeIF框架，解耦前向-后向动力学模型学习过程；VLA-JEPA，从隐空间层面赋予VLA模型高效推理能力等等。



范 略

范略，中国科学院自动化所模式识别实验室助理研究员。研究领域为世界模型驱动的自动驾驶和具身智能。以第一作者和通讯作者身份在 TPAMI、CVPR、ICCV、NeurIPS、ICLR 等国际主流期刊和会议发表论文 30 余篇，获得 3000 余谷歌学术引用。主持国家自然科学基金、北京市自然科学基金培育项目，是新一代人工智能国家科技重大专项核心骨干。曾获 ICRA 具身世界模型国际挑战赛冠军、CSIG 优博、中科院优博等奖励。

报告题目：世界模型：物理与数字空间中的双向探索

报告摘要：本报告首先对主流世界模型进行了梳理，从数据维度的角度分析了当前世界模型领域的发展现状，重点探讨了重建生成一体化世界模型的技术特点和现有局限，分享一种可拓展的通用4D世界模型NeoVerse及其在ICRA 2026 World Model Challenge中的冠军解决方案。报告进一步分享面向视频世界模型全方位评测的最新工作，对下一代通用世界模型的预期特性进行展望。此外，报告将深入介绍MobileGym，一种为移动端智能体打造的数字空间世界模型，在浏览器中实现了完全透明可控的高并发移动设备仿真。



赵 敏

赵敏, 南京大学准聘助理教授, 清华大学“水木学者”, 国产视频生成模型 Vidu 核心作者。主要研究方向为视频生成与可交互视频世界模型。以第一作者 / 共同第一作者身份在 NeurIPS、ICML、ICLR 等国际顶级会议发表论文多篇。提出了 Causal Forcing 系列实时交互生成方法、RIFLEx 系列视频长度外推方法等, 相关技术已被 HunyuanVideo、CogVideoX 等主流开源视频扩散框架集成应用。近期主导开源 minWM 视频世界模型框架, 覆盖数据、训练、蒸馏与推理全流程, 推动实时可交互视频世界模型的系统化开源实践。相关开源项目累计获得 GitHub Star 近 3000 次。

报告题目: 视频生成: 从内容生成到视频世界模型

报告摘要: 视频生成模型正在从离线内容生成走向实时交互的视频世界模型。其关键挑战在于: 如何基于已有高质量双向视频生成基础模型, 构建能够低延迟、流式自回归生成, 并实时响应用户动作与环境变化的交互式世界模型。本报告将围绕这一挑战, 介绍 Causal Forcing 系列方法: 通过帧级单射性约束下的一致性蒸馏与分布匹配蒸馏, 将双向视频基础模型高效转化为少步采样的自回归交互模型, 在保持生成质量的同时提升交互响应速度。最后, 报告将介绍 minWM 开源框架, 覆盖数据、训练、蒸馏与推理全流程, 为构建实时可交互视频世界模型提供系统化开源实践。

分论坛

大模型的高效训练与推理

时间：8月5日下午14:00-17:30 地点：402会议室

国内 容

嘉 宾

主席：包世龙（中国科学院计算技术研究所）、王 伟（北京交通大学）
许倩倩（中国科学院计算技术研究所）

智能座舱的多模态具身交互

马 楠（北京工业大学）

面向高效推理的草稿自拟投机解码方法

车万翔（哈尔滨工业大学）

核心上下文感知Transformer：长上下文高效建模与推理

谭明奎（华南理工大学）

Lion优化器的收敛分析与改进

张利军（南京大学）

大模型压缩与推理加速

张 淼（哈尔滨工业大学（深圳））

论坛主席



包世龙

包世龙，中国科学院计算技术研究所助理研究员，主要从事机器学习与人工智能安全研究，重点包括复杂学习场景下的排序优化、安全可靠的多模态生成等，在 CCF-A 类期刊/会议上发表论文 25+篇，包括TPAMI 9 篇 (IF:20.4)。先后入选中国博士后科学基金国家资助博士后研究人员计划、北京市“高创计划”青年人才托举工程、中国科学院优秀博士学位论文奖、中国科学院特别研究助理项目等人才支持计划；主持国家自然科学基金青年项目（C类）、中国博士后科学基金特别资助、中国博士后科学基金面上等项目，并参与中科院先导 B、国家自然科学基金专项等重大项目。



王 伟

王伟，北京交通大学教授、博士生导师，国家级青年人才计划入选者，现任信息科学研究所副所长。本科毕业于西北工业大学，获南丹麦大学硕士及意大利特伦托大学博士学位，并曾在新加坡国立大学访学、瑞士洛桑联邦理工大学从事博士后研究，之后任意大利特伦托大学助理教授。长期聚焦计算机视觉、机器学习与生成式人工智能，在生成式模型与Transformer架构设计方面开展了系统性工作，涵盖潜在空间编辑、模型轻量化与高效注意力机制等核心技术。他在CVPR、ICCV、NeurIPS、ICML等顶级会议及TPAMI、TIP等权威期刊发表论文近50篇，主持国家自然科学基金、北京市自然科学基金等重点项目，并深度参与了欧盟地平线及瑞士创新署重大研究计划。凭借系列创新成果，荣获ACM MM最佳论文提名奖及意大利计算机视觉与机器学习协会最佳博士论文奖。王伟教授现任NeurIPS领域主席、AAAI高级程序委员，并当选IEEE高级会员。



许倩倩

许倩倩，中国科学院计算技术研究所研究员，博士生导师。研究方向为机器学习、跨媒体计算、知识图谱。主持国家自然科学基金青年科学基金A类、B类、联合基金重点、北京智源学者、中国科学院先导课题等项目10余项。发表CCF-A类论文100余篇（包括TPAMI 20余篇），获2024年ACM MM Honourable Mention Award，并获NeurIPS/CVPR等CCF-A类国际会议竞赛冠军15项。担任国际期刊TMM、TCSVT、ACM TOMM和Multi-media Systems编委，10余次担任CCF-A类国际会议领域主席。现/曾任CSIG青工委副秘书长、CSIG多媒体专委会副秘书长、CCF多媒体技术专委会副秘书长、CAAI深度学习专委会副秘书长。曾获中国图像图形学会自然科学一等奖（第一完成人）、吴文俊人工智能自然科学一等奖（第三完成人）、吴文俊人工智能科技进步二等奖、浙江省技术发明二等奖、茅以升北京青年科技奖、中国图像图形学会石青云女科学家奖、吴文俊人工智能优秀青年奖、中国人工智能学会最佳青年科技成果奖、ACM中国SIGMM新星，并入选首份AI华人女性青年学者榜单。研究成果成功应用于国家安全部门、阿里巴巴、百度等。

论坛讲者



马楠

马楠，北京工业大学教授、博士生导师，人工智能学院院长，国家级高层次领军人才，国家重点研发计划项目首席科学家，具身交互智能北京市重点实验室主任，中国人工智能学会会士、副秘书长，北京市先进工作者，北京市朝阳区“凤凰计划”领军人才。研究方向为交互认知、具身智能、无人驾驶和智能机器人。第一完成人获得中国图像图形学会科技进步一等奖等科技奖项；主持国家、省部级项目10余项；承担北汽、东风等企业委托课题10余项；获得国内外人工智能、无人驾驶赛事冠军20余项；已在IEEE TPAMI、TRO、TIP、Engineering、中国科学·信息科学、AAAI和ICRA等国内外权威期刊和会议发表论文百余篇；先后获得第六届全国教育科学研究优秀成果奖二等奖、北京市教学成果一等奖和北京工业大学教育教学成果特等奖等。

报告题目：智能座舱的多模态具身交互

报告摘要：针对智能汽车发展过程中智能座舱数据结构较为单一，座舱与用户交互模块间关联性有待增强，多模态感知信息获取质量及利用效率尚需提升等问题，团队近年来聚焦基于智能座舱交互多模态数据的用户行为分析与评价方法，包括研究视听触多模态数据智能获取与编码融合方法，实现座舱内嘈杂、遮挡等场景中跨源信息间的配准融合，力求提升智能座舱用户行为感知数据的处理能力；研究基于常规行为字典与开放式行为增量的用户行为交互模型，结合用户行为交互数据以在线的方式微调大语言模型，以离线的方式构造行为字典，实现多模态用户交互行为序列分割与典型交互模式识别，并研究基于对用户微表情、声音、动作的多模态状态感知和建模评价，力求推动智能座舱智能化水平迭代发展。



车万翔

车万翔, 哈尔滨工业大学计算学部长聘教授, 人工智能研究院副院长, 国家级青年人才, 斯坦福大学访问学者。主要研究方向为自然语言处理与大语言模型。现任中国中文信息学会计算语言学专委会副主任兼秘书长, 国际顶级会议 ACL 2025 程序委员会共同主席。承担国家自然科学基金重点等项目, 著有《自然语言处理: 基于大语言模型的方法》等。主导研发的语音技术平台(LTP)已被百度、腾讯、华为等企业付费使用。曾获 2024 年吴文俊人工智能科技进步一等奖(排名第 1)、黑龙江省科技进步一等奖(排名第 2)及 AAAI 最佳论文提名、2025 年 IEEE 信号处理学会最佳论文奖等, 并入选“全球前 2% 顶尖科学家”榜单。

报告题目: 面向高效推理的草稿自拟投机解码方法

报告摘要: 随着大语言模型在各类应用中广泛部署, 其解码效率成为制约端侧推理的重要瓶颈。主流模型依赖自回归逐 Token 生成, 既受限于访存比低的硬件环境, 又难以实现并行加速。投机解码作为新兴方向, 利用小模型生成草稿并由大模型并行验证, 已展现出显著的加速潜力。然而, 现有方法普遍依赖额外草稿模型, 带来显存占用高、兼容性差等问题。为此, 我们提出“草稿自拟”的投机解码思路: 通过加噪训练(MSN)赋予大模型自身并行解码能力, 无需外部草稿模型即可实现 2.28-2.68 倍加速; 同时提出即插即用的 Token Recycling 机制, 仅需不足 2MB 存储即可实现约 2 倍无损加速。该方案在保证生成质量的前提下, 有效缓解资源压力, 具备更强通用性与端侧适配性, 为大模型高效推理提供了新的解决路径。



谭明奎

谭明奎, 华南理工大学教授、博导, TPAMI 及 MIR 编委, “大数据与智能机器人”教育部重点实验室副主任, 专注机器学习与大模型、视觉信息处理等方向研究, 以一作 / 通讯作者发表学术论文 180 余篇、谷歌引用超 2.3 万次。牵头获吴文俊自然科学一等奖、教育部自然科学二等奖, 广东省青年科技创新奖, 世界华人数学家联盟“最佳论文奖”、广东省人工智能产业协会自然科学奖一等奖等。担任人工智能权威会议 NeurIPS、ICML、ICCV、CVPR、AAAI 等会议领域主席。

报告题目: 核心上下文感知Transformer: 长上下文高效建模与推理

报告摘要: 大模型上下文窗口已扩展至百万级, 但传统自注意力机制存在计算开销激增、上下文遗忘、注意力分散等瓶颈, 长序列冗余机理也缺乏系统理论支撑。本报告从理论层面揭示Transformer长序列建模的“维数灾难”本质, 证明注意力权重的稀疏性规律; 提出核心上下文感知 (CCA) 注意力机制, 通过全局语义压缩与局部细粒度保留的互补架构, 可即插即用替换现有注意力模块, 在128K场景下实现7.9倍推理加速、93% KV显存降低, 长文本理解精度优于同类方法。针对免训练落地需求, 进一步提出动态稀疏感知方案, 自适应筛选核心上下文, 实现预填充与解码双阶段加速, 128K场景下推理提速2.8倍、KV缓存缩减61%, 通用任务性能与全注意力机制基本持平。最后展望长上下文大模型的后续研究与产业落地方向。



张利军

张利军，南京大学人工智能学院教授，基金委优青。主要研究大规模机器学习与优化，发表CCF-A类期刊和会议论文120余篇，论文共被引用近8000次。目前担任机器学习权威期刊《Machine Learning》副主编，10余次担任机器学习顶级会议ICML/NeurIPS/ICLR领域主席。主持基金委联合重点项目、国际合作项目等。曾获首届达摩院青橙奖、CCF-A类会议论文奖等荣誉。

报告题目：Lion 优化器的收敛分析与改进

报告摘要：近年来，基于梯度符号的随机优化算法受到广泛关注，并在大模型训练等实践场景中取得了显著成功。然而，其理论研究仍明显滞后于实践发展：已有分析通常依赖较强的假设条件，或仅能建立次优的收敛速率。围绕这一问题，我们在近期工作中研究了该方向的代表性算法 Lion，分别刻画了其在单机和分布式优化场景下的收敛速率；此外，我们进一步引入方差约减技术，从理论上改进了 Lion 的收敛性能。



张 淼

张淼，哈尔滨工业大学(深圳)教授，博士生导师，国家级青年人才，河套学院客座教授，计算机科学与技术学院院长助理。曾就职于丹麦奥尔堡大学计算机学院助理教授，澳大利亚莫纳什大学研究员，博士毕业于澳大利亚悉尼科技大学。主要从事端侧人工智能和高效机器学习方向，涉及领域包括大模型压缩与推理加速，高效模型架构设计、高效多智能体系统等。一作 / 通讯发表 CCF A 类会议或期刊论文 40 余篇，如 IEEE TPAMI, ICML, NeurIPS, CVPR 等。目前担任 IEEE TCSVT Associate Editor, ACL AE, IJCAI SPC 等，主持 6 项国家级和省部级项目和课题。在产业应用与成果转化方面，相关研究成果已在大模型推理优化与多智能体系统方向与多家头部企业开展合作。

报告题目：大模型压缩与推理加速

报告摘要：随着大模型技术的快速发展，模型规模和计算开销持续增长，如何在保证模型能力的同时降低部署成本、提升推理效率，已成为大模型落地应用中的关键问题。大模型压缩与推理加速旨在通过量化、稀疏、高效注意力、KV Cache 优化等技术，减少模型参数量、显存占用和计算延迟，从而提升大模型在云端、边缘端和端侧设备上的部署可行性。本报告将围绕大模型压缩与推理加速的典型方法展开介绍，重点分析低比特量化、激活稀疏、Token 剪枝、KV Cache 压缩等方向的基本思想与团队相关近期工作。

分论坛

博士生论坛

时间：8月5日下午14:00-17:30 地点：202会议室

内 容

嘉 宾

主席：贾 旭（大连理工大学）、杨 鹏（华中科技大学）
张 杰（中国科学院计算技术研究所）、户战选（云南师范大学）

可控世界模拟	马 跃（香港科技大学）
迈向世界模型的三维序贯感知——自表达性约束的理论与实践	邓 辉（西北工业大学）
Not All Actions are Equal:Reliable Value Estimation for Offline-to-Online Reinforcement Learning	夏文科（中国人民大学）
大模型动态评测前沿	陈子健（上海交通大学）
Compositional Intelligence for Embodied Agents: From Modular Methods to Real-World Deployment	秦逸然（香港中文大学（深圳））
多模态数据演进：视觉、语言与代码	宋子帆（同济大学）
正激励噪声：表征、理解与生成	黄思琪（复旦大学）
生成式大模型后门攻击与防御技术	王中琦（中国科学院计算技术研究所）
面向弱网场景下的高保真生成式视频流传输技术	龚天翼（香港中文大学（深圳））
迈向全场景空间智能：空地导航的仿真、感知与决策	张莘芮（上海人工智能实验室）

论坛主席



贾 旭

贾旭，大连理工大学未来技术学院教授，博士生导师，入选国家级青年人才，辽宁省兴辽英才计划拔尖青年人才，大连市高端人才，比利时荷语鲁汶大学博士，长期从事计算机视觉与生成式人工智能领域的研究，曾在华为、商汤和谷歌从事研发工作。在 CCF-A 类期刊和会议发表学术论文 60 余篇，Google 学术引用 13000 余次，已申请和授权国内外发明专利 20 余项，成果获包括诺贝尔奖等权威学者正面评价，入选全球前 2% 顶尖科学家。主持 3 项国家级项目或重点项目子课题，研究成果获得 CCF 自然科学二等奖（序 1）、华为“难题揭榜”火花奖、以及 CVPR 形状恢复挑战赛冠军等多项学术奖励。多次担任 ICLR、ICML、AAAI、ACM MM 等人工智能和计算机视觉领域顶级学术会议的领域主席和高级程序委员会委员。



杨 鹏

杨鹏，华中科技大学电信学院副教授，主要从事多媒体通信、内容生成、边缘智能等方向研究；撰写出版全英文专著一部，在 IEEE TMC、TCSVT、ACM MM 等高水平学术期刊和会议发表第一作者 / 通信作者论文 40 余篇，获“最佳论文奖”3 次；主持国家自然科学基金、湖北省自然科学基金、华中科技大学自主创新基金等项目。任中国图像图形学会青年工作委员会副秘书长，入选中国科协青年人才托举工程，连续入选斯坦福大学“全球前 2% 顶尖科学家年度榜单”，两次指导学生获挑战杯“揭榜挂帅”专项赛全国一等奖。



张 杰

张杰，中国科学院计算技术研究所副研究员、博士生导师。研究方向：多模态大模型安全与可信视觉计算，在知名期刊会议 TPAMI/IJCV/TIP/CVPR/ICCV 等上发表论文 50 余篇，承担中国科学院先导 B 课题、科技部重点研发课题、国家信息安全计划、基金委面上等项目 10 余项，研究成果在华为手机、智能高铁、大模型测评等场景得到规模化应用，授权发明专利 21 项，获 2025 年 CSIG 青年科学家奖、2025 年 CCF 科技进步奖二等奖、2024 年 CSIG 自然科学奖一等奖、2024 年中国铁道学会科技奖二等奖。先后入选北京市科技新星、中科院青促会和微软青年学者铸星计划。担任 IEEE TIP、Pattern Recognition 期刊编委、NeurIPS、ACL 领域主席等。



户战选

户战选，博士，云南师范大学信息学院副教授，入选云南省“兴滇英才”青年人才。主要研究方向为机器学习、开放场景视觉理解等。近年来，在 CVPR、ICCV、ICML、AAAI、IEEE TIP、IEEE TKDE 等会议和期刊发表学术论文 20 余篇。主持国家自然科学基金青年项目、云南省自然科学基金面上项目和云南省“兴滇英才”青年人才项目。长期担任 CVPR、ICCV、AAAI、IEEE-TIP、IEEE-TNNLS、PR 等国际期刊和会议的审稿人。

论坛讲者



马 跃

马跃, 香港科技大学在读博士生, 清华大学硕士, 北京市优秀毕业生。曾于 ETH Zurich、MIT 访问学习, 并在 Meta、腾讯等企业从事研究实习。近五年, 以第一作者 / 共同第一作者在 ICML、NeurIPS、ICLR 等国际顶级会议发表论文 20 余篇, 多篇论文入选 Highlight/Oral, 1 篇论文被精选为 AAAI 最具影响力论文。其研究成果受到学术界与工业界广泛关注, 谷歌学术引用数超过 2,000。

报告题目: 可控世界模拟

报告摘要: 世界模型 (World Model) 致力于学习现实世界中的物理规律、空间结构以及时空动态, 并在虚拟环境中实现对真实世界的建模、预测与推演。随着生成式人工智能的发展, 视频生成模型已经展现出强大的视觉合成能力, 但距离真正的世界模型仍存在重要差距。一个真正的世界模型不仅能够生成逼真的视觉内容, 更应具备对环境、物体和智能体行为的理解能力, 以及支持交互、规划和控制的能力。本次报告围绕“可控世界模拟 (Controllable World Simulation)”展开, 探讨如何从视频生成迈向世界生成。具体而言, 可控世界模拟强调对运动、内容和环境动态的解耦建模, 实现以人为中心的精确控制; 通过构建统一的 4D 时空表示, 保证场景在时间与空间上的一致性; 同时支持灵活交互与实时反馈, 使虚拟世界能够响应用户指令和智能体行为。报告将介绍相关关键技术、最新研究进展及其在机器人训练、自动驾驶、数字孪生、游戏生成和具身智能等领域的重要应用, 展望构建可控、可交互且扎根于真实物理世界的下一代世界模型。



邓 辉

邓辉，西北工业大学电子信息学院在读博士研究生，隶属于西北工业大学陕西省信息获取与处理重点实验室，导师为戴玉超教授。其研究方向为三维视觉，重点关注 3D/4D 重建、深度学习、点云处理、特征提取以及三维模型生成等问题。近年来，邓辉以第一作者(含共同一作)在 TPAMI、ICLR、AAAI 等领域顶级国际期刊与会议发表多篇论文，代表性工作包括构建高性能的序列对序列动态三维重建框架、非刚性三维结构序列结构建模、注意力机制线性化及其应用等。此外，他曾参与国家自然科学基金、重点研发计划等多个重要项目，拥有多项发明专利，并曾获研究生电子设计大赛西北赛区一等奖等荣誉。

报告题目：迈向世界模型的三维序贯感知——自表达性约束的理论与实践

报告摘要：从连续视觉观测中持续感知与推理三维环境，是世界模型研究的一项核心命题。其关键挑战在于：观测逐帧到达，系统需在有限上下文中维护时空一致性——自动驾驶的实时建模、增强现实的空间理解均依赖于此。这引出一个基础问题：连续观测帧之间存在何种结构规律，使得长程一致性得以维系？围绕这一问题，我们以自表达性为核心线索：序列中某一帧的几何状态，可由同窗口内其他帧线性组合表达。我们首先将该约束引入深度框架，通过多头注意力对非刚性形变序列施加整体约束。此后将视角推广至Grassmann流形：隐状态视为子空间表征，状态演化对应流形上的轨迹。在此观点下，自表达性不再依赖训练，偏离时仅需计算窗口内状态相似度即可解析校正。基于上述理论框架，我们在两类重建场景中进行了验证。实验表明，在稀疏非刚性重建中该方法在多个基准上达到领先水平；在流式稠密重建中，解析校正算子能一致抑制长序列累积漂移，在深度、位姿与三维重建上均带来可重复提升，且不增加推理延迟。方法亦有适用边界：极稀疏观测下约束退化；窗口消融表明时序有效半径比预期紧凑。后续方向包括：设计自适应正则机制；将自表达性推广为序列表征学习的通用正则手段。从非刚性重建到流式重建再到通用世界模型，这条线索仍在演进。



夏文科

夏文科，中国人民大学高瓴人工智能学院直博四年级，导师为胡迪副教授，研究方向为具身智能、可泛化物体操纵以及强化学习。以第一作者 / 共同第一作者身份在 RSS、CVPR、NeurIPS、ICRA 等顶级机器学习与机器人会议上发表论文十余篇，聚焦于如何让机器人通过与环境主动交互实现自主矫正与持续进化，并构建在真实场景下可应用的数据飞轮系统。

报告题目：Not All Actions are Equal:Reliable Value Estimation for Offline-to-On-line Reinforcement Learning

报告摘要：离线到在线强化学习通过结合大规模离线策略预训练与真实环境中的在线交互改进，已成为实现通用机器人操作的重要范式。然而，该系统涉及数据构建、价值估计、离线训练与在线探索等多个紧密耦合的环节，使得性能复现与失败诊断仍然具有挑战。其中，价值估计直接影响交互数据的利用方式及其对策略改进的贡献。在报告中，讲者将基于完整的离线到在线强化学习框架，围绕价值函数可靠性这一核心问题展开分析，并结合真机毫米级芯片插入和高泛化积木拆解等任务，探讨可靠价值估计如何提升离线强化学习的扩展性与在线适应的稳定性。



陈子健

陈子健，上海交通大学集成电路学院(信息与电子工程学院)博士研究生，师从张文军院士和翟广涛教授，主要研究方向包括视觉信号处理、多模态大模型、甲骨文信息处理等。其研究主要聚焦于多媒体感知质量评价与大模型能力边界评测与优化。近年来以第一作者在 NeurIPS、ICLR、IEEE TCSVT 以及 npj 系列等顶级会议和期刊发表多篇论文。曾获博士国家奖学金、上海市优秀毕业生，入选中国科协青年科技人才培养工程博士生专项计划。

报告题目：大模型动态评测前沿

报告摘要：随着大语言模型及多模态大模型能力的飞速跃升，传统静态基准逐渐暴露出数据污染严重、易“刷榜”以及难以衡量真实场景能力的局限。本次报告聚焦“大模型动态评测前沿”，结合本人近期的一系列研究工作，探讨大模型评测范式如何向动态生成、对抗博弈与多维感知演进。报告将重点介绍我们在动态交互评测方向的核心探索：首先介绍的是多智能体持续评测网络 MACEval。该工作将传统的单向评测转化为自主“面试”过程，通过多智能体协作实现题目的动态生成与难度自适应，从而有效缓解数据污染并精准探查模型的能力边界。其次，报告将引入对抗性博弈评测框架 SQUID GAME。在“大逃杀的生存”机制下，模型被置于资源受限与信息不对称的挑战中，通过与其他模型对手的交互对抗，深度检验大模型在极端压力和复杂博弈下的推理、规划与决策水平。最后，报告将延伸至多模态评测基准 BioMotion Arena，展示如何突破静态视觉限制，基于人类动作偏好与生物力学特征，量化评估多模态大模型对物理世界与复杂动作的感知与理解能力。通过串联这三项研究，本报告旨在为搭建真实世界复杂度的下一代大模型评测体系提供新视角与实践参考。



秦逸然

秦逸然是香港中文大学（深圳）5年级博士生，同时在牛津大学的Torr Vision Group (TVG)访学。秦逸然目前是NeoteAI公司的数据与算法部门主管，负责触觉传感、具身数据基础设施、数据管理以及机器人操作和具身智能模型训练等方面的工作。曾经在蔚来自动驾驶，上海人工智能实验室，快手可灵实习。研究方向聚焦机器人操作与导航、具身智能体协作，以及视频世界模型在机器人中的应用。近年来，秦逸然以第一或共同第一作者在NeurIPS、ICML、CVPR、ICCV等人工智能领域顶级会议发表十余篇学术论文。

报告题目：Compositional Intelligence for Embodied Agents: From Modular Methods to Real-World Deployment

报告摘要：我的研究聚焦于具身人工智能（Embodied AI），致力于探索智能体如何在开放、动态且充满组合变化的物理世界中实现自主决策与长期任务执行。我认为，具身智能体面临的核心挑战并非单纯的数据规模不足，而在于物理世界本身具有高度的组合性：智能体需要在未见过的物体、场景、约束和协作关系之间进行泛化。围绕这一问题，我提出并实践“组合式智能（Compositional Intelligence）”的研究框架，通过组合泛化、组合智能体、组合约束和组合建模等方向，研究如何利用可重用的知识、技能和模块构建能够适应开放世界的新一代智能体。我的目标是推动具身智能从依赖大规模数据记忆的模式，走向具备系统性泛化、推理与协作能力的自主智能系统。



宋子帆

宋子帆，同济大学计算机科学与技术系本博，师从赵才荣教授，主要研究方向为多模态数据智能、智能体系统。在TPAMI、ICLR、NeurIPS、ICML等国际期刊会议发表论文十余篇，其中以第一作者身份发表六篇，涵盖监督学习训练策略、多模态增强、开放场景视觉建模以及多智能体系统下游任务。曾在微软亚洲研究院、上海人工智能实验室等机构实习，参与大模型基座与世界模型项目科研。

报告题目：多模态数据演进：视觉、语言与代码

报告摘要：随着深度学习和基座模型的发展，数据为中心的研究范式逐步展现不可或缺的研究价值。数据密度、分布特征与监督信号为核心的数据侧演进成为新的研究重点。本报告聚焦于数据为中心的研究路径，从理论建模的迭代化扰动，到基于基座模型的生成式增强，最后探讨智能体驱动的可控型仿真。具体地，（1）迭代化扰动：通过构建关于数据扰动的可微分目标函数，计算微小且精准的图像与标签扰动量，对样本特征分布和监督信号进行联合优化。（2）生成式增强：基于扩散模型与大语言模型，依据属性参考对图像内容和文本描述进行一致性编辑，此外引入后见之明思想实现细粒度数据进化，以扩充定制化多模态样本对，增强模型后训练。（3）可控型仿真：设计数学推理与代码生成能力特化的多智能体框架，支撑可迭代化部署的代码实例和图形引擎驱动的编排仿真。



黄思琪

黄思琪，复旦大学智能机器人与先进制造创新学院博士研究生，师从李学龙教授，主要研究方向包括正激励噪声、图表征学习、多模态大语言模型与多模态生成模型等。其研究聚焦于结构化数据与多模态信息的建模、增强、理解与推理，围绕图对比学习中的数据增强机制、正向有益噪声建模、多模态大模型持续学习以及大模型协同推理等问题开展了系列研究。近年来，黄思琪以第一作者在ICML、AAAI等人工智能领域顶级会议和期刊发表多篇学术论文，其中包括ICML论文1篇、AAAI论文3篇。

报告题目：正激励噪声：表征、理解与生成

报告摘要：噪声通常被视为模型学习中的干扰因素，但近年来的研究表明，经过合理建模与学习的噪声也可以成为提升模型泛化、对齐与适应能力的重要信号。本次报告将围绕“正激励噪声”这一主线，介绍其从理论建模到不同学习范式中的应用扩展。报告首先介绍正激励噪声的基本思想以及其精确定义，说明如何从变分推断角度理解噪声对模型的正向作用，并进一步介绍其在表征学习中的应用。随后，报告将进一步讨论正激励噪声在多模态学习中的扩展。最后，报告还将简要介绍相关方向的最新探索。



王中琦

王中琦，中国科学院计算技术研究所博士研究生，隶属于中国科学院计算技术研究所视觉信息处理与学习（VIPL）研究组及智能算法安全全国重点实验室，导师为山世光研究员。其研究方向重点关注后门攻击与防御、模型评测、多模态大模型安全等问题。近年来，王中琦以一作/学生一作在IEEE TPAMI、IEEE TIFS、ICLR、ECCV等人工智能与计算机视觉领域重要期刊和会议发表多篇论文。此外，他曾获2026年IEEE SaTML后门防御竞赛冠军（序1）等荣誉。

报告题目：生成式大模型后门攻击与防御技术

报告摘要：随着生成式大模型的快速发展，其安全可信问题也日益引发广泛关注。本报告将介绍生成式大模型后门攻击与防御方面的最新研究进展，重点介绍本团队在文生图大模型后门攻防方面的具体工作，包括基于句法触发器的高隐蔽后门攻击方法、针对同化现象的统一防御框架（涵盖后门检测、定位与缓解），以及基于注意力动态异常发现的后门防御方法。最后，报告将对未来研究方向与发展趋势进行总结与展望。



龚天翼

龚天翼，香港中文大学深圳理工学院在读博士生，师从王方鑫教授和崔曙光教授，主要研究方向为生成式编码与多媒体传输技术，在USENIX NSDI/IEEE JSAC/CVPR等期刊会议上发表多篇学术论文。

报告题目：面向弱网场景下的高保真生成式视频流传输技术

报告摘要：视频流量目前已占据互联网总流量的 65% 以上，但在高铁、偏远地区等弱网环境下，用户常常面临画质模糊、卡顿等糟糕的观看体验。传统的视频编解码技术基于像素级的压缩策略在极端带宽受限环境下已接近性能瓶颈，难以在极低带宽下进一步提升画质。近年来，深度学习与生成式 AI 的快速发展为视频压缩与传输带来了新机遇。视觉基座模型展现出卓越的语义压缩和视觉重构能力，为视频传输带来了新的曙光。然而，直接将这些庞大的 AI 模型应用于对实时性要求极高的视频流媒体并非易事，它们通常计算笨重、推理缓慢，且对网络波动极其敏感。本报告围绕上述挑战，针对性地介绍了利用视觉基座模型重构视频传输范式，从而打破传统编码器的压缩极限，并显著提升弱网下的画质表现。



张荻芮

张荻芮，上海人工智能实验室IPEC团队博士研究生，主要研究方向包括具身导航、具身操作与世界模型等。其研究聚焦于全场景导航算法、仿真资产规模化生成及真实场景落地应用，在导航世界模型、三维重建与具身智能系统构建等方面取得了系列成果。近年来，张荻芮以第一作者或共同第一作者在CVPR、ICCV、AAAI、ICRA等计算机视觉与具身智能领域顶级会议发表多篇学术论文，其中包括CVPR论文1篇、ICCV论文1篇、AAAI Oral论文1篇（Top 4.6%）、ICRA Oral论文1篇（Top 11.5%）。

报告题目：迈向全场景空间智能：空地导航的仿真、感知与决策

报告摘要：全场景空间智能要求智能体能够在室内、室外和空中等多样环境中进行感知与决策，是具身智能走向开放真实世界的重要方向。本报告将围绕空地导航中的仿真构建、场景扩展与模型统一展开讨论。报告首先介绍面向全场景导航的高保真仿真构建，涵盖无人机飞行仿真、室内外三维场景重建、真实传感器模拟以及物理交互建模等关键问题，讨论如何通过更逼真的环境表示与任务设计支撑导航智能体的训练和评测。随后，报告将进一步关注导航环境的规模化生成，介绍自动化资产生成、程序化场景合成和多模态生成技术如何降低室内外导航数据的构建成本，并推动导航研究从有限基准走向大规模开放场景。最后，报告将讨论以导航世界模型统一空地导航的研究趋势，探讨如何通过通用空间表征、语言驱动决策和长程环境记忆，实现室内、室外与空中导航任务之间的知识迁移和能力复用。通过上述内容，本报告希望展示空地导航从仿真系统、数据规模化到空间智能基础模型的发展路径，并展望全场景导航智能体在未来具身智能系统中的应用前景。

分论坛

女科学家论坛

时间：8月5日下午14:00-17:30 地点：204会议室

内 容

嘉 宾

主席：董燕妮（武汉大学）、冯 婕（西安电子科技大学）
高联丽（电子科技大学）、相 艳（昆明理工大学）

开放世界空间视觉智能：从三维世界构建到低空场景理解

曹刘娟（厦门大学）

面向复杂营销决策的上下文工程优化

杜军平（北京邮电大学）

多模态文档的理解、生成与编辑

刘世霞（清华大学）

交叉研究：认知启发的视觉计算与推理决策

马惠敏（北京科技大学）

脑启发的高效感知、学习与决策

邬 霞（北京理工大学）

大规模乳腺癌智能筛查实践：从理论突破到技术落地

张 蕾（四川大学）

论坛主席



董燕妮

董燕妮，武汉大学教授、博士生导师，国家优秀青年科学基金获得者。主要从事遥感图像智能解译与地学空间智能应用等研究，已在 IEEE TIP、IEEE TGRS 等国际刊物上发表论文八十余篇，担任 IEEE TGRS 和 IEEE GRSL 副编辑 (Associate Editor)，中国图像图形学会优博俱乐部副主席，IEEE GRSS 武汉分会副主席，湖北省女科技工作者协会理事等；曾获得中国图像图形学会石青云女科学家奖、香江学者奖等；主持国家自然科学基金重点项目、优青项目，应急管理部重点科技计划课题等十余项。



冯 婕

冯婕，西安电子科技大学人工智能学院教授，博士生导师。主要研究方向为遥感图像处理、空天空间智能。发表 CCF A 类和中科院 I 区论文 60 余篇，ESI 高被引 / 热点论文 11 篇，出版专著 2 部。主持国家自然科学基金联合重点项目、军科委基础加强领域基金、教育部联合基金等。入选国家级青年人才、陕西省杰出青年科学基金、中国科协青托等。获得陕西省自然科学二等奖（序 1）等。担任 IEEE TGRS、IEEE J-STARS、Remote Sensing、Journal of Information and Intelligence (信息与智能学报 (英文)) 等期刊编委。目前为 CSIG 学会青工委副秘书长、中国电子学会空间电子学分会青年副主任委员等。



高联丽

高联丽，电子科技大学，教授博导，国家领军人才，国家自然科学基金优青项目获得者。专注于多媒体理解、AI 安全、生成式 AI 与具身智能方面的研究。迄今为止，在人工智能领域重要国际期刊和国际会议发表高水平论文 100+, Google Scholar 引用 1.2 万 +, H 指数 60。主持多项国家级和省部级项目。担任多媒体领域国际期刊 IEEE TMM 编委、多个 CCF-A 类国际会议领域主席（如 CVPR, ICCV, ECCV, NeurIPS, ACM MM, AAAI）、2024 年中国科协青年科学家沙龙 - 通用人工智能青年科学家沙龙执行主席。获 2019 年阿里巴巴达摩院第二届青橙奖（全国仅 10 人）、2020 年国际 IEEE Technical Community on Multimedia Computing Rising Star Award（年度国际唯一）、2025 年 CSIG 青年女科学家奖、2025 年教育部自然科学一等奖、2025 年电子学会一等奖等。



相艳

相艳，昆明理工大学信自学院教授，云南省中青年学术和技术带头人后备人才。主要研究方向为自然语言处理、情感计算、多模态信息处理。主持 / 参与国家重点研发计划、国家自然科学基金、云南省基础研究重点、重大科技专项等国家及省部级项目 10 余项。在 TASLP、NeuNet、ESWA, EAAI, AAAI, WWW, KDD, EMNLP, ICME 等 CCF 推荐期刊和会议发表学术论文 50 余篇，授权发明专利 20 余项。获云南省科技进步特等奖、云南省教学成果一等奖。

论坛讲者



曹刘娟

曹刘娟，厦门大学信息学院副院长，福建省可信人工智能分析与应用工程研究中心主任，教授，博士生导师，国家青年基金 A 类获得者、万人青年拔尖人才、福建省杰青，曾获中国图像图形学会石青云女科学家奖、AI 华人女性青年学者、第二十八届运盛青年科技奖等荣誉。长期从事计算机视觉、三维感知、小目标检测等人工智能领域前沿技术研究，以一作 / 通讯在 TPAMI, IJCV, CVPR 等国际顶级期刊会议上发表论文 60 余篇。主持国家自然科学基金青 A 项目 / 联合基金重点 / 某基础加强重点 / 某部重大专项等国家级项目多项，主持华为 / 腾讯 / 百度等企业合作项目多项，相关成果以第一完成人获 2020 年福建省科技进步一等奖、2023 年厦门市科技进步一等奖和 2025 年教育部科技进步二等奖。

报告题目：开放世界空间视觉智能：从三维世界构建到低空场景理解

报告摘要：随着生成式人工智能、三维视觉和低空智能应用的快速发展，机器视觉正从传统二维感知逐步走向开放世界空间环境的生成、重建、理解与交互。本报告将围绕“开放世界空间视觉智能”展开，介绍课题组近年来在三维世界构建、开放词汇空间理解和低空无人机场景感知等方面的研究进展。首先介绍文本驱动的真实相机轨迹与三维场景生成、高质量秒级三维场景生成，以及视觉几何 Transformer 的训练自由加速方法，探索如何高效构建真实可用的三维空间；随后进一步讨论如何基于三维高斯表达和多视角几何建模，实现语言查询驱动的目标定位与三维指代表达理解，使三维场景从“可重建”走向“可理解、可交互”；最后，面向低空无人机这一典型真实开放空间场景，介绍课题组在旋转目标检测和语言驱动分割方面构建的基准与模型，展示空间视觉智能在复杂视角、小目标、开放环境和真实应用需求下的拓展与落地。



杜军平

杜军平，IEEE Fellow、CAAI/CCF 会士、北京邮电大学教授、博士生导师，中国人工智能学会副理事长、北京邮电大学校学术委员会委员、计算机应用技术中心主任。获国务院政府特殊津贴。研究方向为大数据、机器学习、突发事件预测预警等。主持国家自然科学基金重大项目、国家重点研发计划项目、国家自然科学基金重点项目等。在 IEEE TPAMI、TKDE、TNNLS、VLDBJ、SIGMOD、KDD、AAAI、WWW 等发表论文 700 余篇，H-Index 59，申请发明专利 63 项，已授权 46 项。获吴文俊人工智能杰出贡献奖、CCF 夏培肃奖、2024 年和 2022 年中国电子学会科学技术奖科技进步一等奖、CCF 科技成果奖技术发明一等奖、国家技术发明奖二等奖、高等学校科学研究优秀成果奖(科学技术)技术发明一等奖、北京市科学技术二等奖等。

报告题目：面向复杂营销决策的上下文工程优化

报告摘要：大模型与智能体技术为复杂营销决策带来了机遇，当前面临的难题是如何将大模型的通用推理能力与千变万化、极具商业价值的复杂营销决策场景进行深度融合。本报告围绕基于语义结构化建模的程序性知识挖掘与策略表示、面向复杂营销推理决策的智能体数据演化与合成、基于信息瓶颈理论的复杂营销上下文工程优化以及系统实现等进行阐述，介绍如何应对程序性知识挖掘难、边界样本稀缺、动态推理知识供给失配的行业共性科学挑战。



刘世霞

刘世霞, 清华大学教授, 国家级人才计划入选者, IEEE Fellow。主要研究方向是大数据可视分析与可解释人工智能。获 IEEE 可视化技术成就奖; 入选福布斯中国科技女性 50 和可视化名人堂。担任 CCF A 类会议 IEEE VIS 2016-2017 的论文主席, IEEE VIS 2020-2023 指导委员会委员; 担任 IEEE TVCG 副主编、以及 Artificial Intelligence、IEEE TBD 和 ACM TiiS 的编委。

报告题目: 多模态文档的理解、生成与编辑

报告摘要: 多模态文档由文字、图像、图表、表格和图标等多种元素共同表达信息, 常见形式包括信息图、海报、幻灯片、报表和网页等, 是复杂信息表达、知识传播和辅助决策的重要载体。随着人工智能技术的发展, 如何让机器准确理解、自动生成并灵活编辑这类结构复杂、语义丰富的文档, 成为智能内容创作和文档自动化中的关键问题。本报告聚焦多模态文档的理解、生成与编辑, 介绍我们在数据构建、文档解析和内容生成方面的研究进展, 并探讨相关技术在提升复杂信息表达、理解和决策支持能力方面的作用。



马惠敏

马惠敏, 北京科技大学领军学者, 二级教授, 计算机与通信工程学院副院长、鼎新实验室主任、适人协同智能重点学科与技术研究中心主任。围绕类脑智能与人机混合智能开展研究, 在视觉认知计算领域取得了原创性成果, 作为负责人承担了国家自然科学基金重点项目、国家重点专项等项目, 作为通讯作者在Nature子刊、TPAMI、CVPR等国际高水平期刊和会议发表论文200余篇, 谷歌学术单篇他引4500余次; 获批专利60余项, 6项专利实现了成果转化, 应用于交通安全、航空航天、医疗健康等领域。作为第一完成人获得北京市自然科学一等奖、吴文俊人工智能科技创新一等奖, 中国图像图形学会技术发明一等奖等奖励。

报告题目: 交叉研究: 认知启发的视觉计算与推理决策

报告摘要: 面向适应复杂环境的自动驾驶视觉感知决策等应用, 模拟人的认知机理, 启发智能无人系统实现类人鲁棒视觉计算与推理决策。介绍如何开展认知科学与人工智能的交叉研究, 建立可计算的视觉心智理论认知架构、格式塔完型原则指导下的鲁棒视觉认知方法、基于人类认知双过程理论的高效具身推理决策方法、基于审议实践理论的鲁棒学习方法、基于双过程框架的自动驾驶决策范式。



郭霞

郭霞，北京理工大学计算机学院教授、博士生导师，国家自然科学基金杰出青年基金、优秀青年基金、吴文俊人工智能自然科学一等奖、教育部自然科学二等奖、茅以升北京青年科技奖等获得者。主要研究方向为脑机接口与类脑算法等。近年来，主持承担国家自然科学基金重点项目、国家重点研发计划课题以及北京市自然科学基金重点项目等十余项，以第一/通讯作者在国内外重要学术期刊、会议发表论文百余篇。

报告题目：脑启发的高效感知、学习与决策

报告摘要：随着深度学习模型在复杂任务中的局限性日益凸显，脑认知机制为突破当前人工智能瓶颈提供了重要的科学依据，已成为迈向通用人工智能的关键路径之一。人脑在感知补全、记忆整合与推理决策等方面展现出远超现有模型的卓越能力，为构建新一代智能计算范式提供了深刻启示。为系统阐释脑认知机制与人工智能的交叉融合，本报告将首先梳理脑机制与神经网络发展演进的内在关联。在此基础上，分享脑认知机制启发的智能感知、学习与决策研究前沿进展。最后，对人工智能和人类智能交叉融合、深度互鉴的未来发展趋势进行展望。



张蕾

张蕾，四川大学人工智能学院副院长，教授/博导，在国家优青/杰青等项目资助下，长期致力于人工智能的基础理论及智能医学研究。在相关领域的国际学术期刊和会议上发表论文百余篇，获专利授权20余项，研发多款智能医学系统。获国家自然科学基金二等奖、中国青年女科学家奖、中国人工智能学会吴文俊自然科学一等奖、教育部自然科学一等奖、四川省科技进步一等奖等。

报告题目：大规模乳腺癌智能筛查实践：从理论突破到技术落地

报告摘要：大规模乳腺癌筛查是提升女性健康水平、降低疾病负担的“健康中国-重大民生工程”。面向大规模人群筛查，真正的挑战不仅在于单点诊断模型的准确性，更需要解决跨中心/跨设备/跨专家的复杂数据漂移问题，以及筛查人群分布所导致的群体异质性问题，实现可复制、可推广、可持续的筛查服务。围绕大规模乳腺癌筛查落地中的关键瓶颈，我们开展了乳腺癌智能辅助诊断系列研究，形成了从特异性特征增强、关联性任务建模到多任务统一系统构建的技术路径，提升模型在病灶感知、结构表征、风险分级和临床决策辅助中的稳定性与泛化能力。相关理论成果发表于 Nature 子刊、IEEE TMI、MedIA 等人工智能与医学交叉领域期刊，并搭载于自研的便携式乳腺智能筛查系统，获批国家二类医疗器械认证，已筛查川、藏、云、甘及新疆等西部偏远和少数民族地区适龄女性2万余人次，助力“健康中国一两癌筛查”重大民生工程，实现了大规模乳腺癌智能筛查从理论突破到技术落地。

分论坛

学术新星论坛

时间：8月5日下午14:00-17:30 地点：206会议室

内 容

嘉 宾

主席：阚美娜（中国科学院计算技术研究所）、胡 鹏（四川大学）
邹博超（北京科技大学）、刘笑宏（上海交通大学）、刘 朋（云南大学）

“快思慢想”：类脑持续学习理论与方法	魏 坤（西安电子科技大学）
物理信息驱动的机器学习方法	孙 元（四川大学）
语义驱动的弱监督视频理解	汪 昱（同济大学）
浅谈多模态模型中的表征调和与学习	包世龙（中国科学院计算技术研究所）
SpectraAI: AI-Driven Molecular Identification and Discovery	夏 俊（香港科技大学（广州））
结构化知识赋能类脑分析	梁 科（国防科技大学）
高光谱遥感信息处理：从数据融合到智能解译	宋启亚（湖南师范大学）
多模态大模型的对抗安全问题研究	刘岱宗（武汉大学）
复杂对抗环境下的视觉生成式人工智能安全风险治理	刁云峰（合肥工业大学）

论坛主席



阚美娜

阚美娜，中国科学院计算技术研究所，研究员。研究领域为计算机视觉与具身智能。相关成果在 TPAMI、IJCV、CVPR、ICCV 等相关领域主流国际期刊与会议上面发表论文 60 余篇，获得 CVPR 盲人视觉问答竞赛冠军、IROS 世界模型竞赛冠军、FG PaSC 视频人脸识别竞赛冠军、ACM MM 亲属关系识别竞赛冠军、CVPR 面部关键点定位竞赛亚军。此外，获中国计算机学会优秀博士学位论文奖、中国图像图形学会石青云女科学家奖青年奖，入选北京市科技新星计划、青年科学基金项目（B 类）、承担国家自然科学基金重大项目课题等。



胡 鹏

胡鹏, 四川大学计算机学院教授, 博士生导师, 国家级青年人才, CSIG 青工委副秘书长。主要研究方向为可信机器学习和多媒体分析。2019-2020 年在新加坡科技研究局 (A*STAR) 担任研究员。目前已在 Nature 子刊、TPAMI、IJCV、CVPR 等 JCR 一区期刊及 CCF A 类会议上发表长文 80 余篇, 包括多篇 ESI 热点论文 (前 1%) 和高被引论文 (前 1%)。研究成果获得中国图像图形学会 (CSIG) 青年科学家奖、中国图像图形学会 (CSIG) 自然科学一等奖、四川省自然科学二等奖、中国人工智能学会教学成果激励计划一类成果 (序 1)。担任 TPAMI、TIP、CVPR、ICCV、NeurIPS、ICML、ICLR 等期刊和会议审稿人 / (高级) 领域主席 / 高级程序委员。主持国家自然科学基金 (联合重点、面上、青年) 项目、国家重点研发计划课题、JKW GF 科技创新特区项目、四川省重点研发基金项目等。



邹博超

邹博超, 北京科技大学计算机与通信工程学院副教授, 北理工 - 哈佛大学医学院联合培养博士, 北航 - 中电科联合培养博士后, 适人协同智能重点学科与技术研究中心副主任, 中国图像图形学会理事、副秘书长, 社会安全风险感知与防控大数据应用国家工程研究中心特聘专家。主要从事视觉认知、情感计算、人机协同方向研究, 作为负责人先后承担国家自然科学基金面上、青年项目、联合基金课题、新一代人工智能重大项目子课题, 参与国自然重大仪器项目, 入选北京市青年人才托举工程, 在 IEEE Trans.、ICML、NeurIPS、AAAI、ACM MM 等发表 SCI/EI 论文四十余篇, 授权发明专利十余项, 曾获 2024 年北京市自然科学一等奖, 2021 年中国电子科技集团科学技术一等奖。



刘笑宏

刘笑宏, 上海交通大学计算机学院副教授、博士生导师、上海创智学院全时导师。研究方向为计算机视觉与多媒体信息处理, 现已在顶级国际会议与期刊上发表学术论文 100 余篇, 其中 CCF-A 类或中科院一区论文 80 余篇, 引用 7900 余次, h 指数 43。获中国人工智能学会吴文俊人工智能青年科技奖、华为“难题揭榜”火花奖、美团科研合作卓越创新奖, 入选全球前 2% 顶尖科学家榜单、上海市领军人才 (海外)、上海市浦江人才。主持国家自然科学基金面上、青年项目, 参与重点项目, 并承担包括华为、腾讯、美团、OPPO 等多家公司的产学研合作项目。担任多媒体领域知名期刊 ACM TOMM 编委, 多个顶级会议的领域主席, 中国图像图形学会青工委副秘书长、上海市计算机学会计算机视觉专委会副秘书长。



刘 朋

刘朋, 博士, 云南大学信息学院讲师。主要从事深度学习相关理论与应用研究, 核心聚焦于医学影像智能分析、轻量化模型构建与多源图像融合。近五年已在 Information Processing & Management(IPM)、Engineering Applications of Artificial Intelligence (EAAI)、Knowledge-Based Systems(KBS)、IEEE Transactions on Computational Social Systems(TCSS) 等 JCR 一区 / 二区期刊发表 SCI 10 余篇。担任 IEEE Transactions on Medical Imaging、Medical Image Analysis、Pattern Recognition、IPM、KBS 等期刊审稿人。参与多项省部级课题研究: 国家自然科学基金“云南东巴绘画视觉解析及描述研究”; 云南省教育厅基础研究项目“面向退化图像的神经辐射场 3D 视点间场景重建”。

论坛讲者



魏 坤

魏坤, 西安电子科技大学副教授, 硕士生导师。研究多模态机器学习、持续学习、多智能体等相关方向, 发表 CCF-A 类会议 /SCI 一区期刊 40 余篇, 是 TPAMI、IJCV、CVPR、ICML 等 20 余种 CCF-A 类会议 /SCI 一区期刊审稿人。

报告题目: “快思慢想”: 类脑持续学习理论与方法

报告摘要: 人类与生俱来具有终身不断获取、整合和迁移知识的能力, 这种学习能力被称之为持续学习。在机器学习领域, 特别是深度学习模型提出以来, 持续学习致力于解决模型训练的一个普遍缺陷 -- 灾难性遗忘, 即在新任务上训练时, 在旧任务上的表现通常会显著下降。本报告以人脑学习与记忆机制为引入, 分析探讨现有增量学习主流范式, 并提出增量学习未来发展方向。



孙 元

孙元, 主要研究方向为多模态智能与 AI for PDEs。近年来, 共发表学术论文 50 余篇, 其中以第一作者或通讯作者在 TIP、TKDE、CVPR、ICML 等人工智能领域中科院一区 and CCF-A 类会议上发表论文 30 余篇, 入选 ESI 高被引共 2 篇, 入选 Spotlight/Highlight 共 4 篇。入选四川大学“海纳博士后”资助计划 (15 名), 获得中国电子学会电子信息教学成果二等奖 (序 2)、中国自动化学会科技进步奖二等奖 (序 5)、三等奖 (序 4)、中国兵装集团科技进步二等奖 (序 5)、中国人工智能学会教学成果一等奖 (序 4)、中国自动化学会高等教育教学成果三等奖 (序 5)、四川省计算机学会年度优秀学生论文、成都市技术能手称号。主持四川省自然科学基金项目、美的合作开发项目等共计 6 项, 授权发明专利共 10 余项。担任 Pattern Recognition 的副编辑 (AE), Journal of Artificial Intelligence & Control Systems 期刊青年编委, 多次担任 CCF-A 类会议 NeurIPS/ACL/IJCAI 的领域主席 (AC)/ 高级程序委员 (SPC), 长期担任 ICLR、CVPR、ICML、TPAMI、TIP 等 30 余个顶级期刊 / 会议审稿人。

报告题目：物理信息驱动的机器学习方法

报告摘要：物理信息驱动的机器学习 (Physics-Informed Machine Learning, PIML) 通过将物理规律的先验知识与数据驱动方法深度融合，在偏微分方程 (PDEs) 求解中展现出重要潜力。作为其中的代表方法，物理信息神经网络 (Physics-Informed Neural Networks, PINNs) 虽然已取得较好的应用效果，但其训练过程中仍面临若干关键挑战。一方面，现有 PINN 方法通常对不同难度的采样点采取统一处理策略，尤其忽视了物理边界等关键区域样本的学习差异，容易使模型陷入不良局部最优并导致训练过程不稳定，从而产生预测不平衡问题。另一方面，PINN 损失函数由内部区域、边界条件和初始条件等多类残差共同构成，这些残差在训练中往往存在显著的量级差异与病态梯度冲突，进而引发收敛缓慢、精度受限等问题。针对上述瓶颈，本报告聚焦于提升 PINN 鲁棒性与稳定性的关键技术，围绕不平衡预测和病态梯度问题展开系统研究。报告主要介绍两方面内容：1) 认知物理信息神经网络，旨在增强模型对不同区域、不同难度区域的自适应感知与学习能力；2) 非平衡物理信息神经网络，重点研究如何缓解多目标损失之间的病态梯度优化问题。通过相关方法的分析与探讨，本报告期望为突破 PINN 在复杂 PDE 求解中的性能瓶颈提供新的思路与技术支撑。

**汪 昱**

汪昱，同济大学计算机科学与技术学院副教授，博士生导师，入选上海科技青年 35 人引领计划 (上海 35U35)、上海市启明星计划 (扬帆专项) 人才计划，长期专注于多模态视觉感知、理解与决策等领域，在多模态大模型、开放视觉感知、多模态视频理解与分析等前沿方向从事研究。近五年，以第一 / 通讯作者在 NeurIPS、ICML、CVPR、ICLR、AAAI、ECCV、IEEE TMM、IEEE TCSVT 等 CCF A 类 / 中科院一区国际顶级刊物上发表论文 30 余篇，申请 / 授权国家技术发明专利 20 余项。获 2025 年上海市科技进步二等奖，获人工智能国际顶级挑战赛 8 项冠军 (包括 CVPR 2026 和 CVPR 2025 六项、ACL 2026 一项和 ICDAR 2023 一项)、2 项亚军 (CVPR 2026 两项)。长期担任包括 CVPR、ICCV、ICLR、ICML、NeurIPS、ECCV、IEEE TMM、IEEE TCSVT 等国际顶级刊物的审稿人。主持新一代人工智能国家科技重大专项子课题、国家自然科学基金青年基金、上海市“科技创新行动计划”启明星计划扬帆专项、上海市关键技术研发计划“新一代信息技术”项目，核心参与国家重点研发计划项目、国家自然科学基金重点项目、面上项目 10 余项。

报告题目：语义驱动的弱监督视频理解

报告摘要：视频智能理解是计算机视觉与人工智能的核心研究方向，在城市感知、安防监控、智能交通等领域具备重要应用价值。当前主流全监督方法高度依赖帧级精细标注，面临标注成本高昂、复杂场景泛化能力不足等瓶颈，弱监督与稀疏标注范式已成为领域突破的核心路径。报告人长期深耕视频理解与分析领域，针对弱监督场景中标注稀疏带来的语义定位偏差、异常边界模糊等核心难题，构建了以语义知识为中心的视频理解理论体系。本报告将系统梳理弱监督视频理解的技术发展脉络，深入报告核心研究思路与关键突破，并探讨该方向在城市场景智能化治理中的落地路径与未来挑战。



包世龙

包世龙, 中国科学院计算技术研究所助理研究员, 主要研究方向为生成模型与多模态学习。已在 CCF-A 类国际期刊与会议发表论文 25+, 其中包括 IEEE TPAMI 9 篇 (IF:20.4); 先后入选北京市科协“高创计划”青年人才托举工程, 获中国科学院优秀博士学位论文奖、北京图象图形学学会优秀博士学位论文奖等荣誉, 并获得 ICCV、CVPR 等国际会议竞赛冠军 5 项; 主持或参与国家自然科学基金重大项目、国家自然科学基金青年科学基金项目 (C 类)、中国博士后科学基金特别资助等国家级科研项目。

报告题目: 浅谈多模态模型中的表征调和与学习

报告摘要: 大规模预训练赋予视觉、语言、音频等模态编码器较强的通用表征能力, 使预训练表征成为多模态融合、视觉表征增强与条件生成等任务的重要基础。然而, 预训练目标、表征粒度与任务需求之间的差异, 使其在迁移、组合与复用时仍面临若干适配挑战。本报告围绕“多模态表征调和与学习”展开, 结合我们近期在多模态学习、视觉表征增强与条件图像生成方面的研究, 探讨如何从优化机制层面识别并缓解表征信号间的不一致冲突。报告重点分析多模态融合中的模态竞争、预训练视觉表征中的目标冲突, 以及文本编码器偏差对生成公平性的影响, 旨在说明预训练表征的有效复用不仅依赖模型与数据规模, 也依赖面向下游任务的表征调和与适配。



夏俊

夏俊现为香港科技大学 (广州)、香港科技大学联署助理教授和博士生导师。博士毕业于浙江大学 - 西湖大学联合培养项目计算机科学与技术专业, 导师为讲席教授李子青 (Stan Z. Li, IEEE Fellow); 博士期间在中国电信人工智能研究院 TeleAI 担任见习研究员, 导师为李学龙教授。申请人在与本项目相关的研究工作中有着丰富的积累, 取得了国际领先的科研成果, 在 Nature Methods、ICML、NeurIPS、ICLR 等顶级期刊会议发表 50 余篇论文, 其中以一作和通讯身份发表 31 篇, 一篇一作论文入选 WWW 2022 Most Influential Papers。主持首批国家自然科学基金青年学生基础研究项目 (博士生)、首届中国电子学会 - 腾讯博士生科研激励项目 (全国 17 项) 等, 参与首批科技部新一代人工智能国家重大科技专项“首席科学家负责制”试点项目 (全国 3 项) 等国家重大战略科技项目, 担任 NeurIPS、ICLR、KDD、IJCAI 等 AI 顶级会议的领域主席或高级程序委员会委员, 担任 Nature Communications、TPAMI 等期刊审稿人。曾获 Rising Star in AI by KAUST、DAAD AINet Fellowship、Apple AI/ML Scholar Finalist、西湖大学校长奖章、浙江省优秀毕业生、国家奖学金等。

报告题目: SpectraAI: AI-Driven Molecular Identification and Discovery

报告摘要: Molecular identification and discovery play a pivotal role in biochemical analysis, environmental governance, customs inspection, and other critical fields. Spectroscopic instruments, including mass spectrometry, infrared spectroscopy, nuclear magnetic resonance (NMR) spectroscopy, and X-ray spectroscopy, are fundamental tools for exploring the microscopic molecular world. However, current spectroscopic data analysis methods for molecular identification and discovery primarily rely on database searching and human professional expertise. These approaches face inherent limitations: they struggle to identify and discover novel molecules beyond existing databases, while suffering from low analysis efficiency, poor accuracy, and high costs. To address these challenges, we have conducted systematic research, exploration, and system development across multiple dimensions: data collection and cleaning, the development of a family of spectroscopic large foundation models, the creation of AI Agents for spectroscopic analysis, algorithm design grounded in biochemical principles, and practical application deployment in biomedicine and environmental protection. In this talk, I will elaborate on the research framework of SpectraAI and its latest progress, while highlighting promising research directions for future.



梁 科

梁科, 国防科技大学计算机学院讲师, 硕导。主要研究方向为图机器学习、知识表征与推理、AI4Science。主持国家自然科学基金项目、湖南省自然科学基金项目B类(原省优青)等。在CCF-A发表论文30余篇, 任Pattern Recognition、Information Fusion等国际知名期刊AE, 以及ICML、CVPR等CCF-A类会议AC。获CCF自然科学一等奖、CAAI博士学位论文激励计划等奖项。相关工作有效的应用于军民各类实践, 取得了良好的示范应用效果。

报告题目: 结构化知识赋能类脑分析

报告摘要: 本报告聚焦结构化知识驱动的脑网络分析研究方法, 针对现有任务场景监督信息难获取、多源信息难协同、模型结果难解释三大挑战, 系统介绍本人利用在知识驱动的半监督脑网络数据增强、表型引导的多模态脑网络分析与因果知识增强的可解释脑网络分析三个方面的研究进展。相关工作巧妙利用知识图谱和表型数据等结构化知识, 实现了高质量数据增强、个性化多模预测和可解释疾病诊断, 展示了结构化知识在神经科学应用中的方法价值与潜力。



宋启亚

宋启亚，副教授，硕士生导师，湖南师范大学“世承人才计划”青年优秀人才，博士毕业于湖南大学，主要研究方向为多模信息融合、视频图像处理，近年来在IEEE CVPR、TNNLS、TGRS, SCIS, ACM MM等高水平学术期刊/会议上发表论文20余篇，ESI 高被引论文1篇，授权发明专利15项（其中一项完成转化）；主持或参与了包括青年科学基金、重点研发计划等多个国家级项目；先后获得中国图像图形学会博士学位论文激励计划提名等荣誉称号；在国际顶级会议ACL、ACM以及全球最大竞赛平台kaggle上举办的国际竞赛获得冠军4项、亚军2项；获得国际“互联网+”创新创业大赛全国银奖/铜奖等；长期受邀担任IEEE TPAMI、IJCV、TIP、IJCAI、ACM MM等国内外知名期刊/会议审稿人以及相关专委会执行委员。

报告题目：高光谱遥感信息处理：从数据融合到智能解译

报告摘要：随着高光谱遥感技术的飞速发展，我们获取地物光谱信息的能力呈爆炸式增长，为地物精细识别提供了丰富的数据支撑。然而，高光谱数据固有的高维度与复杂性，也为信息深度挖掘带来巨大挑战。本报告旨在探讨高光谱遥感信息处理中的数据融合与智能解译方法。首先分析高光谱数据的基本特性与核心难题，重点探讨如何将高光谱与多光谱、LiDAR等多源数据进行协同融合，弥补单一数据源的局限性。在此基础上，系统介绍从传统机器学习到深度学习的高光谱智能解译方法，包括光谱特征提取、地物分类与目标识别等关键技术，并结合典型应用案例展望其在精准农业、环境监测等领域的应用前景。



刘岱宗

刘岱宗，武汉大学武汉数学与智能研究院助理教授、研究员、博士生导师。2025年于北京大学王选计算机研究所取得理学博士学位，获评2025年北大优秀博士学位论文，入选2025年全球前2%顶尖科学家榜单。主要研究领域包括多模态学习、三维具身感知、多模态大模型对抗安全，以第一/通讯作者发表TPAMI、NeurIPS、ICML、CVPR等CCF-A类论文30余篇，谷歌学术引用3800余次。长期担任TPAMI、TIP、TIFS、CVPR、ICCV、ACL等多个国际顶级期刊、会议审稿人，担任NeurIPS、ICML、ICLR、IJCAI领域主席。

报告题目：多模态大模型的对抗安全问题研究

报告摘要：多模态大语言模型在视觉理解与智能交互领域取得了显著进展，但其安全性问题日益受到关注。本报告围绕多模态大模型的对抗攻击与安全风险展开，系统介绍从白盒攻击、灰盒攻击到黑盒攻击的发展历程，分析不同攻击场景下的关键技术与挑战。首先，介绍基于模型梯度信息的白盒攻击方法及其优化机制，分析其攻击能力与局限性；随后讨论兼顾攻击效果与实际可行性的灰盒攻击框架，重点介绍如何利用有限模型信息实现高效攻击；进一步探讨真实场景下的黑盒攻击问题，分析对抗样本跨模型迁移的关键挑战与技术路线。重点探讨对抗样本迁移性的提升方法，从数据泛化迁移逐步发展到模型迁移攻击，介绍跨模型、跨架构的对抗攻击机制及最新研究进展。最后，结合当前大模型的发展趋势，展望多模态智能系统面临的安全挑战与未来研究方向。



刁云峰

刁云峰，现任合肥工业大学计算机与信息学院院长助理，安徽省人工智能学会安全人工智能专委会秘书长，CAAI 青工委专委会委员、CCF 多媒体专委会执行委员。研究方向为人工智能安全，媒体内容安全等。目前已在 CCF-A/ 中科院 1 区 Top 期刊发表论文 30 余篇。担任 CCF-A 期刊 TIFS 和 SCI 期刊 DI 和 IJAACS 编委，多个国际会议程序委员会委员和 Workshop 主席，并担任 20 余个国内外知名期刊和会议审稿人，获 ICML Gold Reviewer Award。主持国家自然科学基金、JG 项目、校企横向课题、国家重点实验室开放课题等多项。成果在蚂蚁集团和 JG 单位等得到应用，构建的 AIGC 安全数据集 2024-2026 连续三年入选外滩大会、IJCAI AIGC 安全攻防全球竞赛官方数据集，获 IJCAI Workshop 杰出贡献奖、隐者联盟 2025 年度杰出论文、IROS RODGE Workshop 最佳论文奖、“挑战杯”全国一等奖、全国大学生创新创业实践联盟指导教师优秀事迹等。

报告题目：复杂对抗环境下的视觉生成式人工智能安全风险治理

报告摘要：随着视觉生成式人工智能快速发展，深度伪造、隐私泄露、恶意内容生成与对抗规避等风险日益突出，已对国家安全、社会稳定和公众信任造成挑战。现有 AIGC 安全防护方法往往依赖静态、理想化场景假设，未考虑真实复杂场景中攻击策略的多样性及对抗攻防过程的自适应演化特性，难以应对真实环境中的跨模态伪造、对抗扰动、模型滥用和内容传播风险，防护性能易出现退化甚至失效。本报告面向复杂对抗环境下的视觉生成式人工智能安全防护问题，从数据层、模型层和内容层分析 AIGC 安全风险与防护瓶颈，介绍多模态伪造数据构建、鲁棒鉴伪、生成模型遗忘学习与对抗防御等研究进展，探索内生安全与外生防护协同的全链条治理思路，为构建安全、可信、可控的视觉 AIGC 提供参考。

讲习班

空间基础模型

时间：8月6日上午09:00-12:00 地点：102会议室

内 容

嘉 宾

主席：杨 欣（华中科技大学）、崔兆鹏（浙江大学）

基于4DGS的空间视频生成与编辑

周晓巍（浙江大学）

三维基础模型：从三维重建到通用图像几何表征

姚 遥（南京大学）

迈向具身自我进化

穆 尧（上海交通大学）

论坛主席



杨 欣

杨欣，华中科技大学教授、博导、国家级青年人才；主持国家重点研发计划、国基金联合重点、国际合作等国家级项目，以一作或通讯作者发表 TPAMI、IJCV、CVPR 类期刊会议论文 100 余篇，多次入选斯坦福全球前 2% 顶尖科学家榜单；获湖北省技术发明一等奖、湖北省教学成果特等奖、中国图象图形学会石青云女科学家奖、中国医学科技奖青年科学技术奖、ISMAR 最佳论文提名奖等奖励；指导学生获得挑战杯全国特等奖、中国国际互联网+金奖、CSIG 挑战赛年度总冠军等国内外权威竞赛奖项 10 余次；担任 IEEE-TMI、IEEE-TVCG 等多个期刊编委、IEEE BISP 技术委员会委员、CSIG 青工委副秘书长、CCF 多媒体技术专委会副秘书长等职务。



崔兆鹏

崔兆鹏，浙江大学计算机科学与技术学院“百人计划”研究员、博士生导师，国家级青年人才计划入选者。博士毕业于加拿大西蒙弗雷泽大学，曾任瑞士苏黎世联邦理工学院高级研究员。主要从事三维视觉与物理智能研究，研究内容包括三维重建与理解、三维运动规划、物理建模等。近年来在计算机视觉、机器人、计算机图形学等领域的顶级期刊和会议上发表论文60余篇，主持国家自然科学基金面上项目、专项项目等。现任PR、IEEE RA-L 等国际期刊编委，曾任 CVPR、ECCV 等会议领域主席、SIGGRAPH 程序委员会委员、ICRA 和 IROS 副编委等。曾获 ICRA 2020 机器视觉最佳论文提名、IROS 2021 安全、安保和救援机器人最佳论文提名、3DV 2024 最佳论文荣誉提名。

论坛讲者



周晓巍

周晓巍，浙江大学求是特聘教授，国家级高层次人才，国家重点研发项目首席科学家。研究方向主要为三维视觉与空间智能。近五年在相关领域的顶级期刊与会议上发表论文百余篇，多次入选CVPR最佳论文候选。曾获得浙江省自然科学一等奖，两次获得CCF优秀图形开源贡献奖，入选全球前2%顶尖科学家榜单、爱思唯尔中国高被引学者。培养学生获得CCF优博、国自然优秀学生基金。担任国际顶级期刊IJCV编委、顶级会议CVPR/ICCV领域主席，CSIG 三维视觉专委会副主任，曾任图形学与混合现实研讨会（GAMES）执行委员会主席。

报告题目：基于 4DGS 的空间视频生成与编辑

报告摘要：空间视频（体积视频）是未来数字媒体的重要形式，可广泛应用于远程交流、文化旅游、体育直播、仿真训练等领域，其技术本质是对现实场景动态光场的高保真重建与渲染。相较于传统的网格表达，高斯泼溅在表达能力与渲染真实感等方面展现了巨大的优势，为空间视频提供了新的技术途径，但在采集成本、重建质量、可编辑性等方面仍有局限。本报告将介绍我们在空间视频的生成与编辑等方面的系列工作，并探讨未来还需解决的技术难题。



姚 遥

姚遥，南京大学智能科学与技术学院准聘副教授、南京大学-影石创新智能影像校企联合实验室主任，国家级青年人才、苏州青年科学家。博士毕业于香港科技大学，曾任Altizure创始团队核心成员，2020年随公司收购加入苹果任高级研究员。主要研究方向为三维重建、三维生成与空间智能，代表工作包括MVSNet系列工作、Direct3D系列工作，谷歌学术引用8000余次，多次担任CVPR等会议领域主席，曾获华为火花奖、国际模式识别大会最佳学生论文奖。

报告题目：三维基础模型：从三维重建到通用图像几何表征

报告摘要：近年来，图像三维重建，特别是前馈式三维重建技术得到了快速发展，三维基础模型也朝着通用图像几何表征的方向发展。本次报告将梳理近二十年来的图像三维重建技术发展，包括经典三维重建管线(Classical MVS)、基于深度学习的三维重建网络(MVSNet series)、基于优化的三维重建方法(NeRF/3DGS series)以及基于 transformer 的三维基础模型(3r/VGGT series)。包括也将展望三维基础模型的未来发展方向，特别是其在通用图像几何表征与长时空间一致性记忆方面的重要应用。



穆 尧

穆尧，上海交通大学长聘教轨助理教授，入选国家级青年人才计划，上海市海外高层次青年人才，智源青年学者，EAI-100 2025学术新锐。博士毕业于香港大学计算机系，访学于苏黎世联邦理工学院、新加坡国立大学等。穆尧博士长期从事具身智能的基础研究，深耕多模态具身认知、具身世界模型和具身自主进化系统方向，担任了NeurIPS、ICLR、CORL等国际机器学习顶级会议的领域主席，中国计算机学会智能机器人专委会执委，中国图形图像学会三维视觉专委会专委，Valse执行领域主席，在IJRR、RSS、NeurIPS、ICML、CVPR等计算机领域国际顶级期刊和会议发表论文50余篇，谷歌学术引用超4000余次。曾荣获2026年ICRA最优论文奖提名、2026年ICRA ViTac 研讨会最优论文奖、2025年IROS最优论文奖提名、2024年ECCV协同具身智能研讨会最优论文奖、2024年中国自动化学会自主机器人研讨会奖学金（全国5人）、2021年IEEE ICCAS2020大会最优学生论文奖、IEEE IV2021最优学生论文提名奖等多项奖励。

报告题目：迈向具身自我进化

报告摘要：本报告围绕“具身智能如何从数据驱动走向自主进化”展开。当前具身智能正进入数据规模化与基础模型快速发展的阶段，但单纯依赖模仿学习难以覆盖复杂真实世界中的反事实场景与长尾失败。报告将介绍一种由生成式世界模型驱动的具身智能体自进化框架：通过逼真、多样化的可交互环境提供训练空间，通过可拓展的进展评判模型识别任务进度与失败原因，并结合具备高效学习能力的基础模型和强化学习方法，实现从失败发现、数据生成、策略优化到真实任务验证的闭环演进。报告将重点分享团队在多视角二维—三维联合世界生成、面向力触觉的精准推演、生成式资产构建、具身进展评判基础模型以及大规模强化学习基础模型等方面的研究进展，展示世界模型如何从“预测未来”进一步走向“驱动学习”，支撑机器人在复杂操作任务中持续提升泛化能力、交互能力与任务完成能力。

分论坛

大模型评测

时间：8月6日上午09:00-12:00 地点：101会议室

内 容

嘉 宾

主席：马 珂（中国科学院大学）、王子泰（中国科学院计算技术研究所）
冯 洋（中国科学院计算技术研究所）

大模型压缩加速与边缘算力应用

纪荣嵘（厦门大学）

面向长视频时空细粒度理解的大模型评测

秦 杰（南京航空航天大学）

从精细感知到精细操作

魏秀参（东南大学）

大模型安全评测与可信具身智能验证

叶 茫（武汉大学）

认知反演驱动的大模型安全评测

刘艾杉（北京航空航天大学）

论坛主席



马 珂

马珂，中国科学院大学电子电气与通讯工程学院副教授，博导，近五年在人工智能安全领域展开工作，建立有理论支撑的人工智能算法对抗学习框架。近五年共发表 CCF-A 类论文 20 余篇，其中以第一作者或通讯作者发表 CCF-A 类论文 8 篇，包括第一作者 IEEE TPAMI 3 篇。获中国电子学会优秀博士学位论文、中国科学院优秀博士学位论文、2020 年国际华人数学家联盟最佳论文奖 - 若琳奖、ACL 2025 SAC Highlights Award 以及 China MM 2025 Best Paper Award；获 CCF-A 类会议国际会议竞赛冠军 4 项。主持国家自然科学基金面上，科技创新 2030-“新一代人工智能”重大项目子课题，腾讯犀牛鸟专项研究计划和阿里巴巴学术合作计划。



王子泰

王子泰，中国科学院计算技术研究所特别研究助理，研究方向为机器学习与数据挖掘，包括但不限于复杂评估指标优化、长尾学习、开放域学习、大模型高效学习。已发表 CCF-A 类论文 20 余篇，其中 TPAMI/IJCV/ICML/NeurIPS/CVPR 论文 10 余篇，并有 8 篇入选 Oral/Spotlight，获 CVPR 国际竞赛冠军 2 项；入选中国图像图形学会博士学位论文激励计划、中国科学院优秀博士学位论文；获博士后创新人才支持计划、国家自然科学基金青年科学基金 C 类、北京市自然科学基金 - 海淀原始创新联合基金培育项目等项目资助。



冯 洋

冯洋, 中国科学院计算技术研究所研究员、博士生导师、“新百星人才计划”入选者, 研究方向为人工智能和自然语言处理。担任国际计算语言学学会 (ACL) 秘书长、中国中文信息学会青年工作委员会副主任以及 ACL/EMNLP/IJCAI 等国际会议高级领域主席。主导研发了 BayLing(百聆) 系列大模型, 先后获得 ACL2019 唯一最佳长文奖、中国人工智能学会第七届理事会“杰出贡献奖”、CCF 自然语言处理专委会“青年新锐奖”、中文信息学会“钱伟长中文信息处理科学技术奖 - 汉王青年创新奖”一等奖等。

论坛讲者



纪荣嵘

纪荣嵘, 厦门大学教授、博士生导师, 厦门大学校长助理、人工智能研究院负责人、多媒体可信感知与高效计算教育部重点实验室主任, 主要研究方向为计算机视觉、多媒体技术和机器学习。研究侧重点是开发大数据驱动的计算机智能处理系统, 进行视觉场景理解、视觉行为分析、多媒体内容理解、及视觉内容检索。当前的研究兴趣还包括面向视觉搜索的紧凑描述子、社会多媒体大数据分析、场景深度推理、人群行为理解、以及遥感图像分析等。百余篇科研论文发表于国际期刊 PAMI、IJCV、TIP 和国际会议 CVPR、ICCV、IJCAI、AAAI、ACM Multimedia 上, 曾获得 2007 年微软学者奖、2011 年 ACM Multimedia 最佳论文奖、2015 年黑龙江省自然科学二等奖, 2016 年教育部技术发明一等奖, 2018 年福建省科技进步一等奖, 2020 年福建省科技进步一等奖, 2022 年霍英东青年科学奖。连续两次担任国家科技创新 2030 “新一代人工智能” 重大项目首席科学家 (2022、2026), CVPR 2021-2022、ACM Multimedia 2018-2022 领域主席, 《中国科学: 技术科学》副编辑、Valse(视觉与学习青年学者研讨会) 2017 大会主席、IEEE MMSP 2015、ACM ICMR 2014、IEEE VCIP 2014、ACM MMM 2015、IEEE ISM 2015 等国际会议本地专题领域主席、AAAI 2015、CVPR 2013、ICCV 2013、ACM Multimedia 2010-2015 等四十余个国际会议的程序委员会委员、IEEE Multimedia Communication 技术委员会 Visual Analysis and Content Management for Communications 专家组主席 (2014 至今)。曾任中国图像图形学会青工委副主任、任中国图像图形学学会学术工委副主任、中国自动化学会模式识别与机器智能专业委员会副主任、中国自动化学会粒计算与多尺度分析专委会副主任、中国计算机学会计算机视觉专委会常务委员、中国计算机学会学术工委委员、教育部电子信息类教指委人工智能专业建设咨询委员会委员。2005 年于哈尔滨工程大学获得本科学位, 2011 年于哈尔滨工业大学获得博士学位, 2011 到 2013 年于美国哥伦比亚大学任博士后研究员。2013 年加盟厦门大学信息学院人工智能系, 2014 年获国家优秀青年科学基金, 2017 年国家级青年人才获得者, 2020 年获国家杰出青年科学基金, 2024 年入选享受国务院政府特殊津贴专家。

报告题目：大模型压缩加速与边端算力应用

报告摘要：近年来大模型技术取得了突破性发展和广泛应用。在应用层面上，大模型技术的发展呈现了基于集群云服务的通用强智能系统服务与面向边侧部署的低开销高效应用并重的特点。在线上服务如多模态问答等场景，大多采用强智能、多模态的大模型，而在边侧高效低开销应用，如无人设备巡检任务中，边侧设备多部署具备轻量化、快响应能力的低开销边侧模型。边侧应用对低资源、高效响应的需求引发了近年来对大模型压缩与加速技术的研究热潮。本报告将从模型结构与训练优化、模型权重压缩、动态显存优化三个维度，从怎么训得快且推得快（结构）、怎么存得小（权重）、怎么跑得省（显存）三个阶段介绍大模型压缩与加速技术，覆盖模型结构优化、权重剪枝量化、KV 缓存压缩三个方向内容。



秦 杰

秦杰，南京航空航天大学人工智能学院副院长，教授、博士生导师，脑机智能技术教育部重点实验室副主任，国家级青年人才，江苏省杰青，中国图学学会数字媒体专委会副主任。主要从事计算机视觉、具身智能和多媒体等领域研究，重点关注面向开放动态环境的具身智能感知。发表论文 100 余篇，其中 CCF A 类论文 60 余篇，Google Scholar 引用 7300 余次，H 指数 42。获中国图像图形学会自然科学奖二等奖（排名 1）、CCF A 类会议 ACM MM 2023 唯一荣誉提名奖（1/3072）等。担任 CCF A 类期刊 IEEE TIP 编委、IJCV 客座编委，CCF B 类期刊 Neural Networks 编委，CCF A 类会议 NeurIPS、ICML、ICLR、AAAI、ACM MM 领域主席等。主持国家自然科学基金联合重点、面上项目，新一代人工智能国家科技重大专项课题，江苏省杰出青年基金等项目。

报告题目：面向长视频时空细粒度理解的大模型评测

报告摘要：近年来，视频多模态大语言模型在视频理解、内容生成与复杂推理等任务中取得显著进展，相关评测方法也逐步从单一任务拓展为多维能力刻画。然而，随着长视频分析与视频在线交互需求的不断增加，大模型在长时序、多事件场景下的时空细粒度理解能力仍缺乏系统评测。本报告围绕跨片段语义关联的长视频检索、细粒度时序定位的 AI 生成片段取证，以及实例级空间持续感知的在线目标理解三项任务展开分析，揭示了现有大模型的长视频时空细粒度理解能力与真实应用需求之间的差距，并为后续评测体系构建与模型能力提升路径提供启示。



魏秀参

魏秀参, 东南大学青年首席教授、博士生导师, 国家基金委青年科学基金项目 B 类(原优青)获得者、国家重点研发计划首席青年科学家、江苏省杰青。主要从事计算机视觉、机器学习和机器人方面研究, Google Scholar Citations 逾 10000 次, 目前担任 IEEE TPAMI/TIP/TMM/TCSVT 编委, NeurIPS/CVPR/ICCV/AAAI/IJCAI 等领域主席, 并任中国图像图形学会(CSIG)理事、Fundamental Research/ 电子学报(英文版)青年编委、CSIG 青托俱乐部主席、CSIG 青工委副秘书长等。曾入选 CSIG 青年科学家奖、吴文俊人工智能优秀青年奖等。主编教材《解析深度学习》已重印 8 次并入选“十四五”国家重点出版物出版规划、中国计算机学会推荐教材和省部级重点教材等。

报告题目：从精细感知到精细操作

报告摘要：自然界智能的形成普遍遵循“先感知、后操作”的规律：人类幼儿依靠不断提升的视觉 - 触觉精细分辨能力实现稳定抓握；黑猩猩等灵长类则通过精准感知目标形状与材质完成工具取食等复杂操作——精细感知因此成为可靠动作生成的基础。在具身智能系统中，这一机制同样适用。基于我们在细粒度视觉感知与理解等方面的研究积累，本报告将介绍如何从精细感知过渡到精细操作框架的构建，并展示在精细操作数据生成、基于检索的人机交互及精细抓取等方面的最新进展，为推动感知智能向操作智能的发展提供参考。



叶 茫

叶茫, 武汉大学计算机学院副院长、智能科学系主任、国家高层次青年人才, 科睿唯安全球高被引科学家。长期从事多模态计算等领域研究, 以第一 / 通讯作者发表 CCF-A 类论文 100 余篇, 谷歌学术引用 1.9 万余次。担任 CCF-A 类 IEEE TIP、IEEE TIFS 等期刊编委, CVPR、ICLR、ICML 等顶级会议领域主席等学术职务。主持国家重点研发计划、国自科 - 香港联合基金等科研项目。牵头获中国图像图形学会科技进步一等奖、吴文俊人工智能自然科学二等奖。

报告题目：大模型安全评测与可信具身智能验证

报告摘要：多模态大模型逐渐从视觉语言理解走向具身智能执行，模型安全不再只取决于最终输出是否合规，还需要关注其在复杂场景中的风险识别能力与行动可靠性。本报告将围绕大模型安全如何被测评和验证展开，讨论跨模态输入、长链推理和具身执行中可能出现的安全问题，并介绍我们在大模型安全推理对齐、具身智能模型隐蔽对抗威胁防御方面的探索。最后展望从安全对齐到可信部署的未来方向与机遇。



刘艾杉

北京航空航天大学计算机学院副教授，北京智源青年学者、华为卓越青年学者，长期从事安全可信人工智能、人机价值对齐的研究。累计在 CCF-A 类国际顶会顶刊发表论文 100 余篇，谷歌学术引用 5300 余次，编制国家标准 20 余项，填补我国人工智能算法模型质量标准的空白，获全国信标委人工智能分委会“优秀个人”。率先创办 AdvML CV 系列智能安全国际研讨会议并连续 5 年担任 CVPR、ICML 等 CCF-A 类国际会议的论坛主席、领域主席。研究成果获 CCF-A 类会议 ACM MM 2025 最具影响力演示奖、ACM CCS 2024 杰出研究成果奖、ICML 2025 MAS 杰出论文奖等国内外重要奖励 10 余项，以及省部级科技奖进步一等奖 1 项、二等奖 1 项。

报告题目：认知反演驱动的大模型安全评测

报告摘要：大模型安全是国家战略与产业落地的核心挑战。本报告聚焦认知反演驱动的大模型安全评测新范式，系统性揭示了大模型从“内生安全”（数据投毒、对抗攻击、模型后门）到“衍生安全”（深度伪造、越狱攻击、认知偏见）的多维风险。报告提出三层认知反演评测方法：指令空间中基于视觉感知欺骗的文本扰动与跨模态有害指令基准；意图空间中上下文诱导的后门投毒、渐进式多轮越狱攻击及交互式毒性测量；物理空间中针对具身智能体的时空对抗、环境幻觉攻击及场景感知动态对抗样本生成。在此基础上，构建了“安全沙盒”体系，包括具身智能体安全基准 AgentSafe、GUI 智能体风险诊断 WebSecureArena、自动驾驶仿真推演平台万象，并完成多款主流车型的实车验证，相关成果已支撑多项国家级平台、重大项目，并共建智源 FlagSafe 平台。本报告旨在推动大模型安全从“测得准”向“安全可信防御与自主演进”跃迁，为构建可信人工智能提供系统化技术路径。

分论坛

视觉认知与情感计算

时间：8月6日上午09:00-12:00 地点：201会议室

内 容

嘉 宾

主席：杨 云（云南大学）、胡 斌（北京理工大学）

心理生理计算与精神障碍诊疗	胡 斌（北京理工大学）
从听懂指令到懂你所想：构建机器心智理论的探索与实践	陈恩红（中国科学技术大学）
AI情感陪伴：从语言共情到务实守护	秦 兵（哈尔滨工业大学）
脑连接组深度表征与精准精神疾病解决方案	马 婷（哈尔滨工业大学）
前因聚焦式路怒症多感官刺激调节研究	谷 雨（电子科技大学）
面向复杂场景的指代表理解技术	蔡 毅（华南理工大学）
集成智能关键技术及其在心理健康领域创新应用	杨 云（云南大学）

论坛主席



杨 云

杨云，云南大学软件学院 / 人工智能学院院长，教授，博导。中国计算机学会(CCF)杰出会员，中国人工智能学会(CAAI)高级会员，入选国家级“网信优秀人才”、云南省“云岭学者”专项、省“中青年学术技术带头人”、省“杰青”。担任云南省软件工程重点实验室主任，云南大学软件工程学科与人工智能专业负责人，CAAI 智能融合专委会常委、智慧医疗专委会执委，CCF 人工智能与模式识别专委会执委、开源发展委员会执委、数字医学分会执委，国际学术期刊 Neural Networks 和 Complex&Intelligent Systems 编委。代表性研究成果发表于 TCYB、TNNLS、TIP、TMM、TKDE、TSMC、TCSVT、TII、TAFCC 与 CVPR、SIGKDD、AAAI、MICCAI 等知名学术期刊与会议，获“中国侨界贡献奖”二等奖1项，“云南省自然科学奖”二等奖2项。转化多项发明专利与软著，成果广泛应用于医疗与康养领域，研发多款智慧心理健康产品并产业推广。



胡 斌

胡斌，国家海外高层次人才入选者，973 首席科学家，IEEE/IET/AAIA Fellow, IET Fellow Assessor&Fellowship Advisor，教育部计算机学科教指委委员，科技委委员，IEEE Transaction on Computational Social Systems 主编，国际社会神经科学中国分会副主席等；获 2019 年度中国专利金奖，2018 年度国家技术发明奖二等奖、2016 年度教育部技术发明奖一等奖、2022 年度甘肃省自然科学特等奖，2023 年度全国创新争先奖，(均为第一完成人)等，2020 年全国先进工作者；入选科睿唯安(Clarivate)高被引科学家、世界前 2% 科学家(Stanford and Elsevier)和 0.05% 高水平学者(ScholarGPS)。

论坛讲者



胡 斌

胡斌, 国家海外高层次人才入选者, 973 首席科学家, IEEE/IET/AAIA Fellow, IET Fellow Assessor&Fellowship Advisor, 教育部计算机学科教指委委员, 科技委委员, IEEE Transaction on Computational Social Systems 主编, 国际社会神经科学中国分会副主席等; 获 2019 年度中国专利金奖, 2018 年度国家技术发明奖二等奖、2016 年度教育部技术发明奖一等奖、2022 年度甘肃省自然科学特等奖, 2023 年度全国创新争先奖, (均为第一完成人)等, 2020 年全国先进工作者; 入选科睿唯安(Clarivate)高被引科学家、世界前 2% 科学家(Stanford and Elsevier)和 0.05% 高水平学者(ScholarGPS)。

报告题目: 心理生理计算与精神障碍诊疗

报告摘要: 近年来全球精神健康问题日益凸显, 据世界卫生组织统计: 全世界有9.7亿人患有精神障碍, 占全球总人口的13%。当前精神疾病诊断多以医生访谈和精神疾病诊断量表为主, 缺乏客观量化诊断指标; 治疗方式主要为药物治疗等方法, 但是药物治疗存在副作用大等问题。以人工智能和大数据为代表的前沿信息技术的飞速发展, 为精神疾病的诊疗提供了新契机, 推动精神疾病诊疗由症状描述型向数据驱动型发生转变, 提供了普适化, 精准化和个性化的精神疾病诊疗分析技术, 从而提高了精神疾病临床识别率和治疗有效率。本次报告将介绍人工智能应用于精神障碍诊疗的机遇与挑战。



陈恩红

陈恩红, 中国科学技术大学教授, 博导, IEEE Fellow, CCF、CAAI 会士, 认知智能全国重点实验室副主任, 安徽省计算机学会理事长, 国家级创新领军人才计划, 国家杰出青年基金获得者。主持了国家重点研发计划项目、国家自然科学基金重大仪器研制项目、国家自然科学基金区域联合重点项目等。主要研究方向是人工智能、数据挖掘、个性化推荐系统等, 谷歌学术引用 49000 余次, H-index 108, 获教育部、中国电子学会、中国计算机学会自然科学一等奖, 吴文俊人工智能科技进步一等奖, 及 KDD、ICDM 等的最佳论文 / 学生论文、2024 APWeb-WAIM Test of time 奖等。

报告题目: 从听懂指令到懂你所想: 构建机器心智理论的探索与实践

报告摘要: 人机共生系统中, 智能体亟需从指令执行者进化为“懂你所想”的合作伙伴。然而, 现有智能体难以精准理解人类在多模态交互中的隐性意图。本报告聚焦“心智理论”这一核心能力, 系统介绍我们在三个层面的探索与实践: 通过思维链增强的多模态推理, 从碎片化信息中解析用户意图, 让机器理解人类真实需求; 通过微反应感知与识别技术, 从面部细微运动中揭示隐藏情绪, 让机器洞察人类真实感受; 通过情境感知融合与用户中心交互适配, 依托记忆与反思对齐情绪与偏好, 让机器给予人类可靠共情回应。最后, 浅谈未来三个重要的研究方向。



秦 兵

秦兵，哈尔滨工业大学长聘教授，社会计算与交互机器人研究中心主任。中国中文信息学会常务理事，中国中文信息学会情感计算专委会主任，科技部科技创新 2030-“新一代人工智能”重大项目管理专家组专家。研究方向：自然语言处理及大模型技术等。发表论文近 200 篇，累计被引 3 万余次。连续多年入选爱思唯尔高被引学者榜单，入选 2020 年福布斯中国科技女性榜及 2025 年全球顶尖 2% 科学家。获教育部科学研究优秀成果奖工程技术一等奖，黑龙江省科技进步一等奖、CCF 科技成果奖科技进步一等奖、钱伟长中文信息处理科学技术奖一等奖等奖项。

报告题目：AI情感陪伴：从语言共情到务实守护

报告摘要：当前，AI情感陪伴正在从简单的语言回应，走向面向真实需求的长期陪伴与主动守护。然而，真正有效的情感陪伴并不只是“说得共情”，而是需要系统具备持续理解用户状态、形成个性化支持策略，并通过合适方式转化为实际行动的能力。本报告围绕“感知—规划—表达”三条主线，介绍我们在AI情感陪伴方向的最新研究进展。具体而言，在感知层面，报告将介绍多模态情绪理解与隐性需求识别方法，使AI能够捕捉用户动态情绪变化；在规划层面，讨论如何结合长期记忆与用户画像，形成更稳定、更个性化的陪伴策略；在表达层面，探索语言回应、工具调用、任务协助与具身行为等多种方式，使情感理解真正转化为可执行、可感知、可持续的务实守护。报告将结合青少年心理健康、老年陪护等高需求场景，展示AI情感陪伴系统如何从“语言共情”迈向“行动支持”，并讨论其在安全性、个性化、长期交互与可信部署方面面临的关键挑战。



马 婷

马婷，博士毕业于香港中文大学生物医学工程专业，广东省青年拔尖人才、深圳市基础研究杰青、深圳市地方领军人才。长期致力于神经信息计算、脑机接口、神经 / 精神疾病标志物、智能医学技术等研究，在重大脑疾病早期影像标志物发现与提取技术方面取得了突破性成果，被 Nature Reviews Neurology 特别评论为“重要神经退行疾病早期标志物”；所提出的“神经指纹”技术已成功技术转化，获批医疗器械注册证，并在超过 400 家医院应用；提出首个脑连接组基础模型，收到国际广泛关注与应用。科研成果在 Lancet ID、ADJ、Chem 等高水平期刊 / 会议发表论文 180 余篇。主持科技部重点研发课题 2 项、自然科学基金面上项目 1 项、以及其它省部级项目超过 2000 万元。任 IEEE 生物医学工程学会深圳分会主席、广东省预防医学会人工智能专委会副主任委员、广东省医学会人工智能分会常委、深圳市医院协会人工智能分会副主任委员、深圳市计算机协会人工智能分会常委等。

报告题目：脑连接组深度表征与精准精神疾病解决方案

报告摘要：精神疾病作为全球重大公共卫生问题，其高度临床异质性、客观生物标志物缺失及神经调控治疗反应的个体差异严重制约了精准诊疗发展，脑连接组学与深度学习技术的结合为从海量脑影像数据中挖掘疾病特征、揭示神经机制提供了全新路径。本报告系统介绍本课题组围绕“脑连接组深度表征-影像标志物发现-临床转化应用”主线在精神疾病脑网络分析与精准干预领域取得的系列进展，针对传统方法难以捕捉脑网络复杂时空特征和多中心数据异质性的问题，提出多元距离连接组网络、多尺度上下文Mamba等一系列深度学习模型，实现了多种精神疾病的高精度跨中心辅助诊断且效果显著优于现有主流方法，同时结合可解释性人工智能与孟德尔随机化分析，系统揭示了核心脑网络在精神疾病中的共享与特异性异常，建立了从结构连接组到疾病风险的因果关联，并针对非侵入性神经调控治疗的机制不明和疗效预测难题开展相关临床研究，首次阐明了动态脑网络状态重构在特定精神疾病治疗中的关键作用，发现了基于动态网络转换的治疗反应预测标志物。本研究构建了从脑连接组深度表征到临床转化的完整技术体系，为精神疾病的客观诊断、机制解析和精准干预提供了新的理论基础和技术支撑，未来将进一步整合多模态数据，推动跨诊断精神疾病生物分型与个体化治疗的发展。



谷 雨

谷雨，电子科大“百人计划”研究员、博导，毕业于中科大少年班，先后于微软亚研院、日本筑波大学、日本国立情报学研究所、合肥工业大学从事科研与教学工作。目前研究兴趣人机情感交互，发现并建模路怒中“slow-rise, fast-decay”现象（二阶阻尼震荡模型），提出了前因聚焦式的多感官刺激安抚等方法，在《IEEE 情感计算汇刊》、ACM 人机交互大会等期刊与会议上发表论文 120 余篇，获 IEEE Andrew P.Sage (序 1) 等国际学术奖 6 次；入选四川省高层次人才计划、斯坦福全球 Top 2% 科学家终生影响力榜单；担任《IEEE 情感计算汇刊》编委、中国人工智能学会情感智能专委会副秘书长。

报告题目：前因聚焦式路怒症多感官刺激调节研究

报告摘要：路怒症构成交通安全一大隐患，是一个跨交通、心理学与计算机的交叉学科难题。当前研究范式专注驾驶员愤怒情绪，侧重反应抑制（“治已病”），缺乏预防能力（“防未病”）。为此，我们提出前因聚焦式多感官刺激调节范式，聚焦路怒诱因而非路怒反应，以切断路怒生成链（诱因-注意力-认知-反应），在反应前就介入消弭。本报告将介绍该范式的验证工作：在发现路怒阻尼震荡特征（CHI'26）的基础上，构建并验证了嗅觉多感官刺激循因调节的有效性（CHI'26 Poster），并基于VLM模型构建了实时的诱因发现算法（ICME'26）。



蔡 毅

蔡毅教授现担任华南理工大学软件学院院长，大数据与智能机器人教育部重点实验室主任、华工-金山办公软件联合实验室主任。现为中国计算机学会杰出会员、自然语言处理专委会委员、数据库专委会委员；中国中文信息学会语言与知识计算专委会委员、情感计算专委会执行委员、社交媒体计算专委会委员；广东省大数据专委会副主任；标准规范知识图谱联盟理事成员，广东省特支计划青年拔尖人才。国际学术期刊 IEEE/ACM TASLP 副主编，CMC-Computers, Materials & Continua 副主编，Natural Language Processing Journal 编委。在包括 IEEE TKDE、IEEE TVCG、IEEE/ACM TASLP、IEEE TMM、IEEE Transactions on Affective Computing 等顶级国际期刊和 AAAI、ACL、ACM MM、EMNLP、CIKM、COLING 等顶级国际会议上录用 / 发表论文 230 多篇。

报告题目：面向复杂场景的指代表达理解技术

报告摘要：指代表达理解是自然人机交互领域的重要研究课题，旨在根据人类的自然语言表达在视觉场景中定位到所指代的物体。作为智能机器人抓取操作与人机交互的基础能力，指代表达理解研究对于提升机器人场景适应性和交互自然性具有重要意义。在复杂的现实场景中，人类会根据语言习惯利用丰富的视觉内容、知识常识等信息来指代物体，并存在视觉误差及视角差异等挑战。针对上述挑战，本报告汇报课题组在指代表达理解领域的研究工作，介绍面向场景文本的指代表达理解、知识辅助的指代表达理解等方法，提升机器人在复杂场景中理解人类指代意图的能力。



杨 云

杨云, 云南大学软件学院 / 人工智能学院院长, 教授, 博导。中国计算机学会(CCF)杰出会员, 中国人工智能学会(CAAI)高级会员, 入选国家级“网信优秀人才”、云南省“云岭学者”专项、省“中青年学术技术带头人”、省“杰青”。担任云南省软件工程重点实验室主任, 云南大学软件工程学科与人工智能专业负责人, CAAI 智能融合专委会常委、智慧医疗专委会执委, CCF 人工智能与模式识别专委会执委、开源发展委员会执委、数字医学分会执委, 国际学术期刊 Neural Networks 和 Complex&Intelligent Systems 编委。代表性研究成果发表于 TCYB、TNNLS、TIP、TMM、TKDE、TSMC、TCSVT、TII、TAFRC 与 CVPR、SIGKDD、AAAI、MICCAI 等知名学术期刊与会议, 获“中国侨界贡献奖”二等奖 1 项, “云南省自然科学奖”二等奖 2 项。转化多项发明专利与软著, 成果广泛应用于医疗与康养领域, 研发多款智慧心理健康产品并产业推广。

报告题目: 集成智能关键技术及其在心理健康领域创新应用

报告摘要: 心理健康问题是全球面临的重大挑战。传统诊疗方式高度依赖主观量表与医生访谈, 在客观性、效率与大规模筛查可及性方面存在明显局限, 亟需新技术手段的介入。集成智能通过多源数据融合、多模型协同与知识规则驱动, 恰好契合心理健康问题多源、动态、隐匿的特点, 为客观评估与智能干预开辟了新路径。本报告围绕“关键技术突破—创新应用落地”主线, 系统介绍在多模态时序表征、动态尺度自适应、大模型语义引导、先验知识迁移以及临床知识对齐等方面的研究进展, 并展示从智能筛查、精准评估到个性化干预的系列创新应用成果, 探讨集成智能在心理健康领域的实践价值与发展前景。

分论坛

遥感影像处理

时间：8月6日上午09:00-12:00 地点：207会议室

目 录

嘉 宾

主席：孙 显（中国科学院空天信息创新研究院）、邹征夏（北京航空航天大学）
段普宏（湖南大学）

资源受限遥感影像处理

方乐缘（湖南大学）

面向卫星高光谱影像的潜在空间自监督基础模型

谭 琨（华东师范大学）

文本辅助的遥感影像智能解译

唐 旭（西安电子科技大学）

物理机制嵌入的极化SAR舰船检测网络

殷君君（北京科技大学）

多源遥感图像智能解译

董燕妮（武汉大学）

生成式遥感图像处理

曹相湧（西安交通大学）

论坛主席



孙 显

孙显，中国科学院空天信息创新研究院研究员，国家级领军人才，IET Fellow。长期从事航天遥感图像智能解译方法研究与应用实践。主持国家级科研任务十余项，发表 SCI 论文 100 余篇，出版专著 2 部。获国家级和省部级奖励 5 项。担任 IEEE GRSM、GRSL、JSTARS 等多个国际期刊副主编 / 编委。任中国指控学会副秘书长、中国图像图形学会遥感专委会秘书长。



邹征夏

邹征夏，北京航空航天大学教授 / 博导，国家级青年人才，研究方向为遥感图像处理。在 Proceedings of the IEEE(最受欢迎论文)、Nature 子刊(封面论文)、TPAMI(国家自然科学基金委报道)、TIP、TGRS、CVPR、ICCV 等重要期刊和会议发表论文 100 余篇，谷歌学术论文引用 1.4 万次，担任遥感权威期刊 TGRS 和图像处理权威期刊 TIP 编委；主持国家级和省部级项目十余项，研究成果在航天国防、灾害应急与深空探测领域实现工程化部署与应用，获麻省理工科技评论 AI100 青年先锋奖。



段普宏

段普宏, 湖南大学副教授, 博士生导师, 主要从事高光谱遥感图像处理研究, 主持国家自然科学基金 2 项、国家重大工程子课题等。在 IEEE TGRS、IEEE TNNLS、IEEE TIP 等领域权威期刊上发表 SCI 论文 70 余篇, 入选了湖湘青年英才、中国电子学会青年托举人才等; 获 IEEE 地球科学与遥感协会数据融合大赛冠军、CVPR 图像生成表征学习算法挑战赛冠军、ISPRS 多模态遥感应用算法智能解译大赛冠军等。

论坛讲者



方乐缘

方乐缘, 湖南大学二级教授, 国家杰青, 教务处副处长兼学生创新创业中心主任、中国图像图形学会图像应用与系统集成专委会主任。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。获国家自然科学基金二等奖 2 项 (排 2 和 4)、吴文俊人工智能科学技术奖自然科学一等奖 1 项 (排 1)、湖南省自然科学一等奖 2 项 (排 2 和 3)、IEEE GRSS 最高影响力论文奖 (通讯作者) 等。担任 SCI 期刊 IEEE TIP、IEEE TNNLS、IEEE TGRS、中国科学: 信息科学、Fundamental Research 等期刊编委。研究成果在国际权威期刊和会议发表论文 200 余篇, 其 IEEE TPAMI、IJCV、TIP 等本领域顶级期刊论文 100 余篇, Google scholar 引用 2 万余次, ESI 高被引 25 篇, ESI 热点论文 4 篇。主持国家自然科学基金委杰青、优青、原创探索、联合重点、国家重点研发课题等项目。

报告题目: 资源受限遥感影像处理

报告摘要: 深度学习因其出色的性能在遥感影像处理分析中已被广泛应用。然而, 高性能深度学习模型严重依赖于海量数据、精确标注以及庞大算力。在星载/机载遥感实际应用中, 数据数量、标注质量、计算能力严重受限, 制约着智能模型在遥感领域的应用。本报告首先针对遥感影像数量和质量受限问题, 提出跨图像关联的半监督遥感图像处理方法, 摆脱了模型对原始海量高质量数据的依赖; 其次, 针对遥感影像标注困难、时间代价昂贵的问题, 提出点标签智能解译方法, 保持处理精度的同时, 显著降低标注代价; 最后, 针对遥感计算能力受限的问题, 提出一种高效的网络层压缩方法, 在不影响推理精度的同时, 大幅降低存储和计算代价。



谭 琨

谭琨, 华东师范大学空间人工智能学院教授、副院长, 入选国家高层次青年人才计划、上海市东方英才计划领军项目, 紫江优秀青年学者。担任自然资源部超大城市自然资源时空大数据应用重点实验室副主任, 遥感知名期刊 International Journal of Applied Earth Observation and Geoinformation 副主编。主持国家“慧眼行动”重点项目、国家自然科学基金面上、上海市重点项目等 30 余项国家省部级项目。近年来, 在 ISPRS P&RS、IEEE TGRS 等国内外期刊发表学术论文 100 余篇, 发表的学术论文他引 6000 余次, 连续入选斯坦福大学与 Elsevier 联合发布的全球前 2% 科学家榜单。

报告题目: 面向卫星高光谱影像的潜在空间自监督基础模型

报告摘要: 高光谱遥感影像蕴含丰富的光谱-空间信息, 是精细化地表观测与土地利用/土地覆盖制图的重要数据来源。然而, 不同卫星传感器在光谱覆盖范围、波段设置及观测条件上的差异, 使得高光谱基础模型在跨传感器、跨区域应用中仍面临泛化能力不足的问题。为此, 提出一种面向卫星高光谱影像的潜在空间自监督学习框架, 旨在构建具备可迁移性与可扩展性的高光谱基础表征。该框架采用教师-学生网络结构, 通过可复用的高光谱教师模型提供稳定的潜在语义监督, 并在学生模型学习过程中引入光谱视图扰动, 引导模型在潜在空间中对齐完整光谱语义表征, 实现跨传感器场景的适应能力。相比直接依赖像素重建的预训练范式, 所提出方法更加关注潜在表征层面的语义对齐与跨传感器一致性建模, 可灵活支持不同规模模型与多类下游密集预测任务。实验结果表明, 该方法能够有效提升高光谱基础模型在多传感器场景下的迁移泛化能力。



唐 旭

唐旭, 西安电子科技大学人工智能学院, 教授 / 博导, 国家级青年人才, 副院长。长期从事高分辨率对地观测相关方向的课题研究, 围绕遥感影像表征学习、内容理解等方面开展创新性的研究, 取得了突出的理论与技术成果。近五年, 申请人合作出版专著 2 部; 在本领域权威国际期刊和会议上发表论文 100 余篇, 13 篇论文入选 ESI 高被引文章, 研究成果荣获 2024 年度陕西省自然科学二等奖、2022 中国自动化学会自然科学二等奖。

报告题目: 文本辅助的遥感影像智能解译

报告摘要: 传统视觉单模态的遥感影像解译对复杂场景的语义理解能力有限, 容易停留在“像什么”的层面, 难以准确表达“是什么”和“有什么关系”。在此背景下, 引入文本信息成为一个重要趋势。文本能够提供丰富的语义先验和概念描述, 有助于弥补遥感影像在高层语义表达上的不足。本报告将从多模态融合视角出发, 介绍文本如何为影像提供语义补充与先验知识, 进而提升地物识别与场景理解能力。报告将归纳典型研究工作, 包括基于描述语句的语义对齐、视觉-语言模型在遥感中的应用, 以及文本驱动的检索与分类方法。通过对代表性方法与技术路线的总结, 揭示文本辅助遥感解译的关键思路与发展趋势。



殷君君

殷君君, 教授、博导。2013 年获得清华大学工学博士学位。研究方向为极化雷达应用基础理论、极化合成孔径雷达 (Synthetic Aperture Radar, SAR) 图像理解、紧缩极化 (compact polarimetry) SAR 目标解译、海洋生态遥感及生态环境变化监测等。近年来, 在物理散射机制、图像分割、目标分类以及目标检测等方面, 在 IEEE TGRS, IEEE JSTARS, IEEE GRSL 等国际杂志和 IEEE Radar Conf., IGARSS 等国际会议, 以及国内重要期刊上发表论文 100 余篇, 出版专著 2 部、合著 1 部。获得 2016 年国际发明博览会金奖、国际友人奖以及 2022 年日内瓦国际发明博览会银奖, 获得 2017 年国际三航大会青年科学家奖, 获得 2021 年北京市自然科学奖一等奖。主持国家自然科学基金青 B、面上基金以及其他项目多项; 担任 IEEE TGRS 副主编。

报告题目: 物理机制嵌入的极化SAR舰船检测网络

报告摘要: 深度神经网络由于具有较好的数据拟合效果已经被广泛使用。然而, 在域泛化方面深度神经网络不如理论模型泛化性好, 在训练样本量的需求方面深度学习模型往往需要大量的训练样本来拟合网络, 该两方面严重限制了深度学习的实际应用效果。本报告介绍一种将目标散射物理机制融入到检测网络中的框架, 可以大大提高检测性能、提高泛化性以及降低训练样本的需求量。



董燕妮

董燕妮, 武汉大学教授、博士生导师, 国家优秀青年科学基金获得者。主要从事遥感图像智能解译与地学空间智能应用等研究, 已在 IEEE TIP、IEEE TGRS 等国际刊物上发表论文八十余篇, 担任 IEEE TGRS 和 IEEE GRSL 副编辑, 中国图像图形学会优博俱乐部副主席, IEEE GRSS 武汉分会副主席, 湖北省女科技工作者协会理事等; 曾获得中国图像图形学会石青云女科学家奖、香江学者奖等; 主持国家自然科学基金重点项目、优青项目, 应急管理部重点科技计划课题等十余项。

报告题目: 多源遥感图像智能解译

报告摘要: 随着遥感成像技术的发展, 高光谱、高分辨率等遥感图像数量急剧增加, 海量多源遥感大数据与标签样本难以获取之间的矛盾愈加凸显, 先验标签样本的缺乏成为了遥感图像处理领域中的重要技术瓶颈, 大大限制了多源遥感图像智能解译精度。本报告主要介绍多源遥感图像分类、检测等任务的最新研究进展。



曹相湧

曹相湧, 西安交通大学计算机学院副教授 / 博导, 哥伦比亚大学访问学者, 主要从事底层图像处理、遥感图像解译、生成式大模型等方向研究。在 TPAMI/IJCV/SIAM J.Image Sciences/CVPR/ICCV 等期刊会议以一作 / 通讯发表论文 50 余篇, 主持国家自然科学基金面上、青年、中国电信、华为云等项目, 研究成果落地中国星网、中电科 54 所、航天五院等企业研究所, 获华为难题揭榜火花奖, 指导学生获 CVPR 等国际挑战赛冠 / 亚军 10 余项, Valse 执委会委员, CSIG 青工委委员, CSIG 遥感专委会委员, CCF 人工智能与模式识别专委会委员, 《Pattern Recognition》编委和《电子学报》青年编委。

报告题目：生成式遥感图像处理

报告摘要：面向现有遥感图像处理任务中图像质量差、标签数据少、先验设定难等难题, 本报告汇报课题组近十年来生成式遥感图像处理的系列成果, 涵盖遥感图像处理的底层到高层任务。具体来说, 针对图像质量差的问题, 提出了改进传统手工设定噪声的自动学习噪声方法; 针对标签数据少的问题, 提出改进传统标记数据增广的生成式遥感图像合成方法; 针对先验设定难的问题, 提出了改进传统领域先验建模的基础大模型先验嵌入方法。

分论坛 低空智能视觉

时间：8月6日上午09:00-12:00 地点：302会议室

内 容

嘉 宾

主席：刘 偲（北京航空航天大学）、蒋 铎（北京航空航天大学）、张 聪（山东大学）

低空多模态超宽谱认知遥感

吴启晖（南京航空航天大学）

端侧自进化目标感知基础模型的一点探索

王 鹏（西北工业大学）

低空智能：从感知推理迈向群体具身

朱鹏飞（东南大学）

面向低空智能的大场景理解与自主导航

李冠彬（中山大学）

空地两栖分合式无人系统自主协同技术

屠 展（北京航空航天大学）

论坛主席



刘 偲

刘偲，北京航空航天大学人工智能研究院副院长、教授、博士生导师，国家级领军人才，中国图像图形学会理事，副秘书长，聚焦空天具身智能核心方向，深耕具身智能、多模态内容理解，结合无人机协同导航、在轨维护等空天场景，推动技术从理论创新走向空天领域工程落地。累计发表 IEEE 汇刊或 CCF-A 类论文百余篇，其中 IEEE T-PAMI 18 篇，Google Scholar 引用超 2 万次，近五年连续入选 AI 2000 多媒体领域全球前 100 学者；获 ACM MM 2021 最佳论文、IJCAI 最佳视频奖及 10 余项国际顶级竞赛冠军，成果兼具创新性与实用性。多次担任 ICCV、CVPR 等国际顶会领域主席，兼任 IEEE T-PAMI、T-MM、T-CSVT 等领域顶级期刊编委，主持优青基金、企业联合基金重点项目及科技创新 2030 重大项目课题，荣获国家科技进步二等奖、国防技术发明一等奖，中国图像图形学会自然科学一等奖等多项荣誉。



蒋 铎

蒋铎，北京航空航天大学教授、博导，国家级青年人才（教育部青年 CJ 学者），中国图像图形学会青工委副秘书长。研究方向为计算机视觉、多媒体信号处理、医学影像分析等多个领域，发表高水平论文 60 余篇，其中第一或通讯作者论文 40 余篇，包括 TPAMI、IJCV、TMI、CVPR 等。入选中国科协青年人才托举工程、中组部海外博后引才专项计划，获得北京市优秀博士学位论文、中国图像图形学会优秀博士学位论文、ACMMM 最佳分会论文、CVPR NTIRE 视频质量增强大赛双赛道冠军等多个奖项。作为项目负责人主持国自然青年项目、华为卓越青年学者、阿里巴巴公司 AIR 等 10 余项项目。



张 聪

张聪, 山东大学控制科学与工程学院教授、博士生导师, 国家级青年人才, 机器智能与系统控制教育部重点实验室骨干。博士毕业于香港理工大学, 并先后于香港理工大学、香港中文大学从事研究工作。研究方向主要包括具身智能、模式识别、视觉感知等。作为项目负责人或技术骨干承担国家自然科学基金、香港智慧交通基金、香港创新及科技基金等科研项目多项。已在 IEEE TPAMI、TIFS、TGRS、CVPR 等国际权威期刊与会议发表学术论文 40 余篇, 其中 ESI 高被引论文 4 篇。担任多本国际知名期刊编委及会议专题主席、特邀讲者、程序委员等, 曾荣获日内瓦国际发明奖、CVPR NTIRE 挑战赛季军等。

论坛讲者



吴启晖

吴启晖, 南京航空航天大学, 教授, 博士, 博士生导师, 副校长, 教育部长江学者特聘教授, IEEE Fellow, 国家频谱战略规划政策研究组组长、国家自然科学基金委信息学部咨询委员会委员、中国工程院学科交叉专项总体组成员, 中国电子学会常务理事 / 会士、中国通信学会常务理事 / 会士。主要从事天地一体频谱认知智能、无人机集群网络等方向研究, Google Scholar 引用 20000 余次, ESI 高被引论文 10 余篇, 入选 ESI 工程学与计算机科学领域科学家 (Scientists) 世界前 1%, 获 IET 创新奖等国际高水平奖项 12 项, 制定国际标准 2 项, 承担国家重点研发计划、国家自然科学基金重大仪器等国家重大 / 重点项目 10 余项, 获国家科技进步二等奖 1 项、牵头省部级科技进步一等奖 6 项等。

报告题目: 低空多模态超宽谱认知遥感

报告摘要: 低空经济是战略新兴产业, 无人机巡检 (遥感) 与反无人机监测是低空经济的重要要素。与最典型的模态 (图像、视频、文本和语音等) 不同, 团队研究探索电磁频谱空间的多模态即超宽谱多模态, 包括可见光相机、红外相机、雷达、无线频谱监测等, 利用不连续超宽谱域异构特征与认知智能, 提升目标探测与复杂环境识别水平。在前期研究基础上, 本报告从无人机遥感战略需求出发, 阐述了多模态超宽谱认知的研究背景与科学意义、科学问题与关键技术、研制思路与设计方案、团队成果与演示验证, 探讨了未来趋势与研究展望。



王 鹏

王鹏，西北工业大学教授、博导。2017 年入选中组部青年人才计划，2021 年入选教育部长江学者特岗教授。入选斯坦福“全球前 2% 顶尖科学家”终身影响力榜单。目前担任空天地海一体化大数据应用技术国家工程实验室常务副主任，陕西省科技创新团队负责人。担任西北工业大学人工智能专业负责人。主要从事空天智能感知、具身智能、计算机视觉和机器学习领域的研究工作。主持和参与科技部新一代人工智能重大专项、国家自然科学基金委联合基金等十余个项目。出版中英文专著 2 部，在 TPAMI、IJCV、CVPR、ICCV、ECCV、NeurIPS、ICML、ACL 等权威期刊会议发表学术论文 100 余篇。担任会议 CVPR、AAAI 领域主席、Pattern Recognition 编委。担任中国计算机学会杰出会员、中国计算机学会青工委和教育工委执委(2024-2026)。

报告题目：端侧自进化目标感知基础模型的一点探索

报告摘要：随着边缘智能的快速发展，目标感知模型正从“云端中心化”向“端侧个性化”范式迁移。然而，端侧设备面临的开放动态环境和不同任务要求，对模型的适应性、泛化性与自进化能力提出了严峻挑战。本报告聚焦于端侧自进化目标感知基础模型的构建，从数据、模型、算子、训练等角度进行探讨，为实现具备自适应、可演化能力的下一代边缘目标感知系统提供思路与借鉴。



朱鹏飞

朱鹏飞，东南大学自动化学院教授，雄安国创中心低空智能实验室主任。主要研究方向是智能无人机，构建了大规模无人机视觉开放数据平台 VisDrone。已在 IEEE TPAMI 和 IJCV 等 CCF A 类和 IEEE 汇刊发表论文 100 余篇。获吴文俊人工智能科技进步一等奖等奖励。主持科技创新 2030-“新一代人工智能”重大项目、国家自然科学基金青 B 项目和重点项目 10 余项。

报告题目：低空智能：从感知推理迈向群体具身

报告摘要：低空经济已成为国家新兴支柱产业，人工智能正全方位赋能低空经济以保障其快速稳健发展。报告从低空环境感知、推理决策、具身智能、群体智能四个维度梳理其发展脉络，通过统计机器学习中的数据-模型泛化关联以及大模型时代的数据-模型规模定律介绍低空数据平台构建策略，并分享团队在低空智能方面的前沿进展以及低空城市治理案例。



李冠彬

李冠彬, 中山大学计算机学院副院长, 教授, 博士生导师, 国家优青。主要研究领域为图像视频内容理解与生成。迄今为止累计发表计算机学会 A 类 / 中科院一区论文 200 余篇, 谷歌学术引用超过 22000 次, 入选 ScholarGPS 全球前 0.05% 顶尖科学家榜单。曾获得中国图像图形学会青年科学家奖、吴文俊人工智能优秀青年奖、中国图像图形学会科学技术一等奖、ICCV2019 最佳论文提名奖、CVPR2024 最佳论文候选等荣誉。主持了包括国家自然科学基金优青、面上、重点研发课题、广东省杰青、CCF- 腾讯犀牛鸟科研基金、华为科研合作基金等 20 多项科研项目。担任广东省大数据分析处理重点实验室副主任、广东省图像图形学会计算机视觉专委会主任等职务。担任人工智能领域顶级会议 CVPR、ECCV、AAAI 等领域主席, 获得 10 余项人工智能领域国际顶级会议竞赛冠军, 研究成果应用于华为智驾系统, OPPO 小布助手及智慧医疗诊断等。

报告题目: 面向低空智能的大场景理解与自主导航

报告摘要: 近年来, 低空经济和无人系统的快速发展, 对智能体在复杂城市空间中的自主感知、空间理解与任务执行能力提出了新的需求。与传统地面场景不同, 低空智能面临视角变化剧烈、三维结构复杂、目标尺度稀疏、遮挡与深度歧义突出、长航程决策成本高等挑战。本次报告将围绕面向低空智能的视觉语言感知与三维空间推理展开, 介绍研究组近期在空中视觉语言导航、方向感知路径规划、城市级三维视觉语言理解、主动式视锥聚焦推理以及大场景语言高斯表示等方面的探索。相关工作从高效感知、结构化记忆、空间语义建模到自主导航决策等研究切入点为低空场景中的无人系统自主作业、城市空间理解和人机自然交互提供思路和解决方案。



屠展

屠展, 北京航空航天大学教授 / 博导, 国家级青年人才, 教育部先进长航时无人机系统技术工程研究中心副主任。主要从事飞行器总体设计与协同控制研究, 成果获中国航空学会技术发明一等奖、中国航空学会自然科学二等奖。

报告题目: 空地两栖分合式无人系统自主协同技术

报告摘要: 空地两栖无人平台兼具空中飞行和地面行驶运动模态, 相较传统无人机、无人车等单一运动模态平台越障能力强。传统空地两栖无人平台多为无人机和无人车结构上的简单拼接, 无效负载多、续航能力差、协同性不强, 这些缺点限制了其应用效能。针对上述问题, 开展空地两栖分合式无人系统研究, 空地子系统可自主对接、分离, 充分复用续航能力, 对接状态具备两栖越障能力, 分离状态可实现空地自主协同, 显著提高了空地两栖平台在复杂环境下整体作业能力和智能化水平。

分论坛 AI Coding Agent

时间：8月6日上午09:00-12:00 地点：309会议室

目 录

嘉 宾

主席：刘祥龙（北京航空航天大学）、刘艾杉（北京航空航天大学）
马宇晴（北京航空航天大学）

智能体前沿安全风险与治理	杨 珉（复旦大学）
鸿蒙大模型的构造、优化与落地实践	石恩升（华为）
工业代码大模型发展与挑战	杨 健（北京航空航天大学）
面向知识密集与高约束任务的智能体技术研究	赵才荣（同济大学）
AI Agent Security: Bridging Proximal Flaws and Distal Threats	邵 婧（上海人工智能实验室）

论坛主席



刘祥龙

刘祥龙，北京航空航天大学教授，博士生导师，国家级领军人才，国家重点研发计划项目首席科学家，现任计算机学院党委书记、复杂关键软件环境全国重点实验室副主任、中国仿真学会视觉计算与仿真专委会副主任、国家人工智能标准化总体组治理专题组联合组长。刘祥龙教授长期从事安全可信人工智能的教学和科研，负责《人工智能安全导论》、《数据结构与程序设计》等本科和研究生国家级课程，发表 TPAMI、IJCV、CVPR、ICLR 等人工智能领域国际顶级期刊 / 会议论文 160 余篇，主持国家自然科学基金青年 A 类和 B 类项目、科技创新 2030 重大项目、国家重点研发计划项目等国家级项目。担任 IEEE TNNLS、Pattern Recognition、《智能安全》等期刊编委，ACL、AAAI 等国际会议领域主席。连续入选爱思唯尔中国高被引学者，获中国仿真学会科学技术一等奖、陕西省自然科学一等奖、中国电子学会青年科学家奖、北京市科技新星等。指导学生获得世界人工智能大会云帆奖、CSIG 博士学位论文奖励计划、ACM MM 最受关注演示论文奖、亚太多媒体大会最佳学生论文、IEEE 低功耗计算机视觉挑战赛全球冠军、国家自然科学基金博士生项目、中国科协青年人才托举工程博士生专项计划、字节跳动奖学金、首都挑战杯特等奖等。



刘艾杉

刘艾杉, 北京航空航天大学计算机学院副教授, 北京智源青年学者、华为卓越青年学者, 长期从事安全可信人工智能、人机价值对齐的研究。累计在 CCF-A 类国际顶会顶刊发表论文 100 余篇, 谷歌学术引用 5300 余次, 编制国家标准 20 余项, 填补我国人工智能算法模型质量标准的空白, 获全国信标委人工智能分委会“优秀个人”。率先创办 AdvML@CV 系列智能安全国际研讨会议并连续 5 年担任 CVPR、ICML 等 CCF-A 类国际会议的论坛主席、领域主席。研究成果获 CCF-A 类会议 ACM MM 2025 最具影响力演示奖、ACM CCS 2024 杰出研究成果奖、ICML 2025 MAS 杰出论文奖等国内外重要奖励 10 余项, 以及省部级科技奖进步一等奖 1 项、二等奖 1 项。



马宇晴

马宇晴, 博士, 北京航空航天大学人工智能学院副教授, 入选北京市高创人才计划。主要研究人工智能、开放认知推理等方向, 发表国际顶刊顶会等论文 50 余篇, 申请专利 20 余项; 创办并连续举办国际顶会 IJCAI 开放推理研讨会, 获官媒推荐。获《智能安全》创刊十大优秀论文、医学物理青年学术年会人工智能挑战赛一等奖、北京发明创新大赛铜奖等奖励。

论坛讲者



杨 珉

杨珉, 复旦大学计算与智能创新学院执行院长, 教授、博导, 国务院学位委员会网络空间安全学科评议组成员、国家级领军人才计划入选者、973 计划首席科学家、国家重点研发计划首席科学家、上海市网络空间安全战略研究所执行所长。主要研究操作系统安全、恶意代码检测、漏洞分析挖掘和 AI 系统安全等, 在网络安全四大顶会 S&P、CCS、Security、NDSS 发表论文 60 余篇, 研究发现曾多次获中央领导重要批示, 成果入选 973 重大成果和突出进展、十三五国家科技创新成就。

报告题目：智能体前沿安全风险与治理

报告摘要：随着生成式大模型和智能体技术迅速发展，如何针对以大模型为核心的前沿 AI 系统开展安全风险评测与治理是全球面临的共性难题。一方面，以图、文、音视频生成为主要功能的多模态大模型持续演进，内容安全风险尤为严峻，急需自动化程度高、覆盖面广和靶向性强等动态安全评测技术；另一方面，大模型通过工具、知识库、推理编排框架等基础软件，能够与真实世界自主交互，一旦具备自主复制、欺骗操控等危险红线能力，或对全人类带来灾难性风险。本报告将围绕大模型内容安全和智能体行为安全两大核心议题，分享团队近年来在前沿 AI 系统安全评测与治理方面的研究进展和思考。特别地，团队近期在国际上首次发现十余款大模型能够在无人工干预情况下完成自我复制，得到国内外广泛关注和高度认可，为全球 AI 治理和对话提供了建设性思路。



石恩升

石恩升，华为 AI 算法专家，2024 年通过“天才少年计划”加入华为，现担任鸿蒙代码大模型与智能体构建与优化负责人。他主导推动鸿蒙代码大模型多项性能达到业界领先水平，并成功将技术落地于鸿蒙 CodeMate、CodeGenie、Code Arkify 及 AI 辅助鸿蒙开发等产品，广泛应用于华为内部及鸿蒙生态，成果显著，深受生态伙伴欢迎，获得华为授予员工个人最高荣誉 - 金牌个人奖，并与华为创始人任正非先生单独合影留念。他积极推动高校合作，负责多项胡杨林基金项目及高校技术合作项目，致力于产学研深度融合。近年来，他在人工智能与软件工程领域的顶级会议和期刊（如 ICSE、FSE、ISSTA、ACL、EMNLP 等）发表论文 20 余篇，荣获 ISSTA 2024 杰出论文奖。此前，他作为微软亚洲研究院与西安交通大学联合培养的博士生，受张冬梅、韩石、张洪宇、王焱林和孙宏滨联合指导，聚焦智能化软件工程，论文获得 CCF 2025 软件工程博士学位论文激励计划提名，也荣获西安交通大学授予的博士学位最高荣誉。

报告题目：鸿蒙大模型的构造、优化与落地实践

报告摘要：在国产操作系统与应用——鸿蒙的生态崛起中，AI 大模型成为关键“助推器”。面对数据稀缺与技术空白，通过低资源冷启动、多阶段微调等多项技术和方法，打造性能业界领先的鸿蒙大模型。模型落地后，支持代码续写、生成、转化和理解等多项核心任务，应用于医疗、金融、工具等多个领域的鸿蒙应用开发，大大提升了鸿蒙开发效率并降低了开发者门槛，持续赋能鸿蒙生态。



杨 健

杨健，北京航空航天大学计算机学院副教授，上海 AILab 双聘副研究员，研究方向为代码大模型预训练、后训练、智能体及评测。2023 年获北航计算机博士学位，师从李舟军教授和周明博士，曾为北航与微软亚研院联培博士生，博士期间获 WMT 2021 三项冠军。毕业后以阿里星身份加入阿里通义千问团队，作为 Qwen 及 QwenCoder 系列核心贡献者，负责代码能力建设与模型研发。已在 ICLR、NeurIPS、ACL、EMNLP 等顶会发表论文百余篇，第一 / 通讯作者 30 余篇，谷歌学术引用 25000+ 次。担任 NeurIPS、ARR、AAAI 等国际会议高级领域主席 / 程序委员。与清华、北大、CMU、上海 AI 实验室等保持学术合作，与阿里、字节、美团等企业保持产业合作。

报告题目：工业代码大模型发展与挑战

报告摘要：随着大语言模型技术的飞速发展，代码智能正从通用软件开发向工业领域纵深渗透。工业代码大模型作为人工智能与新型工业化深度融合的前沿方向，旨在让大模型真正理解并生成面向工业控制、嵌入式系统、Verilog 等场景的专业代码。然而，工业代码数据的稀缺性、安全可靠性的极高要求、实时性与边缘部署的资源约束，以及代码与工艺知识的多模态耦合，使其面临远超通用代码大模型的技术挑战。本报告将系统梳理工业代码大模型的发展背景、核心挑战与关键技术路径，介绍其在工业软件开发提效、老旧系统迁移重构等典型场景中的应用实践，并展望从“辅助代码生成”迈向“自主工业智能体”的未来演进趋势，探讨如何通过产学研用协同共建，推动工业代码智能的规模化落地，为制造业高质量发展注入新动能。



赵才荣

赵才荣, 工学博士, 同济大学计算机科学与技术学院教授, 博士生导师。数据科学与人工智能系主任, 上海市计算机学会计算机视觉专委会主任, 中国图像图形学会青工委秘书长。入选东方英才, 担任《Pattern Recognition》、《Visual Intelligence》、《中国图象图形学报》等期刊编委。主要研究领域: 计算机视觉。已在 TPAMI、CVPR 等发表学术论文 60 余篇。主持了国家自然科学基金重点、面上等项目 5 项。研究成果获 2022 年上海市科技进步一等奖, 2023 年上海市自然科学二等奖, 2025 年吴文俊人工智能科技进步二等奖, 2025 年江西省科技进步二等奖。获 2024 年全国人工智能应用场景创新挑战赛总决赛特等奖, 获 2024 年上海市计算机学会教学成果一等奖。指导学生获中国电子学会优秀硕士论文, 获上海市计算机学会优秀博士论文、优秀硕士论文。

报告题目: 面向知识密集与高约束任务的智能体技术研究

报告摘要: 本讲座聚焦智能体研究从实验原型走向生产系统这一核心问题, 展示面向高责任任务的智能体架构设计与产业级转化路径。报告将结合科技金融与智慧座舱两个真实应用场景, 探讨智能体系统如何在知识密集、高约束的任务中形成稳定可追溯的行为与交付。在已产品化落地的科技金融场景中, 报告围绕确定域-涌现域信息空间分离原则展开, 将可验证的信息作为事实基础和约束边界, 引导并校验智能体的多步推理与开放探索, 并进一步支撑从企业尽调、产业穿透到区域推演的智能体矩阵, 形成面向银行授信、产业投资、区域治理等细分应用。在智慧座舱场景中, 讨论面向车规级工程研发的代码智能体, 围绕语音接入、车控逻辑、界面一体化等系统集成, 推动代码生成向规范化、可集成的交付转化, 使智能体能力进入车载软件的研发闭环与质量约束体系。



邵 婧

邵婧, 上海人工智能实验室安全可信 AI 中心联合负责人、青年科学家。博士毕业于香港中文大学 MMLab, 曾任商汤科技研究总监。负责研究自主、可控、可信 AI, 聚焦多种模态大模型及智能体安全评估与价值对齐相关研究工作。在国际顶会 / 期刊累计发表论文 80 余篇, 谷歌引用 15000 余次, 连续两年于 ACL2024、2025 获得 Outstanding Paper Award。2023~2025 连续三年入选斯坦福前 2% 科学家榜单。

报告题目: AI Agent Security: Bridging Proximal Flaws and Distal Threats

报告摘要: AI 智能体技术的快速演进在带来突破性能力的同时, 也引入了新型安全威胁。本报告围绕“近端漏洞发现与远端风险推演”双维度安全框架, 分析多种智能体系统的核心安全问题, 探讨前沿 AI 能力与其潜在风险的关联性, 以及对社会未来演化的影响, 为构建可信 AI、实现安全可控的 AI 智能体生态提供研究思路与可行方向。

分论坛 图像视频生成与编辑

时间：8月6日上午09:00-12:00 地点：401会议室

目 录

嘉 宾

主席：王亚星（吉林大学）、武 宇（武汉大学）、易 冉（上海交通大学）

基于多模态表征的图像和视频世界模型生成

李 玺（浙江大学）

面向生成式内容的质量评价与安全分析研究

刘 婧（天津大学）

可控视觉内容生成

贾 旭（大连理工大学）

生成基座模型的免训练能力拓展

张 驰（西湖大学）

视觉生成与编辑：从像素生成走向物理、几何与逻辑约束

全睿杰（浙江大学）

论坛主席



王亚星

王亚星，吉林大学人工智能学院教授，博士生导师，入选海外高层次人才引进计划青年项目。西班牙巴塞罗那自治大学博士。研究方向为视觉生成模型、迁移学习、无监督学习等。在 TPAMI, ICLR 等期刊会议发表论文 50 余篇。担任 ICLR/NeurIPS/ICML 等领域主席，在国际顶级期刊和会议 TPAMI、ICLR 等多次担任期刊和会议审稿人，ECCV/ICCV 组织多场 workshop。多模态语言翻译国际竞赛中荣获第一名、2022 年粤港澳大湾区（黄埔）国际算法算例大赛（遥感目标检测赛道）亚军（2/116 队伍）。主持 / 参与国家自然科学基金青年项目、三星人脸特定人脸合成、宁波市文旅个性化大模型、重庆长安 AI 智能体等项目。



武 宇

武宇，武汉大学人工智能学院副院长、教授、博导，“十五五”国家重点研发计划实施方案专家组专家，国家高层次青年人才。主持国家新一代人工智能重大项目课题等多项国家级项目。主要从事图像视频生成、多模态相关研究，发表顶会论文 80 余篇，谷歌学术引用 8000 余次，单篇引用超 1000 次。长期担任 NeurIPS、CVPR、ICML 等人工智能国际顶会领域主席，担任 CVPR 2023 大会主席、主要组织者。获 2020 年谷歌博士奖（全球共 10 位获奖者）、2024 年 AAAI 学术新星奖，获 6 项 CVPR/ICCV 国际学术竞赛的冠军。



易冉

易冉，博士，上海交通大学计算机学院副教授、博士生导师。入选中国科协“青年人才托举工程”、上海市扬帆计划。从事计算机图形学、生成式人工智能方向的研究。发表录用100篇论文于IEEE TPAMI、ACM TOG、SIGGRAPH、CVPR等期刊会议，其中CCF-A类论文80篇。获北京市科技进步二等奖，中国图像图形学会高等教育教学成果一等奖，中国图像图形学会石青云女科学家奖(青英组)，CCF-腾讯犀牛鸟基金卓越项目，瑞士Chorafas青年研究奖，数博会十大领先科技成果，ACM SIGAI China新星奖，吴文俊人工智能优秀博士学位论文，中国产学研合作创新成果奖优秀奖，中国计算机学会产学研合作基金优秀项目，中国计算机学会计算机视觉专委会学术新锐奖。担任中国图像图形学会智能图形专委会副秘书长，《Virtual Reality&Intelligent Hardware(VRIH)》期刊青年编委。

论坛讲者



李玺

李玺，男，工学博士，浙江大学求是特聘教授，Fellow of IAPR/IET/AAIA/AIIA，美国国家人工智能科学院 Member，国家杰出青年科学基金项目获得者，国家高层次青年人才入选者，Elsevier 2023-2025 “中国高被引学者”和浙江省特聘专家。从事人工智能和计算机视觉等领域研究，发表高水平学术论文 200 余篇。主持科技部 2030 新一代人工智能重大项目，基金委联合基金重点项目、基金委数学天元基金交叉重点项目、教育部重点规划研究项目、JKW 重大专项项目课题、JKW 基础研究重点项目、浙江省基金重大项目和浙江省杰青项目。荣获中国发明协会发明创新奖一等奖和人物奖、教育部高等学校科学研究优秀成果奖科学技术进步奖一等奖、中国专利优秀奖、中国电子学会科技进步一等奖等八项省部级 / 一级学会科技奖励；此外，还获得 2025 年华为公司优秀技术合作项目奖(全球仅 17 项，浙大首次)，2026 年、2025 年、2024 年和 2022 年共 5 项华为公司“难题揭榜”火花价值奖，2023 年陆增镛 CAD&CG 高科技奖一等奖，2025 年中国视觉智能年度创新人物奖，华为杭研所优秀技术合作成果奖和腾讯鹅厂好专利奖。担任第七届中国图像图形学会理事，国际期刊编委（如 TNNLS 和 TMM）和会议的 Area Chair（如 ICML 和 ICCV）。荣获四项最佳学术论文奖。指导学生获得中国图像图形学会博士学位论文激励计划和浙江省优秀硕士论文奖。

报告题目：基于多模态表征的图像和视频世界模型生成

报告摘要：当今，图像和视频生成是人工智能领域的热点和难点问题，尤其是背后隐含的交互式世界模型建模方式。本报告主要围绕数据驱动的人工智能学习方法，进行多模态高效生成、理解和表达等多维度视角进行了深入剖析。系统地回顾了多模态特征表达和学习领域的不同发展阶段，介绍了近年来我们利用特征学习进行视觉语义分析和理解生成所做的一系列代表性的研究工作及其实际应用。特别关注这些技术在构建视频生成驱动的可实时交互世界模拟器中的应用潜力。世界模拟器旨在通过生成符合物理规律的动态视频序列，模拟真实世界的演化过程，为决策、仿真和内容创作提供基础，这对生成模型的高效性、可控性、时序一致性与物理合理性提出了核心要求。报告的最后将和大家一起探讨一下涉及多模态视觉生成和理解所面临的一些开放性问题 and 难题。



刘 婧

刘婧，天津大学电气自动化与信息工程学院副教授、博士生导师。主要研究方向为多媒体处理和内容理解，曾获 2025 年度吴文俊科技进步一等奖（序 1）、2021 年度天津市科学技术进步特等奖（序 5）、中国图像图形学会优秀博士论文提名奖。先后主持国家自然科学基金面上及青年项目、中国博士后科学基金特别资助、国家重点研发计划子课题等。近年来发表百余篇论文，相关成果获 2024 年国际数字多媒体通信论坛最佳论文奖、2021 年度 IEEE Transactions on Multimedia 最佳论文提名奖等。目前担任 IEEE Transactions on Multimedia、Elsevier Displays 期刊编委。

报告题目：面向生成式内容的质量评价与安全分析研究

报告摘要：生成式人工智能正深刻重塑内容创作与信息交互的模式，但其输出质量的不稳定性与潜在安全漏洞，已成为制约技术规模化可信应用的核心瓶颈。围绕这一痛点，本报告聚焦两大关键挑战：生成内容的感知质量评价与多模态模型的内在安全防御。在质量维度，我们突破传统单图评价局限，面向同主题图像组与生成视频，构建了兼顾整体感知、主题一致性与时序连贯性的主观质量评价新体系。在安全维度，我们针对视觉输入场景下的模型脆弱性，创新性提出多图协同的视觉越狱攻击框架，系统性揭示复杂输入下大模型的潜在风险暴露面。本报告将系统呈现上述两方面的最新探索，旨在为生成式内容的高质量发展与安全治理提供技术支撑。



贾 旭

贾旭，中国图像图形学会会员，大连理工大学未来技术学院 / 人工智能学院教授，博士生导师，入选国家级青年人才，辽宁省兴辽英才计划，大连市高端人才，比利时荷语鲁汶大学博士，从事计算机视觉与生成式人工智能领域的研究。在 CCF-A 类期刊和会议发文发表学术论文 50 余篇，Google 学术引用 13000 余次，已申请和授权国内外发明专利 20 余项，成果获包括诺贝尔奖等权威学者正面评价，入选全球前 2% 顶尖科学家。主持 3 项国家级项目或重点项目子课题，研究成果获得 CCF 自然科学二等奖（序 1）、华为“难题揭榜”火花奖、以及 CVPR、AAAI 挑战赛冠军等多项学术奖励。多次担任 ICLR、ICML、AAAI、ACM MM 等 AI 和 CV 领域重要学术会议的领域主席和高级程序委员会委员。

报告题目：可控视觉内容生成

报告摘要：生成模型已成为推动创意产业数字化转型的核心引擎，为内容创作、艺术设计等实际应用带来了革命性变革。然而，如何实现对生成过程的精准控制、确保输出内容符合用户意图和应用场景需求，仍然是视觉内容生成领域面临的关键挑战。本次报告将主要介绍团队在可控内容生成领域的最新研究成果，涵盖从静态到动态、从可控到可信的多维度内容控制技术。报告将围绕三个核心方向展开：首先介绍定制化视觉内容生成技术，实现对人物身份、视觉效果等个性化元素的精准控制；其次探讨空间位置视觉内容可控生成方法，涵盖图像构图、视频运动轨迹及多镜头叙事等时空层面的精细化控制；最后展示物理可信视觉内容生成的初步探索，使生成内容更加符合真实世界的物理规律。



张 驰

张驰博士现任西湖大学特聘研究员，博士生导师，并担任通用人工智能(AGI)实验室负责人。在加入西湖大学前，曾入选腾讯“技术大咖”头部人才计划担任高级研究员。他长期深耕生成式 AI 与多模态大模型研究，在 CVPR、TPAMI 等国际顶尖会议及期刊发表论文 60 余篇，并连续入选斯坦福大学“全球前 2% 顶尖科学家”榜单。张驰博士致力于产学研的深度融合，同时积极构建开源生态，主导的 AppAgent、MeshAnything、Metric 3D、GaussianEditor 等知名开源项目在 GitHub 累计获星超 20,000，被工业界与学术界广泛采纳为行业基准。

报告题目：生成基座模型的免训练能力拓展

报告摘要：近年来，生成基座模型在图像、视频和三维视觉领域取得了显著进展，但面向复杂应用时，模型仍需要具备更强的可控性、适应性和生成质量。相比依赖重新训练或微调的传统方式，免训练能力拓展通过在推理阶段调控生成过程、重组模型先验、组合外部模块与利用跨模态信息，在不更新模型参数的前提下释放基座模型的潜在能力。本报告将围绕生成基座模型的免训练拓展方法展开，结合团队的相关工作，介绍其在图像生成、视频生成和三维视觉任务中的最新探索。报告将展示如何通过免训练技术提升模型性能与表现，增强生成结果的质量、一致性和可控性，并进一步拓展生成基座模型在多样化视觉任务 and 实际应用中的价值。



全睿杰

全睿杰，浙江大学人工智能学院“百人计划”研究员，国家高层次青年人才。博士毕业于悉尼科技大学，曾于浙江大学和新加坡南洋理工大学担任博士后研究员。研究方向包括生成式人工智能，多模态大模型和 AI for Science。曾在 TPAMI、IJCV、NeurIPS、ICLR、CVPR、ICCV 等 CCF-A 类期刊和会议上发表论文 30 余篇。其长期担任 CVPR、ICCV、NeurIPS、ICLR、ICML、AAAI 等计算机视觉和机器学习顶级程序委员会成员，担任 Frontiers of Computer Science 青年编委。

报告题目：视觉生成与编辑：从像素生成走向物理、几何与逻辑约束

报告摘要：前的生成式视觉大模型虽然在像素层面的高保真度合成上取得了巨大突破，但由于缺乏对现实世界底层规律的显式感知，在面对复杂动态和隐式指令时，极易表现出时空伪影、几何坍塌及逻辑矛盾等瓶颈问题。这表明视觉生成正在经历一场关键的范式转变，即从单纯的文字与像素表面对齐，走向对物理、几何与逻辑约束的深层遵循。本报告将结合团队近期前沿工作，从三个递进的维度分享如何通过引入显式约束突破现有瓶颈：
1) 频域物理约束，针对视频生成中采样轨迹漂移的共性痛点，揭示其谱能量畸变的物理本质，提出在推理阶段通过时空谱解耦与振幅修正，以极低成本恢复视频的时空连贯性与物理真实感；
2) 场景几何约束，针对连续相机运动下的新视角生成问题，摒弃零散的几何注入，提出在视频模型的表征、架构和损失函数层施加统一的几何引导，建立稳健的跨视角三维结构对应，消除视角穿帮与结构重影；
3) 高层逻辑约束，从显式的空间物理延伸至隐式的人类常识认知，构建包含物理、环境、因果等五大维度的推理型图像编辑基准，并分享通过轻量化“诊断-再编辑”机制修正大模型逻辑漏洞的最新尝试。

分论坛 多智能体

时间：8月6日上午09:00-12:00 地点：402会议室

内 容

嘉 宾

主席：吴 飞（浙江大学）、郝建业（天津大学）、吉 炜（南京大学）

多模态协同的目标建模与感知

洪晓鹏（哈尔滨工业大学）

面向物理认知与空间约束的智能体推理及视觉生成

冯 婕（西安电子科技大学）

结构化上下文-有限监督泛化的新路径

姚权铭（清华大学）

大模型内部机制解码与智能体自进化

张 岸（中国科学技术大学）

Agent时代下的博弈决策演进与挑战

杨天培（南京大学）

论坛主席



吴 飞

吴飞，浙江大学求是特聘教授、博导，国家杰出青年基金获得者，现任浙大本科生学院院长、上海高研院常务副院长。主攻人工智能、跨媒体计算与行业大模型，曾任教育部人工智能专家组组长；牵头研发智海系列垂域大模型，主持国家级AI重大项目，主编多套人工智能教材，获吴文俊、教育部科技进步一等奖。



郝建业

郝建业，天津大学教授，国家优青。主要研究方向为强化学习、具身智能和多智能体系统。在Nature Communications/Machine Intelligence和ICML/NeurIPS/ICLR等人工智能领域CCF-A类国际会议发表论文100余篇，获国际会议最佳论文奖4次，NeurIPS大会竞赛冠军4次，中国图像图形学会科技进步一等奖（第一），中国电子学会自然科学一等奖（第二）等。曾先后任华为决策推理实验室主任、大模型算法实验室主任、华为医疗军团副总裁，负责华为决策智能领域技术创新和产业落地，从0到1孵化多项公司级产业项目，团队成果在国产工业基础软件智能化、基础和行业大模型、自动驾驶、5G网络优化、工业物流调度、具身智能等领域广泛落地应用。



吉 炜

吉炜，南京大学副教授，博导，国家级青年人才，研究方向包括多模态理解与生成、多模态检索与推荐、多模态预训练模型及智能体等。目前在 NeurIPS、CVPR、ECCV、SIGIR、AAAI、ACM MM 等国际顶级会议和 TPAMI、TIP、TMM 等顶级期刊发表了 80 余篇论文，谷歌学术引用 6000 余次，担任多媒体领域国际顶会 ACM Multimedia 2025 的大会宣传主席，自然语言处理领域国际顶会 EMNLP 2025 的大会本地主席，NeurIPS、ICLR、ICML、CVPR、ICCV、SIGIR、ACM Multimedia、IJCAI 领域主席，担任 IEEE TIP、IEEE TCSVT、Pattern Recognition、ACM TOMM 等期刊编委，IJCV 客座编委，曾荣获 CVPR 2022 最佳论文提名奖，2024 世界人工智能大会青年优秀论文奖等。

论坛讲者



洪晓鹏

洪晓鹏，哈尔滨工业大学教授，博导。斯坦福大学年度前 2% 顶尖科学家，德国达堡论坛特邀讲者。已在顶级国际和国内刊物和国际会议上发表文章 100 余篇，2 次获得领域内国际权威期刊和会议的优秀论文奖，8 次带队获得国际评测冠军。作为负责人主持了多项国家重点研发计划课题、国家自然科学基金、芬兰信息学会博士后基金等项目。相关工作见诸美国《麻省理工技术评论》等技术媒体专文报道。CSIG 情感计算与理解专委会常务委员和副秘书长，黑龙江省计算机学会学术工作委员会副主任。CVIU 等国际主流期刊编委，AAAI, ACM MM, IJCAI, ICME 等主流国际会议领域主席。目前的研究领域包括：多模态目标感知、深度连续学习、AI 生成检测和智能体规划与协同等。

报告题目：多模态协同的目标建模与感知

报告摘要：本报告将主要汇报在多模态目标感知方面的一些工作，包括 (1) 面向同构模态协同的模态中介与模拟技术、(2) 文图协同的不同应用方式、(3) 多模态智能体构建三个角度，探讨多模态表征学习与建模方法，以及在目标感知和情绪识别等视觉任务上的应用。



冯 婕

冯婕，西安电子科技大学人工智能学院教授，博士生导师。主要研究方向为遥感图像处理、空天空间智能。发表 CCF A 类和中科院 I 区论文 60 余篇，ESI 高被引 / 热点论文 11 篇，出版专著 2 部。主持国家自然科学基金联合重点项目、军科委基础加强领域基金、教育部联合基金等。入选国家级青年人才、陕西省杰出青年科学基金、中国科协青托等。获得陕西省自然科学二等奖（序 1）等。担任 IEEE TGRS、IEEE J-STARS、Remote Sensing、Journal of Information and Intelligence（信息与智能学报（英文））等期刊编委。目前为 CSIG 学会青工委副秘书长、中国电子学会空间电子学分会青年副主任委员等。

报告题目：面向物理认知与空间约束的智能体推理及视觉生成

报告摘要：智能体技术在视觉任务中展现出重要潜力，但现有基础架构在处理真实世界复杂数据时普遍存在局限。首先，智能体难以直接理解场景中的物理属性与潜在交互关系；其次，面对三维空间环境或遥感大尺度数据，智能体难以完成从简短抽象语义到复杂物理布局的精确映射，且缺乏对全局规划与局部细节的联合调度机制。报告将从智能体在物理理解与空间约束上的共性缺陷出发，探讨相应的机制改进及其在不同空间尺度任务中的验证。针对微观三维场景中的物理与几何难题，汇报智能体通过引入先验与跨域证据机制，在物理属性解析、三维高斯场景可供性推断，以及简短文本驱动的室内三维场景生成等工作中的应用进展。针对宏观大尺度数据的调度难题，介绍智能体在宏观布局与微观特征分层控制上的机制探索，及其在遥感图像大图生成中的阶段性工作。



姚权铭

姚权铭，清华大学电子工程系副教授、博士生导师，英国工程技术学会会士（IET Fellow），国家高层次青年人才。本科后于香港科技大学计算机系取得博士学位，博士期间获 Google 机器学习博士奖学金；毕业后任职第四范式首席研究员，牵头组建并带领核心机器学习研究团队。研究方向为机器学习与科学智能，围绕基座模型泛化机理、数据高效学习、动态环境智能体决策开展理论与系统研究，累计发表论文 120 余篇，总引用超 15000 次。获 IEEE、INNS、CAAI、香港科学会早期成就奖，入选福布斯 30Under30、百度全球华人 AI 青年学者榜、蚂蚁集团首届 InTech 科技奖。兼任 ICML、NeurIPS、ICLR 顶会领域主席，IEEE TPAMI, Neural Network, Machine Learning 期刊编委，受邀在 AAAI 2024、IJCAI 2025 作特邀成就报告。指导学生斩获全国挑战杯特等奖等多项赛事最高荣誉，个人获评清华大学刘冰奖、电子系郑君里优秀教师奖。

报告题目：结构化上下文 - 有限监督泛化的新路径

报告摘要：本报告围绕有限监督场景下的泛化问题，探讨结构化上下文作为核心机制的作用。首先分析在少样本、弱反馈及分布变化条件下，传统依赖规模扩展的方法所面临的瓶颈，指出智能体需在缺乏充分监督的情况下实现自适应与自我提升。随后提出一个统一视角：通过结构化上下文，将学习从「依赖标签」转向「依赖结构」。在此基础上，结合多智能体系统中的结构重组、自对抗生成机制以及基于交互的动态监督，展示三类代表性方法如何在有限监督下实现泛化与扩展。最后讨论结构优先的智能扩展范式及未来方向。



张 岸

张岸，中国科学技术大学特任教授、博士生导师，国家级青年人才，荣获2025年Web领域女性新星奖。主要研究领域为大模型驱动的智能体、个性化大模型、可信人工智能，特别关注于智慧校园、大模型安全、以及个性化场景，致力于下一代通用人工智能模型的关键能力与性质研究。相关工作在NeurIPS、ICLR、ICML、WWW、KDD、SIGIR、TOIS、TPAMI等顶级国际会议和期刊发表录用长文40余篇，其中超过3篇论文入选了最高引和最具影响力榜单，Google学术引用超四千次，H-index达30。

报告题目：大模型内部机制解码与智能体自进化

报告摘要：本报告汇报两部分发现：第一部分聚焦大语言模型解释性，从可观测的内部表征与结构机制出发，揭示模型如何预分配推理强度、关键单元与路由机制如何塑造多语与能力差异，并探索层级化调控以实现更可控、可验证的行为。第二部分面向自进化智能体系统，系统介绍三项关键：记忆从事实走向经验与可回访记忆并结合精炼/检索/可视化以支持长期与分布外适应；Tool-Using RL通过自验证、环境扩展与噪声鲁棒提升有效性与泛化；多智能体协同结合自演化rubrics构建“评测—训练—部署—反馈”闭环，实现系统级持续进化。



杨天培

杨天培，南京大学副教授，博士生导师，紫金学者，入选国家高层次青年人才计划（海外）。研究方向包括强化学习与多智能体系统，重点解决强化学习的探索瓶颈、泛化能力弱、不可解释等挑战。科研成果推动了强化学习在国防军事、芯片设计、广告出价等相关领域的应用。近五年在人工智能领域发表高水平论文40余篇，长期担任ICML、NeurIPS、ICLR、AAAI等会议，以及TPAMI、JMLR、TNNLS等国际期刊的审稿人。曾获得CCF人工智能与模式识别专委会多智能体学组优秀博士论文奖。

报告题目：Agent时代下的博弈决策演进与挑战

报告摘要：随着大语言模型向自主智能体跨越，强化学习正经历深刻的范式转移。本报告深入探讨了Agent时代下RL的演进路径与核心挑战，重点解析了从个体到群体训练的变革。在个体训练维度，报告分析了从传统深度强化学习向大模型后训练RL的转变：不同于传统DRL侧重于在封闭环境下通过海量试错学习基础控制逻辑，Agent时代的RL侧重于在已有知识底座上，通过过程奖赏与强化学习算法提升模型的长程规划、思维链推理及自我纠错能力。在智能体协作维度，探讨了从传统多智能体强化学习向基于自然语言通信的Agent集群协作演进，强调在开放环境下通过动态对等通信实现复杂任务的协同求解。最后，报告针对奖赏模型优化、长程信度分配等核心挑战，展望了RL驱动Agent成为真实生产力工具并通往AGI的技术前景。

分论坛 国自然博士生论坛

时间：8月6日上午09:00-12:00 地点：204会议室

内 容

嘉 宾

主席：金 枝（中山大学）、梁孔明（北京邮电大学）

面向开放场景的多模态异常检测	马云峰（湖南大学）
面向复杂场景感知的多源图像融合：从视觉增强到任务自适应	吴冠尧（大连理工大学）
大模型认知优化探索	陈思蕤（同济大学）
以可观测性驱动的coding agent harness自进化	窦士涵（复旦大学）
基于扰动生成的视频压缩预处理	郭苡辰（北京航空航天大学）
工业缺陷开放集检测	杨 震（湖南大学）
科研智能体网络：从单点智能体到可信协作基础设施	郭子晗（中山大学&上海创智学院）
具身智能自我认知的心理学研究	李春一（上海交通大学&南洋理工大学）

论坛主席



金 枝

金枝，中山大学教授，博士生导师，广东省珠江青年学者，深圳市优青，中山大学“逸仙学者”，IEEE Senior Member，中国图像图形学会 CSIG 交通视频专委会副秘书长、青工委副秘书长、多媒体专委会委员，主要从事恶劣环境视觉感知增强、空间智能感知等方面的研究和应用。2016 年博士毕业于英国利物浦大学，博士后分别在深圳大学和德国慕尼黑工业大学进行。在图像处理及计算机视觉领域顶级期刊及会议发表论文 94 篇。作为第一发明人已授权国家发明专利 16 项，已授权软件著作权 10 项。主持国家自然科学基金面上、青年、中德 / 中英国际合作等各类项目 18 项。带领学生连续三年（2024-2026）在 CVPR NTIRE 挑战赛中获全球冠军。



梁孔明

梁孔明, 北京邮电大学副教授, 博士生导师, 中国图像图形学会副秘书长、中国人工智能学会智能服务专业委员会委员, 北京图象图形学学会理事, 入选国家级青年人才计划和北京市科技新星计划。研究方向为视觉概念理解、视觉语言大模型优化及评测、医学影像分析。以第一 / 通讯作者发表 CCF-A 类及领域顶级会议期刊论文 30 余篇。主持国家自然科学基金面上项目和青年科学基金项目, 北京市科技新星计划交叉项目等。获得 2022 年度 MICCAI ISLES 国际竞赛冠军、2023 年度 IEEE IC-NIDC 国际会议最佳论文、2023 年度中国智能技术与大数据会议优秀论文、2023 年度中国图像图形学会技术发明奖一等奖、2024 年度吴文俊人工智能科技进步奖二等奖、2024 年度北京青年优秀科技论文。

论坛讲者



马云峰

马云峰, 湖南大学人工智能与机器人学院在读博士研究生, 师从刘敏教授, 依托机器人视觉感知与控制技术国家工程研究中心, 主要研究方向为工业具身智能、异常检测与持续学习。博士在读期间作为学生负责人全程参与了国家重点研发计划项目, 相关系统已应用于中国航发南方工业有限公司。博士在读期间, 以第一 / 共同第一作者发表 IEEE TPAMI 等 IEEE 汇刊论文 11 篇, 参与制定 2 项国家标准草案, 获“挑战杯”湖南省大学生创业计划竞赛金奖 (2/14)。

报告题目: 面向开放场景的多模态异常检测

报告摘要: 在全球制造业向高端化、智能化转型升级的背景下, 工业异常检测作为保障产品质量与提升生产效率的关键技术, 正逐步从传统的单一模态向多维感知方向发展。然而, 真实的工业生产环境往往呈现出典型的开放场景 (Open Scenarios) 特性, 检测系统需克服高度非平稳数据分布带来的泛化瓶颈, 例如: 突发的传感器异常、高频换线生产带来的零样本/少样本泛化需求, 以及不可预见罕见缺陷模式导致的长尾分布问题。传统的闭集 (Closed-set) 无监督模型在面对这些开放且未知的工况时, 往往会出现泛化能力骤降与误报率激增的现象。报告将紧密围绕“面向开放场景的多模态异常检测”展开。首先, 简要回顾单模态异常检测的技术进展与现有瓶颈。随后, 重点探讨在开放场景下如何突破传统闭集训练的局限, 将依赖于已知数据分布的表征学习, 拓展为面向未知或部分未知分布的开放场景推理范式。最终, 报告旨在探讨如何构建兼具高鲁棒性与强泛化能力的异常检测范式, 为该技术在工业生产环境中的落地应用一些个人思考与初步探索。



吴冠尧

吴冠尧，大连理工大学软件学院博士二年级学生，导师为刘日升教授，主要研究方向为多模态图像融合与图像处理。在 TPAMI、CVPR 等顶级期刊会议发表十余篇论文，其中 6 篇为一作 / 学生一作，3 篇入选 ESI 高被引 / 热点文章，目前谷歌学术引用超 3000 次。主持国家自然科学基金青年学生基础研究项目，入选中国科协青培工程博士专项，曾获校级学生最高荣誉优秀研究生标兵、JAL 国际奖学金、多次国家奖学金等，担任多个会议与期刊的审稿人。

报告题目：面向复杂场景感知的多源图像融合：从视觉增强到任务自适应

报告摘要：多源图像融合旨在综合不同传感器或不同成像条件下的互补信息，为复杂环境中的视觉感知提供更加完整可靠的场景表征。传统图像融合方法主要关注融合结果的视觉质量，但在真实应用中，融合图像不仅需要“看得清”，更需要服务于高层视觉任务。本报告将围绕复杂场景感知中的多源图像融合问题，介绍我们在视觉增强导向到任务自适应导向一些实验探究与发展思路，分析现有多模态图像融合领域的进一步发展方向，并介绍相关初步探索的新型工作。



陈思蕊

陈思蕊，同济大学计算机科学与技术学院博士研究生，导师为赵生捷教授，研究方向为大模型认知智能与大模型智能体。先后在 AAAI、ACL、ICLR、ICML、EMNLP 等人工智能领域权威会议及期刊发表学术论文 10 余篇。主持国家自然科学基金青年学生项目，曾获同济大学优秀毕业生、同济大学优秀学生等荣誉，多次获得本科生、研究生奖学金。

报告题目：大模型认知优化探索

报告摘要：本报告围绕“大模型如何获得更强的认知能力”这一核心问题展开，重新梳理大模型认知的多层内涵，并介绍我们在认知优化方向上的三项探索。针对智能体推理冗长、开销过高的问题，我们提出双效率偏好优化方法，使模型学会兼顾任务成功与表达简洁，在保证能力的同时减少不必要的认知开销，实现更加高效的推理。受人类元认知机制启发，我们进一步将模型对自身知识边界与推理状态的监控信号转化为奖励，以强化其自我校准与深层推理能力。最后，我们探索通过内化世界模型增强语言智能体的前瞻性推理能力，使模型能够在行动之前预演环境变化、预测潜在后果，并基于内部想象完成规划与决策，从而在复杂任务中展现出更稳健、更具远见的行为能力。上述工作共同勾勒出一条从浅层模仿走向认知优化的技术路径。



窦士涵

窦士涵, 复旦大学博士生, 导师为桂韬副研究员、张奇教授和黄萱菁教授, 主要研究方向为语言智能体的上下文能力提升。以第一作者身份在高水平国际会议和期刊发表论文 10 余篇, 谷歌学术总被引用 6400 余次, 曾获 NeurIPS 2023 Instruction Workshop 最佳论文奖, 开源成果被 Hugging Face TRL 项目采纳。曾在米哈游、字节跳动 Seed 等核心基础模型团队实习, 现于腾讯混元实习。曾担任 AAAI 2026 领域主席, 入选中国电子学会 - 腾讯博士生科研激励计划、青源 InnoVibe 2026 最受瞩目学术新星、中国科协青年科技人才培养工程博士生专项计划, 并获批国家自然科学基金青年学生基础研究项目(博士研究生)。

报告题目: 以可观测性驱动的coding agent harness自进化

报告摘要: 随着coding agent进入更为长程的编程任务, 模型能力的释放越来越依赖外部 Harness。传统Harness的迭代依赖人工经验, 难以跟上模型和任务变化。本次报告我们介绍一种全新的harness优化方法, AHE: 一种可观测性驱动的harness自进化框架, 通过组件、轨迹经验和修改决策三层可观测性, 让智能体自动定位失败原因、修改工具, 中间件和记忆模块等并验证效果。实验上, AHE让GPT-5.4在terminal-bench 2上从69.7%提升到77.0%, 超过人类设计的codex harness和多种自动进化基线。进化出的harness在固定后仍能迁移到SWE-bench-verified和多个不同模型, 这证明了进化出来的并非刷榜技巧, 而是一种可复用的框架和经验。



郭苒辰

郭苒辰, 北京航空航天大学博士生, 导师为徐迈教授和蒋铨副教授, 研究方向为面向压缩的图像视频质量优化、全景视频感知与重建等。发表 TPAMI、CVPR 等论文 13 篇, 一作论文 6 篇。主持国家自然科学基金博士生专项、北航博士生卓越学术基金, 获评中国科协青培工程博士生专项、国家奖学金等。

报告题目: 基于扰动生成的视频压缩预处理

报告摘要: 随着高清视频流媒体需求的激增, 视频压缩在缓解带宽压力与保障画质方面发挥着至关重要的作用。然而, 鉴于现有的视频编码器优化与后处理技术已逐渐趋于性能瓶颈, 本研究跳出传统框架, 将优化重心前移至压缩预处理阶段。具体地, 针对实际编解码器的“黑盒”特性带来的深度学习优化难题, 我们提出向输入视频中引入“正向扰动”, 以提前补偿压缩带来的失真。在具体方法上, 本研究设计了一种交替更新机制: 首先, 利用基于攻击的个体优化方法, 直接针对实际编解码器生成定制化的正向扰动; 随后, 通过分布感知的对抗学习进行全局优化, 利用上述扰动训练一个轻量级预处理网络, 使其具备处理任意视频的强泛化能力。该方案有效突破了传统限制, 为改善视频压缩中码率与感知质量的平衡提供了全新思路。



杨 震

杨震，湖南大学人工智能与机器人学院2023级博士研究生，导师为方乐缘教授。主要从事工业缺陷检测技术及应用研究。截至目前以第一作者在IEEE TPAMI、IJCV、IEEE TIP、TMM、TCSVT、TGRS、TIM等权威期刊发表SCI论文15篇。多次获得国家奖学金，获得河南省优秀硕士论文、河南省研究生创新之星、河南省优秀科技论文一等奖、中国研究生电子设计大赛华中赛区一等奖等国家级、省级奖励共7项。作为项目负责人主持了国家自然科学基金青年学生基础研究项目（博士研究生）、湖南省研究生科研创新项目，作为骨干成员参与了国家自然科学基金委杰出青年基金项目、中国移动产业智能联合研究院揭榜挂帅项目等重大项目。

报告题目：工业缺陷开放集检测

报告摘要：工业缺陷检测是智能制造质量管控的关键环节，然而实际产线中未知类缺陷（如新型划伤、非预期形变等）频繁出现，传统闭集模型仅能识别已知类缺陷，难以在保持已知类判别精度的同时实现对未知类的可靠拒判，这已成为开放动态场景下的核心瓶颈。本报告围绕这一问题，开展了三项研究工作：分别从稳健决策边界塑造、高频异常信号增强与类别解耦表征等角度出发，强化已知类的紧致性并提升对未知类的敏感性，形成从特征提取到决策判别的完整开放集识别策略。同时结合典型场景，形成了一套轻量化部署系统，在真实产线中完成了验证。



郭子晗

郭子晗，中山大学与上海创智学院 2025 级博士生，主要研究智能体互联网、大模型多智能体协同与分布式协作系统，其研究与工程实践围绕科研智能体网络、多智能体经济系统、智能体 Skill 系统等方向展开，相关成果已形成二十余篇论文，发表于 Nature Sustainability、Nature Communications、ICML 等国际知名期刊与会议。

报告题目：科研智能体网络：从单点智能体到可信协作基础设施

报告摘要：近年来，AI Research Agent已经能够参与文献综述、实验设计、代码实现、数据分析和论文写作等科研环节。然而，要走向真正的自主科研，系统需要处理的不只是单个智能体的任务执行能力，而是一个多阶段、多角色、多资源、多证据链的协作过程：参与者需要被发现和组织，过程需要被记录和审计，贡献需要被归因，真实数据、算力、仪器、机器人和实验环境也需要被安全接入。本次报告将介绍“科研智能体网络”这一面向 Autonomous Research的协作基础设施设想，讨论如何从单点科研智能体走向可信、可审计、可归因、可接入真实资源的科研智能体网络。



李春一

李春一，上海交通大学 & 南洋理工大学联授博士生，上海人工智能实验室评测专项组成员，师从张文军院士、翟广涛教授、与 Weisi Lin 教授。主要研究方向为多媒体信号处理、具身智能。第一作者发表 SCI 一区期刊 /CCF-A 会议论文多篇，被引三千余次。长期参与全国信标委、上海人工智能行业协会标准制订，在 ACMMM 等国际会议三次牵头举办专题会议与讲习班，担任顶刊 TPAMI/JSAC/TIP 审稿人，ICME 领域主席。曾获博士生国自然科学基金、博士生科协青培、中国电子学会学生会员资助计划、国家奖学金等荣誉。

报告题目：具身智能自我认知的心理学研究

报告摘要：当前，具身智能正从基础的物理感知，向高阶的社会交互发生质的飞跃，但目前的评测体系仍停留在单一机械任务的成功率层面，严重缺乏对智能体在复杂物理世界中认知边界与安全伦理的度量。引入心理学研究，量化评估具身智能体是否具备“自我意识”（如镜面测试、本体边界感知），是打破认知黑盒、确保人机高度协同安全可信的核心。将这一前沿的机器心理学研究转化为标准化的评估体系，旨在为整个具身智能产业提供客观、量化的“认知成熟度”准绳，从根本上防止机器智能在真实世界中的失控风险。

分论坛 托举工程博士生专项论坛

时间：8月6日上午09:00-12:00 地点：206会议室

内 容

嘉 宾

主席：朱翔昱（中国科学院自动化研究所）、徐婧林（北京科技大学）
李华光（云南大学）

从“泛化”到“记忆”：医学图像域泛化的灾难性遗忘	宋禹成（中南大学）
无人系统的自主感知定位与安全环绕控制方法研究	岳晓辉（东北大学）
WeDetect：统一多模态细粒度视觉表征的构建与应用	付盛豪（中山大学）
遥感多模态生成与理解基础模型研究	刘辰阳（北京航空航天大学）
面向真实场景的高效高保真影像增强	陈 铮（上海交通大学）
细粒度场景下的多模态人体动作分析方法	许煌标（福州大学）
从疾病共病因子到临床前风险量化— 多模态机器学习驱动的精准预防新框架	陈 锦（昆明理工大学）
通往全模态行人重识别	左嘉龙（华中科技大学）
DuQuant:Distributing Outliers via Dual Transformation Makes Stronger Quantized LLMs	林浩坤（中国科学院自动化研究所）
Hive：面向数据高效查询式通用声音分离的语义一致数据集	李 凯（清华大学）
面向复杂遥感场景的知识蒸馏体系构建	刘焰锋（武汉大学）
高光谱融合成像方法	刘元晔（湖南大学）

论坛主席



朱翔昱

朱翔昱，中国科学院自动化研究所项目研究员、博士生导师。研究兴趣包括：生物特征识别、三维重建与 AI4Science。入选国家级青年人才计划，国际模式识别协会(IAPR)生物特征识别青年学者奖(YBIA)获得者(两年一次，每次从全球范围内评选 40 岁以下学者一名)。共发表论文 100 余篇，发表文章的 Google Scholar 总引用次数为 12000 余次，入选百度学术全球华人 AI 青年学者榜单(全球 25 人)。获得三次国际竞赛冠军以及四项最佳论文及提名奖。授权国家发明专利 24 项。任国际期刊 T-IFS(CCF:A)、Pattern Recognition(CCF:B)、T-BIOM 编委，国际生物特征识别大会 IJCB 2025 程序主席，中国图像图形学会(CSIG)青托俱乐部副主席，中国人工智能学会模式识别专委会(CAAI-PR)副秘书长。获中国图像图形学会自然科学二等奖(第一完成人)，中国电子学会科技进步二等奖。承担了国家自然科学基金重大项目(培育项目)、面上项目、中科院先导 A 专项课题、北京市联合项目以及华为、腾讯、阿里、商汤、Minimax、美的等校企合作项目。



徐婧林

徐婧林，北京科技大学人工智能学院特聘教授，主要从事图像/视频理解、细粒度运动分析等研究，已发表 TPAMI、IJCV、CVPR 等领域高水平论文 40 余篇。主持国家自然科学基金(青 B、面上、专项、青 C)、北京市自然科学基金面上、腾讯犀牛鸟专项研究计划等项目。入选中国科协青托、小米青年学者，获中国图像图形学会石青云女科学家奖、中国自动化学会自然科学奖一等奖(4/5)、中国图像图形学会自然科学奖二等奖(3/5)等奖励。担任北京图像图形学学会理事、副秘书长，中国图像图形学会青托俱乐部副主席、青工委副秘书长。担任《Machine Intelligence Research》青年编委、《ENGINEERING Information Technology&Electronic Engineering》青年编委、《电子与信息学报》编委、《计算机科学》执行编委等。



李华光

李华光，工学博士，现任云南大学信息学院教师，硕士生导师。长期从事人工智能与计算机视觉研究，聚焦多视角三维重建、散焦估计、医学病灶检测及多模态图像融合等前沿领域，致力于构建面向复杂场景的视觉感知与智能理解理论体系，并推动人工智能技术在智慧医疗、数字文化遗产及跨学科场景的创新应用。在国际高水平期刊和会议发表论文 10 余篇，包括 Knowledge-Based Systems、IEEE Transactions on Instrumentation and Measurement、IEEE Sensors Journal 及 AAAI 等，主持云南省科技厅及教育厅科研项目多项，积极指导本科生和研究生获国家级及省级创新创业竞赛奖项。

论坛讲者



宋禹成

宋禹成，中南大学计算机学院博士，导师廖志芳教授。主要研究方向为计算机与医学影像交叉研究，致力于探索人工智能技术在医学影像分析中的创新应用与理论基础。近年来，围绕深度学习在医学影像任务中的关键挑战展开系统研究，以第一作者或通讯作者身份在 CVPR、IJCAI 等人工智能与计算机视觉领域期刊 / 会议发表学术论文二十余篇。

报告题目：从“泛化”到“记忆”：医学图像域泛化的灾难性遗忘

报告摘要：医学图像分割模型在实际应用中常常受到数据分布差异的影响，导致模型在新医院、新设备或新场景下性能下降。现有域泛化研究主要关注如何提升模型在未知数据上的表现，但往往忽略了一个问题：模型在增强泛化能力的同时，可能会遗忘原有数据上的知识。本次汇报介绍一项面向医学图像分割的域泛化研究。该工作通过生成更多样化的训练数据来模拟潜在的未知场景，并利用协同学习机制同时兼顾“适应新域”和“保持原有知识”两个目标。在此基础上，进一步设计了针对性的引导策略，使模型能够在泛化能力与知识保持之间取得更好的平衡。实验结果表明，该方法不仅能够提升模型在未见数据上的分割性能，还能有效降低源域知识遗忘现象。相关研究为构建更加稳定、可靠的医学图像智能分析系统提供了新的思路。



岳晓辉

岳晓辉，东北大学控制科学与工程专业博士研究生，导师张化光教授。长期致力于智能协同控制、安全强化学习、无人机编队 / 环绕控制等前研方向研究。入选首届“中国科协青年人才托举工程博士生专项计划”，连续两年获得研究生国家奖学金，获评“辽宁省优毕博士”，东北大学五四奖章(十佳研究生)等荣誉称号，主持中央高校基本科研业务费项目 1 项。荣获国际会议论文 CSIS-IAC 2026 最佳论文奖，累计发表 SCI 论文 35 篇，其中以第一作者身份在控制领域顶级期刊 IEEE 汇刊上发表论文 14 篇。目前担任 IEEE TC, IEEE TFS, IEEE TVT, IEEE TASE 等期刊审稿人。

报告题目：无人系统的自主感知定位与安全环绕控制方法研究

报告摘要：相较于轨迹跟踪/协同编队等研究较为成熟的范式，针对非合作目标的自主环绕控制已成为无人系统执行持续监视、目标围捕和区域侦察等任务的重要研究课题。本报告围绕目标信息获取、环绕轨道设计及安全环绕控制三个关键环节展开研究，旨在构建一套面向未知障碍环境的非合作目标自主安全环绕控制理论与方法体系。首先，利用无人系统配备的高精度传感设备，构建基于视线方位信息的精细化目标快速定位机制，有效解决非合作目标信息的快速获取与精准定位难题。其次，为消除传统圆形/椭圆环绕机制在任务灵活性上的固有缺陷，提出任意光滑环绕轨道生成器，确保无人系统能够依据任务需求，灵活规划并精准沿预设环绕轨迹完成对未知非合作目标的动态环绕。最后，引入控制障碍函数，构造兼顾高性能与严苛安全性约束的环绕控制器，实现未知障碍环境下无人系统的安全环绕运行。



付盛豪

付盛豪，现为中山大学计算机学院四年级博士研究生，导师为郑伟诗教授和谢晓华教授。目前，他主要的研究方向包括 2D 目标检测（通用目标检测和开放词汇检测）和多模态大模型（全模态感知与推理），研究产出的多篇论文以第一作者身份发表于 CVPR、ICCV、NeurIPS、ICML、TMM 等高水平会议、期刊，其中一篇论文入选 CVPR2025 Highlight，同时他也担任 CVPR、ICCV、NeurIPS、ICML、ICLR 等 CCF-A 类会议审稿人，并获得 2025 ICCV 杰出审稿人称号。

报告题目：WeDetect：统一多模态细粒度视觉表征的构建与应用

报告摘要：理解与表示图像中的物体是理解图片的关键，本报告围绕基于统一多模态细粒度视觉表征技术展开，WeDetect 系列模型与 ObjEmbed 模型的协同应用，构建了“快速召回-Embedding 精排-Reranker 重排”的三级级联技术链路。在快速召回层，WeDetect/WeDetect-Uni 以检索范式实现开放域目标检测，可为粗糙的类别高效生成高质量目标候选区域，为后续处理提供低延迟、高覆盖的粗匹配结果；在 Embedding 精排层，ObjEmbed 基于多模态对象级嵌入设计，将召回的候选区域转化为兼顾语义信息与定位信息的统一细粒度表征，实现多任务通用的特征级精细化排序；在 Reranker 层，WeDetect-Ref 依托指代理解能力，结合上下文信息对精排结果进行校正与重排，解决语义匹配与用户意图的对齐问题。该系统能够在统一的框架内解决目标检测、前景框提取、局部图像检索、指代表达式理解等细粒度感知与理解任务。



刘辰阳

刘辰阳，北京航空航天大学博士生，导师为史振威教授，研究方向为遥感图像多模态解译，主持国家自然科学基金博士生项目，曾获中国科协博士青培专项、北航沈元奖章（学生最高荣誉）、五四青年奖章等荣誉。在 IEEE TPAMI、GRSM、TGRS 等期刊 / 会议发表论文 20 余篇，其中一作论文 10 篇。5 篇入选 ESI 热点论文，10 篇 ESI 高被引论文。谷歌学术总引用 3000 余次，研究成果被国家自然科学基金委、中国图像图形学学会等报道，服务于 3 型卫星地面系统。受邀担任 IEEE GRSM (IF=16.4)、IEEE TIP (IF=13.7) 等 20 个 SCI Q1 区期刊审稿人。

报告题目：遥感多模态生成与理解基础模型研究

报告摘要：近年来，生成式人工智能与视觉语言模型的发展推动了基础模型研究向多模态生成与理解的演进。遥感数据具有多模态、跨尺度与复杂时空变化等特点，给基础模型的构建带来了新的挑战。本报告围绕遥感多模态生成与理解基础模型展开，介绍本人在遥感场景图像生成和时序图像视觉语言理解方面的研究工作，重点探讨如何构建面向全球尺度的遥感场景统一表示、生成与理解能力。最后分享团队的遥感多模态研究综述，全面回顾该领域的发展历程，为未来研究者提供参考。



陈 铮

陈铮, 上海交通大学计算机学院 2024 级博士研究生, 导师为张宇伦, 研究方向为图像与视频复原、模型加速和多模态模型。已在 NeurIPS、ICCV、ICLR、ICML、CVPR 等 CCF-A 类会议发表论文 21 篇, 其中第一作者论文 10 篇(含 2 篇 Spotlight), Google Scholar 引用 1600 余次, 开源项目获 1700 余次 Star。连续四年担任 CVPR NTIRE Workshop 共同组织者, 并参与相关领域国际顶级会议与期刊审稿工作。曾获研究生国家奖学金、上海交通大学优秀毕业生等荣誉, 并入选中国科协青年科技人才培养工程博士生专项计划和国家自然科学基金青年学生基础研究项目。

报告题目: 面向真实场景的高效高保真影像增强

报告摘要: 真实场景中的图像与视频常面临低分辨率、运动模糊、压缩失真等复合退化, 而移动终端、视频平台和智能感知系统又要求影像增强模型在有限算力下输出清晰、自然且可靠的结果。本报告围绕如何在真实场景中同时实现高保真复原与高效推理这一问题, 介绍面向图像与视频复原的高效影像增强研究。报告将通过长程建模与聚合表征提升复杂内容的细节恢复能力, 通过低比特量化与单步扩散推理降低模型存储计算开销, 并结合退化先验与多模态质量感知增强真实复原结果的可靠性, 分享高效影像增强从基础方法到端侧智能与产业应用的探索。



许煌标

许煌标, 福州大学博士生, 导师为柯道教授, 主要研究人工智能、计算机视觉、人体动作分析。2025 年入选“中国科协青年科技人才培养工程博士生专项计划”, 在 TIP、TMM、ICML、CVPR、ECCV 等国内外顶级期刊会议发表论文二十余篇, 其中以第一或通讯作者在 CCF-A 类、中科院一区等领域顶级期刊会议发表 12 篇论文, 以第一或学生第一发明人授权国家发明专利 5 项, 软件著作权 1 项。作为团队负责人, 曾获 2 项国家级、5 项省部级、9 项院校级人工智能或创新创业竞赛奖项。曾获博士研究生国家奖学金、国家励志奖学金、福建省青年志愿服务者等三十余次个人荣誉和奖项。

报告题目: 细粒度场景下的多模态人体动作分析方法

报告摘要: 人体动作分析是计算机视觉与人工智能领域的重要研究方向, 在智能监控、人机交互、智慧医疗和竞技体育等场景中具有广泛应用价值。然而, 细粒度场景下的人体动作分析面临类内动作差异细微、语义理解复杂、模态信息缺失以及跨场景泛化能力不足等挑战。本报告围绕“细粒度场景下的多模态人体动作分析方法”展开, 介绍在多模态特征融合、不完整多模态学习以及跨场景动作分析等方面的研究工作, 探讨如何充分挖掘多模态信息的互补优势, 提升细粒度动作识别与理解能力, 并构建面向复杂场景的统一分析框架, 为多模态人体动作分析技术的发展提供新的思路与方法。



陈 锦

陈锦, 昆明理工大学灵长类转化医学研究院 / 省部共建非人灵长类生物医学国家重点实验室, 医学统计与机器学习方向博士, 师从王骄健教授。研究方向聚焦神经影像数据分析、机器学习、脑连接组学、影像 - 遗传学关联分析、精神疾病亚型分型。擅长利用统计建模与多模态数据整合揭示疾病的神经 - 遗传 - 环境机制。目前以第一作者在 Nature Communications、Communications Medicine 等期刊发表论文, 并以合作作者身份在 Advanced Science 等期刊发表论文 5 篇。

报告题目: 从疾病共病因子到临床前风险量化—多模态机器学习驱动的精准预防新框架

报告摘要: 精神与躯体疾病共病已成为全球疾病负担的核心挑战, 然而其共同生物学基础及临床前预警机制至今仍不清晰。本报告围绕“共病的机制解析”与“发病前的风险量化”两个递进问题, 汇报两项系统性工作。基于大样本数据库的多模态数据, 通过机器学习与多组学分析, 系统揭示了跨越精神与躯体疾病的一般疾病因子 (d 因子) 的神经-遗传-环境生物学基础, 阐明了共病的共同机制; 在此基础上, 进一步开发了面向临床前健康人群的生物-心理-社会风险评分 (RS4DV), 整合机器学习与多组学验证, 实现了对未来疾病累积与全因死亡的长期精准分层预测, 并以神经影像与蛋白组学为生物学锚点, 揭示了其生物学意义。两项工作共同构建了从机制解析到风险量化的完整研究链路, 为精准预防提供了可解释、可推广的新工具。



左嘉龙

左嘉龙, 华中科技大学人工智能与自动化学院博士研究生, 导师为高常鑫教授和桑农教授。研究方向为行人理解、视频理解与生成以及多智能体系统等, 以第一作者在 CVPR、NeurIPS 等顶级国际会议上发表多篇学术论文。曾获硕士及博士国家奖学金等荣誉, 入选 2025 年度中国科协青年科技人才培养工程博士生专项计划。先后在华为诺亚方舟实验室、上海人工智能实验室和阿里云通义实验室从事算法见习研究。作为第一发起人策划并组织 PRCV2025 首届全模态行人重识别挑战赛; 获 2025 年度中国青年科技创新“揭榜挂帅”擂台赛全国特等奖和一等奖, 获 CVPR 竞赛国际冠亚军。常年担任 CVPR、NeurIPS、ICML、ICCV 等多项国际会议及期刊审稿人。

报告题目: 通往全模态行人重识别

报告摘要: 现实安防场景中的行人重识别往往面临多模态查询的需求, 然而现有方法大多局限于单一或固定两种模态的查询方式。本次报告将介绍全新的全模态行人重识别问题及其解决方案。针对数据集匮乏的挑战, 我们构建了涵盖 RGB、红外、彩铅、草图和文本描述五种模态的高质量数据集 ORBench, 其在多样性上具有显著优势。同时, 我们提出了多模态学习框架 ReID5o, 通过统一编码和多专家路由机制实现任意模态组合的协同融合与跨模态对齐, 在单一模型中实现了全模态行人重识别的统一解决方案。本报告将详细介绍全模态行人重识别的任务背景、数据集构建过程、框架核心设计以及实验验证结果。



林浩坤

林浩坤，中国科学院自动化研究所与香港城市大学联培博士生，获国家奖学金，入选中国科协“青年人才托举工程”博士生专项计划。研究方向聚焦多模态大模型与大模型轻量化，相关成果发表于 CVPR、NeurIPS、ICLR、ICCV、ECCV、ACL 等国际顶级会议。已在 CCF A/B 类国际会议与期刊发表论文 17 篇，Google Scholar 累计引用 1346 次，相关工作累计获得 1000+GitHub Stars。其中，第一作者代表性工作 DuQuant 被 NeurIPS 2024 接收为 Oral Presentation(录取率约 0.4%)，在大模型量化方向产生了较高学术影响力，并已将相关方法成功拓展至 DLLM、DiT、VLA 等模型架构。

报告题目: DuQuant:Distributing Outliers via Dual Transformation Makes Stronger Quantized LLMs

报告摘要: 当今 LLM 中存在非常大的激活值 (outliers)，为低比特量化带来了巨大挑战，我们在实验中发现 LLM FFN 模块中的 down_proj layer 存在明显的 massive outliers，表现为大于几百的异常值并局限于个别的 tokens 中，这些 massive outliers 造成 SmoothQuant 和 OmniQuant 等算法在 4bit 权重激活量化中表现糟糕。为了更好消除 massive outliers 和 normal outliers，DuQuant 通过学习旋转矩阵和平移变换矩阵，在激活矩阵 (Activation) 内部将 outliers 转移到其他通道，最终得到平滑的激活矩阵，从而大幅度降低了量化难度。旋转矩阵和平移变换矩阵都是正交矩阵，保证了权重激活输出 (XW) 的不变性，我们还通过严谨的理论推导证明了两种变换有效降低了量化误差。DuQuant 在 4-bit 权重激活量化 setting 下达到了 SOTA 的效果，后续工作将 DuQuant 拓展至 DiT、DLLM 和 VLA 模型。



李 凯

李凯，清华大学 TSAIL 实验室博士在读。他致力于多模态与单模态语音分离，当前拓展至通用音频理解任务。曾获中国图像图形学会优秀硕士论文，NCMMSC 最佳学生报告。曾参加 Sony 组织的分离竞赛获电影分离第一名。以第一作者身份在 NeurIPS、ICLR、ICML、TPAMI 等会议和期刊发表多篇论文，相关研究聚焦语音分离在真实场景中的应用，特别是在神经科学启发下设计轻量化模型，以实现边缘设备的高效部署。相关项目在 GitHub 上累计获得超过 5k stars。

报告题目: Hive: 面向数据高效查询式通用声音分离的语义一致数据集

报告摘要: 本次报告围绕查询式通用声音分离任务展开。现有方法虽然在模型结构和多模态提示方面取得进展，但训练数据通常来自弱标注的真实场景音频，存在事件共现严重、标签与声音信号不对齐等问题，导致模型容易把背景噪声或伴随事件误学为目标声音的一部分。为缓解这一数据瓶颈，我们提出了一套自动化的数据清洗与合成流程，从野外音频中挖掘高纯度单事件片段，并通过语义一致的混合策略构建Hive数据集。Hive包含约2.4k小时高质量音频，支持训练查询式声音分离模型。实验表明，使用Hive训练的开源模型在数据规模远小于百万小时级基线的情况下，仍能取得具有竞争力的分离效果，并展现出良好的零样本泛化能力。该工作说明，提高监督信号的纯度和语义一致性，是训练高效、可复现的通用声音分离模型的重要路径。



刘焰锋

刘焰锋，武汉大学计算机学院计算机科学与技术专业 2024 级博士研究生，北京中关村学院联合培养博士生，导师为张乐飞教授。研究方向为多模态遥感信息处理。以第一作者 / 学生一作 / 共同一作在 ACM MM、IEEE TGRS 等本领域权威期刊会议上发表论文 11 篇，入选 ESI 热点论文 1 篇、ESI 高被引论文 5 篇，谷歌学术被引近 1000 次。曾获得教育部国家奖学金 5 次、武汉大学研究生学术创新二等奖、2024 年西北工业大学第 22 届研究生学术年会一等奖、2024 年中国图像图形学会硕士学位论文激励计划、2025 年中国图像图形学报研究生学术论坛一等奖、2025 年中国科协青年科技人才培养工程博士生专项计划、2026 年中国电子学会“信号处理研究生激励计划(博士生报告)”。同时，他积极参与学术服务，担任二十余个国际期刊 / 会议审稿人，获 2024 年 IEEE TIM 期刊 Outstanding Reviewers Award，并入选 2025 年 IEEE GRSL 期刊 TOP50 Reviewers。

报告题目：面向复杂遥感场景的知识蒸馏体系构建

报告摘要：当前，面向遥感场景的知识蒸馏研究主要聚焦于模型轻量化，其研究范式通常局限于单一模态、相同尺度、同构任务以及充足标注条件下的知识传递，对于跨尺度、跨模态、跨任务以及低标注场景下的知识迁移机制仍缺乏系统研究，难以应对多源异构数据并存、输入条件多变、任务目标多样化以及标签样本获取受限等现实挑战。围绕“如何突破传统知识蒸馏迁移边界”这一核心问题，本报告构建了一套面向复杂遥感场景的知识蒸馏体系，具体的研究内容如下：

- 1) 针对低分辨率输入下的性能退化问题，研究跨分辨率知识蒸馏机制，实现高分辨率表征向低分辨率模型的有效迁移，提升低分辨率条件下的目标感知能力；
- 2) 针对单一成像模态下目标感知能力下降的问题，研究跨模态知识蒸馏机制，实现多模态表征向单模态模型的高效迁移，增强模态缺失场景下的鲁棒检测性能；
- 3) 针对遥感场景语义利用不足的问题，研究跨任务知识蒸馏机制，实现场景分类信息向显著目标检测任务的有效迁移，提升目标定位精度；
- 4) 针对低标注场景下模型训练受限的问题，研究跨架构知识蒸馏机制，构建视觉基础模型与轻量化半监督模型之间的知识桥梁，显著提高低标注学习效率。

上述研究在遥感目标检测、显著目标检测和道路提取等多个典型任务上显著优于现有技术，验证了所构建知识蒸馏体系的有效性与通用性，为复杂遥感场景下高效、鲁棒、通用视觉理解模型的构建提供了新的研究思路。



刘元晔

刘元晔，湖南大学在读博士，研究方向为高光谱计算成像，以第一作者 / 导师一作（本人二作）在 IEEE CVPR、IJCV、IEEE TNNLS 等国际顶级会议和期刊发表论文 6 篇，其中 ESI 高被引论文 1 篇，授权发明专利 3 项。入选中国科协青年科技人才培育工程博士生专项计划，获得博士研究生国家奖学金、2025 年中国国际大学生创新大赛湖南省赛金奖、2025 年“金种子杯”大学生创业大赛银奖等荣誉。

报告题目：高光谱融合成像方法

报告摘要：高光谱成像能够同时获取目标的空间与光谱信息，在遥感监测、精准农业和医学检测等领域具有重要应用价值。然而，受成像机理与器件制造工艺的限制，高光谱成像长期面临时空分辨率相互制约的难题。针对这一难题，提出基于马赛克编码的高光谱视频融合成像方法，将高光谱视频的空间分辨率提升 8 倍；同时，针对高光谱融合成像计算开销大、效率较低的问题，提出深度低秩张量表示方法，在保证融合质量的基础上显著提升计算效率。

中国图像图形学会简介

中国图像图形学会（China Society of Image and Graphics，缩写CSIG）成立于1990年，是经国家民政部批准成立的国家一级学会，是中国科学技术协会的正式团体会员。

中国图像图形学会自成立以来，积极团结广大电子、计算机、人工智能、自动化、通信网络、生物医学、航空航天等领域从事图像图形研究的科技工作者，推动图像图形基础理论研究与前沿技术创新，促进学科发展和在国民经济各个领域的广泛应用。学会现有5万余名会员，下设14个工作委员会、35个专业委员会，专业覆盖新一代信息技术、人工智能与智能制造、空天科技、图像信息工程、计算机图形学等前沿技术领域。

中国图像图形学会始终以学科发展为根基，锚定国家战略需求，以“学术引领 + 产业赋能 + 生态构建”为特色，系统性开展国内外学术交流、展览展示、技术培训、科学普及、人才举荐、决策咨询、成果鉴定、科技成果转化、团体标准研制、国际合作等工作，编辑出版 Visual Intelligence、《中国图像图形学报》，充分发挥联系科技工作者的桥梁和纽带作用，已成为推动图像图形创新发展的重要社会力量。



扫码注册会员



学会官方订阅号



微信视频号

昆明理工大学简介

昆明理工大学创建于1954年9月1日,时名“昆明工学院”,1995年更名为“昆明理工大学”,1999年原昆明理工大学与原云南工业大学合并组建新的昆明理工大学。学校先后隶属原高等教育部、原冶金工业部、原中国有色金属工业总公司管理,现已建成呈贡、莲华、新迎三个校区,正在全力推进嵩明新校区建设,总占地面积5418亩,主校区为呈贡校区,位于昆明市呈贡大学城。

经过71年的发展,学校现已成为一所以工为主,理工结合,行业特色、区域特色鲜明,多学科协调发展的综合性大学,是云南省规模最大、办学层次和类别齐全的重点大学,在中国有色金属行业和区域经济社会发展中发挥着重要作用。涌现了一大批享誉全国乃至世界的科技创新人才和重要科技成果,其中屈维德教授首创“冲击消震原理”,应用于我国第一颗人造地球卫星的研制;李梦庚教授发明的“粗锡电热连续结晶机除铅铋工艺及设备”,被西方冶金界誉为“20世纪锡冶金最伟大的发明”,与戴永年教授发明的“焊锡真空分离铅锡技术”联合构成了现代火法锡精炼技术,推广至全球炼锡厂;刘北辰教授发明了世界首座倒张拱钢索桥。

2015年1月21日,习近平总书记在考察云南的重要讲话中指出:“云南科教资源比较丰富,既有像昆明理工大学这样的全国著名高校,也有像中科院昆明分院这样的高水平研究机构,省属科研院所也不少。要把这些条件利用好,扎扎实实走出一条创新驱动发展的路子来。要集聚创新人才,坚持培养和引进相结合,补上云南创新人才缺乏的短板。”这既是对昆明理工大学长期立足云南、服务国家和区域经济社会发展的高度认可与赞许,也是对学校为云南经济社会发展再立新功的鞭策和鼓励,更是为学校新时代办学指明方向。遵循习近平总书记的嘱托,昆明理工大学立足云南、服务云南,面向全国、服务行业,辐射南亚东南亚,坚持学术立校、人才强校、质量兴校、开放发展、特色发展、融合发展,做强优势学科,做特支撑学科,做优新兴学科,做精人文学科,以积极有为的人才培养、科学研究、社会服务、文化传承创新和国际交流合作成效,为全国及云南经济社会高质量发展提供人才智力支持。

昆明理工大学现有 1 个学部、38 个学院(含 4 个现代产业学院、中老工程技术学院、卓越工程师学院、昆明理工大学马来西亚理工学院、创新创业学院、数字经济学院、中德学院)、7 个研究院,设有研究生院;有 30 个临床教学基地(含 9 个附属医院、17 个教学医院、4 个实习医院)。现有教职工 4199 人,其中,教授、副教授职称人员 1529 人,“两院院士”、国家级“重大人才工程”入选者等国家级人才 94 人,云南省“兴滇英才支持计划”入选者、“云岭学者”等省部级人才 1031 人。现有学生 52697 人,其中本科生 32424 人,研究生 20273 人。

昆明理工大学工程学、材料科学、化学、环境 / 生态学、植物与动物学、农业科学、临床医学、计算机科学、地球科学、生物学与生物化学、药理学与毒理学、分子生物学与遗传学共 12 个学科先后进入 ESI 全球前 1%, 工程学进入 ESI 全球前 1‰。现拥有国家重点学科 1 个、国家重点培育学科 1 个、省级重点学科 23 个、省级一流建设学科 14 个、省院省校合作共建重点学科 9 个;国家级博士后科研流动站 12 个、省级博士后科研流动站 9 个;一级学科博士学位授权点 18 个、博士专业学位授权点 5 个;一级学科硕士学位授权点 45 个、硕士专业学位授权点 24 个;有 89 个本科招生专业,其中 32 个专业通过国家专业认证(评估),43 个专业获批国家一流本科专业建设点。设有 7 个高等学历继续教育校外教学点,30 个专升本、高起本专业;19 个高等教育自学考试助学中心,4 个本、专科自学考试专业;拥有国家级专业技术人员继续教育基地,智能制造、人工智能、数字化管理师和工业互联网等 4 个国家级数字技术工程师培训资质。

2000 年以来,昆明理工大学创新发展成果亮点纷呈。先后荣获国家级教学成果奖一等奖 1 项、二等奖 10 项,省级教学成果奖 78 项;牵头获得国家科学技术进步奖二等奖 4 项、技术发明奖二等奖 5 项、国际科学技术合作奖 1 项,教育部高等学校科学研究优秀成果技术发明奖一等奖 1 项,省部级科技奖励 608 项;建有真空冶金国家工程研究中心、国家大学科技园;教育部工程研究中心 3 个、教育部重点实验室 2 个、云南省重点实验室 26 个;科技部重点领域创新团队 2 个、教育部创新团队 2 个、云南省顶尖团队 5 个,云南省创新团队 47 个;有甲级资质的综合性昆明理工大学设计研究院。

昆明理工大学是教育部深化创新创业教育改革示范高校、首批国家级创新创业学院建设单位、教育部“卓越工程师教育培养计划”实施高校、国家“中西部高校基础能力建设工程规划”学校、“国家大学生创新性实验计划”项目入选学校。学校已建成国家级一流课程 44 门、教材 16 部、教学团队 3 个、“全国高校黄大年式教师团队”4 个、实验教学示范中心(工程实践教育中心)16 个、国家级新工科(新农科、新文科)研究与实践项目 12 项；在中国国际大学生创新大赛(原中国“互联网+”大学生创新创业大赛)中累计获得国家级奖项 63 项，其中金奖 3 项、银奖 18 项、铜奖 42 项。“挑战杯”全国大学生课外学术科技作品竞赛中获国家级特等奖 1 项、一等奖 2 项、二等奖 7 项、三等奖 32 项；“挑战杯”中国大学生创业计划竞赛中获“优胜杯”2 座，国家级金奖 1 项、银奖 10 项、铜奖 32 项。

昆明理工大学立足学科、地缘优势，不断加强国际合作与交流。目前，学校与 50 个国家和地区的 201 所高校及科研机构签署了 260 余份国际合作协议，与老挝苏发努冯大学、尼加拉瓜国立自治大学(马那瓜)分别合作共建孔子学院，是云南-泰国大学联盟的中方牵头高校，在泰国设置了四个境外办学学院。与马来西亚理工大学合作设立昆明理工大学马来西亚理工学院，为学生提供优质的国际化教育。学校入选国家“高等学校学科创新引智计划”地方高校新建基地，1 名外籍专家荣获国家国际科学技术合作奖和中国政府友谊奖、2 名外籍专家荣获云南友谊奖，在面向周边国家的工程及管理人才培养、国际技术转让，面向发达国家的高水平合作研究与师生交流方面，逐渐形成了自身的特色和影响力。

新的历史时期，昆明理工大学坚持以立德树人为根本任务，秉承“根植红土、情系有色、坚韧不拔、赤诚报国”的精神和“明德任责、致知力行”的校训，坚持内涵发展、开放发展，以团结之心、实干之力、拼搏之劲，立足新发展阶段，贯彻新发展理念，构建新发展格局，加快推进“双一流”创建和特色鲜明的研究型高水平大学建设步伐，为谱写昆明理工大学高质量发展新篇章砥砺前行奋进。

云南师范大学简介

2020年1月20日,习近平总书记考察云南之际,亲临我校西南联大旧址考察调研并指出:“国难危机的时候,我们的教育精华辗转周折聚集在这里,形成精英荟萃的局面,最后在这里开花结果,又把种子播撒出去,所培养的人才在革命建设改革的各个历史时期都发挥了重要作用。这深刻启示我们,教育要同国家之命运、民族之前途紧密联系起来。为国家、为民族,是学习的动力,也是学习的动机。艰苦简陋的环境,恰恰是出人才的地方。我们现在教育的目的,就是要培养社会主义建设者和接班人,培养有历史感、责任感、志存高远的时代新人,不负韶华,不负时代。”这是习近平总书记对学校 and 全体“云师人”的殷殷嘱托,为学校在新时代加快高质量发展指明了奋进方向、提供了根本遵循。

薪火相传,弦歌不辍。88年来,历经代代“云师人”的艰苦奋斗,云南师范大学为国家培养各级各类人才40余万,被誉为“红土高原上人民教师的摇篮”“西南联大精神传承者”“云南教师教育排头兵”“云南基础教育领头雁”“服务南亚东南亚辐射中心建设先行者”,是教育部和云南省人民政府“省部共建”高校、国家中西部基础能力提升工程重点建设高校,是一所区域特色鲜明、多学科协调发展的综合性师范大学。

【历史沿革】学校前身是诞生于1938年的国立西南联合大学师范学院。1946年,组成西南联大的北京大学、清华大学、南开大学三校复员北返,其师范学院留昆独立建校,定名国立昆明师范学院,1950年改名昆明师范学院,1984年更名为云南师范大学。1998年与云南教育学院、云南体育进修学院合并重组云南师范大学。

【学科布局】学校设有哲学、经济学、法学、教育学等12大学科门类;博士学位授权一级学科8个,专业博士学位授权类别2个,博士后科研流动站6个,是全国首批获得国际中文教育专业博士学位授权点10所高校之一;硕士学位授权一级学科26个,专业硕士学位授权类别24个,化学、工程学、环境科学与工程、植物学与动物学4个学科进入ESI全球学科排名前1%;累计共有32个学科(含学科群、重点学科方向)列入云南省重点建设学科,其中8个学科被列为省级一流学科。

【专业结构】学校现有 74 个本科招生专业, 包括 21 个师范类专业、20 个国家级一流专业建设点、27 个省级一流专业建设点、25 个卓越人才计划专业, 2 个专业获省一流专业建设备案, 2 个专业获新兴专业建设备案, 3 个省级高校拔尖学生培养项目立项。

【办学规模】学校设有呈贡校区、一二·一西南联大校区 2 个校区, 设有 2 个学部、22 个学院、67 个省级以上科研机构, 65 所附属、合作、帮扶办学的中小学、幼儿园。现有普通全日制本科生 30000 余名、硕士研究生 8965 名、博士研究生 708 名, 成人继续教育学生 8469 名、国际学生 661 名。

【师资队伍】学校有教职工 2424 人, 高级职称教职工 1209 人、博士学位教职工 1110 人。在编在岗专任教师 2158 人, 外聘专任教师 536 人, 专任教师总数 2426 人。

【办学条件】学校占地面积 2848 亩。设有呈贡主校区图书馆和一二·一西南联大校区图书馆, 两馆总面积 5.73 万平方米。截至目前, 全校藏书量为 424 万册, 电子图书 158 万册, 中外文数据总库 51 个, 子库 96 个。

【人才培养】学校有 4 个国家级实验教学示范中心、2 个国家级虚拟仿真实验教学中心、4 个国家级新文科和新工科改革项目、1 项教育部体育美育浸润行动计划项目、33 门国家级和 90 门省级一流本科课程、3 门国家级精品课程、1 门国家级课程思政示范课程、1 门国家级双语教学示范课程、3 部国家规划教材。获国家级、省级教学成果奖 106 项, 获奖数量位居云南省前列。教师获全国高校教师教学创新大赛二等奖 4 项、三等奖 8 项, 获全国高校青年教师教学竞赛二等奖 1 项、三等奖 7 项。拥有国家级教学团队 3 个、国家级人才培养模式创新实验区 1 个、教育部创新团队 1 个, “全国高校黄大年式教师团队”1 个, 共有 472 项“质量工程”国家级建设项目。与北京大学、清华大学、南开大学合作, 成立西南联合研究生院, 创新开展高层次人才培养和前沿学术研究。学校入选教育部“卓越中学教师培养计划”实施院校、教育部人工智能助推教师队伍建设试点单位。

【科学研究】“十四五”期间, 学校科研创新能力显著增强, 在项目立项、平台建设、成果产出、科研获奖等方面实现系列突破。在项目立项方面, 新增国家社会科学基金 145 项(含重大项目 5 项、重点项目 8 项)和国家自然科学基金 238 项, 以及国家重点研发计划项目 2 项、青年科学家项目 1 项、国家自然科学基金重点项目 1 项、国家(地区)合作交流项目 1 项, 并获国家自然科学基金云南联合基金重点项目 2 项。在平台建设方面, 科研平台建设取得历史性进

展,实现教育部协同创新中心零的突破。在成果产出方面,高水平学术成果丰硕,发表社科类C1以上论文550篇(其中A类6篇、B类200篇,新华文摘转载10篇)和自科类高质量期刊论文3043篇(含中国科学院一区高水平论文356篇)。在科研获奖方面,获教育部高等学校科学研究优秀成果奖(人文社会科学)2项、全国教育科学研究优秀成果三等奖1项、国家民委社会科学研究成果奖(调研报告类)6项、云南省社会科学奖39项(其中一等奖3项)以及云南省科学技术奖18项(其中一等奖5项)。《云南师范大学学报(哲学社会科学版)》进入中国百强报刊,入选教育部“名栏工程”,持续入选CSSCI来源期刊,北大中文核心期刊。《云南师范大学学报(对外汉语教学与研究版)》首次入选CSSCI扩展来源期刊。

【社会服务】学校积极拓展对接边疆人才需求的多元化师资培养途径,持续实施“4+2卓越教师”培养计划,开展公费师范生、优师计划等培养项目,培养扎根农村基层中小学优秀教师和普通完全中学骨干教师,截至2026年2月,培养卓越教师、公费师范生、“优师计划”师范生5200余人。落实“县中托管”帮扶任务,“一校一策”对16所边疆民族地区县中开展办学帮扶,形成了“镇雄模式”“梁河经验”等一系列可推广复制的基础教育帮扶典型。实施“万名校长培训计划”“国培计划”“体育美育浸润行动计划”“新高考九门学科送教下乡”等项目,送教下乡覆盖全省40个县,有效助力云南基础教育高质量发展,为乡村振兴提供智力支持。在国际上率先解决马铃薯自交不亲和与自交衰退等难题,成果得到党和国家领导人重要批示。在生物酶制剂、光伏水泵、太阳能干燥技术、生物质能开发与利用、三七土传病害生物防治等领域实现了科技成果转化和产业化。“守望云心”心理服务平台覆盖全省教师、学生和家長65万人。西南联大干部培训基地累计培训全国各类干部、党员、骨干人才7万余人次。

【文化传承】学校大力传承弘扬西南联大爱国主义精神,高质量建成西南联大博物馆、校史博物馆、民主草坪、人杰园、校歌园等,西南联大物证抢救保护计划征集相关文物档案13000余份,西南联大口述史采集近百位联大老校友口述影像,西南联大文库收集整理完成特色馆藏图书4800多册、专项出版25部,成功推出西南联大组歌、话剧、纪录片、电视剧等。西南联大博物馆入选全国首批科学家精神教育基地、全国关心下一代党史国史教育基地、全国妇女爱国主义教育基地等,年均接待访客100余万人次。

【双创教育】学校是国家级大学生创新创业训练计划高校、全国深化创新创业教育改革示范高校,全国高校创业指导师培训基地,全国高校毕业生就业能力培训基地,毕业生就业工作

进入全国高校 50 强。“十四五”期间,学校建设省级拔尖人才培养项目 3 个,现代产业和专业特色学院 2 个,创新创业国家级一流课程 2 门,全国高校创新创业金课 1 门,“十四五”职业教育国家规划教材 1 本;学生获国家级创新创业大赛奖励 22 项,国家级大学生创新创业类项目立项 281 项。

【国际交流】学校是全国首批来华留学示范基地、教育部出国留学培训与研究中心、国务院侨办华文教育基地、国家汉语国际推广师资培训基地。国务院侨务办公室、云南省人民政府在我校共建云南华文学院,先后承办 8 届“汉语桥”世界中学生中文比赛和 1 届“汉语桥”世界小学生中文秀。中国—老挝国家可再生能源开发与利用联合实验室入选“科创中国”全球百佳技术转移案例。在国外建立 2 所孔子学院、3 所孔子课堂、4 所云华职业学院,与 30 多个国家(地区)的 120 余所大学和科研机构建立合作关系,累计招收 138 个国家(地区)的国际学生 2 万余人次,成为南亚东南亚青年来华留学的重要目的地大学。

近年来,学校先后被中共中央、国务院、教育部、全国总工会、共青团中央、全国妇联等授予全国先进基层党组织、民族团结进步模范集体、依法治校示范校、高校首批“三全育人”综合改革试点单位、教育系统先进集体、工人先锋号、五四红旗团委、全国巾帼文明岗等荣誉。

展望新征程,全体“云师人”将深入学习贯彻习近平新时代中国特色社会主义思想,牢记领袖嘱托,扎根云岭大地,弘扬教育家精神,赓续联大血脉,矢志教师教育,为加快建设西部领先、辐射南亚东南亚的高水平师范大学而不懈奋斗。



云南大学简介

云南大学始建于1922年,1923年正式开学,时为私立东陆大学,1934年更名为省立云南大学,1938年改为国立云南大学,是我国西部边疆最早建立的综合性大学之一。1937年,著名数学家、教育家熊庆来出任校长,一大批著名学者受聘到校任教,奠定了学校较高的发展基础和深厚的学术底蕴,开创了云大办学历史上的第一个辉煌时期。20世纪40年代,云南大学已发展成为一所包括文、法、理、工、农、医等学科在内,规模较大,在国际上有影响的中国著名大学之一。1946年,《不列颠百科全书》将云南大学列为中国15所在世界最具影响的大学之一。

五十年代院系调整,部属云南大学一些重要而有特色的系科,如航空、土木、法律、铁道等划出并入当时的北京航空学院、四川大学、西南政法学院、长沙铁道学院等高校;工、医、农等先后独立建校,并逐步发展为今天的昆明理工大学、昆明医科大学、云南农业大学、西南林业大学等高校。1978年,云南大学被国务院确定为全国88所重点大学之一。

改革开放后,云南大学获得了长足的进步。1996年首批列入国家“211工程”重点建设大学,2001年列入西部大开发重点建设院校,2004年成为教育部和云南省人民政府重点共建高校,2012年成为国家“中西部高校基础能力建设工程”和“中西部高校综合实力提升工程”实施院校,2017年成为国家首批42所“一流大学”建设高校之一,2018年跻身中西部14所“以部为主、部省合建”高校行列,2022年,继续入选第二轮国家“双一流”建设高校。

2023年4月20日,中共中央总书记、国家主席、中央军委主席习近平致信祝贺云南大学建校100周年。贺信指出:100年来,云南大学秉承“会泽百家、至公天下”的办学精神,扎根祖国西南边疆民族地区,培养了大批优秀人才,为促进民族团结进步、服务区域经济社会发展作出了积极贡献。

云南大学下设29个学院(其中,马克思主义学院为全国重点马克思主义学院)、8个研究机构,1个附属医院,1个附属中学,设有本科生院、研究生院。云南大学现有教职员工近3600人(不包括附属医院),其中专任教师2300余人。专任教师中具有高级职称人员1400余人,具有博士学位人员1800余人。学校有全日制本科生近19000人。有硕士研究生近17000人,博士研究生近2600人,学历教育国际学生近1000人。学校占地面积4351.88亩,有呈贡校区和东陆校区,公用校舍建筑面积152余万平方米,教学科研仪器设备资产总值27亿余元,图书馆纸质藏书466万余册。学校是中国教育科研计算机网络(CERNET)云南主节点单位。

学校有本科专业 85 个,其中有国家级和省级一流本科建设专业 68 个。有 7 个专业“菁英班”,10 个专业“卓越班”,45 门“国家级一流本科课程”,获 17 项国家教学成果二等奖;有 22 个一级学科博士学位授权点,4 个专业博士学位授权,43 个一级学科硕士学位授权,30 个专业硕士学位授权,形成了以民族学、生态学、统计学、生物与生物医药、特色资源开发与环境保护,以及边疆问题和区域国别研究为优势特色,学科较为齐全,人才密集的学科专业体系。

学校现有院士、国家高层次人才计划入选者等国家级人才 85 人,省部级人才 925 人,全国专业学位研究生教育指导委员会成员 5 人,全国模范教师 1 人,全国优秀教师 2 人,全国教学名师 4 人,全国教育系统先进工作者 2 人,中宣部文化名家暨“四个一批”人才 1 人,教育部“新世纪优秀人才培养计划”等入选者 10 人,中宣部宣传思想文化青年英才 5 人,国家级教学团队 4 个,科技部重点领域创新团队 1 个,教育部创新团队 1 个。

学校大力实施“学术兴校”战略,科学研究成绩显著。学校先后主持国家“973 计划”“863 计划”项目 6 项,国家水专项重大项目 2 项,国家重点研发计划 11 项,国家科技支撑计划 2 项,国家自然科学基金重大项目 2 项,国家社会科学基金重大招标项目 34 项,国家社会科学基金重大专项 16 项,国家社科基金冷门绝学研究专项团队项目 1 项,教育部人文社会科学重大项目 8 项,马克思主义理论研究和建设工程项目 2 项,国家社科基金成果文库 5 项。荣获国家自然科学基金一等奖、国家自然科学基金二等奖、全国创新争先奖、何梁何利奖、教育部高等学校科学研究优秀成果奖(人文社会科学)、国家科学技术进步二等奖、惠特克杰出生态学家奖、国际青年古生物学家“Hodson Award”奖等多项大奖。荣获云南省科学技术奖杰出贡献奖 3 项、特等奖 3 项。20 余篇论文发表于《Nature》《Science》和《中国社会科学》。学校主办有《思想战线》《云南大学学报》等学术刊物,《思想战线》是首批入选教育部“名刊工程”建设的 11 种期刊之一。

学校有 1 个全国重点实验室,1 个省部共建国家重点实验室,1 个国际联合研究中心,1 个国家技术转移中心,2 个国家级众创空间,1 个教育部哲学社会科学实验室,1 个四部委铸牢中华民族共同体意识研究基地,1 个教育部人文社科重点研究基地,2 个省部共建协同创新中心,2 个教育部国际联合研究中心,1 个教育部工程研究中心,1 个教育部野外科学观察站,3 个省部共建教育部重点实验室,1 个国家 Linux 技术培训与推广中心,1 个教育部中华优秀传统文化传承基地,1 个教育部国家教材建设重点研究基地,1 个国家大学生文化素质教育基地,1 个国家科学决策咨询研究中心。有 4 个国家级人才培养基地,1 个国家级创新创业学院,4 个国家

级实验教学示范中心, 3 项国家级卓越工程师培养计划项目, 2 个国家级实验教学和教学实验室建设研究项目, 1 个国家级应用型、复合型卓越法律职业人才教育培养基地, 4 个国家人才培养模式创新实验区以及国家示范性软件学院。拥有实验动物中心、先进计算中心、电镜中心、1.6 米多通道测光巡天望远镜等一批重大科研设施平台。学校为教育部首批深化创新创业教育改革示范高校。

学校发挥区位优势, 主动融入“一带一路”共建和云南面向南亚东南亚辐射中心建设, 牵头建设南亚东南亚大学联盟, 构建全方位、多领域、多层次的开放办学新格局。与 27 个国家的 100 余所高校和研究机构在人才培养、科学研究、产教融合等领域开展合作交流。与马来亚大学合作设立云南大学云南马来亚学院, 聚焦人工智能、新材料、新能源等战略性领域, 为学生提供优质的国际化教育。学校建有覆盖东南亚、南亚、西亚、非洲和“一带一路”沿线国家较为完备的区域国别研究体系和新型特色智库体系。建设面向南亚东南亚来华留学基地、高层次国际化人才培养创新实践基地和海外调研实习基地, 构建“学生全球胜任力”培育体系和“语言 + 专业”复合型人才培养体系, 打造以汉语教学中心、文化交流中心、教学点、HSK 考点、南亚东南亚国际传播学院为合力支撑的文化遗产创新“西南通道”。

学校积极主动服务国家和区域经济社会发展。“中缅油气管道与中国能源安全”“南方丝绸之路经济大走廊”等成果上升为国家重大决策, “铸牢中华民族共同体意识研究”“阳宗海砷污染治理”“跨境生态安全和国际河流研究”“高原湖泊治理研究”“中国周边外交研究”“政治学与边疆民族问题研究”“古生物和生命起源研究”“生物多样性资源保育与利用研究”“致密天体与高能现象”“银河系与近邻宇宙”“根结线虫生物防治技术”“多年生稻技术”等科研方向和成果直接服务于解决重大现实问题, 创造了良好的效益, 产生了重要影响, 为国家战略和地方经济社会发展作出重要贡献。

站在新的百年征程起点上, 学校将以习近平新时代中国特色社会主义思想为指导, 持续深入学习贯彻党的二十大和二十届历次全会精神, 习近平总书记致云南大学建校 100 周年重要贺信精神, 全面贯彻党的教育方针, 坚守为党育人、为国育才初心使命; 秉承“会泽百家、至公天下”的办学精神, 全面落实立德树人为根本任务, 以高质量发展为主线, 以改革创新为动力, 全面提升学校办学水平和综合实力, 加快建设扎根祖国西南边疆的中国特色、世界一流大学, 为建设教育强国, 为云南经济社会发展, 作出云南大学新的更大贡献。

CSIG 青年工作委员会简介

中国图像图形学会青年工作委员会（以下简称青工委）旨在促进青年学者之间的交流与合作、促进国内图像图形学科发展、促进国内外青年学者间的了解与合作、提供学术交流的平台、提升我国图像图形研究的国际影响力。青工委现有委员 350 人，其中有国家高层次人才计划入选者、IEEE Fellow 等人才 70 余人，国家高层次青年人才计划入选者 150 余人。近年来，多名委员荣获国家科学技术奖、中国青年科技奖、科学探索奖、青橙奖等重要奖项，已成为图像图形学领域的人才高地。青工委组织了以中国图像图形学会青年科学家会议为旗舰的多个品牌学术活动，包括珠峰论坛、CSIG 云讲堂、CSIG 科普中国行、CSIG 校企合作青年学者论坛及企业行等，组织成立了中国图像图形学会青托俱乐部、优博俱乐部。参与协办视觉与学习青年学者研讨会 (Valse)、医学图像计算青年研讨会 (MICS) 等，学术影响不断扩大，已连续 5 年获评中国图像图形学会优秀工委会。



白泽智研
BAIZEZHUYAN

引领 AI for Research 新范式

人类智慧与硅基智能协同探索未知

从科研灵感到可投稿论文

自动化科研管线辅助文献调研、代码实现、GPU实验、数据分析与论文写作



领域知识与专家库



科学假设生成



AutoPaper



具身实验机器人

研究问题 → 文献调研 → Idea生成 → 代码实现 → GPU实验 → 数据分析 → 论文成稿

科学假设自进化



基础设施支持

AutoDL.com

GPUhub.com

AiC 智青春·算未来

第八届全国校园人工智能算法精英大赛

大赛科学顾问

周志华

大赛创始人、中国科学院院士、南京大学副校长、国际人工智能联合会理事会主席

顶尖权威专家顾问

大赛拥有顶尖权威专家顾问团队，由中国科学院院士、南京大学副校长周志华教授担任赛事科学顾问，汇聚全国各大双一流高校学科带头人、重点实验室科研专家、头部AI企业技术专家组成赛事评审委员会与命题专家组，全程把控赛题质量、评审标准与赛事公平性，专业性、权威性位居前列。



官网二维码



AIC公众号

赛道赛题



算法挑战赛



算法创新赛



算法应用赛



算法主题赛



产业命题赛



算法专项赛

构建“基础-技术-应用”全链条

大赛聚焦“人工智能+”知识体系下的人才能力培养及人工智能核心技术应用，围绕国家产业结构调整和发展战略性新兴产业的人才需求设置六大赛道，全方位考查青年学子的算法创新能力、工程实践能力、创新思维能力和团队协作能力，充分彰显了“智青春·算未来”的大赛主题。

联系方式

联系邮箱：office@aicomp.cn

大赛网址：www.aicomp.cn

国赛组委会地址：南京大学鼓楼校区费彝民楼B1013室



一体化新基座，打造智能高效的 教学、实验和科研环境

学院级教学、实训、科研一体化平台——学院“新基座”，融合算力、存力、终端、云管、安全等多维度能力，打造“数智化”基础设施平台，承载学院级共性应用软件、教学应用软件、商业科研软件、开源科研软件，支撑教学、实验、科研多业务高效、智能、安全的运行。

全面赋能教学、实验、科研多场景



学院级新基座一体化承载全院的教学、实验、科研环境



高效、智能、安全的“数智化”基础设施平台



EBFlex 私有部署版·GPU智算管理平台

让团队专注科研，让算力高效流转



CSIG青年科学家会议2026

云安会堂一楼 4&5

科研算力专场

成员自助申请，即时可用

按项目/成员分配配额

内置抢占式调度，盘活闲置卡

管理员一屏掌控全局算力

为科研团队设计



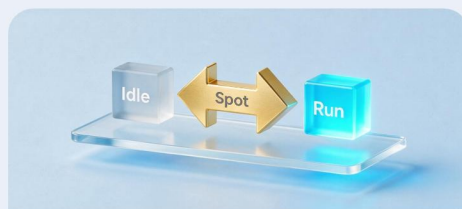
成员自助管理算力

直接申请 GPU，向团队成员分配卡时额度，无需 IT 介入。



组织统一调度

一屏查看全局使用率，为各项目组设置独立配额，消费透明可审计。



内置抢占式调度

自动识别并调度闲置 GPU，高优任务随时抢占，整体利用率显著提升。

兼容主流异构 GPU

NVIDIA

全系列

H100 · A100 · L40S · RTX

原生支持

AMD

Instinct

MI300X · MI250 · MI210

原生支持

国产

自主可控

昇腾 · 天数智芯 · 沐曦

适配接入

混合

异构集群

多型号节点统一纳管

无需换硬件

典型场景



项目组弹性实验

成员自助启动容器，管理者实时查看用量，告别抢卡混乱。



集中预约与上机

提前预约时段，保障团队同时上机，任务结束自动释放。



大模型集中训练

弹性扩容 + 抢占式调度盘活夜间闲置卡，多卡并行训练。



组织级算力报告

自动生成跨项目组报表，数据驱动 GPU 采购与预算决策。

三级管理

组织 > 团队 > 成员

内置抢占式调度

抢占式弹性调度，盘活闲置GPU

支持预约

按需即时或预设时段，错峰使用

全院可视

GPU利用率 · 消费 · 告警



九章云极 DataCanvas 领先的人工智能基础设施及 智算云提供商

九章云极 DataCanvas（以下简称九章云极）是亚洲领先的人工智能基础设施及智算云提供商，国家专精特新重点“小巨人”企业。旗下涵盖九章智算云（Alaya NeW Cloud）、九章智算操作系统（Alaya NeW OS）等业界知名品牌。面对 AI 规模化应用的新阶段，公司依托 AI 工厂驱动，面向 AI 训练和推理提供高性能算力调度、智算云服务，致力于成为智能时代核心要素的规模化生产与交付平台，赋能众多 AI 开发者，为全球客户提供中立、普惠、绿色的智算云服务。

作为全球普惠算力的倡导者，九章云极以“创造智能，探索未知”为使命，以“助力全球企业智能升级”为愿景，自主研发了完善的智算技术栈，首创并推动 DCU（一度算力）统一度量衡体系，形成国内外协同发展的格局，在优化我国算力资源配置、赋能实体经济智能化转型、激发 AI 创新动能等方面发挥积极作用，是数字中国 AI 基础设施领域的核心力量。



算、电、词协同 实现普惠算力战略



关于智慧树

About

智慧树网隶属于上海卓越睿新数码科技股份有限公司，是全球大型的学分课程运营服务平台。服务的会员学校超3000所，已有超1.6亿人次大学生通过智慧树网跨校修读并获得学分。智慧树网帮助会员高校间，实现跨校课程共享和学分互认，完成跨校选课修读。

智慧树网致力于成为中国领先的教育信息化制造商与互联网教育运营商，独特的“平台+内容+服务”，三位一体的业务模式，帮助高等院校完成优质课程的引进和服务配套落地，通过观摩和分享名校名师的优质课程设计，帮助教师完成教学发展培训，协助教师建设新课程，实现教法改革，促动本校教学产生内生动力。在几年的服务过程中，智慧树网也积累了丰富的服务经验，通过完善服务的基础设施，建立全国服务专业团队来应对大范围、大规模教学服务交付的挑战。

The diagram illustrates the integration of various teaching materials and tools into a unified smart course platform. At the top, five categories are listed: 因材施教 (Differentiated Instruction), 能力实训 (Ability Training), 作业批阅 (Homework Review), 学情追踪 (Learning Progress Tracking), and 课程工具箱 (Course Toolkit). Below these, a central blue box labeled 智慧课程 (Smart Course) is connected to five surrounding boxes: 智能体智慧课程 (AI Smart Course), 智慧课程 (Smart Course), 智慧课程 (Smart Course), 智慧课程 (Smart Course), and 智慧课程 (Smart Course). The bottom row shows three specific applications: 学情追踪 (Learning Progress Tracking), 课程工具箱 (Course Toolkit), and 课堂互动 (Classroom Interaction).

能力目标 (Ability Objectives): 能力目标, 子能力, 能力体系

知识目标 (Knowledge Objectives): 知识目标, 知识体系

问题 (Problems): 问题, 问题体系

知识点 (Knowledge Points): 知识点, 技能点, 教学, 知识体系

能力体系 (Ability System): 行业专业导论/就业分析、提取对应岗位能力要求、拆解人才培养方案能力要求、综合对比形成能力模型

问题体系 (Problem System): 将核心问题与相关的子问题、概念和知识点连接起来，形成一个结构化的问题网络，连接知识与能力

知识体系 (Knowledge System): 结构化、可视化知识点等含要素之间的结构关系，与教学资源、与用户建立关联

智慧课程—知识图谱智慧课程

智慧实践课程——虚拟仿真智慧实践课程

A futuristic, dimly lit control room or laboratory with blue and green ambient lighting. In the center, a large screen displays the text "“医有为”教育+医疗大模型" (Medical Large Model for Education and Healthcare). Above the screen is a small circular icon with the letters "AI". Below the main title, a line of smaller text reads: "大模型构建的个性化学习系统，能依据学生知识基础、学习风格 and 塑造目标，精准推送定制学习路径与内容，实现“一生一策”的因材施教" (A personalized learning system built on large models, which can accurately recommend customized learning paths and content based on students' knowledge foundation, learning style, and shaping goals, achieving "one strategy for one student" differentiated teaching). In the foreground, two humanoid robots are visible: one on the left is partially obscured, and one on the right is more prominent, facing the screen. The background shows various control panels, monitors, and structural elements of the facility.



智驭未来 共赴新程



加入长城汽车技术研发团队
共创智能出行新未来

热招岗位



基础AI算法

端到端泊车、泊车Agent开发、强化学习算法工程师、VLA预训练技术负责人、多模态/BEV表征架构负责人、多模态预训练算法工程师、视频时序预训练工程师、BEV/3D表征预训练工程师分布式训练/性能优化工程师



工程技术开发

数据挖掘工程师/专家、无图定位算法工程师、地图引擎、routing、地图众包、AVP开发、泊车感知开发、标定、智驾系统工程师、智能驾驶主动安全算法、行车规划控制算法



云和大数据

后端开发、大数据开发、infra工程师、运维、数据产线、云平台链路



扫码投递简历

让科研算力 触手可及

· 算力赋能科研 · 智能驱动创新 ·

BitaHub 彼塔云平台

一站式AI算力开发平台



算力

弹性高效
即开即用



开发

工具丰富
便捷易用



社区

资源共享
协作共创

AI开发

BitaHub

模型训练

BitaLab

科研创新

BitaLab 科研智能体

科研全流程智能伙伴



文献

快速理解
研究方向



实验

自动完成
环境配置



优化

持续寻优
更优方案



成果

论文输出
报告生成

>>> 从AI开发，到AI科研

>>> AI参与科研，而不仅执行代码



公司介绍

合肥中科类脑智能技术有限公司作为类脑智能技术及应用国家工程实验室迄今唯一的成果转化平台，历经八年异构算力调度、七年电力智能化、三年算电协同的积累，目前正全力打造软硬一体化算电协同体系，围绕“**度电智能**”重构AI基础设施，搭建专业化**Token工厂**，让万千企业、科研机构便捷高效地使用算力Token。



软硬一体
算电协同



度电智能
高效利用



Token工厂
便捷供给



赋能企业
科研创新

有算力就用趋动, 没有算力就用趋动云

关于趋动云

趋动云是面向企业、科研机构和个人 AI 开发者构建的开发和推理训练服务, 也是全球首个基于 GPU 算力池化云的服务。

趋动云的使命是连接算力·连接人:

通过连接全球算力, 趋动云可以为用户提供便宜、好用的AI算力。通过为AI算法开发全流程提供优化服务、构建全球开发者项目和数据社区, 趋动云可以帮助AI开发者接入丰富的生态, 快速实现最佳实践。



便宜、好用、生态丰富的AI算力池化云

趋动云优势



快速切换

无需更改环境或迁移数据,
即可秒级换算力、换节点



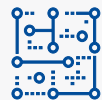
团队管理

管理员拥有管理空间成员金额配
额、资源配额、任务进度等权限



省钱

小实例开发、大实例训练,
秒级切换, 按分钟计费



社区

公开数据集和模型全覆盖,
huggingface镜像站直通



添加小助手微信领取试用算力金

关于趋动科技

趋动科技是全球AI算力资源池化领导者, 2019年成立于北京中关村高新技术园区, 已先后获得国家高新企业、北京市“专精特新”企业认证, 并被评为“REAL 100创新家”、创业邦100未来独角兽、投中2021年度中国人工智能与大数据产业最佳投资案例Top10等。趋动科技的 OrionX 猎户座 AI 算力资源池化软件能够帮助用户提高资源利用率和降低TCO, 提高算法工程师的工作效率。

青软创新科技集团股份有限公司

国内领先的高等教育数字化解决方案，产教融合及人力资源服务提供商

青软创新科技集团股份有限公司创立于2006年，是国内数字化人才培养及服务领域中的领军企业。面向新兴产业发展及产业转型升级中的数字化人才需求，为高校及企业提供：数字化平台及实训室解决方案、产教融合人才培养服务、技术服务及人力资源服务。

 800+ 创新及科技人才	 3/大 产教融合基地	 700+ 合作高校	 6000+ 合作企业	 30000+ 在校生	 25万+ 累计为产业输送人才
---	---	--	---	---	---



关注青软集团公众号

◎ 面向 ICT 方向专业打造实训室及解决方案



AIGC大模型人才培养解决方案

帮助学生掌握预训练、微调等主流大模型开发技能。顺应大模型发展趋势，培养新一代AIGC大模型人才。



鸿蒙特色物联网实训室解决方案

当鸿蒙遇上物联网，培养万物智联信创人才。



人工智能实训室解决方案

适配AI通用开源和国产自主创新技术，注重AI全栈工程能力培养。



特色软件开发实训室解决方案

融入行业主流技术和华为创新技术，培养软件方向信创人才。



大数据实训室解决方案

构建可信大数据环境和规范，构建行业人才培养标准。



集成电路平台及实训室解决方案

拥有5000颗芯片设计经验的集成电路产教融合。



国产化工业软件解决方案

基于工业生产形态的差异性，融入工业互联网等技术，为工业软件人才培养提供遵循流程工艺的五大应用方向的解决方案。



新工科跨学科应用

推动与其他学科的交叉融合，基于真实的产业应用场景，覆盖智能驾驶、工业、医疗等行业，培养高素质科技人才。

◎ 多学科真实岗位实践案例，满足校级AI通识教学

人工智能通用素养系统化赋能	培训赋能	就业赋能	文科类 数据驱动的分析与内容创作	理科类 复杂问题的建模与算法开发	工科类 智能化设计与工程优化	商科类 智能化决策与精准营销	医科类 AI辅助诊断与精准医疗	艺术类 AI驱动的创作表达与生产	农科类 AI精准生产与智能管理	通用类 AI智能办公与高效协作
		综合项目实践模块 提供项目实践课题及考核标准，供老师学生围绕技能应用开展自主探究实践								
		专项技能训练模块								
		AI+文科 文学小说创作 古诗词智学助手 新媒体内容创作 数据新闻记者 数字人文研究员 新媒体运营	AI+理科 事故理赔预测 心理诊断报告生成 生物图片智能监测 AI数据分析师 科研研究员 量化研究员	AI+工科 环境数据监测 工业宣传片生成 工业铸造质检 算法工程师 智能制造工程师 机器人工程师	AI+商科 营销方案生成 候选人评估 财务顾问数字人 金融分析师 智能营销策划师 商业智能分析师	AI+医科 门诊病历生成 电子病历分析 病症诊断模拟 临床医生 AI药物研发工程师 AI医疗影像分析师	AI+艺术 短剧剧本创作 民谣歌曲创作 风格插画生成 AI视觉设计师 动画设计师 AI音乐制作人	AI+农科 作物病理识别 种子智能分型 牧草产量分析 智慧农机操作员 数字农艺师 农场运营	AI+通用 调研报告撰写 科研文献解读 PPT智能生成 学生 教师 企业职员	
		AI通识认知模块 提供理论课程及作业，包含人工智能基础、关键技术、数据分析、大模型技术、智能体等内容								
算力	 区域级 算力中心	 高校 算力集群	 公有云	AI工具	 大模型 (微调) 训练	 小模型 开发与应用	 智能体 开发与应用	 大语言 模型生成	 多模态 AIGC生成	 低代码 商业BI