视觉与学习青年学者研讨会

VISION AND LEARNING SEMINAR

会议手册

主办单位：中国人工智能学会　中国图象图形学学会
中国民族贸易促进会

承办单位：重庆邮电大学

协办单位：重庆大学微电子与通信工程学院
中国图象图形学学会青年工作委员会

支持单位：重庆市科学技术协会　重庆市青年科技领军人才协会
重庆市科学技术局　重庆市教育委员会

2024年5月5日-7日

中国·重庆

目 录

VALSE 2024 欢迎您

欢迎大家来到重庆，享受VALSE给大家带来的学术盛宴。

VALSE发起于2011年，是Vision And Learning SEminar的简写，取法语“华尔兹舞”之意。旨在为计算机视觉、图像处理、模式识别与机器学习研究领域的华人青年学者提供一个自由、平等、低成本的深度学术交流舞台。在这个舞台上，我们恪守并倡导理性批判、勇于探索、实证、创新等科学精神；在这个舞台上，我们倡导自由平等原则下、理性而纯学术的百家争鸣和思想交锋；在这个舞台上，我们期望欣赏到国内青年学者越来越优美的学术华尔兹（VALSE）。通过这个舞台，我们期望促进国内青年学者的思想交流和学术合作，从而在相关领域做出重量级的学术贡献，提升中国学者在国际学术舞台上的学术和影响力。

围绕上述目标，过去十三年来，VALSE逐渐形成了自己的特色社区文化、找准了自己的使命，包括：

1) 创造深层次学术交流与合作的新模式；

2) 搭建经济实用的在线学术交流舞台；

3) 构筑连接学术界和工业界间的桥梁；

4) 践行国际学术规范，倡导先进科研理念。

第十四届VALSE大会于2024年5月5-7日在重庆举行。届时将延续历年传统，呈上3个大会主旨报告、4个大会特邀报告、12个APR报告、4场Tutorial、19场Workshop、407篇顶会顶刊Poster，合计共有百余位知名青年学者共同带来视觉与学习等人工智能领域的一次学术盛会。此外，在会场合作单位展台区还将展示合作单位精彩的演示。

鉴于VALSE不向参会者收取任何费用，故特别感谢视拓云、OPPO、华为、茶思屋、马上消费、腾讯优图、合合信息、阿里妈妈、美团、百度、极视角科技、联想研究院、趋动科技、融科联创（天津）、并行科技、金山办公、思腾合力、美图、爱诗科技、快手、虹软科技、生数科技、思谋科技、重庆长安、迈尔微视、西云算力等企业对本次会议提供的赞助与大力支持。

VALSE 2024 组委会

大会组委会

大会主席 (General Chairs)

高新波	重庆邮电大学
张清华	重庆邮电大学

程序委员会主席 (Program Chairs)

李伟生	重庆邮电大学
肖　斌	重庆邮电大学
王楠楠	西安电子科技大学
谢凌曦	华为
谭晓衡	重庆大学

Tutorial Chairs

左旺孟	哈尔滨工业大学
李冠彬	中山大学
胡　鹏	四川大学
胡　波	重庆邮电大学

Poster Chairs

徐天阳	江南大学
任文琦	中山大学
杨　斌	重庆邮电大学

展览主席

刘夏雷	南开大学
黎俊伟	重庆邮电大学
李　卓	重庆邮电大学

Website Chair

郑海永	中国海洋大学

Finance Chairs

程明明	南开大学
程　一	中科院计算所
王兴珍	重庆邮电大学

Registration Chairs

郭宗辉	中科院计算所
郭春乐	南开大学
郭　坦	重庆邮电大学
宋铁成	重庆邮电大学

Publicity Chairs

马　超	上海交通大学
李　镇	香港中文大学（深圳）
宫　辰	南京理工大学
王梦迪	重庆邮电大学

Sponsorship Chairs

山世光	中科院计算所
姬艳丽	电子科技大学

Workshop Chairs

姬艳丽	电子科技大学
郭裕兰	国防科技大学
江　波	安徽大学
沈　为	上海交通大学
韩晓光	香港中文大学（深圳）
胡建芳	中山大学
欧阳万里	上海人工智能实验室
舒祥波	南京理工大学
张姗姗	南京理工大学
潘金山	南京理工大学
顾　实	电子科技大学
甘　吉	重庆邮电大学
蒲　晓	重庆邮电大学

APR Chairs

郑　乾	浙江大学
崔兆鹏	浙江大学
王福田	安徽大学
袁　霖	重庆邮电大学

本地主席

刘　波	重庆邮电大学
罗甫林	重庆大学
高陈强	重庆邮电大学
徐宗懿	重庆邮电大学
冷佳旭	重庆邮电大学

VALSE 2024 会场分布

VALSE2024 会议总体日程

5 月 5 日		
时间	**内容**	**地点**
08:15-08:55	**开幕式** **颁发赞助商感谢牌**	两江厅
08:55-09:30	**大会主旨报告-1** **报告人：**沈向洋（香港科技大学） **报告题目：**大模型时代的机遇和挑战	
09:30-10:05	**大会主旨报告-2** **报告人：**胡事民（清华大学） **报告题目：**以深度学习框架为牵引促进自主 AI 生态发展	
10:05-10:40	**大会主旨报告-3** **报告人：**李学龙（中国电信人工智能研究院（TeleAI）） **报告题目：**从洞穴的影子到智能的光辉——连接和交互方式的改变塑造未来生活	
10:40-11:05	**铂金企业宣讲** **视拓云宣讲报告：** 题目：破解用卡难复现难新方案 宣讲人：余佳 **华为宣讲报告：** 题目：视觉领域年度探索实践分享 宣讲人：徐航 **马上消费宣讲报告：** 题目：金融场景下的多模态理解与生成 宣讲人：陆全	
11:05-11:50	2023-2024 年度 CV 与 ML 领域重要学术进展	
	年度进展评述（一）	
	讲者：陈华钧（浙江大学） 题目：符号知识和大模型	
	讲者：代季峰（清华大学） 题目：智能体视觉	

12:00-13:30	**午餐**	
13:30-14:00	**大会特邀报告-1** **报告人：**李鸿升（香港中文大学） **报告题目：**图像生成和视频生成若干前沿技术探索	两江厅
14:00-14:30	大会特邀报告-2 **报告人：**马月昕（上海科技大学） **报告题目：**三维场景理解的前世、今生与未来	
14:30-15:00	大会特邀报告-3 **报告人：**吴小俊（江南大学） **报告题目：**多模态视觉融合方法：是否存在性能极限？	
15:00-15:30	大会特邀报告-4 **报告人：**杨易（浙江大学） **报告题目：**混合模型驱动的内容生成与具身智能	
15:30-18:00	**年度进展评述（二）**	
	讲者：谢凌曦（华为） 题目：视觉通用人工智能	
	讲者：卢志武（中国人民大学） 题目：视频生成	
	讲者：林倞（鹏城实验室，中山大学） 题目：面向具身智能的多模态感知与交互	
	讲者：高林（中科院计算所） 题目：三维高斯泼溅（3D Gaussian Splatting）	
	讲者：施柏鑫（北京大学） 题目：神经形态相机视觉计算	
	讲者：王兴刚（华中科技大学） 题目：面向大模型的新型高效率网络架构年度进展报告	
	讲者：赵恒爽（香港大学） 题目：视觉基础大模型	
	讲者：严骏驰（上海交通大学） 题目：世界模型增强的自动驾驶	
	讲者：俞扬（南京大学） 题目：世界模型与具身决策	
	讲者：杨耀东（北京大学） 题目：从偏好对齐到价值对齐与超对齐	
18:30-20:00	**VIP 晚宴**	一楼喜悦厅

	5月6日	
8:30-12:00	**Workshop 9：视觉大模型高效迁移** **组织者**：左旺孟（哈尔滨工业大学）、戴玉超（西北工业大学）、王立君（大连理工大学） **讲者**：丁贵广（清华大学）、程明明（南开大学）、吴建鑫（南京大学）、张伟（百度）、王鑫龙（北京智源实验室）、高鹏（上海人工智能实验室）、张磊（OPPO）、张璐（大连理工大学）	欣悦厅
	Workshop 12：大模型理论与机理 **组织者**：刘勇（中国人民大学）、李崇轩（中国人民大学）、盛律（北京航空航天大学） **讲者**：袁洋（清华大学）、贺迪（北京大学）、张宁豫（浙江大学）、孙若愚 香港中文大学（深圳）、颜航（上海人工智能实验室）	喜悦厅
	Workshop 15：三维重建与生成 **组织者**：郑伟诗（中山大学）、章国锋（浙江大学）、张青（中山大学） **讲者**：王程（厦门大学）、徐凯（国防科技大学）、廖依伊（浙江大学）、许岚（上海科技大学）、齐晓娟（香港大学）、方杰民（华为）	欢悦厅
	Workshop 5：大模型赋能智慧医疗 **组织者**：陈浩（香港科技大学）、夏勇（西北工业大学）、张少霆（上海人工智能实验室） **讲者**：王本友（香港中文大学深圳）、谢伟迪（上海交通大学）、王国泰（电子科技大学）、史淼晶（同济大学）、张晓凡（上海交通大学）、雷柏英（深圳大学）	温德姆宴会厅
	Workshop 6：视觉智能算法防护与模型安全 **组织者**：彭春蕾（西安电子科技大学）、郭宗辉（中科院计算所）、张健（北京大学）、董胤蓬（清华大学） **讲者**：俞能海（中国科学技术大学）、张新鹏（复旦大学）、李斌（深圳大学）、丁守鸿（腾讯优图）、吴保元（香港中文大学深圳）、王奕森（北京大学）、张杰（中科院计算所）	智悦厅
	Workshop 16：艺术智能(AI for Art) **组织者**：霍静（南京大学），金鑫（北京通用人工智能研究院）、秦越（南京艺术学院） **讲者**：姚鸿勋（哈尔滨工业大学）、张克俊（浙江大学）、董未名（中国科学院自动化研究所）、蔡新元（华中科技大学）、薄一航（北京电影学院）、许多（天津音乐学院）、宋睿华（中国人民大学）	博悦厅

12:00-14:00	Poster	一楼大厅
14:00-17:40	**Workshop 4：具身智能的视觉与学习** **组织者**：盛律（北京航空航天大学）、李镇（香港中文大学（深圳））、张瑞茂（香港中文大学（深圳）） **讲者**：卢策吾（上海交通大学）、王鹤（北京大学）、邓富城（极视角）、刘航欣（北京通用人工智能研究院）、弋力（清华大学）	欣悦厅
	Workshop 3：视觉世界模型与视频生成 **组织者**：唐彦嵩（清华大学），舒祥波（南京理工大学），朱霖潮（浙江大学） **讲者**：张兆翔（中国科学院自动化研究所）、黄思远（北京通用人工智能研究院）、赵洲（浙江大学）、王标（阿里妈妈技术）、罗翀（微软亚洲研究院）、刘希慧（香港大学）、魏晓明（美团）、袁粒（北京大学）、白磊（上海人工智能实验室）	喜悦厅
	Workshop 2：面向移动终端的 AI 图像增强 **组织者**：程明明（南开大学）、王云鹤（华为）、李翔（南开大学） **讲者**：杜成（华为）、左旺孟（哈尔滨工业大学）、潘金山（南京理工大学）、顾舒航（电子科技大学）、付莹（北京理工大学）、郭春乐（南开大学）	欢悦厅
	Tutorial 3：开放词汇视觉感知 讲者：李冠彬（中山大学）	温德姆宴会厅
	Tutorial 4：NeRF 和 3DGS 讲者：彭思达（浙江大学）、韩晓光（香港中文大学（深圳））	智悦厅
	Workshop 8：海洋环境智能探测与理解 **组织者**：丛润民（山东大学）、郑海永（中国海洋大学）、李华（海南大学） **讲者**：付先平（大连海事大学）、李胜全（鹏城实验室）、蒲华燕（重庆大学）、孙哲（西北工业大学）、刘日升（大连理工大学）	博悦厅
5 月 7 日		
8:30-12:00	**Workshop 18：多模态大模型** **组织者**：白翔（华中科技大学）、徐天阳（江南大学）、沈为（上海交通大学） **讲者**：刘禹良（华中科技大学）、王士进（科大讯飞）、王文海（香港中文大学）、丁二锐（百度）、师忠超（联想研究院）、张博（浙江大学）	欣悦厅

	Workshop 1：脑启发的视觉与学习 **组织者**：余肇飞（北京大学）、王鑫（清华大学）、郑乾（浙江大学）、顾实（电子科技大学） **讲者**：李国齐（中科院自动化所）、施柏鑫(北京大学)、邓磊（清华大学）、李远宁（上海科技大学）、王林（香港科技大学）	喜悦厅
	Tutorial 1：具身智能及自主智能体 讲者：邱锡鹏（复旦大学）、王鹤（北京大学）	欢悦厅
	Workshop 17：优秀学生论坛 **组织者**：韩晓光（香港中文大学（深圳））、胡建芳（中山大学）、任冬伟（哈尔滨工业大学）、张鼎文（西北工业大学） **讲者**：李志琦（南京大学）、穆尧（香港大学）、李明晗（香港理工大学）、王立元（清华大学）、古祥（西安交通大学）、杨灵（北京大学）、周尚辰（新加坡南洋理工大学）、夏俊（西湖大学）	温德姆宴会厅
	Workshop 14：女科学家成长论坛 **组织者**：董晶（中国科学院自动化研究所）、刘偲(北京航空航天大学)、姬艳丽（电子科技大学） **嘉宾**：张艳宁（西北工业大学）、马惠敏（北京科技大学）、邬霞 （北京理工大学）、姚鸿勋（哈尔滨工业大学）、金琴（中国人民大学）、宋睿华（中国人民大学）、山世光（中科院计算所）、潘纲（浙江大学）	智悦厅
	Workshop 10：异构联邦学习 **组织者**：叶茫（武汉大学）、汪婧雅（上海科技大学）、卢杨（厦门大学） **讲者**：杨强（香港科技大学）、夏勇（西北工业大学）、王乐业（北京大学）、范晓亮 （厦门大学）、潘微科（深圳大学）	博悦厅
12:00-14:00	**Poster**	一楼大厅
14:00-17:40	**Workshop 13：多模态感知与对话** **组织者**：姬艳丽（电子科技大学），胡迪（中国人民大学），田亚鹏（德克萨斯大学达拉斯分校） **讲者**：郑伟诗（中山大学）、李成龙（安徽大学）、罗平（香港大学）、常扬（上海合合信息科技股份有限公司）、陈宸（OPPO）、许华哲（清华大学）、仉尚航（北京大学）	欣悦厅

	Workshop 11：大模型与因果推理 **组织者**：况琨（浙江大学）、宫明明（墨尔本大学） **讲者**：刘礼（重庆大学）、冯福利（中国科学技术大学）、丁效（哈尔滨工业大学）、俞奎（合肥工业大学）、孙鑫伟（复旦大学）、林勇（香港科技大学）	喜悦厅
	Tutorial 2：视频生成的初探及其可控性研究 讲者：王鑫涛（快手）	欢悦厅
	Workshop 7：遥感图像智能解译 **组织者**：赵文达（大连理工大学）、杨曦（西安电子科技大学）、洪丹枫（中国科学院空天信息创新研究院）、王海鹏（海军航空大学） **讲者**：方乐缘（湖南大学）、程塨（西北工业大学）、曹相湧（西安交通大学）、邓良剑（电子科技大学）、谢卫莹（西安电子科技大学）	温德姆宴会厅
	Workshop 19：端到端自动驾驶 **组织者**：李弘扬（上海人工智能实验室）、张兆翔（中国科学院自动化研究所）、杨言超（香港大学）、马超(上海交通大学) **讲者**：陈镭韬（中国科学院香港创新研究院）、李镇（香港中文大学（深圳））、李国法（重庆大学）、张力（复旦大学）、冷佳旭（重庆邮电大学）、杨泽同（上海人工智能实验室）、金鑫（东方理工）、史少帅(滴滴)	智悦厅

Tutorial 与 Workshop 日程一览

时间地点	报告与讲者信息
Tutorial 3: **开放词汇视觉感知** 5 月 6 日 14:00-17:40 温德姆宴会厅	讲者：李冠彬（中山大学） 题目：开放词汇视觉感知
Tutorial 4: **NeRF 与 3DGS** 5 月 6 日 14:00-17:40 智悦厅	讲者：彭思达（浙江大学） 题目：NeRF 的基础及后续扩展
	讲者：韩晓光（香港中文大学（深圳）） 题目：3DGS，三维重建的终点吗？
Tutorial 1: **具身智能及自主智能体** 5 月 7 日 8:30-12:00 欢悦厅	讲者：王鹤（北京大学） 题目：具身智能的 Sim2Real 泛化途径
	讲者：邱锡鹏（复旦大学） 题目：基于大模型的自主智能体
Tutorial 2: **视频生成的初探及其可控性研究** 5 月 7 日 14:00-17:40 欢悦厅	讲者：王鑫涛（快手） 题目：视频生成的初探及其可控性研究
Workshop 9: 视觉大模型高效迁移 5 月 6 日 8:30-12:00 欣悦厅	组织者：左旺孟（哈尔滨工业大学）、戴玉超（西北工业大学）、王立君（大连理工大学）
	讲者：丁贵广（清华大学） 题目：深度学习模型架构设计及微调技术
	讲者：程明明（南开大学） 题目：高效能个性化图像生成
	讲者：吴建鑫（南京大学） 题目：资源高效的视觉大模型微调

时间地点	报告与讲者信息
	企业宣讲：百度 挖掘无标注数据的潜力：半监督学习技术探索与应用（张伟）
	讲者：王鑫龙（北京智源人工智能研究院） 题目：从视觉到多模态基础模型：探索与实践
	讲者：高鹏（上海人工智能实验室） 题目：构建高效的多模态语言大模型
	Panel 嘉宾： 丁贵广（清华大学）、程明明（南开大学）、吴建鑫（南京大学）、王鑫龙（北京智源实验室）、高鹏（上海人工智能实验室）、左旺孟（哈尔滨工业大学）、戴玉超（西北工业大学）、王立君（大连理工大学）、张磊（OPPO）、张璐（大连理工大学）
Workshop 6：视觉智能算法防护与模型安全 5 月 6 日 8:30-12:00 智悦厅	组织者：彭春蕾（西安电子科技大学）、郭宗辉（中科院计算所）、张健（北京大学）、董胤蓬（清华大学）
	讲者：俞能海（中国科学技术大学） 题目：人工智能生成数据水印技术
	讲者：张新鹏（复旦大学） 题目：从深度模型确权到 AIGC 溯源——数字水印新进展
	讲者：李斌（深圳大学） 题目：多媒体取证新场景与新进展
	企业宣讲：腾讯优图 生物特征识别与安全最新进展（丁守鸿）
	讲者：吴保元（香港中文大学深圳） 题目：后门学习发展现状及最新进展
	讲者：王奕森（北京大学） 题目：大模型下的越狱与防御
	讲者：张杰（中国科学院计算技术研究所） 题目：面向视觉模型的对抗攻击与防御方法

时间地点	报告与讲者信息
Workshop 12：大模型理论与机理 5 月 6 日 8:30-12:00 喜悦厅	组织者：刘勇（中国人民大学）、李崇轩（中国人民大学）、盛律（北京航空航天大学）
	讲者：袁洋（清华大学） 题目：定位即智能
	讲者：贺笛（北京大学） 题目：是否所有大模型都具备思维链涌现能力？
	讲者：张宁豫（浙江大学） 题目：大语言模型知识机理与编辑问题
	讲者：孙若愚（香港中文大学） 题目：Transformer 为什么需要 Adam 训练？
	讲者：颜航（上海人工智能实验室） 题目：迈向 1M 上下文长度的大模型
	Panel 嘉宾： 贺迪（北京大学）、袁洋（清华大学）、张宁豫（浙江大学）、颜航（上海人工智能实验室）、孙若愚（香港中文大学）、刘勇（中国人民大学）、李崇轩（中国人民大学）、盛律（北京航空航天大学）
Workshop 15：三维重建与生成 5 月 6 日 8:30-12:00 欢悦厅	组织者：郑伟诗（中山大学），章国锋（浙江大学），张青（中山大学）
	讲者：王程（厦门大学） 题目：基于激光雷达三维视觉的全球定位
	讲者：徐凯（国防科技大学） 题目：鲁棒可扩展的实时三维重建——相机跟踪优化与地图表示学习
	讲者：廖依伊（浙江大学） 题目：面向自动驾驶的高写实仿真：从重建到生成
	讲者：许岚（上海科技大学） 题目：神经渲染和体积视频的一些思考

时间地点	报告与讲者信息
	讲者：齐晓娟（香港大学） 题目：Embodied Environment Generation via Reconstruction, Decomposition and Beyond
	讲者：方杰明（华为） 题目：基于 Gaussian Splatting 的 3D/4D 内容重建和生成
Workshop 5：大模型赋能智慧医疗 5 月 6 日 8:30-12:00 温德姆宴会厅	组织者：陈浩（香港科技大学）、夏勇（西北工业大学）、张少霆（上海人工智能实验室）
	讲者：王本友（香港中文大学深圳） 题目：医疗大语言模型实践和一点儿思考
	讲者：谢伟迪（上海交通大学） 题目：Towards Foundation Model for Medicine
	讲者：王国泰（电子科技大学） 题目：基于医学图像大模型的弱标注和无标注学习
	讲者：史淼晶（同济大学） 题目：深度大模型的兴起以其在医疗场景的应用
	讲者：张晓凡（上海交通大学） 题目：医疗领域大语言模型的训练及智能体应用
	讲者：雷柏英（深圳大学） 题目：大模型赋能 AD 智能诊断
	Panel 嘉宾： 王本友（香港中文大学深圳）、谢伟迪（上海交通大学）、王国泰（电子科技大学）、史淼晶（同济大学）、张晓凡（上海交通大学）、雷柏英（深圳大学）、陈浩（香港科技大学）、夏勇（西北工业大学）、张少霆（上海人工智能实验室）
	组织者：霍静（南京大学），金鑫（北京通用人工智能研究院），秦越（南京艺术学院）

时间地点	报告与讲者信息
Workshop 16：艺术智能 5 月 6 日 8:30-12:00 博悦厅	讲者：姚鸿勋（哈尔滨工业大学） 题目：人工智能的舞蹈创作与情感融入
	讲者：张克俊（浙江大学） 题目：智能艺术与创新设计———智能篆刻
	讲者：董未名（中国科学院自动化研究所）， 题目：绘画中的 AI
	讲者：蔡新元（华中科技大学） 题目：人工智能艺术的观念与路径
	讲者：薄一航（北京电影学院） 题目：计算电影——电影研究的新范式
	讲者：许多（天津音乐学院） 题目：人工智能生成音乐的客观评价方法探析
	Panel 嘉宾： 姚鸿勋（哈尔滨工业大学）、董未名（中国科学院自动化研究所）、蔡新元（华中科技大学）、张克俊（浙江大学）、薄一航（北京电影学院）、许多（天津音乐学院）、宋睿华（中国人民大学）、霍静（南京大学）、金鑫（北京通用人工智能研究院）
Workshop 4：具身智能的视觉与学习 5 月 6 日 14:00-17:40 欣悦厅	组织者：盛律（北京航空航天大学）、李镇（香港中文大学（深圳））、张瑞茂（香港中文大学（深圳））
	讲者：卢策吾（上海交通大学） 题目：具身智能-感知（P），想象（I），执行（E）PIE 方案与具身大模型探索
	讲者：王鹤（北京大学） 题目：通向开放指令操作的具身多模态大模型系统
	企业宣讲：极视角 构建高效 AI：极栈训推一体化平台技术实践（邓富城）
	讲者：刘航欣（北京通用人工智能研究院） 题目：机器人的场景理解与任务运动规划

时间地点	报告与讲者信息
	讲者：弋力（清华大学） 题目：基于人类行为仿真的可泛化人机协作
	Panel 嘉宾： 卢策吾（上海交通大学）、王鹤（北京大学）、刘航欣（北京通用人工智能研究院）、弋力（清华大学）、盛律（北京航空航天大学）、李镇（香港中文大学（深圳））、张瑞茂（香港中文大学（深圳））
Workshop 3：视觉世界模型与视频生成 5 月 6 日 14:00-17:40 喜悦厅	组织者：唐彦嵩（清华大学），舒祥波（南京理工大学），朱霖潮（浙江大学）
	讲者：张兆翔（中国科学院自动化研究所） 题目：基于世界模型的视觉场景生成与应用
	讲者：黄思远（北京通用人工智能研究院） 题目：三维世界模型
	讲者：赵洲（浙江大学） 题目：以人为中心的多模态生成式模型研究
	企业宣讲：阿里妈妈技术 阿里妈妈智能创意最新进展（王标）
	讲者：罗翀（微软亚洲研究院） 题目：探索短视频生成与编辑的前沿技术
	讲者：刘希慧（香港大学） 题目：Towards Controllable and Compositional Visual Content Generation
	企业宣讲：美团 生活服务场景的视觉生成（魏晓明）

时间地点	报告与讲者信息
	讲者：袁粒（北京大学） 题目：Open-Sora Plan 视频生成开源计划 - 进展与不足
	Panel 嘉宾：张兆翔（中国科学院自动化研究所）、黄思远（北京通用人工智能研究院）、赵洲（浙江大学）、罗翀（微软亚洲研究院）、刘希慧（香港大学）、袁粒（北京大学深圳研究生院）、白磊（上海人工智能实验室）
Workshop 2：面向移动终端的 AI 图像增强 5 月 6 日 14:00-17:40 欢悦厅	组织者：程明明（南开大学）、王云鹤（华为）、李翔（南开大学）
	讲者：杜成（华为） 题目：华为手机 Camera 业务难题
	讲者：左旺孟（哈尔滨工业大学） 题目：基于多相机和多曝的 AI 图像增强研究
	讲者：潘金山（南京理工大学） 题目：面向图像视频复原的判别性自注意力机制方法
	讲者：顾舒航（电子科技大学） 题目：底层视觉前沿技术应用实践：AI-ISP
	讲者：付莹（北京理工大学） 题目：基于 RAW 数据的图像增强及应用
	讲者：郭春乐（南开大学） 题目：面向移动终端的极暗场景图像增强技术
	Panel 嘉宾： 程明明（南开大学）、王云鹤（华为）、李翔（南开大学）、杜成（华为）、左旺孟（哈尔滨工业大学）、潘金山（南京理工大学）、顾舒航（电子科技大学）、付莹（北京理工大学）、郭春乐（南开大学）

时间地点	报告与讲者信息
Workshop 8：海洋环境智能探测与理解 5 月 6 日 14:00-17:40 博悦厅	组织者：丛润民（山东大学）、郑海永（中国海洋大学）、李华（海南大学）
	讲者：付先平（大连海事大学） 题目：水下机器人多模式智能全景感知系统及其在海洋牧场中的应用
	讲者：李胜全（鹏城实验室） 题目：水声轨道角动量通信探测技术与开源开放科研平台
	讲者：蒲华燕（重庆大学） 题目：复杂场景智能无人系统感知技术研究进展
	讲者：孙哲（西北工业大学） 题目：水下计算光学成像技术
	讲者：刘日升（大连理工大学） 题目：优化驱动学习的恶劣场景视觉感知
	Panel 嘉宾： 付先平（大连海事大学）、李胜全（鹏城实验室）、蒲华燕（重庆大学）、孙哲（西北工业大学）、刘日升（大连理工大学）
Workshop 18：多模态大模型 5 月 7 日 8:30-12:00 欣悦厅	组织者：白翔（华中科技大学）、徐天阳（江南大学）、沈为（上海交通大学）、刘禹良（华中科技大学）、
	讲者：王士进（科大讯飞） 题目：通用人工智能技术进展和典型应用
	讲者：王文海（香港中文大学） 题目：图文多模态大模型的研究与应用

时间地点	报告与讲者信息
	讲者：丁二锐（百度） 题目：文心·CV 大模型 VIMER 闭环：数据视角
	企业宣讲：联想研究院 联想端侧智能体解决方案（师忠超）
	讲者：张博（浙江大学） 题目：面向实际场景体验的多模态大模型 DeepSeek VL
	讲者：刘禹良（华中科技大学） 题目：通用多模态大模型细节描述能力提升方法
	Panel 嘉宾： 白翔（华中科技大学）、王士进（科大讯飞）、王文海（香港中文大学）、丁二锐（百度）、张博（浙江大学）、沈为（上海交通大学）、徐天阳（江南大学）、
Workshop 1：脑启发的视觉与学习 5 月 7 日 8:30-12:00 喜悦厅	组织者：余肇飞（北京大学）、王鑫（清华大学）、郑乾（浙江大学）、顾实（电子科技大学）
	讲者：李国齐（中国科学院自动化所） 题目：类脑大模型
	讲者：施柏鑫（北京大学） 题目：基于神经形态相机的高速光度属性分析
	讲者：邓磊（清华大学） 题目：脉冲神经动力学网络及硬件实现
	讲者：李远宁（上海科技大学） 题目： Enhancing neural encoding models for naturalistic perception with a multi-level integration of deep neural networks and cortical networks
	讲者：王林（香港科技大学） 题目：大模型浪潮下，基于事件相机的视觉智能研究发展前瞻

时间地点	报告与讲者信息
	Panel 嘉宾：李国齐（中科院自动化所）、施柏鑫(北京大学)、邓磊（清华大学）、李远宁（上海科技大学）、王林（香港科技大学）、顾实（电子科技大学）
Workshop 17：优秀学生论坛 5 月 7 日 8:30-12:00 温德姆宴会厅	组织者：韩晓光（香港中文大学（深圳））、胡建芳（中山大学）、任冬伟（哈尔滨工业大学）、张鼎文（西北工业大学）
	讲者：李志琦（南京大学） 题目：开环端到端自动驾驶的诸多问题
	讲者：穆尧（香港大学） 题目：具身智能大模型与通用机器人系统
	讲者：李明晗（香港理工大学） 题目：从图片到视频分割的通用大模型
	讲者：王立元（清华大学） 题目：生物智能启发的持续学习理论与方法
	讲者：古祥（西安交通大学） 题目：最优传输引导的条件得分扩散模型
	讲者：杨灵（北京大学） 题目： 扩散模型算法与应用
	讲者：周尚辰（新加坡南洋理工大学） 题目：Diffusion Models for Low-Level Vision
	讲者：夏俊（西湖大学） 题目：利用预训练模型解读生物化学密码
	Panel 嘉宾： 胡建芳（中山大学）、任冬伟（哈尔滨工业大学）、张鼎文（西北工业大学）、李志琦（南京大学）、穆尧（香港大学）、李明晗（香港理工大学）、王立元（清华大学）、古祥（西安交通大学）、杨灵（北京大学）、周尚辰（新加坡南洋理工大学）、夏俊（西湖大学）

时间地点	报告与讲者信息
Workshop 14：女科学家成长论坛 5 月 7 日 8:30-12:00 智悦厅	组织者：董晶（中国科学院自动化研究所）、刘偲（北京航空航天大学）、姬艳丽（电子科技大学）
	Panel 1 科研不分性别，科研人有性别
	Panel 2 女性领导力培养
	分组交流：主题研讨
	Panel 嘉宾： 张艳宁（西北工业大学）、马惠敏（北京科技大学）、邬霞（北京理工大学）、姚鸿勋（哈尔滨工业大学）、金琴（中国人民大学）、宋睿华（中国人民大学）、山世光（中科院计算所）、潘纲（浙江大学）
Workshop 10：异构联邦学习 5 月 7 日 8:30-12:00 博悦厅	组织者：叶茫（武汉大学）、汪婧雅（上海科技大学）、卢杨（厦门大学）
	讲者：杨强（香港科技大学） 题目： 联邦学习及大模型的发展趋势及未来发展方向
	讲者：夏勇（西北工业大学） 题目：联邦学习在医学影像分析中的应用
	讲者：王乐业（北京大学） 题目：可信联邦学习评估
	讲者：范晓亮 （厦门大学） 题目：可信联邦学习与行业大模型应用初探
	讲者：潘微科（深圳大学） 题目：跨用户联邦推荐
	Panel 嘉宾： 叶茫（武汉大学）、汪婧雅（上海科技大学）、卢杨（厦门大学）、杨强（香港科技大学）、夏勇（西北工业大学）、王乐业（北京大学）、范晓亮 （厦门大学）、潘微科（深圳大学）

时间地点	报告与讲者信息
Workshop 13：多模态感知与对话 5 月 7 日 14:00-17:40 欣悦厅	组织者：姬艳丽（电子科技大学），胡迪（中国人民大学），田亚鹏（德克萨斯大学达拉斯分校）
	讲者：郑伟诗（中山大学） 题目：多模态行为分析
	讲者：李成龙（安徽大学） 题目：低质多模态融合感知
	讲者：罗平（香港大学） 题目：Efficient Diffusion Transformer for Image and Video Generation
	企业宣讲：上海合合信息科技股份有限公司 多模态产品应用发展与展望（常扬）
	讲者：陈宸（OPPO） 题目：多模态模型端侧轻量化的研究和落地
	讲者：许华哲（清华大学） 题目：大模型赋能具身智能
	讲者：仉尚航（北京大学） 题目：迈向开放世界多模态具身智能感知
Workshop 11：大模型与因果推理 5 月 7 日 14:00-17:40 喜悦厅	组织者：况琨（浙江大学）宫明明（墨尔本大学）
	讲者：刘礼（重庆大学） 题目：面向工业设计的大模型因果推断问题与应用
	讲者：冯福利（中国科学技术大学） 题目：因果大模型赋能推荐初探
	讲者：丁效（哈尔滨工业大学） 题目：大模型因果推理的局限性及自主进化学习
	讲者：俞奎（合肥工业大学） 题目：面向隐私保护数据的联邦因果关系推断方法研究

时间地点	报告与讲者信息
	讲者：孙鑫伟（复旦大学） 题目：Causal Discovery via Conditional Independence Testing with Proxy Variables
	讲者：林勇（香港科技大学） 题目：Spurious Feature Diversification Improves Out-of-distribution Generalization
	Panel 嘉宾： 刘礼（重庆大学）、冯福利（中国科学技术大学）、丁效（哈尔滨工业大学）、俞奎（合肥工业大学）、孙鑫伟（复旦大学）、林勇（香港科技大学）
Workshop 7：遥感图像智能解译 5 月 7 日 14:00-17:40 温德姆宴会厅	组织者：赵文达（大连理工大学）、杨曦（西安电子科技大学）、洪丹枫（中国科学院空天信息创新研究院）、王海鹏（海军航空大学）
	讲者：方乐缘（湖南大学） 题目：开放环境下高光谱图像弱监督解译
	讲者：程塨（西北工业大学） 题目：光学图像小目标检测
	讲者：曹相湧（西安交通大学） 题目：生成式遥感大模型及其应用
	讲者：邓良剑（电子科技大学） 题目：深度嵌入式变分优化方法及其在遥感全色锐化的应用
	讲者：谢卫莹（西安电子科技大学） 题目：多模态遥感图像分布式高效协同解译
	Panel 嘉宾： 方乐缘（湖南大学）、程塨（西北工业大学）、曹相湧（西安交通大学）、邓良剑（电子科技大学）、谢卫莹（西安电子科技大学）、赵文达（大连理工大学）、杨曦（西安电子科技大学）、王海鹏（海军航空大学）

时间地点	报告与讲者信息
Workshop 19：端到端自动驾驶 5 月 7 日 14:00-17:40 智悦厅	组织者：李弘扬(上海人工智能实验室)、张兆翔（中国科学院自动化研究所）、杨言超（香港大学）、马超（上海交通大学）
	讲者：陈韫韬（中国科学院香港创新研究院） 题目：基于视频世界模型的闭环端到端自动驾驶
	讲者：李镇（香港中文大学(深圳)） 题目：3D 车道线检测及多模态点云增强解析
	讲者：李国法（重庆大学） 题目：全工况自动驾驶闭环工具链系统研究
	讲者：张力（复旦大学） 题目：大规模自动驾驶仿真系统研究
	讲者：冷佳旭（重庆邮电大学） 题目：面向自动驾驶的脑启发可信场景分析
	Panel 嘉宾： 陈韫韬（中国科学院香港创新研究院）、李镇（香港中文大学(深圳)）、李国法（重庆大学）、张力（复旦大学）、冷佳旭（重庆邮电大学）、杨泽同（上海人工智能实验室）、金鑫（东方理工）、史少帅（滴滴）、李弘扬 (上海人工智能实验室)、张兆翔（中国科学院自动化研究所）、杨言超（香港大学）

大会报告及讲者简介

沈向洋 香港科技大学

报告题目：大模型时代的机遇和挑战

讲者简介：沈向洋博士，美国国家工程院外籍院士，现任香港科技大学校董会主席、粤港澳大湾区数字经济研究院创院理事长、清华大学高等研究院双聘教授。

沈博士曾任微软公司执行副总裁，负责推动公司中长期总体技术战略及前瞻性研究与开发工作，主管微软全球研究院和人工智能产品线，覆盖人工智能基础设施、服务、应用以及智能助理业务。

他参与创立微软亚洲研究院，并担任院长兼首席科学家，为中国和世界培养了众多一流的计算机科学家、技术专家和企业家。

胡事民 清华大学

报告题目：以深度学习框架为牵引促进自主 AI 生态发展

讲者简介：胡事民，清华大学计算机系教授，可视媒体研究中心主任。 2002 年获得国家杰出青年基金资助，2006 年-2015 年担任国家重大基础研究（973）计划项目 首席科学家，2007 年入选教育部长江学者特聘教授，现为国家自然科学基金委创新群体项目学术带头人。 主要从事计算机图形学、虚拟现实、智能信息处理和系统软件等方面的教学与研究工作， 已在 ACM/IEEE Transactions 和 CVPR 等重要国际刊物和会议上发表论文 100 余篇。 曾担任 PG、SGP、CVM、VR、EG、SIGGRAPH ASIA 等多个国际重要会议的程序委员会主席和委员，曾经和现任 IEEE、Elsevier、Springer 等多个期刊的主编、副主编和编委。

李学龙 中国电信人工智能研究院（TeleAI）

报告题目：从洞穴的影子到智能的光辉——连接和交互方式的改变塑造未来生活

讲者简介：李学龙，中国电信 CTO、首席科学家，中国电信人工智能研究院（TeleAI）院长。从事极致信号处理研究，关注数据基础理论和临地安防技术体系，推动智能应用和生态建设。担任多项型号任务总师，服务于国家重大战略。

李鸿升 香港中文大学

报告题目：图像生成和视频生成若干前沿技术探索

报告摘要：图像生成和视频生成是当前人工智能领域的热点研究课题，其有着广泛的应用价值，但是也面临着生成模型评估困难，生成图像文本依从性不高，数据监督效力弱，生成图像/视频一致性差等若干重要挑战，本报告介绍报告人团队针对这些问题开展的一系列前沿技术探索，并对未来发展方向进行探讨和展望。

讲者简介：李鸿升，现任香港中文大学多媒体实验室副教授，上海人工智能实验室顾问科学家，上海交通大学、中国科学技术大学兼职博士生导师。2006 年获华东理工大学自动化学士学位，2012 年获得美国理海大学计算机科学博士学位。他在计算机视觉、机器学习、医学图像处理顶级期刊和会议上（TPAMI、CVPR、ICCV、ECCV、NeurIPS、ICLR、MICCAI、IPMI 等）发表论文 130 余篇，谷歌学术引用超 4 万次，获得了 2020 年 IEEE 电路与系统协会杰出青年作者奖、2021 年香港中文大学青年学者杰出研究成就奖等奖项。2016 年带领团队参加 ImageNet 2016 国际挑战赛，赢得了视频物体检测项目第一名。他担任国际顶级学术会议 NeurIPS 2021-2023、CVPR 2023、ICCV 2023、ICML 2023-2024、ACM MM 2024 领域主席，AAAI 2022 高级程序委员，国际期刊 IEEE Transactions on Circuits and Systems for Video Technology、Neurocomputing 的副编辑。

杨易 浙江大学

报告题目：混合模型驱动的内容生成与具身智能

报告摘要：本报告将首先介绍混合模型协同在人工智能领域的研究背景，讨论如何在垂直领域应用中实现预训练大模型、先验知识及领域专用模型的高效整合。随后，通过具身智能和内容生成两大关键应用场景的深入分析，本报告将详细阐述混合模型协同在解决特定领域问题时，针对可解释性、计算效率和鲁棒性等方面所展现出的显著优势。最后，本报告将展望人工智能研究中发展混合模型协同技术的前景，探讨其在实际应用中的潜力和价值。

讲者简介：杨易，浙江大学求是讲席教授（二级教授）、国家特聘专家。目前担任浙江大学计算机学院副院长、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。主要研究方向为人工智能及其应用。所发论文 Google Scholar 引用 6 万余次，H-index 123，近 6 年连续入选 Clarivate Analytics 全球高被引学者。获教育部全国优秀博士论文（2010）、澳大利亚基金委青年研究职业奖（2013）、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）、IJCAI 最具影响力论文（2021）、ACM MM 唯一最佳论文奖（2023）等多项 AI 领域国际奖项，以及 20 余次国际科研竞赛世界冠军。

吴小俊 江南大学

报告题目：多模态视觉融合方法：是否存在性能极限？

报告摘要：视觉融合是计算机视觉的重要研究方向。本报告以智慧城市为背景，介绍面向智慧城市的多模态视觉融合方法与研究进展。首先对智慧城市和深度学习进行简单回顾；然后介绍多模态视觉融合的主要框架、方法和研究进展。针对目前性能最好的视觉融合算法，探讨一种增强视觉融合性能的普适方法。同时，本报告将介绍视觉融合在图像质量增强、人脸特征点定位、目标检测、跟踪与识别、行为识别以及融合与视觉上下游任务互促等方面的应用研究。

讲者简介：吴小俊，国际模式识别协会会士（IAPR Fellow）、亚太人工智能协会会士（AAIA Fellow）、国际人工智能产业联盟产业研究院院士（AIIA Fellow）、江南大学至善教授、研究生院院长、Josef Kittler 人工智能研究院院长、教育部装发创新团队负责人、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。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、IETE Gowri Memorial Award 、教育部科技进步一等奖、合作者 Josef Kittler 院士获 2015 江苏省国际科技合作奖和 2016 中国政府友谊奖；担任了 2 项国家重点研发计划项目首席科学家、3 项国家自然基金重点项目和多项 GF 项目负责人。现任 IEEE 智慧城市指导委员会委员、多本国际期刊主编或编委、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。

马月昕 上海科技大学

报告题目：三维场景理解的前世、今生与未来

报告摘要：三维场景理解是基于语言、图像、点云等数据完成对三维物理世界的感知与认知，是机器人、智能车进行安全、高效的行为决策的重要前提。随着视觉-语言大模型的飞速发展，三维场景理解如何突破数据、维度、模态、算力等方面的限制，实现垂直领域的通用模型？本报告将介绍该领域的相关进展、产业应用及探索方向。

讲者简介：马月昕，上海科技大学研究员、助理教授、博导。博士毕业于香港大学，曾到法国国家信息与自动化研究所、美国北卡罗莱纳大学教堂山分校、百度研究院进行访问学习。主要研究方向为三维视觉、机器人，尤其是三维场景理解、多模态感知、自动驾驶、人机协作等。共发表相关领域顶会或顶刊论文 60 余篇，包括 Science Robotics、TPAMI、CVPR、ICCV、ECCV、SIGGRAPH、AAAI 等。获得上海市领军人才（海外）、上海市优秀教学成果（高等教育类）一等奖、国家自然科学基金青年基金项目等。曾获 SemanticKITTI、NuScenes、Argoverse 等多个自动驾驶挑战赛冠军和亚军。

年度进展评述（APR）及讲者简介

谢凌曦 华为

报告题目：视觉通用人工智能年度进展报告

报告摘要：通用人工智能（AGI）是人工智能研究的最高目标。随着 AGI 的火花在自然语言领域被点燃，视觉领域的研究者们也开始论证在更广阔的视觉模态中实现 AGI 的可能性。本报告从 AGI 的定义入手，分析 AGI 研究在自然语言领域率先取得突破的本质原因；随后将上述结论拓展至视觉领域，整理出实现视觉 AGI 的必要条件和主要困难；最后面向这些困难，回顾过去一年的主要研究进展，并展望未来的发展方向。

讲者简介：谢凌曦，华为公司的高级研究员。分别于 2010 年和 2015 年于清华大学获得本科和博士学位，并且于 2015 年至 2019 年期间在美国加州大学洛杉矶分校和约翰霍普金斯大学担任博士后研究员。研究兴趣覆盖计算机视觉的各个方向，主要包括统计学习方法和深度学习模型与基础视觉任务的结合，并积极推动自动机器学习算法和视觉基础模型在上述领域的应用。在国际顶级的学术会议和期刊上发表超过 100 篇论文，谷歌学术引用超过 1.5 万次。

代季峰 清华大学

报告题目：智能体视觉年度进展报告

报告摘要：在过去的几十年里，计算机视觉作为人工智能领域的一个重要分支，已经取得了显著的进展。它使计算机能够从图像或视频中识别、处理并理解视觉信息。然而，随着技术的不断进步，特别是在通用人工智能（AGI）领域的研究加深，我们逐渐认识到，仅仅使计算机“看到”还不够；我们需要的是使计算机能够理解并对视觉信息做出智能响应。这就引出了“智能体视觉”的概念，即在计算机视觉与通用智能体技术结合的基础上，发展出的一种能使机器不仅仅识别图像，而且能够理解图像内容并在此基础上做出决策和行动的能力。本报告旨在探讨计算机视觉与通用智能体技术的结合，以及这种结合如何推动智能系统的发展。结合计算机视觉和通用智能体技术，智能体视觉不仅关注图像的表面特征，更重要的是，它能够理解这些视觉信息背后的上下文和含义。从自动驾驶汽车到高级机器人助理，智能体视觉使机器能够在复杂、动态的环境中做出快速而准确的决策。通过对计算机视觉与通用智能体技术结合的前沿介绍，本报告旨在为读者提供一个全面的视角，理解智能体视觉如何塑造我们与机器交互的未来。

讲者简介：代季峰，清华大学电子工程系副教授，博士生导师，上海人工智能实验室双聘领军科学家。主要研究领域为多模态基础模型和视觉基础模型。在 2009 年和 2014 年于清华大学自动化系分别获得工学学士和博士学位，博士导师周杰教授。2014 年至 2019 年在微软亚洲研究院视觉组工作，担任首席研究员、研究经理。2019 年至 2022 年在商汤科技

研究院工作，担任执行研究总监，二级部门长。2022 年 7 月全职加入清华大学电子工程系。他在相关领域发表国际期刊、会议文章 50 余篇，论文总引用 3 万余次。以可变形卷积为代表的多篇论文成为物体识别领域里程碑式的成果，被选入深度学习权威框架 PyTorch 成为标准算子。他连续两年获得物体识别领域权威的 COCO 比赛冠军，之后历届冠军系统也使用了他提出的算法。他提出的算法获得自动驾驶感知领域权威的 Waymo 2022 竞赛冠军，获得 CVPR 2023 最佳论文奖。他是视觉领域顶刊 IJCV 的编委，和视觉领域顶会 NeurIPS , ICCV，CVPR，ECCV, ICLR 的领域主席，ICCV 2019 的宣传主席。代季峰，清华大学电子工程系副教授，博士生导师，上海人工智能实验室双聘领军科学家。主要研究领域为多模态基础模型和视觉基础模型。在 2009 年和 2014 年于清华大学自动化系分别获得工学学士和博士学位，博士导师周杰教授。2014 年至 2019 年在微软亚洲研究院视觉组工作，担任首席研究员、研究经理。2019 年至 2022 年在商汤科技研究院工作，担任执行研究总监，二级部门长。2022 年 7 月全职加入清华大学电子工程系。他在相关领域发表国际期刊、会议文章 50 余篇，论文总引用 3 万余次。以可变形卷积为代表的多篇论文成为物体识别领域里程碑式的成果，被选入深度学习权威框架 PyTorch 成为标准算子。他连续两年获得物体识别领域权威的 COCO 比赛冠军，之后历届冠军系统也使用了他提出的算法。他提出的算法获得自动驾驶感知领域权威的 Waymo 2022 竞赛冠军，获得 CVPR 2023 最佳论文奖。他是视觉领域顶刊 IJCV 的编委，和视觉领域顶会 NeurIPS , ICCV，CVPR，ECCV, ICLR 的领域主席，ICCV 2019 的宣传主席。

陈华钧 浙江大学

报告题目：符号知识和大模型年度进展报告

报告摘要：大型语言模型和知识图谱都是表示和处理知识的手段。大模型重规模，知识覆盖广泛，语言理解能力强，但训练代价大，可靠可控问题突出。知识图谱重表示，知识正确可靠，逻辑推理能力强，但规模扩展难，不容易泛化迁移。本报告首先分析了大模型与传统符号知识表示的优缺点，然后从知识增强预训|练、知识增强提示指令、知识检索增强、大模型知识编辑等多方面探讨了符号知识对于增强大模型能力的最新研究进展，最后从“规模+表示”相融合的视角展望了符号知识与大模型的未来融合发展趋势。

讲者简介：陈华钧浙江大学计算机科学与技术学院教授，博士生导师， 主要从事知识图谱、大语言模型、AI for Science 等领域的研究工作，在 Nature Machine Intelligence、Nature Communications、NeurIPS、ICML、ICLR、IJCAI、AAAI、ACL、EMNLP、KDD、Proc. IEEE 等国际人工智能会议或学术期刊上发表多篇论文。牵头发起中文开放知识图谱 OpenKG，主持构建阿里巴巴藏经阁知识引擎。曾获国际语义网会议 ISWC 最佳论文奖、国际知识图谱联合会议 IJCKG 最佳论文奖、浙江省科技进步二等奖、教育部技术发明一等奖、中国中文信息学会钱伟长科技奖一等奖、阿里巴巴优秀学术合作奖、华为优秀技术合作项目奖、中国工信传媒出版集团优秀出版物一等奖等。担任浙江省数智科技研究会副会长、中国人工智能学会知识工程专委会副主任、中国中文信息学会语言与知识计算专委会副主任、新华社媒体融合国家重点实验室学术委员会委员、Elsevier Big Data Research 主编等学术服

务工作。全球前 2%顶尖科学家榜单（人工智能领域），浙江省有突出贡献中青年专家，浙江省高层次人才特殊支持计划科技创新领军人才。

卢志武 中国人民大学

报告题目：视频生成年度进展报告

报告摘要：不同于图像生成，视频生成在内容一致性、长视频生成、计算资源消耗等方面均面临巨大的挑战。但是，视频生成仍然在 2023 年取得了飞速的发展，涌现出 Stable Video Diffusion、Runway Gen-2、Video Diffusion Transformer、Sora 等优秀模型。本报告首先介绍当前视频生成面临的挑战，然后详细介绍最新的视频生成优秀模型，最后还对视频生成的技术发展进行展望。

讲者简介：卢志武，博士，中国人民大学高瓴人工智能学院教授，博士生导师。2005 年毕业于北京大学数学科学学院信息科学系，获理学硕士学位；2011 年毕业于香港城市大学计算机系，获 PhD 学位。研究方向为机器学习与计算机视觉。设计首个中文通用多模态预训练模型文澜 BriVL。发表多模态领域首篇 Nature 子刊论文。早于 OpenAI 发布类 Sora 的视频生成底座 VDT。

林倞 鹏城实验室，中山大学

报告题目：面向具身智能的多模态感知与交互年度进展报告

报告摘要：具身智能是目前智能科学前沿的方向，被认为是实现通用人工智能的必经之路。它是一种基于物理身体进行感知和行动的智能系统，通过与环境的持续交互来获取信息、理解问题、做出决策并实现行动，进而具备高效解决复杂任务的能力。作为具身智能的重要组成部分，多模态感知与交互受到了研究人员的广泛关注和探索，通过整合视觉、听觉、触觉等多模态信息，促使具身智能体能够精准感知环境状态以及对物理世界进行高效交互。本次报告将概述近年来具身智能多模态感知与交互领域的主要进展，并对未来研究方向做出展望。

讲者简介：林倞，鹏城实验室多智能体与具身智能研究所所长，中山大学二级教授，国家杰出青年科学基金获得者，IEEE/IAPR Fellow，曾任商汤科技首席研发总监/研究院执行院长。长期从事多模态人工智能、大规模机器学习等领域的应用基础研究，作为首席科学家/项目负责人，承担国家 2030 科技创新重大项目，入选国家万人计划。在国际顶级学术期刊和会议发表论文 300 余篇，论文被引用 3 万余次(谷歌学术统计)，多次获得最佳学术论文奖；指导博士生获得 CCF 优秀博士论文奖、ACM China 优秀博士论文奖及 CAAI 优秀博士论文奖；带领团队获得吴文俊人工智能自然科学奖、中国图象图形学会科学技术一等奖、省级自然科学一等奖等。

高林 中科院计算所

报告题目：三维高斯泼溅（3D Gaussian Splatting）年度进展报告

报告摘要：高斯泼溅极大地加速了三维场景新视角合成的渲染速度。与神经辐射场隐式表示的三维场景表示不同，高斯泼溅利用一组高斯椭球对场景进行建模，通过将高斯椭球光栅化来实现高效渲染。除了快速的渲染速度外，高斯泼溅的显式几何表示有利于动态重建、几何编辑和物理仿真等任务。考虑到该领域研究的快速增长和不断变化，本报告从三维重建、三维编辑和其他下游应用三个方面对近期高斯泼溅工作进行介绍。

讲者简介：高林，中科院计算所研究员，博士生导师，国科大岗位教授，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 80 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 GDC 2024 大会联合程序主席，SGP 2023 大会联合主席，China3DV2023 程序委员会联合主席，SIGGRAPH 2023-2024 技术论文程序委员会委员，NeurIPS 2024 领域主席，IEEE TVCG 编委，CSIG 智能图形专委秘书长，入选国家自然科学基金委优青，北京市杰青，英国皇家学会牛顿高级学者，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，CCF CAD&CG 开源软件奖等奖励。

施柏鑫 北京大学

报告题目：神经形态相机视觉计算年度进展报告

报告摘要：脉冲相机（spike camera）、事件相机（event camera）等神经形态相机（neuromorphic camera），对比逐帧成像的普通相机，拥有独特的优势，尤其是其对于高速运动物体和高动态范围场景的感知能力，在日常拍照和机器视觉应用中发挥了重要的作用。本次报告将对 2023 年以来两类主流的神经形态相机在应对诸多视觉计算任务中的进展进行回顾，涵盖影像重构、辅助拍照、智能感知、下游应用等多个方面，并对未来工作进行展望。

讲者简介：施柏鑫，北京大学计算机学院视频与视觉技术研究所副所长，视频与视觉技术国家工程研究中心、多媒体信息处理全国重点实验室研究员、长聘副教授、博士生导师（“博雅青年学者”）；北京智源人工智能研究院青年科学家。2013 年博士毕业于日本东京大学，曾先后在麻省理工学院媒体实验室、新加坡科技设计大学、南洋理工大学、日本国立产业技术综合研究所从事研究工作。研究方向为计算摄像学与计算机视觉，发表论文 200 余篇（包括 TPAMI 论文 23 篇，计算机视觉三大顶级会议论文 82 篇）。论文获评国际计算摄像会议（ICCP）2015 最佳论文提名（Best Paper - Runner Up）、国际计算机视觉会议（ICCV）2015 最佳论文候选（入选 IJCV 特刊 Best Papers from ICCV 2015），2021 年获得日本大川研究助成奖。科技创新 2030—“新一代人工智能”重大项目首席科学家，国家自然科学基金重点项目负责人，国家级青年人才计划入选者。担任国际顶级期刊 TPAMI、IJCV 编委，

顶级会议 CVPR、ICCV、ECCV 领域主席。IEEE、CCF、CSIG 高级会员，APSIPA 杰出讲者。更多信息请访问“相机智能”实验室主页：http://camera.pku.edu.cn.

王兴刚 华中科技大学

报告题目：面向大模型的新型高效率网络架构年度进展报告

报告摘要：得益于 Transformer 网络的强大建模能力和高度可并行的特性，基于 Transformer 的多模态大模型在人工智能领域取得了巨大的成功，然而 Transformer 的时空复杂度会随着输入序列的长度成二次方的增长，严重限制了 Transformer 在精确建模长序列方面的应用。针对这些问题，本次报告中将介绍近年来新兴的一些高效率网络架构（如 RWKV、RetNet、Mamba 等）及其在大语言模型（LLM）、多模态 LLM、计算机视觉、语音处理等领域的应用，总结和分析面向大模型的新型高效率网络架构领域现阶段取得的成就和面临的挑战。

讲者简介：王兴刚，华中科技大学电信学院教授博导，国家“万人计划”青年拔尖人才，现任 Image and Vision Computing 期刊（Elsevier, IF 4.7）共同主编。主要从事基础模型、视觉表征学习、目标检测分割跟踪等领域研究、在 IEEE TPAMI、IJCV、CVPR、ICCV、NeurIPS 等顶级期刊会议发表学术论文 60 余篇，谷歌学术引用 2.6 万余次，其中一作/通讯 1000+引用论文 5 篇，入选 Elsevier 2023 中国高被引学者。担任 CVPR, ICCV, ICIG 等会议领域主席，Pattern Recognition、Machine Vision and Application 等期刊编委。入选了中国科协青年人才托举工程，获湖北青年五四奖章、CSIG 青年科学家奖，吴文俊人工智能优秀青年奖，CVMJ 2021 最佳论文奖，湖北省自然科学二等奖等，指导学生获 2022 年全国“互联网+”大赛金奖、2023 年挑战杯“揭榜挂帅”专项赛全国一等奖。

赵恒爽 香港大学

报告题目：视觉基础大模型年度进展报告

报告摘要：随着深度学习模型能力的增强和海量数据的高效获取利用，视觉基础大模型的构建获得广泛关注并得以实现。与传统单任务单领域的模型相比，视觉基础大模型展现出强大的通用泛化能力并能同时处理复杂视觉场景多任务跨领域的理解，这对于众多下游应用如自动驾驶、智能机器人、智慧医疗、智能制造、社会公益等方面具有广泛的应用价值。本年度回顾将介绍近期视觉基础大模型的主要进展，包括框架设计、表征学习、微调机制等相关技术，并对未来工作进行展望。

讲者简介：赵恒爽博士是香港大学计算机科学系助理教授。此前，他曾在麻省理工学院和牛津大学担任博士后研究员，博士毕业于香港中文大学。他的研究兴趣涵盖计算机视觉、机器学习和人工智能等广泛领域，特别着重于构建智能视觉系统。他和他的团队获得过多次国际学术竞赛的冠军，包括 ImageNet 场景解析竞赛等。他获得了世界人工智能大会的新星奖，并被 AI 2000 评为计算机视觉领域最具影响力的学

者之一。他曾担任 CVPR、ECCV、NeurIPS 和 AAAI 的领域主席，研究工作被引用超 26,000 次。

严骏驰 上海交通大学

报告题目：世界模型增强的自动驾驶年度进展报告

报告摘要：OpenAI 的视频生成模型 Sora 的爆火，引发了对世界模型的热议。自动驾驶，作为一个综合了感知、预测、决策的复杂任务，对于样本的数量级与多样性均有着极高的要求，而且真实世界自动驾驶数据的采集存在安全性与成本问题，二者都成为了自动驾驶纯数据驱动范式的巨大障碍。因而，如何为自动驾驶任务构建世界模型，以及如何使用世界模型的推演能力来提升自动驾驶任务性能，降低对真实世界数据的需求，成为了业内热议的话题。本报告主要介绍最新的关于自动驾驶的世界模型的研究进展，并分析未来发展趋势。

讲者简介：严骏驰，上海交通大学计算机系教授、支部书记、CCF 优博及导师/杰出会员/演讲者、英国工程技术学会会士（IET Fellow）。国家自然基金委优青、教育部 AI 资源建设深度学习首席专家。科技部 2030 新一代人工智能重大项目、基金委下一代人工智能重大研究计划项目负责人。曾就职于 IBM 研究院 7 年，任首席研究员。研究方向为机器学习及其交叉应用。发表 Nature 子刊、CCFA 类第一/通讯作者论文过 200 篇，引用逾万次（爱思唯尔高被引学者），获陕西省自然科学一等奖、AAAI21/IJCAI23 最具影响力论文等荣誉。任机器学习三大会议（ICML/ICLR/NeurIPS)、及 SIGKDD/CVPR/AAAI/IJCAI 等顶级会议（高级）领域主席、Pattern Recognition、ACM TOPML 等期刊（创刊）编委。指导学生获 CCF 优博/CV 新锐奖、挑战杯特等奖、华为量子计算冠军、交大学术之星、丘成桐杯金奖，及自然基金委本科生青年基金项目。毕业学生多人在 MIT 等机构从事博后研究，以及头部高校教职。

俞扬 南京大学

报告题目：世界模型与具身决策年度进展报告

报告摘要：世界模型的研究旨在让智能体构建关于外部环境的内部模型，用于支撑推理、规划和想象，因此也被认为是具身决策系统的关键组成。近期世界模型受到高度关注，在机器人控制、自动驾驶等领域取得了进展。报告将介绍近一年此方向的研究进展，探讨世界模型在具身决策中的应用。

讲者简介： 俞扬，南京大学人工智能学院教授。主要从事强化学习的研究工作，工作获 5 项国际论文奖励和 3 项国际算法竞赛冠军。入选国家青年人才计划、IEEE“国际人工智能十大新星”，获 CCF-IEEE 青年科学家奖，首届亚太数据挖掘“青年成就奖”，并受邀在国际人工智能联合大会 IJCAI 2018 上作“青年亮点报告”

杨耀东 北京大学

报告题目：从偏好对齐到价值对齐与超对齐年度进展报告

报告摘要：从偏好对齐到价值对齐与超对齐。大语言模型的训练离不开基于人类反馈的强化学习技术。面向下一代对齐方法，尤其是从偏好对齐到价值对齐，乃至面向超级智能体的对齐，仍有许多挑战。本报告将会介绍价值对齐中的难点，价值系统的刻画及挑战，以及从内外对齐算法设计中的思考，同时也会介绍在安全对齐中的一些可行方法。最后，本报告也会提出面向超对齐研究的一些思考。

讲者简介：杨耀东，博士，北京大学人工智能研究院研究员（博导）、AI 安全与治理中心执行主任。国家高层次留学人才计划、国家高层次青年人才项目、中国科协青年托举计划、北大博雅青年学者获得者。重点研究通用多智能体系统构建、博弈交互与价值对齐等问题，科研领域包括强化学习、博弈论和多智能体系统。本科毕业于中国科学技术大学，随后在伦敦帝国理工大学、伦敦大学学院获得硕士及博士学位（论文获学校唯一提名 ACM SIGAI 优博奖）。曾于伦敦国王大学信息学院任助理教授。发表 AI 领域顶会顶刊论文一百余篇，谷歌引用四千余次，主持国自然、科技部、市科委、校企实验室等项目经费超三千万元。曾获国际计算机视觉会议 ICCV’23 最佳论文奖入围（Best Paper Initial List）、机器人学习会议 CoRL’20 最佳系统论文奖（Best System Paper）、多智能体系统会议 AAMAS’21 最具前瞻性论文奖（Best Blue-Sky Paper）、世界人工智能大会（WAIC’22）云帆奖璀璨明星、ACM SIGAI China 新星奖。

Tutorial 报告及讲者简介

时间	内容	地点
5 月 6 日 14:00-17:40	**Tutorial 3：开放词汇视觉感知** 讲者：李冠彬（中山大学） 题目：开放词汇视觉感知	温德姆宴会厅
5 月 6 日 14:00-17:40	**Tutorial 4: NeRF 与 3DGS** 讲者：彭思达（浙江大学） 题目：NeRF 的基础及后续扩展 讲者：韩晓光（香港中文大学（深圳）） 题目：3DGS，三维重建的终点吗？	智悦厅
5 月 7 日 8:30-12:00	**Tutorial 1：具身智能及自主智能体** 讲者：邱锡鹏（复旦大学） 题目：基于大模型的自主智能体 讲者：王鹤（北京大学） 题目：具身智能的 Sim2Real 泛化途径	欢悦厅
5 月 7 日 14:00-17:40	**Tutorial 2：视频生成的初探及其可控性研究** 讲者：王鑫涛（快手） 题目：视频生成的初探及其可控性研究	欢悦厅

Tutorial 1: 具身智能及自主智能体

邱锡鹏 复旦大学

报告题目：基于大模型的自主智能体

报告摘要：长期以来，人类致力于寻找能够达到人类水平的智能体，而智能体则被认为是探寻道路上的一个非常有前景的媒介。智能体可以感知环境，做出决策，并采取行动。从上世纪以来，科学家们在开发智能体上进行了很多努力，这些努力大多集中在对于算法或训练策略上的改进，从而提升其在特定任务上能力的提升。而事实上，社区缺乏一种通用的且强大的模型来作为设计通用智能体的起点。大语言模型已经展示出了非常强大的通用能力，被认为是通往通用人工智能（AGI）的火花。因此，很多研究者开始基于大语言模型构建智能体，并取得了极为显著的进展。在这次报告汇总，我们将对基于大语言模型的智能体进行详细的介绍。我们首先从哲学起源开始，追溯了智能体的发展脉络，阐明其定义，并解释了为何大语言模型适合作为智能体的基座。基于这些分析，我们提出了一个基于大语言模型的通用智能体概念框架，包括三个主要部分：控制器、感知器和行为器。这个框架可以通过裁剪来适应下游应用。特别地，我们从三个方面探索了基于 LLM 的 agent 的应用场景：单智能体场景，多吃能提场景和人机协作场景。之后，我们介绍了智能体社会，探索了基于大语言模型的智能体的社会行为与人格，还结合搜啊了智能体社会中涌现出来的社会现象以及他们可以给予人类社会的启示。最后我们讨论了该领域内的几个重要话题和开放性问题。

讲者简介：邱锡鹏，复旦大学计算机学院教授，担任中国中文信息学会大模型大搜索与生成专委会副主任、上海市计算机学会自然语言处理专委会主任等，主要研究方向为自然语言处理模型和算法，发表 CCF A/B 类论文 100 余篇，被引用 1.9 万余次，连续入选爱思唯尔中国高被引学者、2023 全球前 2%顶尖科学家榜单。获得 ACL 2017 杰出论文奖（CCF A 类）、CCL 2019/2023 最佳论文奖、《中国科学：技术科学》2021 年度高影响力论文奖，有 5 篇论文入选 ACL/EMNLP 等会议的最有影响力论文，主持研发的 MOSS 已经成为国内影响力最大的开源大型语言模型之一。曾获中国科协青年人才托举工程、国家优青等项目，2020 年获第四届上海高校青年教师教学竞赛优等奖，两次获得上海市计算机学会教学成果奖一等奖，2022 年获钱伟长中文信息处理科学技术奖一等奖，2023 年入选教育部“高校计算机专业优秀教师奖励计划”。培养学生多人次获得国家一级学会优博。

王鹤 北京大学

报告题目：具身智能的 Sim2Real 泛化途径

报告摘要：具身智能（Embodied AI）是当下计算机视觉、机器人学等的前沿交叉领域，致力于发展可以与真实世界进行泛化物理交互的机器人通用智能体。具身智能当前发展的瓶颈在于缺少海量场景中具身数据，因此无法将一系列操作技能泛化到任意新场景。本报告将围绕 Sim2Real 的途径，即通过高逼真渲染和物理仿真生成具身智能需要的海量数据，并且直接迁移到真实世界、并保持极高的性能。报告将介绍我们的系列工作，通过 Sim2Real 达成了任意材质泛化物体抓取、跨物体类别泛化零件操作、机器人移动导航和操作、灵巧手泛化操作等，重点探讨 Sim2Real 的条件和方法。

讲者简介：王鹤博士，北京大学计算机学院前沿计算研究中心（CFCS）助理教授，博士生导师。他创立并领导了北大具身感知与交互实验室(EPIC Lab，主页：https://hughw19.github.io)，研究目标是通过发展具身技能及具身多模态大模型推进通用具身智能。他同时担任北大-银河通用具身智能联合实验室主任和北京智源人工智能研究院具身智能研究中心主任。他已在计算机视觉、机器人学和人工智能的顶级会议和期刊（CVPR/ICCV/ECCV/TRO/RAL/ICRA/NeurIPS/ICLR/AAAI 等）上发表五十余篇工作，其论文获得 ICCV2023 最佳论文候选，ICRA2023 最佳操纵论文候选，2022 年世界人工智能大会青年优秀论文（WAICYOP）奖，Eurographics 2019 最佳论文提名奖。他担任了 CVPR2022 和 WACV2022 的领域主席，Image and Vision Computing 的副主编和诸多顶会的审稿人、程序委员。在加入北京大学之前，他于 2021 年从斯坦福大学获得博士学位，师从美国三院院士 Leonidas. J Guibas 教授，于 2014 年从清华大学获得学士学位。

Tutorial 2：视频生成的初探及其可控性研究

王鑫涛 快手

报告题目：视频生成的初探及其可控性研究

报告摘要：视频生成正日益受到学术界和工业界的关注。该报告将介绍视频生成的相关背景和最新进展，以及我们在开源视频基础模型 VideoCrafter 系列工作的初步研究，包括文生视频、图生视频以及视频生成的标准化评测等。在视频生成的应用过程中，可控性是一个重要的方面。本报告也将介绍视频可控性方面的进展，并分享我们对视频运动可控性的研究。最近，随着 OpenAI Sora 的发布，我们更加清晰看到了视频生成的潜力和挑战，报告将对 Sora 所采用的技术进行简要讨论，并分享我们在其中的简单探索(Mira)，以及 Sora 对我们研究工作带来的思考。

讲者简介：王鑫涛，快手专家研究员，本科毕业于浙江大学，博士毕业于香港中文大学 MMLab，曾任腾讯 ARC Lab 和 AI Lab 专家研究员。主要研究视觉生成，包括图像、视频和 3D 的生成与编辑。在国际顶级会议期刊发表多篇文章，包括 T2I-Adapter, PhotoMaker, ESRGAN, GFPGAN 等，论文 Google Scholar 引用 12600 余次，被评选为世界前 2%顶尖科学家。

Tutorial 3: 开放词汇视觉感知

李冠彬 中山大学

报告题目：开放词汇视觉感知

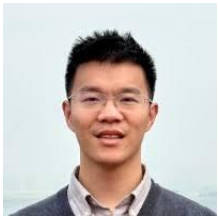
报告摘要：近年来，以 CLIP 为代表的大规模视觉语言预训练模型在零样本识别、开放场景视觉识别以及视觉语言多模态匹配任务中取得了显著的性能突破，并表现出了极高的潜能和优势。视觉感知与理解的研究技术关注点也逐渐由针对封闭类别词汇的单一任务感知模型设计及训练算法优化拓展到以大模型提示微调和适配器设计为代表的大规模开放词汇感知算法设计。在感知模型的泛化性研究上，也逐渐由跨数据分布泛化鲁棒性推广到跨任务泛化性的研究。在本次课程讲座中，我们将首先介绍开放词汇视觉感知的定义及常见的技术实现范式，分别介绍针对图像分类、物体检测、语义分割、多模态内容理解等各层次任务的开放词汇视觉感知任务的经典工作及研究方法，最后介绍单模态与多模态大模型驱动的开放视觉感知的发展现状及技术趋势。希望通过本次课程讲座能和大家共同探讨开放词汇视觉感知前沿研究课题和技术创新思路。

讲者简介：李冠彬，中山大学计算机学院教授，博士生导师，国家优秀青年基金获得者。主要研究领域为跨模态视觉感知、理解与生成。 迄今为止累计发表 CCF A 类/中科院一区论文 150 余篇，Google Scholar 引用超过 11000 次。曾获得吴文俊人工智能优秀青年奖、ACM 中国新星提名奖、中国图象图形学学会科学技术一等奖、ICCV2019 最佳论文提名奖、ICMR2021 最佳海报论文奖等荣誉。主持了包括国自然优青、面上、青年、广东省杰青、CCF-腾讯犀牛鸟科研基金、华为 Mindspore 学术基金、美团北斗科研合作基金等 10 多项科研项目。担任 CSIG 青工委副秘书长、CCF YOCSEF 广州候任主席、广州计算机学会副秘书长、VALSE 执行副主席等。担任 CVPR、ECCV、AAAI 等顶级会议领域主席（AC）或高级程序委员。担任 The Visual Computer 编委，获得 8 项 CVPR、NeurIPS、ACM MM 等国际顶级会议竞赛冠军。

Tutorial 4: NeRF 与 3DGS

彭思达 浙江大学

报告题目：NeRF 的基础及后续扩展



报告摘要：近年来，三维建模与渲染领域因为神经场景表示技术的兴起而受到学术界的热切关注。NeRF 作为一种早期的神经场景表示，与传统的点云和网格等三维表示相比，在精细度、模型表达灵活性以及渲染的真实感方面实现了巨大的飞跃。这种技术已经被广泛应用于各种三维重建、渲染与生成等场景。在本课程中，我们将深入探讨 NeRF 的基础理论和关键技术，并介绍近几年来 NeRF 的后续扩展工作。

讲者简介：彭思达，浙江大学软件学院“百人计划”研究员。在 2023 年获得浙江大学计算机科学与技术博士学位。研究方向为三维计算机视觉，代表工作为 4K4D、Neural Body、PVNet。至今在 TPAMI、CVPR、ICCV 等期刊或会议发表三十余篇论文，谷歌学术引用 3200 余次，其中一篇一作论文获得 CVPR 最佳论文提名，在 GitHub 开源获得超过上万个 stars。曾获得 2023 年全球 IMC 三维重建挑战赛冠军，2023 年世界人工智能大会云帆奖-明日之星、2022 Apple Scholar、2020 年 CCF-CV 学术新锐奖、2021 年中国 CCF 图形开源软件奖。

韩晓光 香港中文大学（深圳）

报告题目：3DGS，三维重建的终点吗？

报告摘要：近期，三维建模与渲染领域由于 3DGS 的引入而迎来了新的发展机遇。3DGS 技术相较于 NeRF 在多个关键方面展现出了显著的优势，引起了学术界和工业界的广泛关注。与 NeRF 相比，3DGS 在渲染速度、训练速度以及三维场景表达能力等方面实现了突破性的进展，为三维场景的创建和渲染提供了更为高效的解决方案。这种技术已经在实时渲染、增强现实、虚拟现实以及 SLAM 等领域得到了成功应用。在本报告中，我们将详细介绍 3DGS 的基础知识，包括其数学模型和工作原理，并介绍基于 3DGS 的后续拓展工作。

讲者简介：韩晓光博士，现任香港中文大学（深圳）理工学院助理教授，校长青年学者。他于 2017 年获得香港大学计算机科学专业博士学位。其研究方向包括计算机视觉和计算机图形学等，在该方向著名国际期刊和会议已发表论文近 100 篇，包括顶级会议和期刊 SIGGRAPH(Asia), CVPR, ICCV, ECCV, NeurIPS, ACM TOG, IEEE TPAMI 等。他曾获得吴文俊人工智能优秀青年奖，广东省杰出青年基金资助，香港中文大学（深圳）青年科研奖。目前担任 CVPR2023/2024，NeurIPS 2023 以及 ECCV2024 领域主席，同时也担任 IEEE TVCG 的编委。他的工作曾两次获得 CCF 图形开源数据集奖（DeepFashion3D 和 MVImgNet），曾两次入选 CVPR 最佳论文列表。

Workshops 报告及讲者简介

Workshop 1: 脑启发的视觉与学习

组织者：余肇飞（北京大学）、王鑫（清华大学）、郑乾（浙江大学）、顾实（电子科技大学）

时间：5 月 7 日（周二）8:30-12:00　　地点：喜悦厅

时间	主持人	内容
8:30~8:35		Workshop 开场
8:35~9:05	余肇飞	讲者：李国齐（中国科学院自动化所） 题目：类脑大模型
9:05~9：35	王鑫	讲者：施柏鑫（北京大学） 题目：基于神经形态相机的高速光度属性分析
9:35~10:05	王鑫	讲者：邓磊（清华大学） 题目：脉冲神经动力学网络及硬件实现
10:05~10:35	顾实	讲者：李远宁（上海科技大学） 题目：Enhancing neural encoding models for naturalistic perception with a multi-level integration of deep neural networks and cortical networks
10:35~10:45	中场休息	
10:45~11:15	顾实	讲者：王林（香港科技大学） 题目：大模型浪潮下，基于事件相机的视觉智能研究发展前瞻
11:15~11:50	郑乾	Panel 嘉宾： 李国齐（中科院自动化所）、施柏鑫(北京大学)、邓磊（清华大学）、李远宁（上海科技大学）、王林（香港科技大学）、顾实（电子科技大学）

李国齐 中科院自动化所

报告题目：类脑大模型

报告摘要：类脑智能是受人脑信息处理机制启发，基于神经元和神经环路的结构和功能，以更通用人工智能为目标构建计算系统的技术总称。近年来脉冲神经网络在通用场景逼近传统深度学习的主流网络性能，展现出引领未来智能技术的潜力。本报告介绍脉冲神经网络的模型、算法及其硬件部署以及基于类脑脉冲神经的大模型的科研进展。

讲者简介：李国齐，中国科学院自动化所研究员、国家杰出青年基金获得者， 在 Nature、Nature 子刊、Science 子刊、Proceedings of the IEEE、IEEE TPAMI 等期刊和会议上发表论文 200 余篇；主持国家自然科学基金重点项目、联合重点项目、科技部重点研发项目、 JKW 基础加强项目、北京市科委项目等 20 余项；担任 IEEE TNNLS、IEEE TCDS、清华大学学报-自然科学版等多个国际期刊编委；曾入选中国科学院百人计划、北京市杰青、北京市智源学者，中国算力十大青年先锋人物、中国智能计算科技创新人物。

施柏鑫 北京大学

报告题目：基于神经形态相机的高速光度属性分析

报告摘要：神经形态相机对比逐帧成像的普通相机拥有独特的优势，尤其是其对于高速运动物体和高动态范围场景的感知能力。已有研究充分展示了神经形态相机在图像去模糊、高动态范围重建、高速检测识别等计算机视觉任务上发挥的优势，然而神经形态相机的光度成像模型尚未得到充分的分析。本次报告将分享围绕光强高速变化触发事件的光度成像过程进行建模与分析的一系列研究进展：对光强变化瞬间的事件频率进行分析得到高可靠的场景辐射度估计，对高速运动的光源构建实时光度立体视觉进行表面法线估计，以及对高速遮挡物投射的阴影变化进行捕捉进而实现直接-间接光照分离的方法。

讲者简介：施柏鑫，北京大学计算机学院视频与视觉技术研究所副所长，视频与视觉技术国家工程研究中心、多媒体信息处理全国重点实验室研究员、长聘副教授、博士生导师（“博雅青年学者”）；北京智源人工智能研究院青年科学家。2013 年博士毕业于日本东京大学，曾先后在麻省理工学院媒体实验室、新加坡科技设计大学、南洋理工大学、日本国立产业技术综合研究所从事研究工作。研究方向为计算摄像学与计算机视觉，发表论文 200 余篇（包括 TPAMI 论文 23 篇，计算机视觉三大顶级会议论文 82 篇）。论文获评国际计算摄像会议（ICCP）2015 最佳论文提名（Best Paper - Runner Up）、国际计算机视觉会议（ICCV）2015 最佳论文候选（入选 IJCV 特刊 Best Papers from ICCV 2015），2021 年获得日本大川研究助成奖。科技创新 2030—“新一代人工智能”重大项目首席科学家，国家自然科学基金重点项目负责人，国家级青年人才计划入选者。担任国际顶级期刊 TPAMI、IJCV 编委，顶级会议 CVPR、ICCV、ECCV 领域主席。IEEE、CCF、CSIG 高级会员，APSIPA 杰出讲者。更多信息请访问“相机智能”实验室主页：http://camera.pku.edu.cn

邓磊 清华大学

报告题目：脉冲神经动力学网络及硬件实现

报告摘要：脉冲神经动力学网络是重要的脑启发计算模型，是类脑计算领域的前沿课题，其丰富的时空动力学被认为具有处理时序传感信号的巨大潜力。然而，目前对于脉冲神经动力学网络模型成分的深入理解较为匮乏，如何利用动力学特性处理时序计算任务获取高应用性能仍未充分解决。本次报告将系统性分析脉冲神经动力学网络的关键模型成分，进而构建集成树突时域异质性的新型脉冲神经动力学网络模型，揭示多尺度时域特征的融合机制，展示其在处理语音信号、视觉信号、脑电信号、机器人位置信号等宽频谱时序传感信号中的准确率和鲁棒性优势以及结合类脑计算芯片实现低功耗信息处理的效率优势，探讨该领域的机遇和挑战。

讲者简介：邓磊，清华大学精密仪器系，副教授，博士生导师，入选国家高层次青年人才计划。从事类脑计算领域研究超过 10 年，在 Nature、Nature Communications、PIEEE、JSSC、ICML、ICLR 等期刊和会议发表论文近百篇，包含 3 篇封面论文、1 篇 ESI 0.1%热点论文和 2 篇 ESI 1%高被引论文，担任 Frontiers in Neuroscience 期刊副主编、中国人工智能学会脑机融合与生物机器智能专委会委员，曾任多个国际会议分论坛主席和程序委员会委员。个人曾获全球前 2%顶尖科学家、北脑青年学者、吴文俊人工智能优秀青年奖和《麻省理工科技评论》中国籍 35 岁以下科技创新 35 人等荣誉，代表性成果曾获中国科学十大进展、世界互联网领先科技成果和中国计算机学会技术发明一等奖、日内瓦国际发明展金奖等奖项。

李远宁 上海科技大学

报告题目：Enhancing neural encoding models for naturalistic perception with a multi-level integration of deep neural networks and cortical networks

报告摘要：Cognitive neuroscience aims to develop computational models that can accurately predict and explain neural responses to sensory inputs in the cortex. Recent studies attempt to leverage the representation power of deep neural networks (DNNs) to predict the brain response and suggest a correspondence between artificial and biological neural networks in their feature representations. However, typical voxel-wise encoding models tend to rely on specific networks designed for computer vision tasks, leading to suboptimal brain-wide correspondence during cognitive tasks. To address this challenge, this work proposes a novel approach that upgrades voxel-wise encoding models through multi-level integration of features from DNNs and information from brain networks. Our approach combines DNN feature-level ensemble learning and brain atlas-level model integration, resulting in significant improvements in predicting whole-brain neural activity during naturalistic video perception.

Furthermore, this multi-level integration framework enables a deeper understanding of the brain's neural representation mechanism, accurately predicting the neural response to complex visual concepts. We demonstrate that neural encoding models can be optimized by leveraging a framework that integrates both data-driven approaches and theoretical insights into the functional structure of the cortical networks.

讲者简介：李远宁，上海科技大学生物医学工程学院研究员、助理教授、博导、计算认知与转化神经科学实验室主任，卡内基梅隆大学神经计算与机器学习博士，加州大学旧金山分校神经外科博士后，曾获美国 NIH 神经科学杰出学者奖，入选国家高层次青年人才计划，主持国自然面上项目等。主要研究方向为计算和认知神经科学，主要基于高密度颅内皮质脑电技术等解析语言、视觉认知过程的神经机制，并运用机器学习和人工智能方法建立神经编码和解码的计算模型，开发语言脑机接口，代表性研究成果发表在 Nature Neuroscience, Nature Communications, Science Advances, PNAS 等期刊。

王林 香港科技大学

报告题目：大模型浪潮下，基于事件相机的视觉智能研究发展前瞻

报告摘要：事件相机是受生物启发（类脑）的传感器，异步捕捉每个像素的强度变化，并产生编码时间、像素位置和极性（符号）的事件流。与传统的基于帧的相机相比，事件相机具有许多优势，如高时间分辨率、高动态范围、低延迟等。事件相机能够在具有挑战性的视觉条件下捕捉信息的能力，有潜力克服传统计算机视觉和机器人算法的局限性。近年来，深度学习已被引入到这个新兴领域，并激发了挖掘其潜力的积极研究。视觉学习与智能系统实验室（VLIS LAB）长期从事基于生物驱动传感器智能感知研究。本报告旨在分享近年来 VLIS LAB 在此研究领域的进展和成果。本报告重点分享以下内容：１）事件相机引导的视觉模太融合方法；２）VLM-引导的事件基础模型；３）多模态基础模型（包含事件模态）研究。最后，本报告探讨此新兴领域的发展前景及困境，与诸同仁共享。

讲者简介：王林老师目前为香港科技大学（广州）信息枢纽人工智能学域助理教授、香港科技大学计算机科学与工程系联署助理教授，本科事务主任及视觉学习与智能系统实验室负责人。2020-2021 年期间在英国帝国理工学院访学并与 T.-K. Kim 教授开展合作研究。王老师于 2021 年在韩国科学技术院（KAIST， QS 2022 Top 40）获得工学博士学位，主要从事人工智能驱动的机器人传感和感知研究。凭借出色的研究成果，于 2021 年被 KAIST 授予“最佳博士研究奖”。王老师课题组代表性研究方向为智能系统中的传感 (sensing) 与 2D/3D 感知 (perception) 研究，主要体现在：1）生物驱动传感器的视觉感知; 2) 多传感器视觉基础模型；3) 传感与 3D 建图；4) 具身智能系统（XR，机器人）应用。曾获得 IEEE VR (CCF-A) 最佳论文奖, CCF-腾讯犀牛鸟基金“优秀专利项目奖”，CVPR 2021“优秀审稿人”等荣誉。近年来，王老师在人工智能、机器人及人机交互（如 CVPR，ICRA, IEEE VR）等顶尖学术会议及期刊发表论文 60 余篇,并担任多个学术会议及期刊的学术委员会委员。相关研究成果请见：https://vlislab22.github.io/vlislab/。

顾实 电子科技大学

讲者简介：顾实，电子科技大学计算机科学与工程学院教授，脑与认知实验室负责人。美国宾夕法尼亚大学（University of Pennsylvania）应用数学与计算科学专业博士，清华大学数理基础科学专业学士，2017 年入选第十三批国家青年特聘专家，同年入选福布斯中国“30 岁以下 30 人“榜单。主要研究方向为计算神经科学与类脑人工智能，当前主要研究兴趣是基于人工智能的脑认知表达和调控模型。计算神经科学方面，提出并发展了脑网络上的控制理论分析，拓展了脑系统发育过程中的认知调控系统成熟化理论，相关研究发表在 Nature Communications 和 PNAS 等国际期刊上；类脑人工智能方面，建立了传统人工神经网络向脉冲神经网络转换的误差分析模型，并建立了与之相应的高效转换算法，极大地降低了转换算法的延迟，相关研究发表在 ICLR、ICML，NeurIPS 等人工智能国际会议上。

Workshop 2：面向移动终端的 AI 图像增强

组织者：程明明（南开大学）、王云鹤（华为）、李翔（南开大学）

时间：5 月 6 日（周一）13:30-18:00　　地点：欢悦厅

时间	主持人	内容
14:00~14:30	程明明	讲者：杜成（华为） 题目：华为手机 Camera 业务难题
14:30~15:00	李翔	讲者：左旺孟（哈尔滨工业大学） 题目：基于多相机和多曝的 AI 图像增强研究
15:00~15:30	王云鹤	讲者：潘金山（南京理工大学） 题目：面向图像视频复原的判别性自注意力机制方法
15:30~15:50	中场休息	
15:50~16:20	李翔	讲者：顾舒航（电子科技大学） 题目：底层视觉前沿技术应用实践：AI-ISP
16:20~16:50	程明明	讲者：付莹（北京理工大学） 题目：基于 RAW 数据的图像增强及应用
16:50~17:20	王云鹤	讲者：郭春乐（南开大学） 题目：面向移动终端的极暗场景图像增强技术
17:20~17:40	王云鹤	Panel 嘉宾： 程明明（南开大学）、李翔（南开大学）、杜成（华为）、左旺孟（哈尔滨工业大学）、潘金山（南京理工大学）、顾舒航（电子科技大学）、付莹（北京理工大学）、郭春乐（南开大学）

杜成 华为

报告题目：华为手机 Camera 业务难题

报告摘要：华为相机在过去数年中依靠软硬结合的技术创新为消费者提供了较好的拍照体验，但仍存在一些技术难题亟待解决，例如高层语义理解方向中的困难场景识别的准确率和稳定性以及自动对焦点的选择，底层图像恢复中的噪声细节、颜色还原、亮度和对比度、运动模糊、flicker&banding、摩尔纹和跨模组的一致性问题，AI 模型轻量化中的低比特量化和扩散模型的压缩，以及未来的多模态理解和可控生成技术。

讲者简介：杜成博士，毕业于清华大学，现任华为终端 BG Camera 部部长，自 2011 年加入华为公司后，一直从事华为相机的研发工作，主导了华为相机双摄、三摄、多摄和可变光圈等相机解决方案，以及 3A、大光圈、夜景、长焦、RAW 域恢复和增强等华为相机算法。

左旺孟 哈尔滨工业大学

报告题目：基于多相机和多曝的 AI 图像增强研究

报告摘要：当前的主流智能手机通常已预装多个相机并支持多曝拍摄，因而可为手机改善成像质量和拍摄体验提供新的契机。本报告将主要汇报团队近年来基于多相机和多曝的 AI 图像增强方面的研究进展，具体包括基于多相机机的图像超分辨和数字变焦方法，以及基于多曝图像的自监督去噪/去模糊、自监督高动态成像、以及多复原任务统一框架。

讲者简介：左旺孟，哈尔滨工业大学计算机学院教授、博士生导师。主要从事底层视觉、视觉生成、视觉理解和多模态学习等方面的研究。在 CVPR/ICCV/ECCV/NeurIPS/ICLR 等顶级会议和 T-PAMI、IJCV 及 IEEE Trans.等期刊上发表论文 200 余篇。曾任 ICCV、CVPR 等 CCF-A 类会议领域主席，现任 IEEE T-PAMI、T-IP、中国科学-信息科学等期刊编委。

潘金山 南京理工大学

报告题目：面向图像视频复原的判别性自注意力机制方法

报告摘要：得益于 Transformer 中自注意力机制在图像视频非局部特征建模中的有效性，基于该方法的图像视频复原取得了显著的进展。然而，目前大多数解决图像视频复原的自注意力机制方法不能有效地建模图像视频的细节特征并且在特征建模过程中存在判别性弱的问题。针对这一问题，我们将介绍具有判别性的自注意力特征建模方法，实现图像结构和细节特征的协同建模，提升图像视频复原质量。

讲者简介：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。目前在国际权威期刊和会议上发表论文 100 余篇。所发表论文在 Google Scholar 中被引用 13000 余次。研究工作获得 2018 年度中国人工智能学会优秀博士学位论文奖、辽宁省优秀博士学位论文奖以及 2019 年度国家优秀青年科学基金资助。担任 CVPR、ECCV、ICML、ICLR 等权威国际会议的领域主席。目前主持国家自然科学基金委-联合基金重点项目、面上项目等国家级科研项目。

顾舒航 电子科技大学

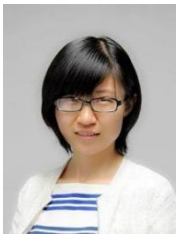
报告题目：底层视觉前沿技术应用实践：AI-ISP

报告摘要：近年来，基于深度学习的图像复原与增强技术取得了突破性的进展，促进了其在产业界的广泛应用。图像信号处理器（ISP）是成像设备的重要组成部分，需要实时解决光学系统局限性、电子器件局限性以及视觉感知非线性带来的多方面挑战，对于图像增强方法的鲁棒性、灵活性以及高效性具有很高的要求。本报告结合报告人近年来在“基于人工智能的图像信号处理器”（AI-ISP）方面的工程实践，介绍如何通过设计专用硬件架构、高效增强网络结构、多源数据协同训练、弱监督训练、网络结构优化等技术，实现终端平台的鲁棒、灵活、高效图像增强。

讲者简介：顾舒航，电子科技大学教授，国家级青年人才。博士毕业于香港理工大学，后加入苏黎世联邦理工学院任博士后研究员。长期从事人工智能与计算机视觉方面的研究，主要围绕图像复原与增强、图像压缩、图像生成等底层视觉问题展开研究。发表中科院一区期刊/CCF-A 类会议论文 40 余篇，包括 IEEE TPAMI, IJCV, CVPR, ICCV, NeurIPS 等。谷歌学术 12000 余次，获得吴文俊人工智能奖自然科学一等奖（第三完成人），入选斯坦福大学全球顶尖科学家榜单。曾参与/主导多个底层视觉相关的产业界项目，2020-2022 年主管 OPPO 芯片子公司哲库多媒体 AI 算法部门，作为核心人员参与 6 纳米制程的影像外挂芯片与 4 纳米制程的手机系统级芯片项目。

付莹 北京理工大学

报告题目：基于 RAW 数据的图像增强及应用



报告摘要：长期以来，计算机视觉领域的主要研究对象为经过图像信号处理器（ISP）处理之后得到的标准 RGB 图像。尽管这类图像在查看、存储、传播上十分便利，但受限于 ISP 过程中一系列不可逆的退化操作，如模糊、量化等，这些信息损失使得标准 RGB 图像在非理想光照条件下的使用受到限制。RAW 图像作为传感器直接采集到的未经退化的数据，具有无损、高数据位宽、线性响应等特

征，因此近年来被越来越多的研究者采纳，用于更加极端光照环境下的图像质量增强。本报告将关注这类典型的极端光照环境场景，利用 RAW 数据特有的特征和优势实现图像高质量增强，并进一步介绍 RAW 数据的下游应用。

讲者简介：付莹，北京理工大学教授，博士生导师。2017 年入选海外高层次青年人才引进计划。主要从事计算机视觉、计算摄像、图像/视频处理等相关领域研究，近五年以第一/通讯作者在中科院一区期刊和计算机机学会推荐 A 类会议发表论文 50 余篇，授权发明专利 30 余项。获计算机学会推荐 A 类国际会议 ICML 杰出论文奖和国际会议 PRCV 最佳论文奖，主持国家自然科学基金重点项目等国家级项目/课题 10 余项，获中国指挥与控制学会科学技术进步奖一等奖、人工智能学会教学成果奖励计划一等奖、图象图形学会石青云女科学家奖。

郭春乐 南开大学

报告题目： 面向移动终端的极暗场景图像增强技术

报告摘要：在当前的数字时代，智能手机等移动设备的普及使得摄影成为了日常生活的重要组成部分。然而，低光环境下拍摄的图像往往由于噪声过多、细节缺失和对比度低，造成视觉感知质量不佳，同时也影响视觉感知算法性能。这一系列的问题与挑战激发了研究者对于在移动设备上实现高效的低光图像增强技术的兴趣，并吸引了移动设备生产商的广泛关注。本报告将探讨如何利用深度学习技术，来改进移动设备上极暗场景图像质量。首先，在基于 RGB 的图像亮色增强方面，本报告将讨论基于增强曲线的增强算法在低光场景中的一些成功应用，并分析其优劣及应用前景。接着，深入数字图像成像底层，在基于 RAW 数据的极暗场景图像增强方面，结合传统图像信号处理（ISP）技术思想，讨论 AI ISP 算法的设计与实现。此外，结合工业应用场景，讨论基于 RAW 数据增强模型与特定传感器的耦合问题，以及在跨设备快速迁移上的一些初步尝试。最后，本报告将展望面向移动终端的极暗场景图像增强技术的研究前景。

讲者简介：郭春乐，南开大学，南开国际先进研究院（深圳福田），副教授。主持包括国家自然科学基金、华为、三星等资助的多项科研项目，累计获得资助 300 余万元，相关多项专利技术完成成果转化。他的主要研究内容包括计算成像、图像增强与复原、图像生成与编辑，交互式分割等。作为第一作者（通讯作者/共同一作）在 TPAMI、TIP、CVPR 等国际学术期刊及会议上发表论文而 30 余篇，其中 3 篇论文入选 ESI 高被引论文、1 篇论文入选 ESI 热点论文，谷歌学术引用 5000 余次，其中一作论文单篇最高引用 1200 余次，10 篇引用过百，多篇会议论文入选 Oral/Highlight/Spotlight。曾任 VALSE2022 宣传主席，BMVC2022 领域主席。现担任 SCI 二区期刊 IEEE Journal of Oceanic Engineering 编委。

Workshop 3：视觉世界模型与视频生成

组织者：唐彦嵩（清华大学）、舒祥波（南京理工大学）、朱霖潮（浙江大学）

时间：5 月 6 日（周一） 13：30-18：00 地点：喜悦厅

时间	主持人	内容
14:00-14:30	舒祥波	讲者：张兆翔（中国科学院自动化研究所） 题目：基于世界模型的视觉场景生成与应用
14:30-15:00	舒祥波	讲者：黄思远（北京通用人工智能研究院） 题目：三维世界模型
15:00-15:30	舒祥波	讲者：赵洲（浙江大学） 题目：以人为中心的多模态生成式模型研究
15:30-15:40	舒祥波	企业宣讲：阿里妈妈技术 阿里妈妈智能创意最新进展（王标）
15:40-16:10	唐彦嵩	讲者：罗翀（微软亚洲研究院） 题目：探索短视频生成与编辑的前沿技术
16:10-16:40	唐彦嵩	讲者：刘希慧（香港大学） 题目：Towards Controllable and Compositional Visual Content Generation
16:40-16:50	唐彦嵩	企业宣讲：美团 生活服务场景的视觉生成（魏晓明）
16:50-17:20	唐彦嵩	讲者：袁粒（北京大学） 题目：Open-Sora Plan 视频生成开源计划 - 进展与不足
17:20-18:00	朱霖潮	Panel 嘉宾： 张兆翔（中国科学院自动化研究所）、黄思远（北京通用人工智能研究院）、赵洲（浙江大学）、罗翀（微软亚洲研究院）、刘希慧（香港大学）、袁粒（北京大学）、白磊（上海人工智能实验室）

张兆翔 中国科学院自动化研究所

报告题目：基于世界模型的视觉场景生成与应用

报告摘要：报告首先概述了生成式世界模型的最新进展，以及其在通用领域、自动驾驶、机器人等领域的应用情况。随后，重点介绍了我们团队在自动驾驶领域的生成式世界模型研究工作。我们首次提出了一种名为 Drive-WM 的全新多视图世界模型，旨在增强端到端自动驾驶规划的安全性。Drive-WM 模型通过整合文本、3D 场景布局以及自车的运动信息等多种条件，实现了对未来情景的多样化生成控制。该模型利用多视图世界模型，能够预测不同规划路线的未来情景，并根据视觉预测结果实时获取奖惩反馈，从而优化当前路线选择，提升自动驾驶规划的安全性。

讲者简介：张兆翔，博士，中国科学院自动化研究所模式识别国家重点实验室与智能感知与计算研究中心研究员、博士生导师，中国科学院大学岗位教授，中国科学院脑科学与智能技术卓越创新中心骨干。张兆翔博士的研究兴趣包括：模式识别、计算机视觉与深度学习，具体研究方向包括：视觉认知计算、类脑学习和面向开放环境的视觉感知与理解，在本领域国际主流期刊与会议上发表论文 200 余篇，近五年来在 JMLR、IEEE T-IP、IEEE T-NN 等顶级期刊与 CVPR、ICCV、ECCV、NIPS、AAAI、IJCAI 等顶级会议发表论文 60 余篇，申请专利 20 余项，已授权 10 项，承担了国家自然科学基金重点项目、国家重点研发项目课题、北京市新一代人工智能重点研发项目等多项国家级科研项目和企业合作项目。张兆翔博士是 IEEE 高级会员，VALSE 常务 AC，中国计算机学会 CCF 杰出会员、中国人工智能学会 CAAI 杰出会员、中国人工智能学会 CAAI 副秘书长、中国人工智能学会 CAAI 理事、中国人工智能学会 CAAI 模式识别专委会秘书长、中国图象图形学学会会员发展与服务委员会副主任。张兆翔博士是 IEEE T-CSVT、Patten Recognition、NeuroComputing 编委，担任 CVPR、ICCV、AAAI、IJCAI、ACM MM、ICPR、ACCV 等国际会议的领域主席（Area Chair）。张兆翔博士入选“教育部长江学者”、“国家万人计划青年拔尖人才”和“教育部新世纪优秀人才支持计划”，曾获得中国人工智能学会吴文俊技术发明奖二等奖（排名第一）。

黄思远 北京通用人工智能研究院

报告题目：三维世界模型

报告摘要：人类的智能在很大程度上来自于对三维世界的理解以及与物理环境的深入互动。例如，婴儿通过抓取、投掷和敲击等动作，在与物理世界的互动中学习并获得常识，这些互动不断丰富他们对三维世界的理解。同样地，要想让机器达到甚至超过人类智能，使其理解三维空间至关重要。通过教学和互动完成具身任务，机器能够学习常识，这是推动人工智能进步的关键一步。本次报告将重点介绍我们最近在三维世界模型构建（SceneVerse）以及利用三维世界模型提升具身任务（LEO）的相关工作，并讨论物理常识对三维世界模型以及通用人工智能的重要性。

讲者简介：黄思远博士是北京通用人工智能研究院（BIGAI）的研究科学家，并担任通用视觉实验室负责人。他在加州大学洛杉矶分校（UCLA）统计系获得博士学位。他的研究旨在构建一个能够理解和与三维环境交互的类人通用智能体。为实现这一目标，他的研究在以下方向做出了贡献：（1）开发可泛化的视觉表征以用于三维重建和语义落地，（2）

建模并模仿人类与三维世界的复杂交互，（3）构建擅长与三维世界和人类交互的具身智能体。他的研究发表于四十余篇 CVPR/ICCV/NeurIPS/ICML 等会议及期刊论文，并曾获得 ICML Workshop 最佳论文。

赵洲 浙江大学

报告题目：以人为中心的多模态生成式模型研究

报告摘要：多模态生成式模型最近取得了巨大的突破，用户可以输入自然语言生成图像、视频、音频，甚至是 3D 内容。本报告首先介绍针对以人为中心的多模态生成式模型存在的一些挑战和现有工作的最新进展，包括可泛化说话人视频合成和低延时说话人语音合成等。其次，本次分享一些我们面向数字说话人方面的多模态生成式模型的工作，包括可泛化说话人视频合成 GeneFace（ICLR23）、单图三维重构说话人 Real3D-Portrait（ICLR24）和低延时说话人并行化语音合成 FastSpeech 1/2（NeurIPS19/ICLR21）、歌声合成 DxiffSinger（AAAI22）和零样本合成 MegaTTS（ICLR24）。

讲者简介：赵洲，浙江大学计算机学院教授/博士生导师，主要研究方向为多媒体计算和交互式生成模型，在国际期刊 TPAMI 和机器学习和视觉计算会议 NeurIPS、ICML、ICLR、CVPR 和 ICCV 等上发表 60 余篇论文，包括：低延时伪数值扩散模型推理算法 PNDM（ICLR22）、并行化语音生成算法 FastSpeech 1/2 和 DiffSinge（NeurIPS19/ICLR21/AAAI22）、单图三维重构算法 GeneFace（ICLR23/ICLR24）和零样本语音合成 MegaTTS（ICLR24）等，谷歌学术引用 12720 次，应用于微软、字节、Stability AI、华为等公司，获 2021 年度中国电子学会科技进步一等奖、2022 年度教育部科技进步一等奖，入选斯坦福大学发布的“全球前 2%顶尖科学家榜单”和爱思唯尔高被引学者。

罗翀 微软亚洲研究院

报告题目：探索短视频生成与编辑的前沿技术

报告摘要：近期，OpenAI 发布的 Sora 文本到视频生成模型成为了业界瞩目的焦点，激发了广泛的讨论与关注。该模型不仅能够根据简单的文本描述生成高度真实且充满创意的视频场景，而且在视频的持续时间、动作流畅性以及物体连贯性等关键技术上取得了显著的进步。视频生成技术在开放领域的迅速崛起，仅用了几年时间便取得了令人瞩目的成就。本次讲座将首先对视频生成技术的发展历史进行简单回顾，梳理并阐述关键技术的演进路径。接下来，我将详细介绍我们研究团队在短视频生成和视频内容编辑领域的两项创新性研究（均获 CVPR 2024 的收录）。这两项研究不仅对视频生成和编辑的流程进行了优化，而且引入了创新的关键技术支持，极大地提升了视频内容生成与编辑的可控性和创造性的平衡，实现了既符合美学标准又满足用户需求的视频内容创作。在报告的最后部分，我将对视频生成技术的未来发展趋势进行展望，并探讨其在不同领域的潜在应用前景。

讲者简介：罗翀，博士，微软亚洲研究院智能多媒体组首席研究经理，担任中国科学技术大学和西安交通大学兼职教授、博士生导师，IEEE 高级会员。主要研究方向为计算机视觉、多媒体理解、多媒体生成等。现任 IEEE 电路与系统学会多媒体系统与应用技术委员会委员，现任权威多媒体期刊 IEEE Transactions on Multimedia 编委，担任多个计算机视觉与人工智能领域顶级会议领域主席。已在顶级会议和多份 IEEE 期刊发表论文近百篇，总被引 6000 余次。曾获 2016 年上海市计算机学会网络领域最有影响力论文奖和 2023 年 ICLR 优秀论文奖。

刘希慧 香港大学

报告题目：Towards Controllable and Compositional Visual Content Generation

报告摘要：Visual content generation has achieved great success in the past few years, but current visual generation models still lack controllability and compositionality. In real applications, we desire highly controllable visual generation models which allow users to control the generated contents in a fine-grained manner. We also desire models which can effectively compose objects with different attributes and relationships into a complex and coherent scene. In this talk, I will introduce our several works towards controllable and compositional visual content generation. I will introduce T2I-CompBench for benchmarking compositional text-to-image generation. I will also introduce our recent works on drag-based video editing, controllable 3D generation, and flexible diffusion transformers (FiT). I will discuss some of our ongoing efforts towards controllable and compositional video generation.

讲者简介：Xihui Liu is an Assistant Professor at the Department of Electrical and Electronic Engineering (EEE) and Institute of Data Science (IDS), The University of Hong Kong, affliated with HKU-MMLab. Before joining HKU, she was a postdoc Scholar at UC Berkeley, advised by Prof. Trevor Darrell. She obtained her Ph.D. degree from Multimedia Lab (MMLab), the Chinese University of Hong Kong, supervised by Prof. Xiaogang Wang and Prof. Hongsheng Li, and received her bachelor's degree from Tsinghua University. Her research interests cover computer vision, machine learning, and artificial intelligence, with special emphasis on visual synthesis, generative models, vision and language, and multimodal AI. She has published over 30 papers on top conferences (CVPR/ICCV/ECCV/NeurIPS/ICLR/ICML/…) and serves as the area chair for CVPR 2024. She was awarded Adobe Research Fellowship 2020, MIT EECS Rising Stars 2021, and WAIC Rising Stars Award 2022.

袁粒 北京大学

报告题目：Open-Sora Plan 视频生成开源计划 - 进展与不足

报告摘要：文生视频模型 Sora 能生成高质量视频内容，但是该技术和模型并未开源开放给大众；本次报告主要介绍北大-兔展 AIGC 联合实验室共同发起的 Open Sora Plan, 旨在聚集开源社区力量共同复现出一版类 Sora 架构的文生视频模型，并将该框架开发给所有人使用。该开源计划一经发起后获得开源社区广泛支持，同时也获得国产 AI 算力的支持，目前已经开源第一版模型和算法细节。本次报告将详细介绍该开源项目发起初衷、技术路线、当前进展与不足，以及未来计划。开源计划项目链接：https://github.com/PKU-YuanGroup/Open-Sora-Plan。

讲者简介：袁粒，北京大学博雅青年学者、博导，北京大学深圳研究生院助理教授，入选国家海外高层次青年人才计划、国家优秀留学生奖(归国类)、福布斯亚洲 30U30 名单等，主持国家科技重大专项课题和国自然青年基金等。研究方向为多模态深度学习，代表性学术工作包括 VOLO, T2T-ViT 等深度神经网络框架和知识蒸馏相关工作，一作论文单篇被引用千余次，代表性应用工作包括 ChatExcel，ChatLaw 等大模型垂直领域应用。

唐彦嵩 清华大学

个人简介：唐彦嵩，清华大学深圳国际研究生院助理教授、博士生导师，于清华大学自动化系获得工学学士和博士学位，并先后在美国加州大学洛杉矶分校、微软亚洲研究院、英国牛津大学从事访问学者和博士后研究工作。主要从事计算机视觉等领域的相关工作，在国际权威期刊和会议上发表论文 40 余篇，主持国家重点研发计划课题等国家级项目，以及中国人工智能学会-华为 Mindspore 学术奖励基金、中国计算机学会-腾讯犀牛鸟等校企联合项目。获得吴文俊人工智能优秀博士学位论文，入选第八届中国科协青年人才托举工程和微软亚洲研究院“铸星计划”，担任中国人工智能学会模式识别专业委员会副秘书长等学术职务。

舒祥波 南京理工大学

个人简介：舒祥波，南京理工大学计算机科学与工程学院/人工智能学院院长助理、教授、博士生导师、国家优秀青年基金获得者、江苏省杰出青年基金获得者、CCF/IEEE 高级会员。近年主要研究方向为视频内容分析，在 TPAMI、TIP、TNNLS、CVPR、ICCV、AMM 等国际期刊/会议上发表学术论文近 100 篇，其中 ESI 高被引论文 7 篇；获中国电子学会自然科学一等奖、ACM MM 2015 最佳论文提名、MMM 2016 最佳学生论文奖、江苏省优秀博士论文奖、中国人工智能学会优秀博士论文奖；主持或参与国家自然科学重点/面上/青年项目、国家重点研发课题、国防基础科研项目等国家级项目。担任 CSIG 青工委副秘书长，以及 IEEE TNNLS、IEEE TCSVT、Information Sciences、The Visual Computer 等期刊编委，获 2022 年度 IEEEE TNNL、IEEE TMI 杰出审稿人。

朱霖潮 浙江大学

个人简介：朱霖潮，浙江大学百人计划研究员、博士生导师。获首届谷歌学术研究奖（2021）。主要研究方向为跨媒体智能及其应用、通用基础模型、人工智能科学计算等。曾获 CVPR MABe 多智能体行为建模竞赛(2022)等 8 项国际冠军。曾担任 IEEE MLSP 领域主席(2021)、ICIP 领域主席（2024）、ECCV 领域主席（2024），并多次组织在 CVPR 等国际会议的专题研讨会。在 IEEE T-PAMI、IJCV、CVPR、ICCV 等高水平学术期刊及会议发表论文 70 余篇，含 9 篇大会口头报告。

Workshop 4: 具身智能的视觉与学习

组织者：盛律（北京航空航天大学）、李镇（香港中文大学（深圳））、张瑞茂（香港中文大学（深圳））

时间：5 月 6 日（周一）13:30-18:00　　地点：欣悦厅

时间	主持人	内容
14:00~14:30	盛律	讲者：卢策吾（上海交通大学） 题目：具身智能-感知（P），想象（I），执行（E）PIE 方案与具身大模型探索
14:30-15:00	盛律	讲者：王鹤（北京大学） 题目：通向开放指令操作的具身多模态大模型系统
15:00-15:10	盛律	企业宣讲：极视角 题目：构建高效 AI：极栈训推一体化平台技术实践（邓富城）
15:10-15:40	中场休息	
15:40-16:10	李镇	讲者：刘航欣（北京通用人工智能研究院） 题目：机器人的场景理解与任务运动规划
16:10-16:40	李镇	讲者：弋力（清华大学） 题目：基于人类行为仿真的可泛化人机协作
16:40-17:10	张瑞茂	Panel 嘉宾： 卢策吾（上海交通大学）、王鹤（北京大学）、刘航欣（北京通用人工智能研究院）、弋力（清华大学）、盛律（北京航空航天大学）、李镇（香港中文大学（深圳））

卢策吾 上海交通大学

报告题目：具身智能-感知（P），想象（I），执行（E）PIE 方案与具身大模型探索

报告摘要：该讲座介绍讲者具身智能 PIE 方案。P（Perception）介绍讲者机器人全感知与交互感知工作。I（Imagination），介绍讲者的物理世界概念驱动仿真推理框架。E（Execution）介绍讲座通用元操作技能设想与工作。基于上述三个模块，介绍具身 PIE 大模型探索与初步成果。最后介绍具身认知智能工作，如何验证脑神经行为与身体行为稳定隐射关系。

讲者简介：卢策吾，上海交通大学教授，博士生导师，2016 年或海外高层次青年引进人才，2018 年被《麻省理工科技评论》评为 35 位 35 岁以下中国科技精英（MIT TR35），2019 年获求是杰出青年学者，2020 年获上海市科技进步特等奖（第三完成人），2022 年获教育部青年科学奖，IROS 最佳论文之一（6/3579），2023 年获机器人顶会 RSS 最佳系统论文提名奖（共四项），科学探索奖。以通讯作者或第一作者在《自然》，《自然 ·机器智能》，TPAMI 等高水平期刊和会议发表论文 100 多篇；担任 Sicence 正刊，Nature 子刊，Cell 子刊等期刊审稿人，NeurIPS，CVPR，ICCV ，ECCV，IROS ，ICRA 领域主席。研究兴趣包括具身智能，计算机视觉。

王鹤 北京大学

报告题目：通向开放指令操作的具身多模态大模型系统

报告摘要：本体层、技能层和大模型层构成的三层级具身多模态大模型系统是实现通用机器人的一种方案。本报告将讨论通过三维视觉打造多个泛化的移动和操作技能，包括抓取、铰接类物体操作、柔性物体操作和建图导航等等。而大模型层则负责大脑的能力，本报告将展示 GPT-4V 为代表的非具身多模态大模型进行视觉感知、任务规划和调用中层的三维视觉技能，实现从家用电器泛化操作到开放指令物体摆放的能力。最后，报告将展望端到端具身多模态大模型，讨论其中的机会和挑战。

讲者简介：王鹤博士是北京大学计算机学院前沿计算研究中心（CFCS）的助理教授和博士生导师。他创立并领导了北大具身感知与交互实验室（EPIC Lab，主页：https://hughw19.github.io），研究目标是通过发展具身技能及具身多模态大模型推进通用具身智能。他同时担任北大-银河通用具身智能联合实验室主任和北京智源人工智能研究院具身智能研究中心主任。他已在计算机视觉、机器人学和人工智能的顶级会议和期刊（CVPR/ICCV/ECCV/TRO/RAL/ICRA/NeurIPS/ICLR/AAAI 等）上发表五十余篇工作，其论文获得 ICCV2023 最佳论文候选，ICRA2023 最佳操纵论文候选，2022 年世界人工智能大会青年优秀论文（WAICYOP）奖，Eurographics 2019 最佳论文提名奖。他担任了 CVPR2022 和 WACV2022 的领域主席，Image and Vision Computing 的副主编和诸多顶会的审稿人、程序委员。在加入北京大学之前，他于 2021 年从斯坦福大学获得博士学位，师从美国三院院士 Leonidas. J Guibas 教授，于 2014 年从清华大学获得学士学位。

刘航欣 北京通用人工智能研究院

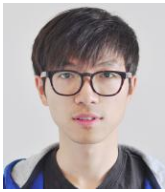
报告题目：机器人的场景理解与任务运动规划

报告摘要：本报告介绍了一种基于虚拟运动链条（virtual kinematic chain）的机器人建模思路，该方法能够将机器人的移动底座、机械臂和被操作的物体建模为一个整体，更全面地表示机器人的姿态和运动，从而使机器人在困难环境中的产生更协调、灵活的动作。这个建模思路进一步驱动了一种以体现运动信息为目标的场景重建方法，该方法可以从 RGB-D 输入中构建一个能够与虚拟运动链条思路结合的场景图表征，实现了一个新的机器人感知和规划框架。该框架在多步骤复杂移动操作任务中展现了机器人更强的移动操作和抓取能力。

讲者简介：刘航欣，北京通用人工智能研究院研究员、机器人实验室负责人。2021 年获加州大学洛杉矶分校计算机科学博士学位，2016 年获弗吉尼亚理工大学机械工程与计算机科学双学士学位。研究重点包括机器人感知、认知、学习，任务和运动规划、人机交互等。博士期间参与了美国 DARPA XAI（人工智能可解释性）、DARPA SIMPLEX（机器人感知及认知学习）、ONR MURI（任务理解）、ONR Cognitive Robot（认知机器人）等重大研究项目。在机器人和人工智能领域的国际顶级会议和期刊上发表论文 30 余篇，包括 Science Robotics/Engineering/IJCV/RA-L/ICRA/IROS/AAAI 等，并获得了 2019 年 ACM 中国图灵大会（TURC）获得了最佳论文奖、2023 年 IROS 移动操作方向最佳论文提名。

弋力 清华大学

报告题目：基于人类行为仿真的可泛化人机协作



报告摘要：具身智能很重要的研究目标在于使机器人能够与人类进行交互和协作。近年来，尽管在教授机器人无需人类参与的操作技能方面已取得了重大的技术进展，但在可扩展地学习人机协作技能以应对各种任务和人类行为方面仍存在滞后。现实世界中针对人机协作的机器人训练成本高昂且风险较大，从可扩展性的角度来看，这种训练方法在实际应用中并不实际。因此，在将机器人部署到现实世界之前，有必要在虚拟环境中模拟人类行为并对机器人进行训练。在本次报告中，我将讨论我们近期在采集大规模人物交互数据集、模拟能够推广到新环境和任务的逼真人类行为、以及利用可扩展的人物仿真实现可泛化人机协作方面所做的努力。通过在多样化的场景中模拟人类交互，我们创建了以人为中心的机器人仿真器。通过采用动态任务和动作规划来生成高质量的示例，我们可以训练可泛化的人机协作技能。我们相信，这种方法为推进真实世界的人机协作提供了一种强大的范式。

讲者简介：弋力博士，现任清华大学交叉信息研究院助理教授，国家优青（海外）。他在斯坦福大学取得博士学位，导师为 Leonidas J. Guibas 教授，毕业后在谷歌研究院任研究科学家。在此之前，他在清华大学电子工程系取得了学士学位。他近期的研究兴趣涵盖三维视觉和具身人工智能，他的研究目标是使智能机器人具备理解三维世界并与之互动的能力。他在计算机视觉、计算机图形学以及机器学习领域的顶级会议发表论文六十余篇，并担任 CVPR 2022-2024、IJCAI 2023、NeurIPS 2023 领域主席。他的工作在领域内得到广泛关注，引用数 20000+，代表作品包括 ShapeNet Part，光谱图 CNN，PointNet++等。

盛律 北京航空航天大学

个人简介：盛律，博导，北京航空航天大学“卓越百人”副教授，入选北航青年拔尖计划。2011 年获浙江大学学士学位，2017 年获香港中文大学博士学位，同年加入香港中文大学多媒体实验室从事博士后研究。主要研究方向为三维视觉、多媒体分析和具身智能。在 IEEE TPAMI/IJCV/TIP 以及 CVPR/ICCV/NeurIPS/ICLR/ECCV/AAAI/ACM MM 等重要国际期刊和会议发表论文超过 50 篇, Google Scholar 显示被引用数超 4700 次。获 CVPR 2021 年 ScanRefer 三维定位挑战赛第一名。组织 ICCV 2021 SenseHuman 等多个国际会议研讨会和竞赛。现任 ACM Computing Surveys 副编辑，CVPR 2024、ECCV 2024 和 ACM Multimedia 2024 领域主席，以及多个领域顶会顶刊审稿人和程序委员。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。

李镇 香港中文大学（深圳）

个人简介：李镇博士现任香港中文大学（深圳）理工学院/未来智联网络研究院助理教授，校长青年学者。李镇博士获得香港大学计算机科学博士学位（2014-2018 年），他还于 2018 年在芝加哥大学担任访问学者。李镇博士荣获 2021 年中国科协第七届青年托举人才，2023CVPR HOI4D 竞赛第一名，2022 年 SemanticKITTI 语义分割竞赛第一名，2023 年 IROS 最佳论文 Finalist，ICCV2021 Urban3D 竞赛第二名，CASP12 接触图预测全球冠军等。李镇博士还获得了来自于国家、省市级以及工业界的科研项目。李镇博士领导了港中深的 Deep Bit Lab（https://mypage.cuhk.edu.cn/academics/lizhen/），其主要的研究方向是 3D 视觉解析及应用 (包括但不限于点云解析，多模态联合解析)，深度学习等基础理论算法研究，并致力于将 2D/ 3D 人工智能算法推广应用于交叉学科，自动驾驶，工业视觉等场景中，在该方向著名国际期刊和会议发表论文 60 余篇，包括顶级期刊 Cell Systems, Nature Communications, T-PAMI, TMI, TVCG，TNNLS 等，以及顶级会议 CVPR, ICCV, ECCV, NeurIPS, ICLR，IROS，ACM MM，AAAI, IJCAI, MICCAI 等。李镇博士担任 IEEE Transactions on Mobile Computing、IROS 副编以及众多顶刊顶会的审稿人，李镇博士还是广东院士联合会脑科学与类脑智能专委委员，VALSE、MICS、中国图象图形学学会机器视觉专委会，3DV 专委会等学术组织的委员。

张瑞茂 香港中文大学（深圳）

个人简介：张瑞茂，现为香港中文大学（深圳）数据科学学院的副研究员。张博士于中山大学获得学士和博士学位。后在香港中文大学多媒体实验室担任博士后研究员。他的研究兴趣包括计算机视觉和具身智能。在 TPAMI、IJCV、ICML、ICLR、CVPR、ICCV 等顶级会议和期刊上发表论文 50 余篇，谷歌学术论文引用 4500 余次。张博士曾获得多项国际比赛的奖项，例如 2017 年 Youtube 8M 视频分类挑战赛金奖、2020 年 AIM 学习图像信号处理管道挑战赛第一名。2021 年，张博士被评委 NeurIPS 优秀审稿人。张博士也是多媒体领域知名期刊 ACM Transactions on Multimedia Computing、Communications and Applications 的副编辑。

Workshop 5: 大模型赋能智慧医疗

组织者：陈浩（香港科技大学）、夏勇（西北工业大学）、张少霆（上海人工智能实验室）

时间：5 月 6 日（周一）8:30-12:00　　地点：温德姆宴会厅

时间	主持人	内容
8:30~9:00	陈浩	讲者：王本友（香港中文大学深圳） 题目：医疗大语言模型实践和一点儿思考
9:00~9:30	夏勇	讲者：谢伟迪（上海交通大学） 题目：Towards Foundation Model for Medicine
9:30~10:00	张少霆	讲者：王国泰（电子科技大学） 题目：基于医学图像大模型的弱标注和无标注学习
10:00~10:30	陈浩	讲者：史淼晶（同济大学） 题目：深度大模型的兴起以其在医疗场景的应用
10:30~10:40		中场休息
10:40~11:10	夏勇	讲者：张晓凡（上海交通大学） 题目：医疗领域大语言模型的训练及智能体应用
11:10~11:40	张少霆	讲者：雷柏英（深圳大学） 题目：大模型赋能 AD 智能诊断
11:40-12:00	陈浩	Panel 嘉宾： 王本友（香港中文大学深圳）、谢伟迪（上海交通大学）、王国泰（电子科技大学）、史淼晶（同济大学）、张晓凡（上海交通大学）、雷柏英（深圳大学）、夏勇（西北工业大学）、张少霆（上海人工智能实验室）

王本友 香港中文大学（深圳）

报告题目：医疗大语言模型实践和一点儿思考

报告摘要：一年来，OpenAI 的 ChatGPT 大大推进了人工智能应用，但是其技术闭源和国内的封锁也影响了国内用户的使用，甚至高端 GPU 芯片也被禁售，也将加大了技术的差距。本次分享将会介绍我们团队先后开发的华佗 GPT，AceGPT 阿拉伯语大模型以及最近的多模态大模型，这些实践主要集中在如果把语言模型适配到中文医疗，特别是多模态的场景。此外还将展望一下未来大模型的发展方向。

讲者简介：王本友，香港中文大学（深圳）数据科学学院助理教授和医疗健康信息研究中心副主任、深圳市大数据研究院研究科学家。迄今为止，他曾获得了 SIGIR 2017 最佳论文提名奖、NAACL 2019 最佳可解释 NLP 论文、NLPCC 2022 最佳论文和华为火花奖。其领导的研究团队开发的大模型包括多语言大模型凤凰、医疗健康垂直领域大模型华佗 GPT 和阿拉伯语大模型 AceGPT。

谢伟迪 上海交通大学

报告题目：Towards Foundation Model for Medicine

报告摘要：近年来，基于图像-文本联合学习的 foundation model 展现出了前所未有的性能，在各个领域中越来越受到关注。其中，基于自监督学习的模型编码了大量人类总结的知识与常识，使得模型具备了超强的表征能力和泛化能力。在此背景下，知识驱动的多模态表征学习逐渐成为了研究的热点，如何进一步向广义基础模型进行医疗专业知识的注入仍然存在巨大挑战。本次报告中，我将介绍我们组近期关于医学知识增强的多模态基础模型的相关研究。从数据、模型和下游任务三个角度展开。其中包括：大规模医疗图文数据集的构建，医疗语言基础模型构建，多模态基础模型的训练，通用分割模型训练，以及知识增强的疾病诊断与临床应用。

讲者简介：谢伟迪，上海交通大学长聘轨副教授，上海人工智能实验室青年科学家。国家级青年人才，上海市（海外）高层次人才计划获得者，科技部科技创新 2030 — “新一代人工智能”重大项目青年项目负责人。博士毕业于牛津大学视觉几何组，Google-DeepMind 全额奖学金获得者，主要研究领域为 Computer Vision，AI4Medicine，共发表论文超 50 篇，包括 CVPR，ICCV, NeurIPS, ICML, Nature 子刊等，Google Scholar 累计引用超 9500 次，多次获得国际顶级会议研讨会的最佳论文奖和最佳海报奖，最佳期刊论文奖；担任计算机视觉和人工智能领域的旗舰会议 CVPR ，NeurIPS，ECCV 的领域主席。

王国泰 电子科技大学

报告题目：基于医学图像大模型的弱标注和无标注学习

报告摘要：报告人将分别介绍预训练大模型在三维医学影像分割和病理图像分类中的应用。首先，针对三维医学图像分割任务，我们开发了一个大规模自监督预训练模型，提出一种新的基于图像区域融合系数预测的自监督学习方法，在一个 11 万量级的无标注三维 CT

图像数据集上进行预训练后，在多种下游任务如 CT 和 MRI 图像中多种正常器官和病灶的分割中均展现出明显的性能提升。其次，将分享利用 CLIP 等大模型实现无标注病理图像分类。由于直接用大模型在下游任务上预测的性能有限，我们利用预测不确定性和特征相似性筛选高质量伪标签，在无需人工标注的情况下，将该模型有效迁移到下游分类任务中。

讲者简介：王国泰，电子科技大学教授。2011 年取得上海交通大学生物医学工程和智能科学与技术双学位，2018 年取得伦敦大学学院（UCL）博士学位，2023 年入选国家级青年人才计划。主要从事医学图像计算、人工智能与计算机视觉方面的研究,在领域顶级期刊及会议 IEEE TPAMI、IEEE TMI、Medical Image Analysis, NeuroImage、IPMI、MICCAI、AAAI 等发表高水平论文 60 余篇，谷歌学术引用量 9000 余次，6 篇论文入选 ESI 高被引，入选斯坦福大学发布的《年度科学影响力排行榜》全球前 2%。指导学生多次在 MICCAI 医学图像分割挑战赛中获得国际冠军、亚军。多次担任担任 MICCAI 和 ISBI Area Chair，近年担任 Medical Physics 期刊 Associate Editor、Medical Image Analysis 期刊 Guest Editor。

史淼晶 同济大学

报告题目：深度大模型的兴起以其在医疗场景的应用

报告摘要：本次报告我将首先介绍大模型的兴起，从自监督式大模型，到大语言模型，以及多模态大模型，穿插其中我将介绍团队将其应用到医疗场景的最新工作，包括生理信号估计，手术器械分割，病症检测识别，诊断报告生成等，相关研究聚焦大模型提示词，适配器，预训练以及多模态等关键技术，通过一系列具体案例来展示大模型向医疗场景垂直迁移时所面临的挑战与解决之策。

讲者简介：史淼晶，同济大学电子与信息工程学院教授、伦敦国王学院客座教授，国家高层次人才。博士毕业于北京大学，历任法国国家信息与自动化研究院研究员，英国伦敦国王学院信息系助理教授，副教授。主要研究计算机视觉及其在医学图像、遥感图像的应用。先后主持中国自然科学基金项目，英国工程与自然科学研究理事会项目，欧洲研究理事会地平线项目等多项国家级项目课题。

张晓凡 上海交通大学

报告题目：医疗领域大语言模型的训练及智能体应用

报告摘要：结合医学知识与海量数据的医疗领域大语言模型能为医疗领域提供一个高效、可靠、易用的智能辅助工具。我们开发的大语言模型 PULSE 与 2023 年在 OpenMEDLab 发布，它由包含继续预训练、指令微调、奖励模型训练、工具使用在内的四类数据集，共计 270 亿 token 的训练语料训练而成。在此基础之上，我们引入外部知识库、长程记忆、搜索计算器等各类工具，构造医学智能体，并积极探索以语言模型为核心的多模态大模型的开发，为复杂诊疗任务的完成提供了可能。

讲者简介：张晓凡现任上海交通大学电院清源研究院长聘教轨副教授，博士生导师。北京航空航天大学学士，美国北卡罗莱纳大学夏洛特分校博士学位。曾任京东硅谷研究院计算机视觉研究员、商汤科技北美智慧医疗实验室高级研究员。2021 年 6 月加入上海交通大学清源研究院。研究领域：医学图像、语言大模型，多模态决策。在国际顶级期刊及会议如 MedIA、TMI、CVPR、ACL、MICCAI 等发表多篇论文。Google Scholar 近五年引用千余次。申报美国专利 3 项。

雷柏英 深圳大学

报告题目：大模型赋能 AD 智能诊断

报告摘要：人工智能技术在人体健康诊断方面取得了显著的进展，尤其是基于深度学习的大模型（Large-scale Model），如 BERT、GPT 等，展现出了强大的数据驱动和自我学习能力。针对阿尔茨海默病（AD）智能诊断难以满足临床需求的困境，基于大模型的新型研究思路有望促进该研究的发展。大模型主要基于 Transformer 网络架构，此外涉及脑网络构建、早期生物标志物发掘、多模态多组学数据融合、多中心学习等多方面的研究基础。因此我们将对该研究思路的相关研究基础进行详细介绍，并对大模型在 AD 智能诊断中的应用进行展望。

讲者简介：雷柏英，国家级青年人才入选者，深圳大学特聘教授，西安电子科技大学客座教授，博士生导师，深圳市海外高层次人才（孔雀计划）、获新加坡南洋理工大学博士学位。主要研究方向为医学图像处理和人工智能。在 IEEE TMI 和 Medical Image Analysis 以第一/通讯作者（含共同)发表 SCI 论文 100 余篇(6 篇 ESI 高被引，1 篇热点论文)。谷歌学术总引用超 9500 次，H 指数 48。主持国家自然科学基金联合基金重点 1 项，面上 2 项等项目 20 余项（含国家级 7 项）。现任 IEEE TNNLS、TCYB、TMI、JBHI 等 SCI 期刊编委。IEEE BISP、BIIP、BSP 等技术委员会委员，医学图像顶级学术会议 MICCAI 领域主席（2021-2023）。IEEE 广州分部 WIE 主席，人工智能 A 类会议 AAAI、IJCAI 程序委员会委员，入选美国斯坦福大学发布的“全球前 2%顶尖科学家”（2020-2023），获 “强国青年科学家”提名（2022，全国共 40 人），CSIG 石青云女科学家奖（2022）

陈浩 香港科技大学

个人简介：陈浩，香港科技大学计算机科学与工程系和化学与生物工程系助理教授，研究兴趣包括人工智能，医疗图像分析，深度学习等。他领导的人工智能医疗实验室（Smart Lab），专注于可信赖人工智能技术在医疗领域的前沿研究与转化应用。陈博士于 2017 年获得香港中文大学博士学位。在 MICCAI、IEEE-TMI、MIA、CVPR、ICCV、AAAI、IJCAI、Radiology、Lancet Digital Health、Nature Machine Intelligence、JAMA 等顶级期刊和会议发表论文 100 余篇（谷歌学术引用次数达到 24400 余次，h-index 63），入选 2022 年和 2023 年斯坦福大学全球排名前 2%科学家名单。此外，陈博士还具有丰富的工业研究和产业转化经验，拥有二十余项人工智能和图像分析方面专利。 曾获得 2023 年亚洲青年科学家、国家教育部高等学校科学研究优秀成果二等奖、北京市科技进步一等奖、世界人工智能大会最高奖

项卓越人工智能引领者(SAIL)奖 、2019 年人工智能医学影像顶级会议 MICCAI 青年科学家影响力奖、Elsevier-MICCAI 最佳论文奖、医学影像与增强现实会议最佳论文奖、福布斯中国 30 岁以下 30 位精英、香港资讯及通讯科技银奖等奖项，担任包括 IEEE TNNLS、J-BHI、CMIG 和 Medical Physics 等期刊编委，担任 ACM Multimedia、MICCAI 2021-2023、ISBI 2022、MIDL 2022-2024、CVPR 2024 等多个人工智能与医学影像分析国际会议的领域主席和程序委员，曾带领团队获得 15 余项国际医学图像分析的挑战赛冠军。

夏勇 西北工业大学

个人简介：夏勇，西北工业大学计算机学院长聘教授、博导、空天地海一体化大数据应用技术国家工程实验室成员。研究方向为医学影像智能计算，近 5 年在 JAMA Network Open、Radiology、IEEE-TPAMI/TMI/TIP、MedIA、NeurIPS、CVPR、ECCV、MICCAI、IJCAI 发表论文 70 余篇，先后在 BraTS2020、KiTS21、KiPA22、SegRap2023 等 10 余项国际学科竞赛中获得前三名；担任中国体视学学会理事、中国计算机学院数字医学分会常委、中国图象图形学学会视觉大数据专委会常委等。

张少霆 上海人工智能实验室

个人简介：张少霆博士现担任上海人工智能实验室智慧医疗中心主任及领军科学家、商汤科技集团副总裁及研究院执行院长。其本硕博分别毕业于浙江大学、上海交通大学、美国罗格斯大学，此后于美国北卡罗莱纳大学夏洛特分校计算机系担任教职至终身副教授。他的研究课题得到了包括多项 NSF 在内的数百万美元的资助，其论文成果多次获得领域内顶级会议的青年科学家奖和最佳论文奖、美国橡树岭大学联合会青年教授奖等。他在《柳叶刀 数字健康》、《自然 机器智能》、《自然 通讯》等顶级期刊上以第一作者或通讯作者发表文章数十篇，总引用 1.7 万次，H-Index 60，并入选美国斯坦福大学发布的全球前 2%顶尖科学家“终身科学影响力排行榜”。应邀将担任医学图像分析顶会 IPMI’25 及计算机视觉顶会 CVPR’26 的程序委员会主席。

Workshop 6: 视觉智能算法防护与模型安全

组织者：彭春蕾（西安电子科技大学）、郭宗辉（中科院计算所）、张健（北京大学）、董胤蓬（清华大学）

时间：5 月 6 日（周一）8:30-12:00　　地点：智悦厅

时间	主持人	内容
8:30-8:35	张健	Workshop 开场
8:35-9:05	张健	讲者：俞能海（中国科学技术大学） 题目：人工智能生成数据水印技术
9:05-9:35	彭春蕾	讲者：张新鹏（复旦大学） 题目：从深度模型确权到 AIGC 溯源——数字水印新进展
9:35-10:05	彭春蕾	讲者：李斌（深圳大学） 题目：多媒体取证新场景与新进展
10:05-10:15	彭春蕾	企业宣讲：腾讯优图 生物特征识别与安全最新进展（丁守鸿）
10:15-10:25	中场休息	
10:25-10:55	董胤蓬	讲者：吴保元（香港中文大学深圳） 题目：后门学习发展现状及最新进展
10:55-11:25	董胤蓬	讲者：王奕森（北京大学） 题目：大模型下的越狱与防御
11:25-11:55	郭宗辉	讲者：张杰（中国科学院计算技术研究所） 题目：面向视觉模型的对抗攻击与防御方法
11:55-12:00	郭宗辉	Workshop 总结

俞能海 中国科学技术大学

报告题目：人工智能生成数据水印技术

报告摘要：以 ChatGPT，Stable Diffusion 为代表的生成式人工智能，凭借其惊艳的生成能力，在全世界掀起了一波人工智能领域的技术巨浪。然而生成式人工智能技术快速发展也带来了包括虚假信息传播、知识产权窃取等安全问题，亟需对生成数据嵌入标识，实现有效检测。本报告将从数字水印的角度介绍人工智能生成数据的主动标识方法，包括数据生成过程水印和生成数据处理水印，并分享我们团队关于生成数据性能无损水印方面的尝试和思考。

讲者简介：俞能海，中国科大网络空间安全学院执行院长，讲席教授，博士生导师，教育部高等学校网络空间安全专业教指委副主任委员，中国图象图形学学会副理事长，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。主要研究领域为视觉智能与安全、信息隐藏与数据安全等。发表高水平学术论文 300 余篇，获得国家网络安全先进个人，国家网络安全优秀教师，宝钢全国优秀教师奖，军队科技进步一等奖，安徽省自然科学一等奖，省教学成果特等奖，安徽省教学名师等。

张新鹏 复旦大学

报告题目：从深度模型确权到 AIGC 溯源——数字水印新进展

报告摘要：随着 AI 技术的不断进步，大模型在不同领域展现出巨大的潜力，但同时也面临各种风险问题。一方面，大模型昂贵的开发成本，对模型非法窃取与传播将严重损害模型持有者的正当权益。另一方面，利用 AIGC 技术生成逼真的虚假照片、视频和声音等信息信息可能被别有用心者用于操纵公众舆论、捏造虚假宣传、抹黑商业对手、诋毁个人名誉甚至实施精准欺诈。模型水印技术作为一种主动防御手段，近年来被广泛应用于深度模型的版权保护和生成内容溯源。本报告将介绍深度学习模型水印的发展历程，探讨判别式模型水印和生成式模型水印的核心问题和面临的挑战，展望 AIGC 水印发未来发展。

讲者简介：张新鹏，国家杰出青年科学基金获得者，二级教授。入选上海市东方英才计划领军项目、上海市优秀学术带头人、上海市曙光人才计划、上海市“东方学者”跟踪计划、上海市浦江人才计划、上海市“青年科技启明星”跟踪计划。曾赴美国纽约州立大学宾汉顿分校访问一年，受德国洪堡基金会资助作为资深研究员赴德国康斯坦茨大学访问 14 个月。主持国家自然科学基金重点项目、国家重点研发计划项目、国家 863 计划等科研项目 40 余项。发表论文 400 余篇，被引 18000 余次，2014 年—2023 年连续十年入选“爱思唯尔”中国高被引学者榜单，2020 年入选“科睿唯安”全球高被引科学家。申请发明专利 40 余项，授权近 30 项。获上海市自然科学奖一等奖、安徽省自然科学奖一等奖、国家级教学成果二等奖。担任 IEEE Trans. on Information Forensics and Security （IEEE T-IFS）等国际学术期刊的 Associate Editor、ACM IH&MMSec 和 IEEE WIFS 等国际学术会议的主席。

李斌 深圳大学

报告题目：多媒体取证新场景与新进展

报告摘要：多媒体虚假内容屡见不鲜、伪造方法层出不穷，容易造成经济损失和负面社会影响。多媒体鉴伪取证技术应运而生，在政务、电商、传媒等众多领域有十分旺盛的需求。随着信息处理和人工智能技术的迅猛发展，产生了一些新的多媒体取证需求。本报告将分享我们团队通过利用机器学习方法在多媒体来源辨识和处理历史取证等方面的探索尝试，讨论所面临的挑战，展望大模型时代多媒体取证的发展。

讲者简介：李斌，深圳大学教授，博导，广东省智能信息处理重点实验室主任、深圳市媒体信息内容安全重点实验室主任、深大-软牛人工智能联合创新中心主任，广东省自然科学基金杰出青年基金获得者，深圳市高层次专业人才，深圳市优秀教师，IEEE/CCF/CSIG 高级会员，连续多年入选斯坦福大学全球前2%科学家榜单。长期从事多媒体信息取证与安全的理论与方法研究，近年在视听媒体取证、图像隐写、深度伪造攻防等方面取得成果进展。发表包括国际权威刊物 IEEE T-IFS、IEEE T-IP、IEEE T-CSVT、CVPR、IJCAI 等 SCI/EI 论文 150 多篇，谷歌学术引用 6400 余次，参与制定多媒体取证行标/团标 3 件，指导学生获最佳论文奖或学科竞赛奖 10 余项。主持国家自然科学基金联合基金重点项目、深圳市基础研究重点项目、企业委托鉴伪取证项目等，组织承办阿里安全伪造图像对抗攻击竞赛。获中国计算机学会科学技术奖自然科学一等奖、深圳市自然科学二等奖、广东省计算机学会自然科学二等奖等。担任广东省图象图形学学会副理事长、IEEE WIFS 2024 共同主席、IEEE ICIP/ACM IH&MMSEC/ChinaMFS 程序委员会主席、领域主席、宣传主席、IEEE Transactions on Information Forensics and Security AE/SAE 等学术职务

吴保元 香港中文大学（深圳）

报告题目：后门学习发展现状及最新进展

报告摘要：本次讲座将从后门攻击和防御两个方面介绍后门学习的发展现状与最新进展。在后门攻击方面，将对现有后门攻击方法进行结构化的梳理，厘清该领域的发展历程、现状和趋势，然后介绍我们在后门攻击、防御问题上的最新成果，包括后门学习与大模型的交叉研究。最后，还将介绍我们最新创建的后门学习基准平台 BackdoorBench，集成了大量的主流后门攻防方法，并提供了全面的评测和分析，相关信息详见 https://backdoorbench.github.io/。

讲者简介：吴保元博士现任香港中文大学（深圳）数据科学学院终身副教授、深圳市模式分析与感知计算重点实验室（筹）副主任、龙岗区智能数字经济安全重点实验室主任、腾讯 AI Lab 智能决策组顾问。其研究方向包括可信人工智能、机器学习和计算机视觉，在人工智能的顶级期刊和会议上发表论文 80 多篇，并曾入选人工智能顶级会议 CVPR 2019 最佳论文候选名单。其担任人工智能领域国际期刊 Neurocomputing 编委、第五届中国模式识别与计算机视觉大会 PRCV 2022 组委会主席、国际会议 CVPR 2024、NeurIPS 2022/2023/2024、NeurIPS Datasets and Benchmarks Track 2023/2024、ICLR 2022/2023/2024、ICML 2023/2024、AAAI 2022/2024 领域主席、中国自动化学会模式识别与机器智能专委会副秘书长，入选斯坦福大学“全球前 2%顶尖科学家”2021、2022 年度榜单。作为项目负责人承担广东省自然科学基金杰出青年项目 1 项，科技部重点研发计划重点专项课题 1 项，国家自然科学基金面上项目 1 项， 深圳市

优秀科技创新人才优秀青年基础研究项目 1 项，CCF-腾讯犀牛鸟基金 1 项，CCF-海康威视斑头雁基金 1 项，CAAI-华为 MindSpore 学术奖励基金 1 项，腾讯犀牛鸟研究专项基金 2 项。

王奕森 北京大学

报告题目：大模型下的越狱与防御

报告摘要：近期大模型在多种任务上取得了巨大的成功，但其也面临着新的安全性挑战：攻击者可以通过设计恶意的提示词（prompt）诱导大模型越狱（jailbreak）以生成不当内容，如有害信息或仇恨言论等。本报告将探讨大模型特有的上下文学习能力（In-context Learning）在操纵大模型对齐能力方面的表现，并提出一系列攻击和防御方法，期望为大模型安全和对齐提供新视角。

讲者简介：王奕森，北京大学助理教授，博士生导师。主要研究方向为机器学习理论和算法，目前重点关注大模型的理论、安全等。已发表机器学习三大顶会 ICML/NeurIPS/ICLR 文章 50 余篇，多篇被选为 Oral 或 Spotlight，获 ECML 2021 最佳机器学习论文奖、ICML 2021 Workshop 最佳论文银奖、CVPR 2021 竞赛第一等，研究成果被麻省理工科技评论（MIT Technology Review）和中国中央广播电视总台（CCTV）专题报道。主持基金委“下一代人工智能”重大研究计划项目、科技创新 2030“新一代人工智能”重大项目课题。担任 NeurIPS 2024 Senior Area Chair 等。

张杰 中国科学院计算技术研究所

报告题目：面向视觉模型的对抗攻击与防御方法

报告摘要：深度学习模型在计算机视觉任务中展现出了卓越的性能。然而，随着其应用的广泛扩散，深度学习系统面临的安全威胁也日益增加，其中最为严重的威胁之一是对抗攻击，即模型很容易受到对抗噪声的影响，输出错误的结果。本报告首先介绍对抗攻击的基本原理与常见类型，以及对抗防御的基本策略与方法；然后介绍我们在对抗攻防方面的最新研究进展，包括自适应扰动的对抗攻击方法、面向大模型的对抗防御方法以及高效的对抗训练方法等。

讲者简介：张杰，中国科学院计算技术研究所副研究员、硕导，主要从事算法攻防、人脸识别、图像分割等基础研究和应用研究。主持中国科学院先导 B 课题、科技部青年科学家课题、基金委面上、青基等项目，在知名期刊和会议 TPAMI、TIP、TIFS、ICCV、CVPR 等上发表论文 20 余篇，授权专利 6 项，获国际竞赛 1 次冠军，4 次亚军；部分成果在华为多款手机、平安多款人脸识别产品、高铁动车组故障监测、工业视觉检测设备中得到规模化应用，连续两年荣获华为优秀合作奖。先后入选了北京市科技新星计划、中科院青促会和 MSRA“铸星计划”。

彭春蕾 西安电子科技大学

组织者简介：彭春蕾，西安电子科技大学副教授，主要从事视觉身份可信识别方面的理论和应用研究，在国内外权威期刊会议上发表论文40 余篇。主持国家自然科学基金面上项目、重点项目课题负责人、CCF-腾讯犀牛鸟基金等。担任国际期刊 Visual Computer 编委。入选中国科协青年人才托举工程。获吴文俊人工智能优秀青年奖、湖北省技术发明一等奖、中国图象图形学学会自然科学二等奖、陕西省优博、中国图象图形学学会优博、ACM 西安优博奖、ACMMM 专题论坛最佳论文奖。指导研究生学位论文入选中国电子学会硕士学位论文激励计划。

郭宗辉 中科院计算所

组织者简介：中国科学院计算技术研究所智能信息处理重点实验室博士后，主要从事图像合成与伪造检测方向研究，以第一作者在 IEEE TPAMI、CVPR、ICCV 等期刊和会议上发表学术论文 7 篇，申请发明专利 2 项，获中国海洋大学研究生卓越奖学金，主持国家自然科学基金面上项目，任 VALSE EACC，VALSE 年度会议注册主席 (2019 至今)。

张健 北京大学

组织者简介：张健，博士，北京大学深圳研究生院助理教授/研究员、博士生导师，深圳市海外高层次人才。哈尔滨工业大学数学与应用数学理学学士（2007）、计算机科学与技术工学硕士（2009）及计算机应用工学博士（2014），并分别在北京大学、香港科技大学和沙特国王科技大学做博士后访问研究。主要从事图像处理与计算机视觉方面的基础理论和应用研究，在在 SPM/TPAMI/TIP/CVPR/ECCV/ICCV 等高水平国际期刊会议上发表论文 90 余篇，Google Scholar 引用超过 3700 次，SCI 他引超过 1700 次（其中单篇第一作者期刊论文 Google Scholar 最高引用超过 600 次;单篇第一作者会议论文 Google Scholar 最高引用超过 500 次）。相关研究成果申请中国专利 10 余项。获得 IEEE 视觉通讯与图像处理（VCIP）国际会议 2011 年度最佳论文奖以及该国际会议 2015 年度最佳学生论文奖、IEEE 多媒体 IEEE MultiMedia 国际期刊 2018 年度最佳论文奖、中国多媒体大会 ChinaMM 2021 年度最佳论文奖、深圳市科学技术协会 2021 年优秀自然科学学术论文奖（共 16 篇，信息领域唯一 1 篇）、深圳市海外高层次人才、2020 及 2021 连续两年入选人工智能与图像处理领域全球前 2%顶尖科学家“年度影响力”榜单。

董胤蓬 清华大学

组织者简介：董胤蓬，清华大学计算机系博士后研究员，导师为朱军教授。于 2017 年和 2022 年在清华大学计算机系分别获得学士和博士学位。主要研究方向为机器学习与计算机视觉，聚焦深度学习鲁棒性的研究，先后发表 TPAMI、IJCV、CVPR、NeurIPS 等顶级国际期刊和会议论文四十余篇。担任 VALSE 执行 AC，担任 TPAMI、IJCV、ICML、NeurIPS、CVPR 等期刊和会议审稿人。曾在 ICML2021、AAAI2022 等国际会议上组织了对抗机器学习相关研讨会。曾获得 CCF 优秀博士学位论文激励计划、CCF-CV 学术新锐奖、微软奖学金、百度奖学金、字节奖学金、清华大学“水木学者”计划、博新计划等。

Workshop 7：遥感图像智能解译

组织者：赵文达（大连理工大学）、杨曦（西安电子科技大学）、洪丹枫（中国科学院空天信息创新研究院）、王海鹏（海军航空大学）

时间：5 月 7 日（周二）13:30-18:00　　地点：温德姆宴会厅

时间	主持人	内容
14:00-14:35	赵文达	讲者：方乐缘（湖南大学） 题目：开放环境下高光谱图像弱监督解译
14:35-15:10	赵文达	讲者：程塨（西北工业大学） 题目：光学图像小目标检测
15:10-15:45	赵文达	讲者：曹相湧（西安交通大学） 题目：生成式遥感大模型及其应用
15:45-15:55	中场休息	
15:55-16:30	洪丹峰	讲者：邓良剑（电子科技大学） 题目：深度嵌入式变分优化方法及其在遥感全色锐化的应用
16:30-17:05	洪丹峰	讲者：谢卫莹（西安电子科技大学） 题目：多模态遥感图像分布式高效协同解译
17:05-17:35	洪丹峰	Panel 嘉宾： 方乐缘（湖南大学）、程塨（西北工业大学）、曹相湧（西安交通大学）、邓良剑（电子科技大学）、谢卫莹（西安电子科技大学）、赵文达（大连理工大学）、杨曦（西安电子科技大学）、王海鹏（海军航空大学）

方乐缘 湖南大学

报告题目：开放环境下高光谱图像弱监督解译

报告摘要：高光谱遥感技术作为非接触式的“星-空-地高光谱观测平台”,使得高光谱遥感技术在智慧城市、智慧农业、环境监测、工业检测以及军事侦察等领域的应用不断得到拓展。随着国家高分专项任务以及民用高光谱遥感卫星的快速发展，更加开放复杂的应用场景为高光谱遥感数据的解译任务带来了诸多挑战。本报告围绕开放环境下高光谱图像的弱监督解译，重点介绍本团队在监督信息不完全分布不一致以及模态不确切等受限环境下的高光谱解译研究进展，并展望了开放环境下高光谱遥感数据的智能化解译任务的未来发展。

讲者简介：方乐缘，湖南大学岳麓学者特聘教授，国家优青，科睿唯安（Clarivate Analytics）全球“高被引科学家”、爱思唯尔中国高被引学者，湖南省创新领军人才。获得国家自然科学二等奖 1 项、IEEE GRSS 最高影响力论文奖、湖南省自然科学一等奖 2 项等奖项。担任 SCI 期刊 IEEE Transactions on Image Processing、IEEE Transactions on Neural Networks and Learning System、IEEE Transactions on Geoscience and Remote Sensing、Neurocomputing 等期刊编委。现主要从事深度学习、弱监督学习以及在遥感图像处理与分析等方面的研究。研究成果在国际权威期刊和会议发表论文 160 余篇，其中 SCI 期刊发表论文 100 余篇（IEEE TPAMI、IJCV、TIP 等本领域顶级期刊论文 70 余篇），国际权威会议论文 30 篇，Google scholar 引用 14000 余次，ESI 高被引 22 篇，ESI 热点论文 4 篇。主持国家自然科学基金联合重点、国家重点研发课题等项目。

程塨 西北工业大学

报告题目：光学图像小目标检测

报告摘要：小目标检测旨在利用视觉算法对图像和视频中的小尺寸目标进行识别及定位，在公共安全、工业自动化、智慧交通和军事国防等领域应用广泛。本次报告中，我们首先从研究背景切入，阐述光学图像小目标检测任务面临的三大挑战。接着，从数据层面和算法层面分别对领域内的发展现状进行回顾。随后详细介绍我们推出的针对光学图像小目标检测的大规模数据集 SODA，并对其数据特性以及相关基准方法的性能进行阐述，同时介绍我们针对小目标检测任务所设计的算法。最后，对小目标检测领域的发展趋势进行展望。**讲者简介：**程塨，西北工业大学教授，国家级青年人才，连续 4 年入选科睿唯安“全球高被引科学家”和爱思唯尔“中国高被引学者”，获得中国图象图形学学会青年科学家奖。主要研究方向为光学遥感图像理解、计算机视觉等。以第一作者/通讯作者发表论文 60 余篇，包括 Proceedings of the IEEE、TPAMI、IJCV、CVPR、ICCV 等，谷歌总引用 1.8 万余次，5 篇论文单篇引用大于 1000 次，获得 2021 年度 IEEE TCSVT 最佳论文奖、2021 年度和 2023 年度 IEEE 地球科学与遥感学会最高影响力论文奖（IEEE GRSS Highest Impact Paper Award）、中国百篇最具影响国际学术论文等学术奖励，获得中国电子学会自然科学一等奖、陕西省科学技术一等奖等省部级科技奖励，担任 IEEE GRSM、IEEE TGRS、JRS 等多个国际期刊编委。

曹相湧 西安交通大学

报告题目：生成式遥感大模型及其应用

报告摘要：生成式模型是目前大模型时代的热门研究领域，其中代表性的技术是扩散模型。本报告主要介绍我们团队将扩散模型应用到遥感图像智能解译领域的初步探索，具体内容包括：1.扩散模型在遥感图像底层处理任务（如去噪、超分、修复、融合）的应用； 2.基于扩散模型的高光谱图像生成； 3.多条件可控生成式遥感大模型。

讲者简介：曹相湧，西安交通大学计算机学院副教授，博士生导师，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。已在 IJCV, SIAM J. Image Sciences, TIP, TGRS, CVPR, ICCV, ECCV 等期刊和会议发表论文 50 余篇，多篇论文入选 ESI 高被引，主持国家自然科学基金面上、青年项目、参与重大、重点、军工等项目，研究成果在大唐华银、平高集团、比亚迪等公司落地，担任 PRCV 2023 领域主席，Valse 第七届执委会委员，中国图像图形学会遥感专委会委员。

邓良剑 电子科技大学

报告题目：深度嵌入式变分优化方法及其在遥感全色锐化的应用

报告摘要：本报告主要探讨如何将深入学习先验嵌入到传统变分优化模型框架，从而有效提升全色锐化方法的精度、泛化性和解释性。报告内容主要包括两个方面：1）提出嵌入深度先验的深度先验融合项，通过手动设计的自适应连接矩阵搭建起传统变分优化模型与新兴深度学习方法的桥梁，最终提出 VO+Net 方法框架；2）通过构建全色图像与多光谱图像之间的隐式保真关系，提出可学习的非线性映射保真项，并嵌入传统变分优化模型框架，有效克服传统全色图像与多光谱图像线性表示关系引起的精度问题。

讲者简介：邓良剑，电子科技大学数学科学学院研究员。长期从事应用数学和人工智能领域的交叉研究，主要研究方向为：机器学习、变分图像处理、图像融合等。分别于 2010 年和 2016 年获得电子科技大学理学学士和理学博士学位。曾作为联合培养博士生在美国凯斯西储大学学习一年，赴香港浸会大学进行博士后研究工作一年。曾在对方资助下短期访问剑桥大学牛顿数学科学研究所、香港浸会大学。主持国家自然科学基金项目 2 项、省部级项目 1 项，作为研究骨干参与国家重点研发等国家级项目多项。近五年，以第一或通讯作者身份在 SIAM J. Imag. Sci.、Int. J. Comput. Vis.、Info. Fusion、IEEE 汇刊发表期刊论文 20 余篇，在 CCF-A 类学术会议发表论文 10 余篇。曾获四川省数学会应用数学一等奖(独立)等奖励。担任多个 SCI 期刊编委/客座编辑。

谢卫莹 西安电子科技大学

报告题目：多模态遥感图像分布式高效协同解译

报告摘要：多模态遥感图像分布式协同解译领域的发展仍面临特征计算不完备、通信能力受限等挑战。本报告从通信与计算联合优化的角度出发，着眼于最大程度地实现多源数据的一致融合，并探讨在有限资源约束下，通过通信内容压缩策略解决分布式边缘节点的通信瓶颈问题，以促进多模态遥感图像分布式协同解译的发展。

讲者简介：谢卫莹，教授、博士生导师，国家优秀青年科学基金获得者，IEEE Senior Member。主持国家自然科学基金、科技委领域基金、ZF

项目、博士后特别资助等多项项目。以第一/通讯作者身份发表 IEEE Trans.系列中科院一区 TOP 期刊及 CCF A 类会议论文 50 篇，其中 ESI 高被引论文 7 篇，热点论文 1 篇，h 指数为 31。获陕西省自然科学优秀学术论文奖、“天智杯”人工智能挑战赛全国冠军及 100 万项目奖励、“互联网＋”大赛全国金奖、“强芯健魂，铸基智能”计算平台挑战赛全国二等奖及 369 万项目奖励。以第一发明人获授权国家发明专利 10 件已完成转化应用。入选全球前 2%顶尖科学家榜单、全国优秀创新创业导师、中国科协青年人才托举工程、中国科协优秀中外青年交流计划。

赵文达 大连理工大学

组织者简介：赵文达，大连理工大学副教授，博导。研究方向包括多模态图像分析，如图像融合、分类、目标检测等。在包括 IEEE TPAMI，IEEE TIP，IEEE TNNLS，IEEE TMM, IEEE TGRS，IEEE TCSVT 等本领域顶级期刊，以及 CVPR，ECCV，AAAI 等本领域顶级会议上发表学术论文 40 余篇。授权发明专利 10 余项。主持包括国家自然科学基金面上、青年项目，博士后特助、面上项目等 10 余项。获得 IEEE MMTC 2020 和 ISAIR 2018 最佳论文奖，山东省科技进步一等奖等。大连市“科技之星”，大连理工大学“星海骨干”计划入选者。担任中国人工智能学会智能融合专业委员会副秘书长，中国指挥与控制学会青年工作委员会委员，中国指挥与控制学会多域态势感知与认知专委会委员，以及视觉与学习青年学者研讨会议（VALSE）执行 AC。

杨曦 西安电子科技大学

组织者简介：杨曦，西安电子科技大学教授、博导。主要从事计算机视觉、机器学习、多媒体信息处理等领域的研究工作，主持装发预研项目、部队示范应用项目、国家自然科学基金面上项目、陕西省重点研发计划等课题。担任《The Visual Computer》等期刊编委、IEEE Senior Member、中国图象图形学学会青年工作委员会委员等。近年来在 IEEE T-IP、T-NNLS、CVPR、ICCV 等领域内重要期刊和学术会议上发表论文 80 余篇（第一/通讯作者论文 60 余篇，IEEE Trans.论文 40 余篇，ESI 高被引 4 篇），出版教材 2 部，授权专利 10 余项。获得陕西省杰青、中国科协青年人才托举工程、陕西省青年科技新星、陕西省优秀博士学位论文、中国图象图形学学会自然科学二等奖等荣誉。

洪丹枫 中国科学院空天信息创新研究院

组织者简介：洪丹枫，中国科学院空天信息创新研究院研究员，博士生导师，遥感与数字地球重点实验室副主任，国家优秀青年基金（海外）项目获得者，科睿唯安 2022/2023 年度全球“高被引科学家”。曾为德国宇航中心（DLR）的研究员兼“光谱视觉”课题组组长、法国傅立叶实验室（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。主要研究方向为人工智能、多模态遥感、基础大模型、高光谱成像、大尺度地学应用、智能感知等。研究成果在 IEEE TPAMI、IEEE TIP、RSE、ISPRS、CVPR、NeurIPS、

ECCV 等顶级期刊/会议发表 180 余篇， Google Scholar 引用 12000+，获得 IEEE GRSS 杰出青年奖（Early Career Award）、国际高光谱顶级会议 WHISPERS 杰出论文奖（Jose Bioucas Dias 奖）等，担任 IEEE TGRS、ISPRS JP&RS、Information Fusion 等期刊副主编/编委。

王海鹏 海军航空大学

组织者简介：王海鹏，海军航空大学信息融合研究所副所长、教授、博导。长期从事海洋态势感知与认知方向的技术研究与工程应用，主持国家自然科学基金、山东省重大专项、军委 XX3 等重点课题 10 余项，在海上弱小目标跟踪、目标融合识别、态势评估、博弈对抗等方面取得技术创新，为更好的解决复杂场景海上目标看不见、跟不上、辨不明、猜不透等卡脖子问题提供了重要技术支撑，相关成果已应用于国防和公共安全领域，在多次重大任务中发挥了重要作用。获山东省科技进步一等奖 2 项（均排名 1），授权国家发明专利 40 余项；出版著作 5 部，在 Information Fusion、IEEE Transactions on NNLS、TGRS 等发表论文 30 余篇。被评为青年长江学者、军队拔尖人才、山东省青年泰山学者等，享受国务院特殊津贴，获求是杰出青年奖、山东省青年科技奖，荣立二等功、三等功各 1 次。兼任教育部“电子信息类”教学指导委员会副秘书长、中国指挥与控制学会青工委副主任委员、多域态势感知与智能认知专委会副主任委员、中国人工智能学会智能融合专委会常委等。

Workshop 8: 海洋环境智能探测与理解

组织者：丛润民（山东大学）、郑海永（中国海洋大学）、李华（海南大学）
时间：5 月 6 日（周一）14:00-18:00　　地点：博悦厅

时间	主持人	内容
14:00-14:35	丛润民	讲者：付先平（大连海事大学） 题目：水下机器人多模式智能全景感知系统及其在海洋牧场中的应用
14:35-15:10	郑海永	讲者：李胜全（鹏城实验室） 题目：水声轨道角动量通信探测技术与开源开放科研平台
15:10-15:45	李华	讲者：蒲华燕（重庆大学） 题目：复杂场景智能无人系统感知技术研究进展
15:45-15:55	中场休息	
15:55-16:30	李华	讲者：孙哲（西北工业大学） 题目：水下计算光学成像技术
16:30-17:05	郑海永	讲者：刘日升（大连理工大学） 题目：优化驱动学习的恶劣场景视觉感知
17:05-17:40	丛润民	Panel 嘉宾： 付先平（大连海事大学）、李胜全（鹏城实验室）、蒲华燕（重庆大学）、孙哲（西北工业大学）、刘日升（大连理工大学）

付先平 大连海事大学

报告题目：水下机器人多模式智能全景感知系统及其在海洋牧场中的应用

报告摘要：我国发展海洋经济和国防安全在领海和毗邻区均具有巨大的水下探测需求。结合水下机器人、无人船等多智能体，汇报近浅海水下复杂环境下光学全景清晰成像及大面积海域水下小目标识别技术；针对水下环境数据特征提出水下多模态数据对齐方法、构建水下环境感知专用大模型，提高水下机器人环境感知能力。介绍弱光信号的多域协同自适应增强传输方法，提升水下机器人移动作业场景中光信号数据传输的鲁棒性。介绍辽宁省水下机器人工程研究中心的科研成果及其在渤海和黄海的海洋牧场的应用情况。

讲者简介：付先平，现任大连海事大学信息科学技术学院院长，教授，博士生导师、计算机科学与技术学科带头人，辽宁省水下机器人工程研究中心主任、辽宁省高水平创新创业团队负责人。2005 年获大连海事大学博士学位，2007.12-2008.02 清华大学计算机系博士后，2008-2009 年期间在美国哈佛大学做博士后高级研究员，并获得美国 RPB(Research to Prevent Blindness)国际研究学者奖。主持国家重点研发计划政府间合作项目、工业和信息化部“新一代人工智能产业创新重点任务”—“水下捕捞机器人”重点项目。大连市领军人才、中国计算机学会智能机器专委常委、多媒体专委委员、人机交互专委委员，中国图象图形学学会多媒体专业委员会、计算机视觉专业委员会委员，大连计算机学会常务理事等职务。共发表学术论文 140 余篇，其中 CCF-A/B 类期刊与会议、中科院一、二区权威期刊和国内权威中文期刊论文 40 余篇。

李胜全 鹏城实验室

报告题目：水声轨道角动量通信探测技术与开源开放科研平台

报告摘要：水下目标成像和水声信息传输面临着水声信道带宽窄、成像分辨力不高、水声通信速率低、硬件系统复杂、成本昂贵等难题。针对这些问题，我们提出了采用水声轨道角动量（OAM）实现高通量三维成像的方法：首先描述涡旋声场声学成像的数学物理模型，继而根据理论研制了基于 OAM 的三维成像系统原理样机，最后介绍了原理样机湖试和海试的情况。提高水声通信速率技术主要有高阶调制技术、多输入多输出水声通信技术、同时同频全双工水声通信技术等高频谱利用率技术，但受限于通信带宽，上述方案能提高的频谱利用率有限，我们采用 OAM 通信方式，研究了高通量信号传输方法。此外，为解决水声行业门槛高问题，我们建设了端云协同的水下技术开放共享试验系统，通过标准化试验载体、声学传感设备以及通信定位等软硬件接口，在通信网络支持下为用户提供远程申请和访问功能，使研究人员能够远程在试验节点本体部署和更新算法，在线监测试验全过程工作状态，突破试验场地建设、试验载体研发等成本壁垒，推动各类水下试验数据大规模回收整理和开放共享。

讲者简介：李胜全，博士，鹏城实验室海洋信息通信研究所所长、研究员，博士生导师，中山大学博士生导师。长期从事海洋空间环境感知与信息系统工作，聚焦信号处理和水声电子的海洋地形地物测量识别装备，融合海洋环境信息和物标感知需求的水下定位、测绘导航、海洋调查技术，面向海洋物理属性、气象水文现象和 GIS 应用的海洋时空信息表达方法，在海洋导航、水下集群、水下精细探测方面取得显著成绩。获省部级技术发明一等奖、科技进步二等奖 4 项、三等奖 5 项。主持国家重点研发计划子课题 1 项、国家自然科学基金委重大仪器专项子课题 1 项、鹏城实验室重大攻关任务 2 项以及某科技委重大基础加强、预研、某装预研、某部重大专项等项目。

蒲华燕 重庆大学

报告题目：复杂场景智能无人系统感知技术研究进展

报告摘要：感知能力是智能无人系统理解周围环境，并与环境产生交互的基础，智能无人系统在复杂场景下自主作业能力的发展离不开感知能力的不断提升。本团队长期围绕复杂场景智能无人系统感知技术开展深入研究，突破了非结构化动态环境感知、陆地非合作目标感知、空中低小慢目标感知、海面弱观测目标感知和结构化工业场景特征感知关键技术，形成了以相机、激光雷达、红外、雷达等多传感器融合的感知技术体系，取得了系列化创新技术成果，应用于多类型复杂场景下多型智能无人系统的环境和目标感知任务，包括：应用于两栖无人车自主跨域非结构环境感知，应用于士兵等陆地非合作目标、无人机等空中低小慢目标和舰船等海绵弱观测目标智能识别，获包括国家科技进步奖二等奖在内的科技奖励多项。

讲者简介：蒲华燕，重庆大学高端装备机械传动全国重点实验室教授，国家杰青，全国青年岗位能手，上海市三八红旗手标兵。国家海洋环境与岛礁可持续发展重点专项总体组专家。长期从事复杂工况服役机器人及智能无人系统研究，研究成果两次获得国家技术发明奖二等奖（2019 年排 2、2016 年排 4），以第一完成人获得上海市技术发明奖一等奖、重庆科技进步奖一等奖、中国自动化学会科技进步奖一等奖。

孙哲 西北工业大学

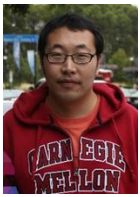
报告题目：水下计算光学成像技术

报告摘要：水下计算光学成像技术作为一种新型成像方法，近年来在水下探测、水下光学导引等领域展现出了广阔的应用前景。该技术利用计算机视觉与光学成像原理，通过算法优化与图像处理，实现对水下环境的精确感知与成像。报告介绍了高能量副载波调制激光成像雷达和水下智能计算光学成像技术，分析了其在水下复杂环境中的成像特点与优势，并展望了在水下光学导引领域的应用。

讲者简介：西北工业大学副教授，德国耶拿大学/亥姆霍兹研究所博士后，国家重点研发计划“深海和极地关键技术与装备”专项主任设计师，中国科学院西部青年学者。长期从事智能光学探测/成像领域研究，主持国家重点研发计划专项课题，中德博士后交流项目，德国联邦教育及研究部和德国科学基金会资助的德国精英集群计划，中国科学院西部之光，陕西省自然基金面上项目等。发表论文 40 多篇；申请/授权专利 15 项。

刘日升 大连理工大学

报告题目： 优化驱动学习的恶劣场景视觉感知



报告摘要：智能视觉感知是支撑无人系统完成各种复杂作业任务的关键技术。然而，在海洋环境等恶劣场景下，解决该问题面临“外在”和“内在”两个方面的耦合挑战。本报告重点汇报团队近期在恶劣场景视觉感知相关方向的研究进展，包括优化驱动学习基础理论，恶劣场景视觉感知关键技术，以及在水下机器人敏捷作业中的应用探索。

讲者简介： 刘日升，大连理工大学软件学院教授（破格），日本立命馆大学兼职教授。近年来在计算机视觉、优化方法、深度学习等领域发表 IEEE 汇刊及 CCF 推荐 A 类会议论文 100 余篇（含 10 篇一作 TPAMI）。获各级自然科学成果奖 4 项，CCF 推荐学术会议论文奖 7 篇。

丛润民 山东大学

组织者简介：丛润民，山东大学教授、博士生导师，入选中国科协“青年人才托举工程”、人社部“香江学者”计划、山东省“泰山学者”青年专家、CSIG 优博等。主要研究方向包括计算机视觉、模式识别、多媒体信息处理等。主持、参与了包括国家自然科学基金、国家重点研发计划、科技创新 2030 在内的多项科研项目。发表 CCF-A、IEEE/ACM Trans 论文 70 篇，其中 ESI 热点论文 2 篇、ESI 高被引论文 15 篇；担任担 CSIG 青年工作委员会常务副秘书长，IEEE TNNLS、Neurocomputing、IEEE Journal of Oceanic Engineering 等 SCI 期刊编委，荣获 CSIG 自然科学奖二等奖（第 1 完成人）、IEEE ICME 最佳学生论文奖亚军、ACM SIGWEB 中国新星奖等。VALSE SAC。

郑海永 中国海洋大学

组织者简介：郑海永，中国海洋大学教授、博士生导师。开展任务驱动混合课程教学改革，推行多元化全过程学生学习评价机制，主持山东省本科高校教学改革研究项目 1 项、山东省研究生教育质量提升计划建设项目 2 项，负责山东省一流本科专业和一流本科课程，发表教学论文 2 篇。聚焦智能信息感知与处理研究，在视觉合成、视频分析、水下视觉清晰化、海洋生物识别等方面取得系列成果，主持包括国家自然科学基金在内的科研项目 10 余项，发表包括计算机视觉顶级期刊 TPAMI、IJCV 和三大顶级会议 CVPR、ICCV、ECCV 以及多媒体顶级会议 ACM MM 在内的学术论文 80 余篇，授权国家发明专利 10 多项。担任 IEEE Journal of Oceanic Engineering 和 Intelligent Marine Technology and Systems 期刊编委，IEEE 海洋工程学协会水下光学与视觉技术委员会主席（2023-2024），中国图象图形学学会青年工作委员会委员，山东电子学会理事会理事，北太平洋海洋科学组织 PICES 工作组 WG-48 成员，VALSE 首批常务领域主席，IEEE 和 CSIG 高级会员。荣获生物信息学国际权威会议 InCoB 2017 最佳论文奖、山东省省级教学成果奖、山东省优秀硕士学位论文指导奖，以及中国海洋大学本科教学优秀奖、天泰优秀人才奖、东升课程教学卓越奖、五四青年奖等，入选山东省泰山学者青年专家。VALSE LAC。

李华 海南大学

组织者简介李华，海南大学计算机科学与技术学院副教授、博士生导师，海南省高层次 D 类人才，2023 年获第一批海南省“南海新星”科技创新人才。主要研究方向包括计算机视觉、海洋环境感知、水下图像处理、人工智能、深度学习等。主持/参与了包括国家自然科学基金、海南省自然科学基金等多项科研项目。近五年，在计算机视觉领域以第一或通讯作者在 ICCV、ACM MM、IEEE TMM、IEEE TII、IEEE ICME 等高水平国际期刊和会议上发表多篇 SCI/EI 文章。担任 IEEE TMM、IEEE TIE、IEEE JOE、ACM MM 等期刊和会议的审稿专家。2021 年、2023 年荣获 IEEE 杰出领导力奖。2022 年获 IEEE Journal of Oceanic Engineering 杰出审稿人奖。

Workshop 9: 视觉大模型高效迁移

组织者：左旺孟（哈尔滨工业大学）、戴玉超（西北工业大学）、王立君（大连理工大学）

时间：5 月 6 日（周一）8:30-12:00　　地点：欣悦厅

时间	主持人	内容
8:30-9:00	左旺孟	讲者：丁贵广（清华大学） 题目：深度学习模型架构设计及微调技术
9:00-9:30	左旺孟	讲者：程明明（南开大学） 题目：高效能个性化图像生成
9:30-10:00	戴玉超	讲者：吴建鑫（南京大学） 题目：资源高效的视觉大模型微调
10:00-10:1	戴玉超	企业宣讲：百度 挖掘无标注数据的潜力：半监督学习技术探索与应用（张伟）
10:10-10:20	中场休息	
10:20-10:50	戴玉超	讲者：王鑫龙（北京智源人工智能研究院） 题目：从视觉到多模态基础模型：探索与实践
10:50-11:20	王立君	讲者：高鹏（上海人工智能实验室） 题目：构建高效的多模态语言大模型
11:20-11:50	王立君	Panel 嘉宾： 丁贵广（清华大学）、程明明（南开大学）、吴建鑫（南京大学）、王鑫龙（北京智源实验室）、高鹏（上海人工智能实验室）、左旺孟（哈尔滨工业大学）、戴玉超（西北工业大学）、张磊（OPPO）、张璐（大连理工大学）

丁贵广 清华大学

报告题目：深度学习模型架构设计及微调技术

报告摘要：深度学习模型的端侧部署已经成为人工智能应用的主要形式之一，如智能手机、智能摄像头、自动驾驶等场景，然而深度学习模型复杂度高、参数量大，给端侧部署带来了巨大挑战，如何在不损失或较少损失模型精度的前提下，减小模型的计算复杂度是人工智能领域研究的重要方向之一。本次报告将介绍深度学习模型的压缩优化技术，介绍“训大用小” 的深度学习训练和部署方法论，并探索多模态大模型的微调技术以适配不同的下游任务应用等。

讲者简介：丁贵广教授，毕业于西安电子科技大学，博士后工作于清华大学，现为清华大学软件学院教授、博士生导师，北京信息科学与技术国家研究中心副主任，中国人工智能学会理事、智能光学成像专委会主任。长期从事计算机视觉相关研究，作为负责人先后承担国家“863 计划”、国家自然科学基金、国家重点研发计划、清华-京东联合研究中心、清华-OPPO 联合研究中心等项目 20 余项。多媒体内容理解、深度学习模型压缩优化方法等成果成功应用于快手、OPPO、京东、凌云光等单位，部分成果曾获国家科技进步二等奖。

程明明 南开大学

报告题目：高效能个性化图像生成

报告摘要：以大模型为代表的多模态图像生成技术可以有效地根据文本信息生成高质量的图像。然而，现有多模态生成技术在模型训练和个性化生成方面表现出较低的效率。例如，作为最近 AI 顶流的 Sora 模型虽然可以生成数十秒的流畅视频，但其训练代价相当高。Sora 核心组件 Difussion Transformer (DiT) 经常需要数十万次地迭代训练才能生成高质量的图像。此外，在图像生成中引入个性化的信息虽然富有吸引力，但是经典通过模型微调的形式经常耗费数十分钟才能得到高质量的结果。这些问题给生成式模型的大规模推广造成了障碍。本报告将介绍如何通过引入结构信息建模能力和个性化信息编码能力，有效地避免上述问题，并将该领域主流方法的性能提升 2 个数量级以上。

讲者简介：程明明，南开大学杰出教授，计算机系主任。主持承担了国家杰出青年科学基金、优秀青年科学基金项目、科技部重大项目课题等。他的主要研究方向是计算机视觉和计算机图形学，在 SCI 一区/CCF A 类刊物上发表学术论文 100 余篇（含 IEEE TPAMI 论文 30 余篇），h-index 为 80，论文谷歌引用 4 万余次，单篇最高引用 4700 余次，多次入选全球高被引科学家和中国高被引学者。技术成果被应用于华为、国家减灾中心等多个单位的旗舰产品。获得教育部自然科学一等奖 2 项、其他省部级科技奖 2 项。培养的 3 名博士生获得省部级优秀博士论文奖。现担任中国图象图形学学会副秘书长、天津市人工智能学会副理事长和顶级期刊 IEEE TPAMI, IEEE TIP 和《中国科学：信息科学》编委。

吴建鑫 南京大学

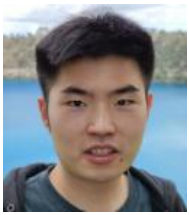
报告题目：资源高效的视觉大模型微调

报告摘要：训练一个视觉大模型需要海量的资源：训练用时、训练数据，以及模型参数、训练时的显存占用等，因此微调是模型复用与迁移的常用范式。然而，即便仅微调视觉基础模型仍然需要大量的训练时间与显存占用，也需要更新模型的所有参数。近来的参数高效微调方法（如 LoRA）是参数且训练数据高效的（即对这两种资源的需求显著减少），但对于其他资源（微调的训练用时和显存占用）仍然低效。本报告将讨论我们在针对各种资源均高效的视觉大模型微调方面的一些研究进展。

讲者简介：吴建鑫于南京大学获计算机科学与技术学士与硕士学位，于佐治亚理工学院获计算机科学博士学位，现任南京大学人工智能学院/计算机软件新技术全国重点实验室教授。曾任 CVPR、ICCV、ECCV、AAAI、IJCAI 等会议的资深领域主席或领域主席，IEEE TPAMI 编委，并担任 CVPR 2024 程序主席。在相关领域的重要学术期刊、会议发表了 100 多篇论文。具体来说，目前的研究兴趣为计算、数据资源受限情况下的深度学习与计算机视觉。

王鑫龙 北京智源人工智能研究院

报告题目：从视觉到多模态基础模型：探索与实践



报告摘要： 语言基础模型率先取得突破，如何构建通用的视觉和多模态基础模型，成为现在视觉领域关注的热点问题。本次报告将围绕视觉上下文学习、大规模图文对比学习、生成式多模态预训练等前沿技术，介绍视觉大模型、视觉提示、生成式多模态大模型的最新研究进展。

讲者简介：王鑫龙，智源研究院视觉模型研究中心负责人。本科毕业于同济大学，博士毕业于澳大利亚阿德莱德大学，师从沈春华教授。他的研究兴趣是计算机视觉和基础模型，近几年研究工作包括视觉感知 （SOLO, SOLOv2），视觉表征 (DenseCL, EVA)，视觉通才模型(Painter, SegGPT)，多模态表征(EVA-CLIP, Uni3D)，多模态通才模型(Emu, Emu2)。入选 Google PhD Fellowship、国家海外高层次青年人才。

高鹏 上海人工智能实验室

报告题目：构建高效的多模态语言大模型

报告摘要：多模态语言大模型能够连接视觉和语言大模型，是当前最重要的研究领域之一。本报告介绍将介绍 LLaMa-Adapter、SPHINX 的多模态语言大模型训练方法细节，分析如何高效构建高效多模态大模型，并且进行视觉多模态大模型的高效迁移。同时本报告将分析开源与闭源多模态大模型差距，展望多模态大模型未来的发展趋势。

讲者简介：高鹏，上海人工智能实验室青年研究员。主要从事视觉语言模型、多模态大模型和文生图大模型等基础模型研究和相关的高效微调方面的研究。在相关领域发表论文 40 余篇，其中第一作者或者通讯作者发表论文 20 余篇，Google Scholar 总引用量为 5000。开发 LLaMa-Adapter、SPHINX 和 Lumina-T2X 等基础多模态生成和理解大模型，组内开源工作 Github Star 总数超过一万。

左旺孟 哈尔滨工业大学

个人简介：左旺孟，哈尔滨工业大学计算机学院教授、博士生导师。主要关注生成式对抗网络、迁移学习模型及其在底层视觉、图像生成、视觉跟踪、图像分类等领域的应用。在 CVPR/ICCV/ECCV 等顶级会议和 T-PAMI/IJCV、IEEE Trans.等顶级期刊上发表论文 100 余篇，谷歌学术引用 15,000 余次，谷歌学术 H-Index 为 62。发表 ESI 高被引论文 8 篇。提出的 DnCNN 模型被正式收入 MATLAB 2017b 及后续版本的 Image Processing 和 Deep Learning Toolbox，并被 CV-News Magazine 作为 Main Story 专题报道，目前最高单篇引用 1900 余次。担任人工智能学会模式识别专委会常委、中国图象图形学会机器视觉专委会常委、中国图象图形学会青工委执委，国家自然科学基金重点项目负责人。担任 ICCV 2019、CVPR 2020/2021 等 CCF-A 类领域主席及 IJCAI、AAAI 等 CCF-A 类会议高级程序委员，担任 ECCV AIM 2020 竞赛单元主席之一。

戴玉超 西北工业大学

个人简介：戴玉超，西北工业大学电子信息学院教授、博士生导师，国家级青年人才，院长助理，陕西省信息获取与处理重点实验室主任及国际联合研究中心副主任。于 2005 年、2008 年和 2012 年分别获得西北工业大学学士、硕士和博士学位。2012 年至 2014 年在澳大利亚国立大学从事博士后研究工作，2014 年至 2017 年在澳大利亚国立大学工程研究院担任 ARC DECRA 学者。主要研究方向是智能视觉感知、三维场景重建、多视角几何、显著性检测、无人驾驶等。已在 IEEE TPAMI、IJCV、ICCV、CVPR、NeurIPS、ECCV 等国际期刊和会议上发表论文 70 余篇，谷歌学术引用 10000 余次，获陕西省自然科学一等奖、IEEE CVPR 2012 最佳论文奖（大陆高校首次）、IEEE CVPR 2020 最佳论文奖提名等多项学术奖项，担任 CVPR、ICCV、ACM-MM 等国际顶级会议领域主席。

王立君 大连理工大学

个人简介：王立君，大连理工大学未来技术学院副教授，硕士生导师。主要研究方向聚焦于图像深度估计、目标检测分割与跟踪等。主持国自然联合重点、面上基金等国家级纵向项目 3 项，省部级纵向项目 3 项，企业横向项目 2 项，入选人社部“博士后创新人才支持计划”和大连市“科技人才创新支持计划”。在本领域顶级学术会议和期刊发表论文三十余篇，谷歌学术总引用量 7000 余次。相关研究成果获得辽宁省科技进步一等奖，中国图象图形学会自然科学二等奖，教育部自

然科学二等奖，中国图象图形学学会优秀博士论文奖，以及辽宁省优秀博士论文奖。连续三年获得 VOT 国际视觉目标跟踪竞赛 (2020-2022) RGB-D 赛道冠军。担任多个国际顶级会议和期刊审稿人，VALSE 第 6-7 届执行委员会 (EACC) 委员，CCF-CV 与 CSIG-MV 专委会执行委员，以及《计算机技术与发展》期刊专题编委

张磊 OPPO

个人简介：张磊教授（IEEE Fellow）于 2006 年加入香港理工大学电子计算学系，2017 年起任职讲座教授，长期致力于计算机视觉、图像处理、模式识别等方向的研究，是底层视觉方面的国际权威学者，担任 IEEE Trans. on Image Processing（TIP）的高级编委，IEEE Trans. on Pattern Analysis and Machine Intelligence（TPAMI）、SIAM Journal of Imaging Sciences 等多个国际期刊的编委。从 2015 年至 2023 年，张教授连续被评为 Clarivate Analytics Highly Cited Researcher。张磊教授现担任 OPPO 研究院 AI 科学家，从事 AI 影像前沿技术的研发。

张璐 大连理工大学

个人简介：张璐，大连理工大学信息与通信工程学院副教授，博士生导师，研究方向为计算机视觉、大模型高效迁移、生成式人工智能。在人工智能领域的国际顶级会议和期刊上发表论文 20 余篇，授权两项国际专利。主持或参与国家自然科学基金重大项目、重点项目、青年基金项目等多项科研项目。获得辽宁省自然科学奖二等奖（排名第三）、中国图象图形学学会优秀博士学位论文、大连理工大学优秀博士学位论文，入选 2021 年中国博士后创新人才支持计划（“博新计划”），大连市高层次人才项目“高端人才”。

Workshop 10: 异构联邦学习

组织者：叶茫（武汉大学）、汪婧雅（上海科技大学）、卢杨（厦门大学）
时间：5 月 7 日（周二）08:30-12:00　地点：博悦厅

时间	主持人	内容
8:30~9:10	叶茫	讲者：杨强（香港科技大学） 题目：　联邦学习及大模型的发展趋势及未来发展方向
9:10~9:40		讲者：夏勇（西北工业大学） 题目：联邦学习在医学影像分析中的应用
9:40~10:10		讲者：王乐业（北京大学） 题目：可信联邦学习评估
10:10~10:20	中场休息	
10:20~10:50	汪婧雅	讲者：范晓亮　（厦门大学） 题目：可信联邦学习与行业大模型应用初探
10:50~11:20		讲者：潘微科（深圳大学） 题目：跨用户联邦推荐
11:20~12:00	卢杨	Panel 嘉宾： 叶茫（武汉大学）、汪婧雅（上海科技大学）、杨强（香港科技大学）、夏勇（西北工业大学）、王乐业（北京大学）、范晓亮　（厦门大学）、潘微科（深圳大学）

杨强 香港科大

报告题目：联邦学习及大模型的发展趋势及未来发展方向

报告摘要：联邦学习作为一种保护数据隐私的分布式学习方法，其在解决数据隐私保护方面起着关键作用。大模型技术的崛起虽为各种复杂任务带来了突破性的性能提升，同时也带来了计算和数据隐私保护方面的挑战。本次演讲将围绕大模型及联邦学习技术，讨论如何运用大模型+联邦学习的思路来应对挑战，同时将重点介绍大模型在金融领域的实际应用案例及其可能为未来金融智能化提供的重要启示和方向。

讲者简介：杨强（Qiang Yang），男，加拿大皇家科学院院士，加拿大工程院院士，电气与电子工程师协会会士（IEEE Fellow），美国科学促进会会士，中国人工智能学会会士，美国计算机协会会士（ACM Fellow），国际模式识别协会会士，香港科技大学计算机科学系讲座教授，微众银行首席人工智能官，中国移动通信集团有限公司独立非执行董事。

夏勇 西北工业大学

报告题目：联邦学习在医学影像分析中的应用

报告摘要：在医学影像分析领域，数据驱动的深度学习方法正日益显示出其巨大的潜力和价值。然而，由于数据隐私保护、多源异构数据整合以及分布式计算环境的需求，传统的中心化学习方法在医学影像分析中面临诸多挑战。为了解决这些问题，联邦学习作为一种新兴的分布式机器学习范式，正在逐渐受到医学影像分析领域的关注和研究。本报告将从如何减少数据标注负担、实现域适应和提升性能公平性这三个方面探讨联邦学习在医学影像分析中的应用，介绍我们所做的一些工作，以展示联邦学习在医学影像分析中的有效性和实用性。最后，我们将对联邦学习在医学影像分析中的未来发展趋势和挑战进行展望。

讲者简介：夏勇，西北工业大学计算机学院 长聘教授、博导、空天地海一体化大数据应用技术国家工程实验室成员。研究方向为医学影像智能计算，近 5 年在 JAMA Network Open、Radiology、IEEE-TPAMI/TMI/TIP、MedIA、NeurIPS、CVPR、ECCV、MICCAI、AAAI、IJCAI 发表论文 70 余篇，谷歌引用 1.2 万余次，指导学生先后在 BraTS2020、KiTS21、KiPA22、SegRap2023 等 10 余项国际学科竞赛中获得前三名。个人主页：https://teacher.nwpu.edu.cn/yongxia.html

王乐业 北京大学

报告题目：可信联邦学习评估

报告摘要：在可信联邦学习系统的研究领域，目标多种多样，包括提高预测的准确性、优化通信效率以及增强隐私保护等。这些目标的多样性使得对联邦学习研究的评价面临固有的挑战。报告将首先对当前联邦学习领域的主要研究评估目标进行梳理，并详细分析每个目标所对应的评估指标。通过对比分析，旨在揭示如果评估过程缺乏全面性，可能会导致结论的误导性和不完整性。此外，还将介绍 FedEval 开源联邦学习评估平台。该平台提供了一套标准化和系

统化的评估框架，专门用于衡量联邦学习算法在性能、效率和安全性等多个维度上的全方位表现。通过使用 FedEval，研究人员能够更加便捷地对联邦学习算法进行全面而深入的评估，并降低评估工作所需投入的时间和精力。

讲者简介：王乐业，北京大学计算机学院、高可信软件技术教育部重点实验室研究员、助理教授，入选国家级青年人才计划。长期从事面向城市智能应用的泛在数据治理研究，聚焦于隐私保护下的合规数据采集、计算、建模、共享等数据全生命周期处理技术。发表国际会议、期刊论文 80 余篇，累计引用 6000 余次。

范晓亮 厦门大学

报告题目：可信联邦学习与行业大模型应用初探

报告摘要：大模型训练的模型安全和数据隐私等问题受到广泛关注。同时，通用大模型的落地面临“行业数据不出域”的严峻挑战，亟需平衡大模型精度、安全与效率。可信联邦学习、参数高效微调等技术作为新型数据安全流通范式，可为大模型落地行业提供支撑。本次报告将介绍可信联邦学习算法及其在智慧交通、高等教育等领域的落地实践，并讨论大模型高效安全微调等领域的发展思考。

讲者简介：范晓亮，法国巴黎第六大学计算机科学博士，厦门大学信息学院高级工程师，数字福建城市交通大数据研究所常务副所长。研究兴趣:可信联邦学习、行业大模型应用。主持 3 项国家自然科学基金和百度、腾讯、厦门地铁等产学研项目，牵头国内首个高等教育隐私计算标准。TKDE、TSC、TMC、AAAI、IJCAI、WWW 等期刊会议发表论文 70 余篇。获 CCF 服务计算青年才俊奖(2022)、福建省科技进步一等奖(2018)。IEEE 高级会员、IEEE 教育数据挖掘工作组副主席，CCF 厦门会员活动中心执委，CCF 高级会员，CCF 服务计算专委会委员。

潘微科 深圳大学

报告题目：跨用户联邦推荐

报告摘要：智能推荐技术已成为众多在线服务平台的重要引擎和标准配置。近年来，随着用户隐私保护意识的增强，以及《通用数据保护条例》和《互联网信息服务算法推荐管理规定》等法律法规的颁布，一类新的被称为“联邦推荐”的隐私敏感推荐技术引起了学术界和工业界的广泛关注。本报告将介绍如何将一个已有的中心化的推荐模型升级为跨用户（分布式）的联邦版本，进而实现原始数据不离开用户本地的隐私保护目的。

讲者简介： 潘微科，博士，深圳大学计算机与软件学院教授，博士生导师。2005 年毕业于浙江大学获学士学位，2012 年毕业于香港科技大学获博士学位。主要研究方向为迁移学习、联邦学习、推荐系统和机器学习，已发表 100 多篇科研论文，出版推荐技术教材一本，曾获 ACM TIIS 2016 年度最佳论文奖和 SDM 2013 最佳论文提名奖。先后主持国家自然科学基金项目 3 项。担任 AAAI 2021 线上会议的副主席、ACM TORS 的创刊编委、Neurocomputing 的编委等。

Workshop 11：大模型与因果推理

组织者：况琨（浙江大学）、宫明明（墨尔本大学）

时间：5 月 7 日（周二）13:30-18:00　　地点：喜悦厅

时间	主持人	内容
14:00-14:30	况琨	讲者：刘礼（重庆大学） 题目：面向工业设计的大模型因果推断问题与应用
14:30-15:00	况琨	讲者：冯福利（中国科学技术大学） 题目：因果大模型赋能推荐初探
15:00-15:30		讲者：丁效（哈尔滨工业大学） 题目：大模型因果推理的局限性及自主进化学习
15:30-16:00	况琨	讲者：俞奎（合肥工业大学） 题目：面向隐私保护数据的联邦因果关系推断方法研究
16:00-16:10	中场休息	
16:10-16:40	宫明明	讲者：孙鑫伟（复旦大学） 题目：Causal Discovery via Conditional Independence Testing with Proxy Variables
16:40-17:10	宫明明	讲者：林勇（香港科技大学） 题目：Spurious Feature Diversification Improves Out-of-distribution Generalization
17:10-17:40	宫明明	Panel 嘉宾： 刘礼（重庆大学）、冯福利（中国科学技术大学）、丁效（哈尔滨工业大学）、俞奎（合肥工业大学）、孙鑫伟（复旦大学）、林勇（香港科技大学）

刘礼 重庆大学

报告题目：面向工业设计的大模型因果推断问题与应用

报告摘要：工业设计（如汽车造型设计）对工业产品安全性和市场竞争影响巨大。然而，传统设计过程效率低、成本高、创意主观，难以满足用户需求。引入生成式人工智能，将为新能源汽车造型等工业设计带来革命性变革。大模型以其先进的学习和生成能力不仅能生成高质量设计图，为设计提供灵感和创意资源，但仍不能完全满足设计过程中的约束条件，对可调控、可解释的能力有限。因此如何将因果推断融入大模型是推动工业设计高效化的关键，也是适应大模型人工智能时代的必然趋势。

讲者简介：重庆大学教授、博士生导师。主持国家科技部“工业软件”重点项目课题，国家“新一代人工智能”重大科技专项课题等 20 余项；已在 AAAI、MM、ICDE、TSE、TNNLS、TLT 等国内外顶级学术期刊和会议上发表论文近 200 篇；获省部级奖 1 项；入选 2022 年度 AI2000 人工智能最具影响力学者，著有《因果论》、《因果漫步》等书籍。

冯福利 中国科学技术大学

报告题目： 因果大模型赋能推荐初探

报告摘要： 因果推理和生成式大模型是推荐系统领域近年来广受关注的方向，在推荐纠偏、推荐公平、推荐泛化等问题上广泛应用。本报告围绕因果推理和大模型，介绍推荐系统领域相关的核心问题与最新进展，包括三方面内容：首先介绍因果推理在推荐系统领域的发展脉络和前沿进展；其次介绍推荐大模型的主要技术路线和关键问题；最后展望推荐视角下因果大模型的可能形态并展望因果大模型的未来发展方向。

讲者简介： 冯福利，中国科学技术大学特任教授。研究领域：信息检索、数据挖掘、机器学习等，发表国内外顶级会议和期刊论文 100 篇，谷歌学术引用 10000 次，承担科技部重点研发计划项目课题、基金委面上项目等国家级项目，研究成果在多家公司的商业系统应用。曾获 SIGIR 2021 最佳论文提名奖、WWW 2018 最佳演示论文奖。担任 ACM TORS 编委（AE），众多顶级期刊审稿人，会议 SPC/PC，包括 SIGIR、WWW、SIGKDD、NeurIPS、ICML、ICLR、ACL、TOIS、TKDE、TPAMI、JASA、Nature Sustainability。

丁效 哈尔滨工业大学

报告题目：大模型因果推理的局限性及自主进化学习

报告摘要：大语言模型在诸多自然语言处理任务上均取得了令人瞩目的表现，其智能涌现的机理也备受关注。因果推理能力是检验大语言模型智能程度的一个重要任务，当前以 ChatGPT 为代表的大语言模型在因果推理任务上表现的到底如何？是否存在这一定的局限性，在本次报告中将进行详细的实证剖析。在此基础上，本次报告还将介绍让大模型自主进行因果推理能力提升的全新任务框架和学习方法。

讲者简介： 丁效，哈尔滨工业大学教授，博士生导师。主要研究方向为人工智能、自然语言处理、知识计算、因果推理。在 TKDE、ACL、AAAI、IJCAI 等人工智能领域的顶级国际期刊和会议上发表相关论文 70 余篇，承担国家部委项目、黑龙江省优青项目、“新一代人工智能”重大项目课题、国家自然科学基金重点项目课题、面上项目等多项省部级以上项目。荣获国家级教学成果二等奖、黑龙江省科技进步二等奖、SemEval 2020 国际语义评测“检测反事实陈述”任务第一名，入选 2022 年 AI 2000 全球人工智能最具影响力学者、华为云 AI 名师奖等，担任中国中文信息学会社会媒体处理专委会秘书长等职务。

俞奎 合肥工业大学

报告题目：面向隐私保护数据的联邦因果关系推断方法研究

报告摘要：理解大数据内蕴含的因果关系，揭示数据的产生机制与规律，是构建高可解释性与强泛化性的学习模型的关键。近年来，由于数据滥用和隐私泄露事件频发，为保护数据隐私，多个组织、群体之间的数据相互隔离，形成数据孤岛问题，严重阻碍因果驱动的大数据理解与分析技术的发展与应用。在联邦学习范式下，本报告首先介绍团队将传统的 PC 算法的分层学习思想与联邦学习的分布式框架巧妙融合从而提出的全局联邦因果关系学习算法。然后，针对全局联邦因果关系学习存在的计算效率和学习准确度问题，分别介绍团队提出的自适应空

间选择的全局联邦因果关系推断方法和局部-全局的可扩展的联邦因果关系推断方法。

讲者简介：俞奎，合肥工业大学计算机与信息学院教授，博士生导师，研究方向为因果推断与机器学习。2013 年-2018 年分别在加拿大和澳大利亚从事全职研究工作，2018 年入职合肥工业大学。 在 IEEE TPAMI、IEEE TKDE、IEEE TNNLS、ICML、KDD、AAAI 等国际权威期刊与国际顶级学术会议发表学术论文 60 多篇，出版《因果推断导论》专著一本。曾获中国计算机学会（CCF）优秀博士学位论文与加拿大 PIMS 博士后奖。目前主持科技创新 2030 新一代人工智能重大项目课题 1 项，子课题 1 项，以及国家自然科学基金面上项目 2 项。目前任安徽省人工智能学会副理事长，因果与认知智能专委会主任；IEEE TETCI 等国际期刊的副主编和多个人工智能领域国际顶级会议的领域主席。

孙鑫伟 复旦大学

报告题目：Causal Discovery via Conditional Independence Testing with Proxy Variables

报告摘要：Distinguishing causal connections from correlations is important in many scenarios. However, the presence of unobserved variables, such as the latent confounder, can introduce bias in conditionalindependence testing commonly employed in constraint-based causal discovery for identifying causal relations. To address this issue, existing methods introduced proxy variables to adjust for the bias caused by unobserveness. However, these methods were either limited to categorical variables or relied on strong parametric assumptions for identification. In this talk, we introduce a novel hypothesis-testing procedure that can effectively examine the existence of the causal relationship over continuous variables, without any parametric constraint. Our procedure is based on discretization, which under completeness conditions, is able to asymptotically establish a linear equation whose coefficient vector is identifiable under the causal null hypothesis. Based on this, we introduce our test statistic and demonstrate its asymptotic level and power. We validate the effectiveness of our procedure using both synthetic and real-world data.

讲者简介：Xinwei Sun (孙鑫伟) is currently an assistant professor with the School of Data Science, Fudan University. He received his Ph.D in school of mathematical sciences, Peking University in 2018. His research interests mainly focus on high-dimensional statistics, causal inference, with their applications on medical imaging and machine learning. His work has been published on journals in statistics and mathematics, such as JRSSB, JASA, ACHA, TSP, and also on machine learning conferences such as ICML, NeurIPS.

林勇 香港科技大学

报告题目：Spurious Feature Diversification Improves Out-of-distribution Generalization

报告摘要：In this talk, we focus on out-of-distribution (OOD) generalization problems. The OOD community has put huge efforts into learning invariant features that are stable under distributional shifts and avoiding spurious features that are unstable. However, researchers have found that invariant feature learning methods have inferior empirical performance, probably due to many strong conditions and assumptions. Instead, we propose a new perspective on OOD generalization, namely, spurious feature diversification, which incorporates a more diverse set of spurious features. Spurious feature diversification can weaken the

individual contribution of each spurious, resulting in significantly better OOD performance. The success of spurious feature diversification lies in three key factors: (1) there are a large number of spurious features, (2) different spurious features exhibit certain randomness under distributional shifts, and (3) a DNN trained by ERM with SGD would learn a subset of spurious features. Moreover, spurious feature diversification can be achieved through many cheap methods, e.g., model averaging or feature concatenation, and it works effectively on many large-scale benchmarks (e.g., ImageNet variants) with large neural networks such as CLIP and Large Language Models (LLMs).

讲者简介： Yong Lin is a final year PhD student advised by Professor Tong Zhang at the Hong Kong University of Science and Technology and in-coming postdoc fellow at Princeton. His main research focus is on the trustworthiess of machine learning, including out-of-distributional (OOD) generalization and the alignment of large language models (LLMs) with human preferences. He concentrates on in-depth theoretical analysis as well as comprehensive empirical investigations. Yong has published a series of papers at NeurIPS, ICML, ICLR, and CVPR, with several of them being selected as highlight papers (oral/spotlight). He was awarded the Apple AI/ML PhD fellowship and the Hong Kong PhD fellowship. Before starting his PhD studies, he worked as a Machine Learning Engineer at Alibaba, gaining experience with industrial machine learning applications and the inherent instabilities of DNNs.

Workshop 12：大模型理论与机理

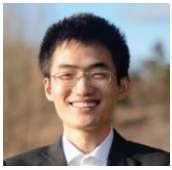
组织者：刘勇（中国人民大学）、李崇轩（中国人民大学）、盛律（北京航空航天大学）

时间：5 月 6 日（周一）08:30-12:00　　地点：喜悦厅

时间	主持人	内容
8:30-9:05	刘勇	讲者：袁洋（清华大学） 题目：定位即智能
9:05-9:40	刘勇	讲者：贺笛（北京大学） 题目：是否所有大模型都具备思维链涌现能力？
9:40-10:15	刘勇	讲者：张宁豫（浙江大学） 题目：大语言模型知识机理与编辑问题
10:15-10:20	中场休息	
10:20~10:55	刘勇	讲者：孙若愚（香港中文大学深圳） 题目：Transformer 为什么需要 Adam 训练？
10:55-11:30	刘勇	讲者：颜航（上海人工智能实验室） 题目：迈向 1M 上下文长度的大模型
11:30-12:00	刘勇	Panel 嘉宾： 贺迪（北京大学）、袁洋（清华大学）、张宁豫（浙江大学）、颜航（上海人工智能实验室）、孙若愚（香港中文大学）、李崇轩（中国人民大学）、盛律（北京航空航天大学）

袁洋 清华大学

报告题目：定位即智能



报告摘要：OpenAI 提出了著名的“压缩即智能”观点，但没有明确其与传统文件压缩算法的差异，也没有给出其与常用算法的联系。相比之下，我认为“定位即智能”是一个更好的视角，不仅拥有优美的理论支撑（范畴论），还可以和大量预训练、多模态算法构建起联系。本报告将介绍在此视角下，关于对比学习与大模型算法的理论分析。

讲者简介：袁洋，清华大学交叉信息学院助理教授。2012 年毕业于北京大学计算机系，2018 年获得美国康奈尔大学计算机博士学位，师从 Robert Kleinberg 教授。他于 2018-2019 年前往麻省理工学院大数据科学学院（MIFODS）做博士后。袁洋的主要研究方向是智能医疗、AI 基础理论、应用范畴论，在 NeurIPS，ICLR，ICML 等计算机和人工智能领域顶级会议上发表论文三十余篇。曾获得福布斯中国 2019 年 30 Under 30、2019 年北京智源青年科学家等荣誉。

贺笛 北京大学

报告题目：是否所有大模型都具备思维链涌现能力？

报告摘要：如何对长序列进行建模是当前自然语言处理中的一个热点问题。长序列建模面临诸多挑战，效率是其中一个重要的瓶颈。针对效率问题，过去已经有许多研究工作提出多种 Transformer 的高效变体，但这些变体模型是否存在理论缺陷？面临具体实际问题时模型结构应当如何选择？到底哪些变体模型能真正完美地取代 Transformer？在这个 talk 中，我将围绕着团队最近的工作，试图对上述问题进行理论层面的回答。

讲者简介：贺笛，北京大学智能学院助理教授，前微软亚洲研究院主管研究员。主要从事机器学习模型、算法与理论方向的研究工作，已发表 ICML、NeurIPS、ICLR 等重要期刊/会议论文 50 余篇，谷歌引用数超过 8000，指导学生 2 次在图神经网络国际顶级评测竞赛上取得冠军。所设计的模型、算法多次被 DeepMind、OpenAI、微软、Meta 等国际顶尖研究机构使用。获得机器学习顶级国际会议 ICLR 2023 杰出论文奖。

张宁豫 浙江大学

报告题目：大语言模型知识机理与编辑问题

报告摘要：掌握知识一直是人工智能系统发展的核心追求。在这方面，大语言模型展示了巨大的潜力并在一定程度上掌握和应用了广泛的知识。然而，我们对于大语言模型如何内在地习得、存储知识等方面的理解仍然非常有限，我们也无法及时对大语言模型内部的错误及有害知识进行修正。在本次 Talk 中，我将基于团队最近的研究成果，探讨大语言模型的知识机理与编辑问题，并尝试在理论层面上提供解答。

讲者简介：张宁豫，浙江大学副教授，浙江大学启真优秀青年学者，在高水平国际学术期刊和会议上发表多余篇论文，5 篇入选 Paper

Digest 高影响力论文，1 篇被选为 Nature 子刊 Featured Articles。主持国家自然科学基金、计算机学会、人工智能学会多个项目，获浙江省科技进步二等奖，IJCKG 最佳论文/提名 2 次，CCKS 最佳论文奖 1 次，担任 ACL、EMNLP 领域主席、ARR Action Editor、IJCAI 高级程序委员，主持开发大语言模型知识编辑工具 EasyEdit (1.2k)。

孙若愚 香港中文大学（深圳）

报告题目： Transformer 为什么需要 Adam 训练?

报告摘要：Transformer 的训练一般使用 Adam 而不是 SGD，然而其背后的原因尚不明确。在这项工作中，我们通过 Hessian 矩阵的角度，提供 SGD 在 Transformer 上失败的一个解释：(i) Transformer 是“异质性”的：不同参数块的 Hessian 谱差异巨大，这一现象我们称之为“块异质性”；(ii) 异质性阻碍了 SGD：SGD 在存在块异质性的问题上表现不佳。为验证异质性影响了 SGD 的表现，我们检查了各种 Transformer、CNN、MLP 和二次函数优化问题，发现 SGD 在没有块异质性的问题上工作良好，但当存在异质性时表现不佳。我们的初步理论分析表明，SGD 失败是因为它对所有块使用单一学习率，因而无法处理块间的异质性。

讲者简介：孙若愚现为香港中文大学（深圳）数据科学学院长聘副教授、博士生导师。此前他任伊利诺伊大学香槟分校（UIUC）助理教授、博士生导师。曾任脸书人工智能研究所访问科学家，曾任斯坦福大学博士后研究员。他在美国明尼苏达大学电子与计算机工程系获得博士学位，北京大学数学科学学院获得本科学位。研究方向包括神经网络理论和算法、生成模型、大规模优化算法、学习优化等。他曾获得 INFORMS (国际运筹与管理协会) George Nicolson 学生论文竞赛第二名，以及 INFORMS 优化协会学生论文竞争荣誉奖。目前担任 NeurIPS, ICML, ICLR, AISTATS 等会议的领域主席，Transaction on Machine Learning Research 的 action edtor。

颜航 上海人工智能实验室

报告题目：迈向 1M 上下文长度的大模型

报告摘要：过去一年基于 Transformer 的语境长度从 2k 一路狂飙到 M 级别，平均每个月语境长度就翻了一倍。最常见的无损超长语境方案是通过调整相对位置编码进行长度外推，那么外推方案是否可以无限进行外推呢？如何确定一个模型能够支持的最大语境长度？如何寻找一个可以理论上无限进行外推的位置编码方案？在本次 Talk 中我将基于团队在位置编码相关研究，试图对上述问题进行回答。

讲者简介：上海人工智能实验室青年科学家。研究兴趣包括信息抽取、开源 NLP 工具建设、大规模预训练模型等。开源平台 OpenLMLab 主要贡献者，设计并开发了 InternEvo、fastNLP、fitlog 等开源工具，负责了浦江实验室 InternLM 大模型的预训练相关工作。在 ACL、TACL、EMNLP、NAACL 等会议或杂志上发表了多篇论文，2022 年获钱伟长中文信息处理科学技术奖一等奖（4/5）。

刘勇 中国人民大学

个人简介：刘勇，中国人民大学高瓴人工智能学院副教授，博士生导师。从事机器学习研究，特别关注统计机器学习、图表示学习、自动机器学习等。发表高水平论文 80 多篇，其中以第一作者或通讯作者发表高水平文章 50 余篇，涵盖机器学习领域顶级期刊 TIT、JMLR、TPAMI、Artificial Intelligence 和顶级会议 ICML，NeurIPS，ICLR 等。曾获得中国科学院“青年创新促进会”会员（院人才）以及中国科学院信息工程研究所“引进优秀人才”称号。

李崇轩 中国人民大学

个人简介：李崇轩，中国人民大学高瓴人工智能学院准聘副教授、博士生导师，2010-2019 年获清华大学学士和博士学位。主要研究机器学习、深度生成模型，代表性工作 Analytic-DPM、DPM-Solver 作为核心采样技术部署于 DALL·E 2、Stable Diffusion 等。获国际会议 ICLR 杰出论文奖、吴文俊优秀青年奖、吴文俊人工智能自然科学一等奖、中国计算机学会优秀博士论文、ACM SIGAI 中国新星奖等。担任 ICLR 2024 领域主席。

盛律 北京航空航天大学

个人简介：盛律，博导，北京航空航天大学“卓越百人”副教授，入选北航青年拔尖计划。2011 年获浙江大学学士学位，2017 年获香港中文大学博士学位，同年加入香港中文大学多媒体实验室从事博士后研究。主要研究方向为三维视觉、多媒体分析和具身智能。在 IEEE TPAMI/IJCV/TIP 以及 CVPR/ICCV/NeurIPS/ICLR/ECCV/AAAI/ACM MM 等重要国际期刊和会议发表论文超过 50 篇, Google Scholar 显示被引用数超 4700 次。获 CVPR 2021 年 ScanRefer 三维定位挑战赛第一名。组织 ICCV 2021 SenseHuman 等多个国际会议研讨会和竞赛。现任 ACM Computing Surveys 副编辑，CVPR 2024、ECCV 2024 和 ACM Multimedia 2024 领域主席，以及多个领域顶会顶刊审稿人和程序委员。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。

Workshop 13：多模态感知与对话

组织者：姬艳丽（电子科技大学）、胡迪（中国人民大学）、田亚鹏（德克萨斯大学达拉斯分校）

时间：5 月 7 日（周二）14:00-17:30　　地点：欣悦厅

时间	主持人	内容
14:00-14:30	姬艳丽	讲者：郑伟诗（中山大学） 题目：多模态行为分析
14:30-15:00		讲者：李成龙（安徽大学） 题目：低质多模态融合感知
15:00-15:30		讲者：罗平（香港大学） 题目：Efficient Diffusion Transformer for Image and Video Generation
15:30-15:40		企业宣讲：上海合合信息科技股份有限公司 多模态产品应用发展与展望（常扬）
15:40-15:50	中场休息	
15:50-16:10	胡迪	讲者：陈宸（OPPO） 题目：多模态模型端侧轻量化的研究和落地
16:10-16:40		讲者：许华哲（清华大学） 题目：大模型赋能具身智能
16:40-17:10		讲者：仉尚航（北京大学） 题目：迈向开放世界多模态具身智能感知

郑伟诗 中山大学

报告题目：多模态行为分析

报告摘要：近年来，围绕第一视觉行为建模、行为质量评估、机械臂自由抓取行为等任务，我们做了一些多模态数据处理与分析。这次报告将展示我们如何通过协同 egocentric 特征和 exocentric 特征以建模第一视觉行为中的交互、如何用 RGB 与音频等模态设计行为质量评估网络、如何将二维 RGB 和三维点云数据结合提升机械臂自由抓取的精度。

讲者简介：郑伟诗教授，教育部“长江学者奖励计划”特聘教授、英国皇家学会牛顿高级学者，现任教育部机器智能与先进计算重点实验室主任。长期研究协同与交互分析理论与方法，解决人体建模和机器人行为的视觉计算问题。发表 CCF-A/中科院 1 区/Nature 子刊 论文 150 多篇，谷歌学术引用约 2 万次。担任国际人工智能顶级期刊 IEEE T-PAMI 、Artificial Intelligence Journal 等期刊的编委。主持承担国家级重点类项目和人才项目 5 项、以及广东省自然科学基金委卓越青年团队（负责人）项目等。获中国图象图形学学会自然科学奖一等奖、广东省自然科学奖一等奖等、国家教学成果奖二等奖等。

李成龙 安徽大学

报告题目：低质多模态融合感知

报告摘要：多模态融合感知以多种来源的视觉传感器数据为基础，通过多模态表征、融合、学习等关键技术提供更全面、更准确、更鲁棒的目标感知和理解能力，在智能交通、公共安全、无人系统等场景下有着巨大的应用需求。然而，多模态成像系统一般由不同的传感器组合而成，会带来多模态数据的空间不对齐、信息不完整和质量不均衡等低质问题，为多模态融合感知研究与应用带来诸多挑战。报告将介绍近些年在低质多模态融合感知方面的研究进展。

讲者简介：李成龙，安徽大学人工智能学院教授、博士生导师、系主任，中国图像图形学学会视觉大数据专委会委员，安全人工智能安徽省重点实验室副主任，安徽省计算机学会奖励工作委员会副主任，安徽省人工智能学会多源信息融合专委会副主任。主要研究领域包括多模态视觉和深度学习，发表 IEEE TPAMI、IEEE TIP、IEEE TNNLS、CVPR、ECCV、AAAI 等高水平期刊和会议论文 70 余篇，谷歌学术引用 5000 余次，获得授权国家发明专利 21 项，其中 6 项实现成果转化或技术应用，主持国家自然科学基金（3 项）、安徽省杰出青年基金等科研项目，获得安徽省科技进步二等奖、安徽省计算机学会青年科学家奖等学术奖励，入选 2023 年度斯坦福全球前 2%顶尖科学家榜单。

罗平 香港大学

报告题目：Efficient Diffusion Transformer for Image and Video Generation

报告摘要：This talk will introduce a series of our recent work on the model, data, and computing efficiency of image and video generation developed from 2022 to 2024, such as RAPHAEL (NeurIPS’23), Video DiT (ICLR’24), PixArt-alpha (ICLR’24), PixArt-delta (arXiv:2401.05252), PixArt-sigma (arXiv:2403.04692), GenTron (CVPR’24), DiffAgent (CVPR’24), and implicit prompting (arXiv:2403.02118), with emphasis on the technical details, philosophy, and experience in conducting these researches.

讲者简介：罗平是香港大学计算机科学系副教授，港大数据科学研

究院（HKU IDS）副主任和香港大学与上海人工智能实验室联合研究室副主任。他于 2014 年在香港中文大学信息工程系获博士学位，师从汤晓鸥教授（商汤科技创始人）和王晓刚教授。2019 年加入香港大学之前，他曾是商汤科技的研究总监。他在 TPAMI、ICML、ICLR、NeurIPS 和 CVPR 等国际会议和期刊上发表了 100 余篇论文，Google Scholar 上引用超过 50,000+ 次。他曾获得 2015 年 AAAI Easily Accessible Paper Award，被提名为 2022 Computational Visual Media Journal 年度最佳论文,获得 2022 ACL 杰出论文奖,2023 世界人工智能大会（WAIC）优秀青年论文奖，并且是 ICCV’23 最佳论文候选。2020 年，他被《麻省理工科技评论》（MIT TR35）评为亚太地区 35 岁以下的创新者之一。他指导了 30 名博士生，其中许多人获得了包括 Nvidia 奖学金、百度奖学金、WAIC 云帆奖等重要奖项。

陈宸 OPPO

报告题目：图文多模态模型在端云场景的落地实践

报告摘要：大模型时代的模型参数和算法复杂度与日俱增，导致推理的成本和时延越来越高。本次报告将介绍图文检索，文生图，多模态语言模型的优化研究和应用落地。图文检索模型将介绍模型蒸馏，向量引擎和端侧部署的优化方案，赋能端侧智慧搜索；文生图基础模型将介绍中文文化迁移，采样加速，模型小型化等技术方案，实现端侧秒级高质量文生图；多模态语言模型将介绍多模态信息总结、图文并茂生成等应用优化，助力办公娱乐场景的用户需求。

讲者简介：陈宸，香港科技大学博士，OPPO AI 中心大模型算法部高级算法工程师。主要研究方向为视频编解码、鲁棒机器学习、多模态语言模型、视觉理解和图像生成等。在相关领域发表论文 40 余篇。在 OPPO 负责文本引导的图像生成与编辑、图文跨模态检索、多模态语言模型等基础技术的研究和应用落地。

许华哲 清华大学

报告题目：大模型赋能具身智能

报告摘要：探讨大模型在具身智能中的用处，以促进机器人和其他智能系统在物理世界中的交互能力。谈及大模型，大家首先想到的是用于任务分拆和规划。但大模型对于具身智能的底层控制也有重要的推动作用。本报告主要讨论如何用大语言模型生成具身智能环境、多模态大模型生成奖励函数等重要方向。

讲者简介：许华哲博士现为清华大学交叉信息研究院助理教授，博导，清华大学具身智能实验室负责人。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。其研究领域是具身人工智能（Embodied AI）的理论、算法与应用，具体研究方向包括深度强化学习、机器人学、基于感知的控制（Sensorimotor）等。其科研围绕具身人工智能的关键环节，系统性地研究了视觉深度强化学习在决策中的理论、模仿学习中的算法设计和高维视觉预测中的模型和应用，对解决具身人工智能领域中数据效率低和泛化能力弱等核心问题做出多项贡献。顶级智能机器人会议 CoRL'23 最佳系统论文得主，在 IJRR, RSS, NeurIPS 等发表顶级期刊/会议论文五十篇，代表性工作曾被 MIT Tech Review，Stanford HAI 等媒体报道。曾在 IJCAI2023、IJCAI2024、ICRA2024 担任领域主席/副主编。

仉尚航 北京大学

报告题目：迈向开放世界多模态具身智能感知

报告摘要：虽然机器视觉为各个领域带来巨大成功，但已有具身智能感知往往针对封闭环境，存在闭集假设和大样本假设等局限。而现实世界中的具身智能体往往面对开放环境，存在以下关键挑战：1）开放环境中存在大量数据域偏移，已有方案难以适应新数据域、对新场景进行准确理解；2）开放环境中新的类别动态出现，无法及时获得标注，已有方案难以在少量标注下准确识别新事物。本次分享将针对上述挑战，介绍一系列增强开放世界具身智能感知的泛化能力，使其自动适应新环境、识别新事物的研究工作。尤其针对 Corner Case 等问题提出新型持续泛化学习范式和多模态大模型解决方案。

讲者简介：仉尚航，北京大学计算机学院研究员、博士生导师、博雅青年学者。致力于开放环境泛化机器学习理论与系统研究，在人工智能顶级期刊和会议上发表论文 80 余篇，Google Scholar 引用数 8800 余次。荣获世界人工智能顶级会议 AAAI’ 2021 最佳论文奖。作为编辑和作者由 Springer Nature 出版英文书籍《Deep Reinforcement Learning》，至今电子版全球下载量超二十万次，入选中国作者年度高影响力研究精选。于 2018 年入选美国“EECS Rising Star”，于 2023 年入选“全球 AI 华人女性青年学者榜”、“中国科协青年百人会”。曾获国际人脑多模态计算模型响应预测竞赛第一名，ICCV 持续泛化学习竞赛第一名。曾多次在国际顶级会议 NeurIPS、ICML 上组织 Workshop，担任 AAAI 2022&2023&2024 高级程序委员。仉尚航于 2018 年博士毕业于美国卡内基梅隆大学，并于加州大学伯克利分校从事博士后研究。

姬艳丽 电子科技大学

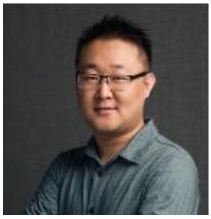
个人简介：姬艳丽，电子科技大学教授，博导，博士毕业于日本九州大学，东京大学客员研究员。研究方向:围绕人的音视觉多模态信息理解、情感认知。作为主要作者发表中科院 JCR 二区以上 SCI 期刊及 CCF A 国际会议论文 50 余篇，申请三十余项发明专利。获得 28th Australasian Database Conference 国际会议 Best Paper Award，第二十届中国虚拟现实大会最佳论文提名奖。入选 2023 年度 AI 华人女性青年学者榜。参与承办包括 ACM MM2021 等国际会议 5 次，国内学术峰会 VALSE 及 ACM SIGAI CHINA symposium in TURC 等 18 余次。中国计算机学会高级会员(CCF Senior Member)，中国图象图形学会青年工作委员会执行委员、副秘书长，VALSE 委员会 SAC 主席。

胡迪 中国人民大学

个人简介：胡迪，现任中国人民大学高瓴人工智能学院准聘副教授，博导，受中国科协青年人才托举工程资助。主要研究方向为机器多模态感知、交互与学习，以主要作者身份在领域顶级国际会议及期刊上发表论文 30 余篇，如 TPAMI、NeurIPS、CVPR、ICCV、ECCV 等。代表性工作如多模态场景理解算法 DMC；平衡多模态学习理论，机制与方法；和以运动学为引导的可泛化铰链物体操纵等。攻博期间曾入选 CVPR Doctoral Consortium；荣获 2020 中国人工智能学会优博奖，2021 陕西省优博奖；荣获 2022 年度吴文俊人工智能优秀青年奖；百度全球顶尖人工智能人才计划等。

受邀为多个国际高水平会议及期刊审稿，担任 AAAI、IJCAI SPC/Session Chair 等，并主办/协办多场国际顶级会议的多模态学习讲习班（Tutorial）。

田亚鹏 德克萨斯大学达拉斯分校



个人简介：田亚鹏，德克萨斯大学达拉斯分校助理教授、博士生导师。本科就读于西安电子科技大学，硕士毕业于清华大学，在美国罗彻斯特大学获得博士学位。目前已在 CVPR/ECCV/ICCV/NeurIPS/ICLR/AAAI/TPAMI/TIP 等国际顶级会议与期刊发表 30 多篇论文，相关工作在视听多模态学习和图像视频复原领域起到引领性作用，作为主要组织者之一在 WACV 和 CVPR 参与举办两届视听场景理解讲座。 曾获得 Cisco Faculty Research Award，入选 CVPR 2022 博士论坛和 AAAI 2023 New Faculty Highlights。

Workshop 14：女科学家成长论坛

组织者：董晶（中国科学院自动化研究所）、刘偲（北京航空航天大学）、姬艳丽（电子科技大学）

时间：5 月 7 日（周二）8:30-12:00　　地点：智悦厅

时间	主持人	内容
8:30-8:40	刘偲	破冰问答
8:40-9:20		Panel 1 科研不分性别，科研人有性别
9:20~10:20	董晶	Panel 2 女性领导力培养
10:20-10:30	中场休息	
10:30-11:30	姬艳丽	分组主题研讨

Panel 嘉宾：
张艳宁（西北工业大学）、马惠敏（北京科技大学）、邬霞（北京理工大学）、姚鸿勋（哈尔滨工业大学）、金琴（中国人民大学）、宋睿华（中国人民大学）、山世光（中科院计算所）、潘纲（浙江大学）

张艳宁 西北工业大学

个人简介：张艳宁，西北工业大学教授、博士生导师，教育部“长江学者”特聘教授、中组部首批“万人计划”科技创新领军人才、973 技术首席。现任西北工业大学副校长兼研究生院院长。先后主持和承担某基础加强重点项目、国防 973 项目、国家自然科学基金重点项目、国家/国防 863、总装预研等国家级项目 40 余项。兼任中国图象图形学学会副理事长等。在 IEEE TPAMI、IEEE TIP、IJCV、CVPR、ICCV 等国内外本领域权威期刊和重要国际会议上发表论文百余篇，出版专著 3 部，获国家/国防授权发明专利 50 余项，以第一完成人获国家技术发明二等奖、国防技术发明一等奖、国家教学成果二等奖、陕西省科技进步一等奖等 8 项。她主要从事图像处理、模式识别、计算机视觉与智能信息处理等研究，长期致力于图像处理、模式识别、计算机视觉与智能信息处理等的研究，并与航天、航空等方面的国家重大需求相结合。其团队在空间态势感知、空天地海一体化大数据应用技术上发展出了丰富成果。

马惠敏 北京科技大学

个人简介：马惠敏，北京科技大学教授、计算机与通信工程学院物联网与电子工程系主任、北京市“三八红旗奖章”获得者，中国图象图形学学会副理事长兼秘书长、中国女科技工作者协会理事、CSIG 脑图谱专业委员会副主任、CCF-CV 专委执行委员。主要研究方向为认知决策，将计算机视觉与认知心理学结合，在复杂环境仿真、视觉认知、智能决策方面取得了原创性成果，以第一完成人获 2016 年吴文俊人工智能科技创新一等奖、2020 年中国图象图形学学会技术发明一等奖、2017 教育部技术发明奖二等奖、日内瓦国际发明展银奖，并多次在国际最大的自动驾驶数据集(KITTI)评测中获得第一名。主持国家自然科学基金联合基金重点项目、专项重点项目等 20 余项，作为通讯作者发表论文 100 余篇，包括 TPAMI、IJCV、TIP、TITS、CVPR、NIPS、ICCV 等，获批专利二十余项，两项专利完成了科研成果转化。

邬霞 北京理工大学

个人简介：邬霞，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、教育部新世纪优秀人才、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获得者。主要研究方向为脑信号智能分析、类脑算法等。近年来，主持承担国家自然科学基金重点、国家重点研发计划等项目十余项、以第一/通讯作者在国内外重要学术期刊、会议发表论文 100 余篇。

姚鸿勋 哈尔滨工业大学

个人简介：姚鸿勋，哈尔滨工业大学计算学部长聘教授，黑龙江省政府特殊津贴专家，教育部“新世纪优秀人才”，中国图象图形学学会第七、第八届常务理事，现任中国图象图形学学会情感计算与理解专业委员会主任，曾任第一届国际情感计算会议 ACII2005 大会共同主席，国际互联网多媒体计算与服务会议 ACM ICIMCS2010 大会主席，2021 世界人工智能大会哈尔滨分会论坛主席，2023 年中国图象图形学学会青年科学家大会共同主席、2023 第一届情感智能大会主席。主要研究领域为计算机视觉智能、多媒体数据分析与理解、情感计算等。已发表 ICCV，CVPR，ACM MM 等顶级国际会议及 IJCV，TPAMI，TIP，TMM 等高影响因子国际期刊文章学术论文 300 余篇，H 指数 >50，Google 引用量 10000+，入选全球人工智能 TOP 2000 单学者榜单，中国人工智能 100 位榜单人物。获国家发明专利 30 余项，出版教材 6 部。主持或负责完成国家自然基金重点项目 4 项，主持新一代人工人智能国家科技重大专项课题 1 项，主持完成国家自然基金面上项目 8 项，完成国家 863、973 项目以及国际合作项目多项。获国家及省部级自然科学奖项 4 项，获中国计算机专业优秀教师奖 (2021)。已培养成长为“国家级高层次人才”、“国家级青年人才”、“拔尖人才”等教授、副教授十余名，省、校、行业“优博”“优硕”称号 5 人，省级优秀毕业生十余名。

金琴 中国人民大学

个人简介：金琴，中国人民大学信息学院教授，博士生导师。分别于清华大学计算机科学与技术系获得学士与硕士学位，美国 Carnegie Mellon University 获得博士学位。主要研究领域包括多媒体智能计算，人机交互。在多模态情感计算、视觉语言描述生成、跨模态交互等研究与应用中取得了突出成果，在国际顶级学术会议发表论文百余篇，包括 CVPR，ICCV，NeurIPS，ACM Multimedia，ACL，AAAI 等。蝉联多项国际权威竞赛冠军，包括：2017-2023 年 TRECVID Video-To-Text（VTT）评测冠军； 2018-2020 年 CVPR ActivityNet Dense Video Captioning 竞赛冠军；2017-2019 年 ACM Multimedia Audio-Visual Emotion Challenge (AVEC) 竞赛冠军；2019 年之江杯全球人工智能大赛视频内容描述生成冠军等。荣获 ACM Multimedia 2017 Best Grand Challenge Paper Award，ICMR 2018 Best Paper Runner-up。主持多项国家与省部级科研项目。担任 CCF A 类国际会议 ACM Multimedia 2022 Technical Program 程序委员会主席。担任国际期刊 ACM Transactions on Multimedia Computing Communications and Applications 的 Associate Editor 副主编。

宋睿华 中国人民大学

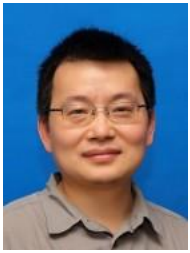
个人简介：宋睿华，中国人民大学高瓴人工智能学院长聘副教授，国家高层次人才特聘专家。曾任微软亚洲研究院主管研究员、微软小冰首席科学家。近期研究兴趣为多模态理解、创作和交互。发表学术论文 100 余篇，申请专利 30 余项。曾获 WWW 2004 最佳论文提名奖，AIRS 2012 最佳论文奖，和 CLWS 2019 优秀论文奖，2022 年度教育部自然科学一等奖。她的算法完成了人类史上第一本人工智能创作的诗集《阳光失了玻璃窗》。2021-2022 年作为学术带头人，发布文澜系列多模态预训练大模型，并成功落地快手、OPPO

等企业。2023 年，参与发布玉兰大语言模型，完成从基础模型、对话模型到智能体模型的自研训练。曾担任 SIGIR 短文和讲习班主席，ACL 的领域主席，EMNLP 的资深领域主席，和 Information Retrieval Journal 的主编。

山世光 中科院计算所

个人简介：山世光，中国科学院计算技术研究所研究员/博导，智能信息处理重点实验室主任，智能算法安全重点实验室（中国科学院）副主任，IEEE Fellow。研究领域为计算机视觉、模式识别和机器学习。已发表论文 400 余篇，其中 CCF A 类论文 180 余篇，论文被谷歌学术引用 3.7 万余次。研究成果获 2005 年度国家科技进步二等奖、2015 年度国家自然科学二等奖、2021 年度北京市科技进步二等奖、2022 年度中国图象图形学学会自然科学一等奖。他是国家特支计划领军人才，基金委优青，国务院特殊津贴专家，北京市科技新星，人社部国家百千万人才工程有突出贡献中青年专家，CCF 青年科学家奖获得者，中科院青促会优秀会员，腾讯科学探索奖获得者。他是中国人工智能学会(CAAI)模式识别专委会副主任，CAAI 情感智能专委会副主任，中国计算机学会(CCF)青年工作委员会副主任，CCF 计算机视觉专委会常务委员，VALSE 的联合发起人和运营者。

潘纲 浙江大学



个人简介：潘纲，浙江大学计算机学院教授、博导，脑机智能全国重点实验室主任，国家杰出青年基金获得者，入选国家“万人计划”科技创新领军人才，计算机系统结构与网络安全研究所所长，中国人工智能学会会士、常务理事、脑机融合与生物机器智能专委会主任委员，中国自动化学会机器人智能专业委员会副主任委员，中国神经科学学会类脑智能分会副主任委员、脑机接口与交互分会副主任委员。主要研究方向为人工智能、脑机接口、类脑计算、计算机视觉、普适计算等。已发表学术论文 200 多篇，获授权发明专利 60 多项。获 CCF-IEEE CS 青年科学家奖(2016)、IEEE TCSC Award for Excellence (Middle Career Researcher, 2018)，入选 2016 年中国高等学校十大科技进展、2020 年世界互联网领先科技成果等。获两次国际会议或期刊的时间考验论文奖（Test-of-Time Paper Award），指导学生获 CCF-A 类国际会议的最佳论文奖、最佳论文提名奖多次。目前担任 IEEE Trans. Neural Networks and Learning Systems、Cognitive Neurodynamics 等多个国际期刊编委。

董晶 中科院自动化所

组织者简介：董晶，中科院自动化所智能感知与计算研究中心研究员。现为中国图象图形学学会(CSIG)理事、副秘书长，CSIG 女科技工作者委员会秘书长，北京图象图形学学会常务理事、青工委主任，中国人工智能学会(CAAI)理事、杰出会员，IEEE/CCF/CSIG 高级会员，中国科学院青年创新促进会会员，IEEE 亚太区执委（2017-2022），IEEE 信号处理协会全球会员发展主席（2022-2024），IEEE 亚太区人道主义科技活动委员会主席（2019-2022），IEEE 亚太区女工程师委员会主席（2017-2018）。担任 IAPR《Newsletter》主编，Elsevier《Journal of Information Security and Application》国际期刊副主编。曾获 2016 年度 1BM 教职人员奖、2018 年度国际模式识别大会最佳科技论文奖、2019 年度中国人工智能学会杰出贡献会员奖、2020 年度 CSIG 石青云女科学家奖（青年组）、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。主要从事计算机视觉、生物特征识别、多媒体内容取证与安全等 AI 前沿方向的技术研究，已在国际权威期刊及学术会议上发表学术论文 60 余篇，申请发明专利 30 余项，其中已授权 18 项中国专利含 3 项美国专利。主持或主要参与了国家 863 计划 973 计划，科技支撑计划、重点研发计划、国家自然科学基金、北京市杰青等 20 余项国家和省部级科研项目。

刘偲 北京航空航天大学

个人简介：刘偲，北京航空航天大学教授，博导。国家级青年人才。担任中国图象图形学学会理事、副秘书长。主持企业创新发展联合基金重点支持项目，担任科技创新 2030—重大项目课题负责人。研究方向是跨模态智能分析、目标检测和跟踪。共发表了 CCF A 类论文 80 余篇，含 IEEE TPAMI 10 篇。Google Scholar 引用 12000+ 次。获国家科技进步二等奖(9/10)，中国图象图形学学会自然科学奖一等奖（1/5），CCF-腾讯犀牛鸟专利奖、吴文俊人工智能优青奖、CSIG 石青云女科学家奖。获多媒体领域顶会 ACM MM 2012 最佳技术演示奖，ACM MM 2013、ACM MM 2021 最佳论文奖，以及人工智能领域顶会 IJCAI 2021 最佳视频奖。获 ChinaMM 2018 最佳学生论文奖和 PRCV 2020 最佳论文提名奖。多次担任 ICCV、CVPR、ECCV、ACM MM、NeurIPS 等顶级会议领域主席（AC）。担任 IEEE TMM、IEEE TCSVT、CVIU 编委（AE）。获得 10 余项 CCF A 类会议竞赛冠军。

Workshop 15: 三维重建与生成

组织者：郑伟诗（中山大学）、章国锋（浙江大学）、张青（中山大学）

时间：5 月 6 日（周一）8:30-12:00　　地点：欢悦厅

时间	主持人	内容
8:30-9:00	郑伟诗	讲者：王程（厦门大学） 题目：基于激光雷达三维视觉的全球定位
9:00-9:30		讲者：徐凯（国防科技大学） 题目：鲁棒可扩展的实时三维重建—相机跟踪优化与地图表示学习
9:30-10:00	章国锋	讲者：廖依伊（浙江大学） 题目：面向自动驾驶的高写实仿真：从重建到生成
10:00-10:30		讲者：许岚（上海科技大学） 题目：神经渲染和体积视频的一些思考
10:30-10:40	中场休息	
10:40-11:10	张青	讲者：齐晓娟（香港大学） 题目：Embodied Environment Generation via Reconstruction, Decomposition and Beyond
11:10-11:40		讲者：方杰明（华为） 题目：基于 Gaussian Splatting 的 3D/4D 内容重建和生成

王程 厦门大学

报告题目：基于激光雷达三维视觉的全球定位

报告摘要：全球定位在数字经济中占据核心地位，但城市复杂环境限制了卫星定位的应用。三维激光扫描技术，凭借精确的三维感知能力，正成为城市定位的新曙光。本报告将介绍厦门大学 ASC 实验室在激光雷达视觉全球定位方面的研究进展。首先，我将解释基于隐式表达的激光雷达视觉定位的基本原理。接着，将介绍从深度回归到几何编码的高效定位方法，展示国际首个达到亚米级定位精度的大范围激光雷达全球定位成果，汇报获得 CVPR2024 满分评分的扩散模型激光雷达视觉定位。最后，我将总结在该领域的努力，并展望传统的三维视觉任务与感知全球定位的叠加发展趋势。

讲者简介：厦门大学南强重点岗位教授，教务处处长，国家级人才计划基金获得者，入选国家“万人计划”科技创新领军人才，IET 会士。担任福建省智慧城市感知与计算重点实验室主任。研究兴趣包括三维视觉、激光雷达、遥感数据智能处理。发表 200 余篇论文，包括 Nature Communication，CVPR，NeurIPS，ICLR，AAAI 等期刊和会议，谷歌引用 10000+，h-index=55。他还是 ISPRS 的多传感器集成与融合工作组的主席，中国计算机学会 YOCSEF 厦门分论坛 (创始)主席，中国图象图形学会常务理事。曾获得国际摄影测量与遥感学会（ISPRS）Giuseppe Inghilleri 奖，获得省部级科技进步一等奖等奖励 5 项。

徐凯 国防科技大学

报告题目：鲁棒可扩展的实时三维重建——相机跟踪优化与地图表示学习

报告摘要：作为空间智能的重要使能技术，实时 RGB-D 稠密三维重建已在增强现实、具身智能等领域广泛应用。以往方法大多面临相机跟踪鲁棒性和在线建图扩展性等难题。相机跟踪方面，过快的相机运动将导致相机跟踪失败和重建错误。我们提出基于随机优化的方法，仅依靠深度信息，实现了快速相机运动下的实时三维重建。该方法在没有全局位姿优化和回环检测的情况下，达到了 SOTA 的相机跟踪和三维重建精度。我们将该方法拓展应用于深度-惯导 SLAM，实现了更快速（线速度达 5 m/s）的相机跟踪和更稳定的在线重建。在线建图方面，现有的显式体地图表示存储开销大，难以支撑大场景的在线重建。我们提出一种多神经隐式场表示，随着扫描进程动态、增量式地分配神经子地图，并基于局部神经集束调整进行学习。在后端，多个子地图可以联合学习以实现全局回环优化。为提高重建鲁棒性，我们首次在神经隐式重建中实现了基于随机优化的相机跟踪。最后简要介绍我们在显式-隐式融合表示的实时三维重建方面的最新工作。

讲者简介：徐凯，国防科技大学教授，国家杰青、优青。普林斯顿大学访问学者。研究方向为计算机图形学、三维视觉、具身智能、数字孪生等。在国际上较早开展了数据驱动三维感知、建模与交互工作，提出面向复杂三维数据的结构化感知、建模与交互理论方法系统。相关成果已成功应用于发射场三维数字孪生建设，基于 AI+3D 视觉的重工智能制造产线核心技术及软件系统研发等。发表 TOG/TPAMI/TVCG 等 A 类论文 90 余篇，其中图形学顶会 SIGGRAPH 论文 29 篇（第一作者 10 篇）。担任图形领域顶级国际期刊 ACM Transactions on Graphics 的编委，以及多个领域重要会议的主席。任中国图象图形学会三维视觉专委会副主任、中国工业与应用数学学会几何设计与计算专委会副主任。获湖南省自然科学一等奖 2 项（排名 1 和 3）、中国计算机学会自然科学一等奖（排名 3）、军队科技进步二等奖、军队教学成果二等奖等。

廖依伊 浙江大学

报告题目：面向自动驾驶的高写实仿真：从重建到生成

报告摘要：高写实度仿真平台对自动驾驶具有重要价值，而人工制作的仿真平台通常成本高昂且无法逼真还原现实世界。本次汇报将分享系列利用神经渲染技术从现实世界构建高写实度驾驶场景的工作，从现实场景逼真重建出发，到无限街景三维高效生成，构建具有逼真外观、丰富语义和多自由度控制能力的城市场景。

讲者简介：廖依伊，浙江大学信电学院特聘研究员。2018 年获浙江大学博士学位，2018 年到 2021 年在德国马普所从事博士后研究。研究兴趣为三维视觉，包括三维生成、场景重建、沉浸式媒体编码等。在 TPAMI、CVPR、ICCV、NeurIPS 等顶级期刊和会议发表文章三十余篇，谷歌学术引用超 3300 次，其中论文获欧洲 AIGameDev 最具科学前景奖， 牵头搭建的 KITTI-360 数据集受到广泛应用且获 ESI 高被引。担任 3DV 2025 程序主席，CVPR2023、NeurIPS 2023、CVPR 2024 领域主席，IEEE 视频处理与通信技术委员会（VSPC）委员。获浙江大学启真优秀青年学者、2023 百度 AI 华人女性青年学者。

许岚 上海科技大学

报告题目：神经渲染和体积视频的一些思考

报告摘要：深度学习和神经表示技术的发展，为高质量的动静态场景重建、渲染和生成都带来新突破。并且进一步随着 5G 网络、虚拟数字人、元宇宙等应用场景的蓄势待发，其需求也变得越发迫切。本次报告结合过去一年课题组在浙西方面的科研进展，重点探讨基于神经表示的体积视频技术和以人为中心的生成技术，并且展望未来体积视频渲染和生成相结合的潜在方向。

讲者简介：许岚博士，上海科技大学信息科学与技术学院助理教授、研究员、博士生导师，MARS 实验室主任。他的研究方向聚焦于计算机视觉、计算机图形学和计算摄像学，致力于光场智能重建理论与技术的研究，突破了动态神经辐照场和虚拟数字人的一批核心关键技术，率团队研制了系列光场装置，为人工智能推动的超写实数字人提供了新范式。他在 CVPR、SIGGRAPH、IEEE TPAMI 等顶级刊物发表数十篇文章，并多次担任人工智能顶级会议 CVPR、ICCV、AAAI 等领域主席。

齐晓娟 香港大学

报告题目：Embodied Environment Generation via Reconstruction, Decomposition and Beyond

报告摘要：Building intelligent agents that can accurately perceive their environment, reason, and plan accordingly is a long-term goal of the field of AI. Undoubtedly, these abilities are acquired by human beings through observing and interacting with the physical world and learning from past experiences. However, training intelligent agents such as autonomous vehicles and robots using real-world data is challenging due to the high costs and potential safety risks involved. In this presentation, I will discuss our recent efforts in developing 3D decomposed reconstruction techniques for creating interactive 3D environments. Additionally, I will introduce our initiative to leverage Google Street View images to build a platform for evaluating and developing embodied agents.

Finally, I will also touch upon future research directions in the field of generative AI for simulation and embodied intelligence.

讲者简介：齐晓娟博士（https://xjqi.github.io）是香港大学电气与电子工程系的助理教授。她从香港中文大学获得博士学位，并曾在多伦多大学、牛津大学和英特尔视觉计算组工作和交流。她致力于赋予机器在开放世界中感知、理解和重构视觉世界的能力，并推动它们在具身智能中的应用。她在 CVPR、ICCV、NeurIPS 等顶级计算机视觉和机器学习会议上发表了 60 余篇论文，其中多篇论文受邀进行口头报告。她曾受邀担任 ICCV 2021、CVPR 2021、AAAI 2021、AAAI 2022、CVPR 2023、NeurIPS 2023 和 CVPR 2024 的领域主席。

方杰明 华为

报告题目：基于 Gaussian Splatting 的 3D/4D 内容重建和生成

报告摘要：3D Gaussian Splatting (GS) 在快速建模并实时渲染 3D 场景上取得了巨大的成功，相比于以 NeRF 为代表的神经渲染方法，GS 具有更强的显式性和更高的渲染效率，在 3D 领域颇具潜力。本报告将分享讲者围绕 GS 在 3D/4D 内容重建和生成等任务中做出的一系列探索，包括 4D 场景实时渲染 (4D Gaussian Splatting)、文本生成 3D (GaussianDreamer)、文本编辑 3D (GaussianEditor) 等研究工作。

讲者简介：方杰民 博士，华为云高级研究员、3D 视觉方向助理首席专家，分别于 2023 年和 2018 年在华中科技大学电信学院获得博士和学士学位。其主要研究方向为 3D 计算机视觉、神经架构搜索/设计，于 TPAMI, CVPR, ICLR, SIGGRAPH Asia 等期刊/会议发表论文 20 余篇，谷歌学术引用 1000 余次；获 CVM 期刊 2021 年度最佳论文奖；作为负责人获第八届中国国际“互联网+”创新创业大赛全国金奖；获 2022 年度“中国大学生自强之星标兵”；获华中科技大学十大研究生“学术新星”。

郑伟诗 中山大学

个人简介：郑伟诗博士，中山大学计算机学院教授/副院长、教育部“长江学者奖励计划”特聘教授、英国皇家学会牛顿高级学者，现任教育部机器智能与先进计算重点实验室主任。长期研究协同与交互分析理论与方法，解决人体建模和机器人行为的视觉计算问题。担任 IEEE T-PAMI 等期刊的编委。主持承担国家级重点类项目和人才项目 5 项、以及广东省自然科学基金委卓越青年团队(负责人)项目等。获中国图象图形学学会自然科学奖一等奖、广东省自然科学奖一等奖等、国家教学成果奖二等奖等。

章国锋 浙江大学

个人简介：章国锋，浙江大学教授，博导，国家优秀青年科学基金获得者。主要从事三维视觉与增强现实方面的研究，尤其在同步定位与地图构建（SLAM）和三维重建方面取得了一系列重要成果，研制了一系列相关软件 （http://www.zjucvg.net），并开源了一系列 SfM/SLAM 系统或关键模块算法的源代码 (https://github.com/zju3dv/)，是 OpenXRLab 扩展现实开源平台（https://openxrlab.org.cn/）的主要发起人。获全国优秀博士学位论文奖、计算机学会优秀博士学位论文奖、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（排名第 4）、浙江省自然科学奖一等奖（排名第 2）、浙江省技术发明奖一等奖（排名第 4）、2023 年度吴文俊人工智能科学技术奖科技进步奖一等奖（排名第 7）以及混合现实和增强现实领域国际顶级会议 ISMAR 2020 唯一最佳论文奖。目前为《Virtual Reality & Intelligent Hardware》、《计算机辅助设计与图形学学报》和《应用科学学报》编委，《中国图象图形学报》青年编委，中国图象图形学学会三维视觉专委会副主任，浙江省人工智能学会增强现实分会副会长，增强现实核心技术产业联盟副理事长；曾担任 VALSE 第二届和第三届资深 AC 委员会主席，VALSE 2019、2021 大会程序委员会主席，VALSE 2022 大会主席，ChinaVR 2021 大会程序委员会主席，CVPR 2021、2023 领域主席，ISMAR 2019-2023 以及 VR 2021-2023 程序委员会委员。

张青 中山大学

个人简介：张青，中山大学计算机学院副教授。主要从事计算摄像学和计算机图形学等方面的研究，特别是图像生成与编辑，以及基于图像的三维建模和绘制等。在 ACM TOG、TPAMI、IJCV 等期刊以及 SIGGRAPH (Asia)、CVPR 等会议发表论文 40 余篇，谷歌学术引用 3100 余次。主持或参与多项国家级科研项目。获 2019 年湖北省自然科学二等奖, 2022 年世界人工智能大会优秀青年论文奖, CCF CAD&CG 2021 最佳论文奖。

Workshop 16: 艺术智能

组织者：霍静（南京大学）、金鑫（北京通用人工智能研究院）、秦越（南京艺术学院）

时间：5 月 6 日（周一）8:30-12:00　地点：博悦厅

时间	主持人	内容
8:30-9:00	霍静	讲者：姚鸿勋（哈尔滨工业大学） 题目：人工智能的舞蹈创作与情感融入
9:00-9:30		讲者：张克俊（浙江大学） 题目：智能艺术与创新设计———智能篆刻
9:30-10:00		讲者：董未名（中国科学院自动化研究所） 题目：绘画中的 AI
10:00-10:30		讲者：蔡新元（华中科技大学） 题目：人工智能艺术的观念与路径
10:30-10:40	中场休息	
10:40-11:10	金鑫	讲者：薄一航（北京电影学院） 题目：计算电影——电影研究的新范式
11:10-11:40		讲者：许多（天津音乐学院） 题目：人工智能生成音乐的客观评价方法探析
11:40-12:10	秦越	Panel 嘉宾： 姚鸿勋（哈尔滨工业大学）、董未名（中国科学院自动化研究所）、蔡新元（华中科技大学）、张克俊（浙江大学）、薄一航（北京电影学院）、许多（天津音乐学院）、宋睿华（中国人民大学）、霍静（南京大学），金鑫（北京通用人工智能研究院）

姚鸿勋 哈尔滨工业大学

报告题目：人工智能的舞蹈创作与情感融入

报告摘要：随着科技的发展，智能化视觉生成已经深入到艺术创作与娱乐生活中，不仅可以应用于一些现实场景，还可以满足很多社交媒体中的娱乐趣味性需求。报告将重点介绍基于音乐驱动的自动编舞生成和基于情感表达的舞蹈动作生成两个跨模态动作生成研究工作。通过基于音乐和情感的人工智能舞蹈创作，探索人体动作的复杂性，讨论舞蹈动作生成相关领域面临的问题挑战和发展趋势。

讲者简介： 姚鸿勋，哈尔滨工业大学计算学部长聘教授，黑龙江省政府特殊津贴专家，教育部"新世纪优秀人才"，中国图象图形学学会第七、第八届常务理事，现任中国图象图形学学会情感计算与理解专业委员会主任，曾任第一届国际情感计算会议 ACII2005 大会共同主席，国际互联网多媒体计算与服务会议 ACM ICIMCS2010 大会主席，2021 世界人工智能大会哈尔滨分会论坛主席，2023 年中国图象图形学学会青年科学家大会共同主席、2023 第一届情感智能大会主席。主要研究领域为计算机视觉智能、多媒体数据分析与理解、情感计算等。已发表 ICCV,CVPR, ACMMM 等顶级国际会议及 IJCV, TPAMI, TIP,TMM 等高影响因子国际期刊文章学术论文 300 余篇，H 指数>50，Google 引用量 10000+，入全球人工智能 TOP 2000 榜单学者榜单，中国人工智能 100 位榜单人物。获国家发明专利 30 余项,出版教材 6 部。主持或负责完成国家自然基金重点项目 4 项，主持新一代人工人智能国家科技重大专项课题 1 项，主持完成国家自然基金面上项目 8 项，完成国家 863、973 项目以及国际合作项目多项。获国家及省部级自然科学奖项 4 项，获中国计算机专业优秀教师奖(2021)。已培养成长为“国家级高层次人才”、“国家级青年人才”、“拔尖人才”等教授、副教授十余名,省、校、行业“优博”“优硕”称号 5 人，省级优秀毕业生十余名。

张克俊 浙江大学

报告题目：智能艺术与创新设计———智能篆刻

报告摘要：张克俊，教授、博士生导师。浙江大学计算机科学与技术学院工业设计系系主任、国际设计研究院副院长。专注于智能艺术与创新设计、人工智能与情感计算，承担国家自然科学基金、国家重点研发计划课题、国家文化与旅游科技创新等项目，研究成果发表于 IEEE/ACM 汇刊、Design Studies、SCIENCE CHINA-Information Science 等重要期刊和会议，曾获教育部高等学校优秀科学研究成果奖二等奖、世界人工智能大会最高奖卓越人工智能引领者奖。专注于科技设计人才培养，主持教育部“设计+X”创新创业教育改革虚拟教研室，主持国家级一流课程 2 门，指导学生获红点、IF 奖、中国智造大奖、好设计奖 10 余项及教育部“互联网+”大赛国赛季军 1 项（金奖 5 项），并获浙江省首届高校教师教学创新大赛特等奖以及国家级教学成果奖二等奖 3 项。

讲者简介：张克俊，教授、博士生导师。浙江大学计算机科学与技术学院工业设计系系主任、国际设计研究院副院长。专注于智能艺术与创新设计、人工智能与情感计算，承担国家自然科学基金、国家重点研发计划课题、国家文化与旅游科技创新等项目，研究成果发表于 IEEE/ACM 汇刊、Design Studies、SCIENCE CHINA-Information Science 等重要期刊和会议，曾获教育部高等学校优秀科学研究成果奖二等奖、世界人工智能大会最高奖卓越人工智能引领者奖。专注于科技设计人才培养，主持教育部“设计+X”创新创业教育改

革虚拟教研室，主持国家级一流课程 2 门，指导学生获红点、IF 奖、中国智造大奖、好设计奖 10 余项及教育部“互联网+”大赛国赛季军 1 项（金奖 5 项），并获浙江省首届高校教师教学创新大赛特等奖以及国家级教学成果奖二等奖 3 项。

董未名 中国科学院自动化研究所

报告题目：绘画中的 AI

报告摘要：AI 绘画是人工智能生成内容（AIGC）领域热门的方向之一。近期，随着多模态大模型和扩散模型技术的迅速发展，由人工智能生成的绘画作品在艺术性和内容丰富度方面都有了极大的提升。本次报告将回顾 AI 绘画技术的发展历程，介绍图像/视频风格迁移、文字引导的艺术图像/视频生成和多模态信息引导的艺术图像/视频生成等 AI 绘画技术的基本原理，并展示由相关技术生成的艺术作品。另外，还将探讨 AI 绘画与人类艺术家创作之间的关系，并对 AI 绘画技术未来的理论研究和应用发展方向进行展望。

讲者简介：董未名，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，博士生导师，中国电影美术学会理事，CCF 计算艺术分会常务委员。长期从事计算艺术研究，在包括 ACM TOG、IEEE TVCG、IEEE TIP、SIGGRAPPH 和 CVPR 等重要国际期刊和国际会议发表学术论文百余篇。主持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课题等国家项目以及腾讯、快手、中文在线和爱奇艺等企业合作项目。成果应用于腾讯天天 P 图、快手魔法滤镜、爱奇艺秀场和 Follow 相机等多项产品中。获中国电影美术学会学术理论贡献奖。

蔡新元 华中科技大学

报告题目：人工智能艺术的观念与路径

报告摘要：人工智能是国家战略的重要组成部分，是未来国际竞争的焦点和经济发展的新引擎。在艺术设计领域，人工智能的介入不仅带来了创意工具的转化，还带来了艺术与设计观念上的变革，“人工智能艺术”（Artifcial Intelligence Art，AI Art）应运而生，其具备与现代主义设计截然不同的创意生产规律、步骤、工具和审美法则。我们的学科、行业、产业、社会服务等都面临新的“人工智能+”构建。本次报告通过拆解新时代下技术与艺术的关系，探讨新语境下的、设计思维和逻辑的转向，剖析新模式下人才培养机制的变革。同时，依托 AI 绘本、AI 展演、AI 元宇宙、AI 光影秀等新一代人工智能示范应用场景，全面阐释人工智能新质生产力的领域拓新、产业落实、行业培育路径。

讲者简介：蔡新元，男，博士，华中科技大学建筑与城市规划学院教授、博导，副院长；光影交互服务技术文化和旅游部重点实验室主任，数字光影技术湖北省工程研究中心主任，教育部动画、数字媒体专业教学指导委员会委员，中组部国家特聘专家。

薄一航 北京电影学院

报告题目：计算电影——电影研究的新范式

报告摘要：计算电影，作为计算人文学科新的研究分支，是人工智能时代下电影研究的新范式。随着机器学习、模式识别、计算机视觉、计算机听觉以及自然语言处理等相关人工智能技术的成熟与进步，面对今天数字电影中多模态的数据类型，诸多算法和计算工具可以应用于电影的研究当中。计算电影针对电影本体、电影创作与制作过程以及电影市场和观众三个研究对象展开研究，借助机器学习、算法设计、模式识别等计算机技术，旨在优化和改进电影制作流程，提升创作效率与质量，拓展电影本体研究的理论和方法。结合传统的电影研究方法，计算电影的出现会给新时代的电影研究注入更加新鲜血液。

讲者简介：薄一航，北京电影学院美术学院副教授，硕士生导师。北京交通大学与 UC Irvine 联合培养博士生，分别于中科院自动化所模式识别国家重点实验室与 Boston College 完成博士后研究。任中国电影美术学会理事、中国图象图形学会女工委委员、中国图象图形学会人机交互专委会委员以及中国人工智能学会情感智能专委会委员。研究方向：人工智能、计算机视觉、交互艺术设计、情感计算等。近几年主持并完成国家自然科学基金青年基金、国家社科基金艺术学项目、北京市教委自然科学项目等多个科研项目，发表 SCI、EI 检索及国内核心期刊学术论文多篇，出版专著《数字图像程序基础——从一个矩形画起》、《虚拟空间交互艺术设计》和《虚拟空间技术》、《艺术程序设计讲义》等多部。

许多 天津音乐学院

报告题目：人工智能生成音乐的客观评价方法探析

报告摘要：黑格尔将美分为主观和客观，对于音乐而言，虽然纯艺术的美以主观感性认知为主要衡量标准，但当声音具有某种和谐、对称或秩序的特征时，它们通常具有某种客观的美感。现在人工智能生成的音乐数量巨大，良莠不齐，本组研究将提供一种自动衡量的方法以节省人工筛选成本。

讲者简介：天津音乐学院艺术管理系副教授，曾获留学基金委资助“艺术类人才培养特殊项目”赴美做神经美学研究。现任中国音乐家协会会员、中国自动化学会普及委员会副秘书长、中国计算机学会高级会员、计算艺术专委会委员，中国技术经济学会神经经济管理专业委员会常务委员，国家自然基金函评评委等。主持国家自然科学基金、教育部人文社科项目等国家级与省部级项目 6 项，主研国家社科基金项目等国家与省部级项目 10 项。在 SCI、EI 检索及《人民音乐》、《音乐艺术》等国内外核心期刊发表文章 30 余篇。

霍静 南京大学

个人简介：霍 静 ，南京大学计算机科学与技术系准聘副教授，分别于 2017 年、2011 年在南京大学计算机科学与技术系获得博士学位以及在南京师范大学强化培养学院获得学士学位。曾分别在英国曼彻斯特大学，香港大学等高校进行学术访问交流。2022 年入选国家级青年人才项目，2018 年获得江苏省计算机学会优博，2018 年获得江苏省科学技术奖二等奖。研究方向为新型机器学习技术，包括视

频图像生成式模型及其智能创意应用、强化学习及其智能感知与决策应用。主持国家自然科学基金面上基金、青年基金各 1 项，江苏省自然科学基金青年基金 1 项，参与国家自然科学基金重大项目 1 项，科技部 2030 新一代人工智能项目 2 项等。在相关研究领域的期刊会议发表论文 50 余篇，包括 CVPR，ICCV，AAAI，IJCAI，ACM MM，TPAMI，TIP，TMM，TNNLS、TCYB、PR 等。

金鑫 北京电子科技学院/北京通用人工智能研究院

个人简介：金鑫，北京电子科技学院副教授，可视计算与安全实验室（victory-lab）主任，北京通用人工智能研究院 AI 音乐项目负责人。北京航空航天大学 CS 博士，师从赵沁平院士；莲花山计算机视觉研究院访问学生，师从朱松纯教授；清华大学访问学者，师从戴琼海院士与刘烨斌教授。主要研究方向为计算美学、计算机视觉、人工智能安全。在计算机视觉、图像处理、人工智能、多媒体技术等领域的重要学术期刊与会议发表论文 80 余篇，包括 TIP、TMM、TITS、TOMM、SC-IS、ACM MM、AAAI、IJCAI、CVPR、ECCV 等。获国家发明专利授权 43 项、2023 金海鸥影视荣誉作品杰出技术贡献奖、2023 年中国电影美术学会荣誉盛典学术理论贡献奖、ROSENET2018/19 最佳论文奖、ISAIR2017/23 最佳论文奖等，参与了全国信标委国家标准《信息技术 计算机视觉术语》的制定。主持或参与国家级科研项目 9 项，省部级与企业合作项目 20 余项（含华为、阿里、兵器工业、中石化等）。CCF 高级会员，担任 ISAIR 常务委员、CSIG 视觉大数据专委会常务委员、CIE 虚拟现实分会副秘书长、CCF 计算机视觉专委会、计算艺术分会执行委员、中国电影美术学会虚拟空间专委会委员等。

秦越 南京艺术学院

个人简介：秦越，南京艺术学院现代音乐与科技学院专业教师，博士后,主要研究方向为：音乐科技。江苏高校“青蓝工程”优秀青年骨干教师，“海粟人才”青年骨干人才。中国音乐家协会电子音乐学会、中国人工智能学会艺术与科技专委会会员。科研方面，先后主持关于音乐人工智能方面的江苏省社会科学基金项目、江苏省教育科学规划项目、江苏高校哲学社会科学项目、高峰计划科研项目等近 10 项，在核心期刊等发表学术论文近 10。艺创方面，担任第 35 届中国电影金鸡奖开幕式作曲，北京奥运会博物馆主题音乐作曲，国家艺术基金大型舞剧《十二秒》作曲。多首原创歌曲被新华社、中国教育电视台、学习强国平台等官方媒体转载。先后创作院线级电影音乐、音乐剧、舞剧音乐作品二十余部，编创歌曲近百首。

Workshop 17: 优秀学生论坛

组织者：韩晓光（香港中文大学（深圳））、胡建芳（中山大学）、任冬伟（哈尔滨工业大学）、张鼎文（西北工业大学）

时间 5 月 7 日（周二）8:30-12:00　　地点：温德姆宴会厅

时间	主持人	内容
8:30-8:40	任冬伟	讲者：李志琦（南京大学） 题目：开环端到端自动驾驶的诸多问题
8:40-8:50		讲者：穆尧（香港大学） 题目：具身智能大模型与通用机器人系统
8:50-9:00		讲者：李明晗（香港理工大学） 题目：从图片到视频分割的通用大模型
9:00~9:10		讲者：王立元（清华大学） 题目：生物智能启发的持续学习理论与方法
9:10~9:20		讲者：古祥（西安交通大学） 题目：最优传输引导的条件得分扩散模型
9:20-9:30		讲者：杨灵（北京大学） 题目： 扩散模型算法与应用
9:30-9:40		讲者：周尚辰（新加坡南洋理工大学） 题目：Diffusion Models for Low-Level Vision
9:40-9:50		讲者：夏俊（西湖大学） 题目：利用预训练模型解读生物化学密码
9:50-10:20	胡建芳	科研经验交流
10:20~10:30	中场休息	
10:30~11:00	张鼎文	圆桌分组交流
11:00~11:40	韩晓光	主题分享讨论
11:40-12:00		Panel 嘉宾： 胡建芳（中山大学）、任冬伟（哈尔滨工业大学）、张鼎文（西北工业大学）、李志琦（南京大学）、穆尧（香港大学）、李明晗（香港理工大学）、王立元（清华大学）、古祥（西安交通大学）、杨灵（北京大学）、周尚辰（新加坡南洋理工大学）、夏俊（西湖大学）

李志琦 南京大学

报告题目：开环端到端自动驾驶的诸多问题

报告摘要：端到端自动驾驶是实现高阶自动驾驶的一条可信赖路径，然而现阶段在真实数据上进行端到端自动驾驶训练仅能依靠开环方案。本次我将分享现有开环自动驾驶方案的诸多不足，希望引起社区重新思考未来端到端的发展方向。

讲者简介：南京大学博士生，研究方向为计算机视觉和自动驾驶感知，在 CVPR/ICCV/ECCV 等会议上发表多篇论文。论文 BEVFormer 入选 2022 年引用最高的 100 篇论文。获得 2022 年 Waymo 自动驾驶挑战赛纯视觉赛道冠军，2023 年 CVPR 自动驾驶挑战赛占用网格赛道冠军。获得 2024 年英伟达奖学金。

穆尧 香港大学

报告题目：具身智能大模型与通用机器人系统

报告摘要：随着 AIGC 的迅速发展，具身智能与通用机器人系统已成为研究的前沿领域。通过整合大模型、CV 和机器人控制等先进技术，我们正朝着更智能、自主和高效的机器人系统迈进，并在多领域发挥重要作用。本次报告将聚焦于面向开放世界，拥有具身认知、规划、执行的能力的具身智能大模型 RoboCodeX 及 RoboScript 通用机器人代码生成评测平台。RoboCodeX 采用树状结构，将复杂的人类指令细化为多个以对象为中心的操作单元。RoboScript 则致力于通过代码生成，实现机器人操作的快速部署；不仅验证了代码及仿真的准确性，还揭示了不同大模型在处理复杂物理交互时的性能差异。

讲者简介：穆尧，香港大学在读博士生, 师从罗平教授，共在 NeurIPS, ICML, ICLR, CVPR 等顶会顶刊发表论文 16 篇, 累计发表文章 30 余篇, 曾获 ICCAS2020 大会最优学生论文奖，IEEE IV2021 最优学生论文提名奖等多项学术奖励, 于 2021 年在清华大学取得硕士学位，荣获香港博士政府奖学金，香港大学校长奖学金，国家奖学金，清华大学优秀硕士毕业生，清华大学优秀硕士论文奖等荣誉称号。研究方向: 具身智能、强化学习、机器人控制和自动驾驶。个人主页：yaomarkmu.github.io

李明晗 香港理工大学

报告题目：从图片到视频分割的通用大模型

报告摘要：图像分割的通用模型因其能够处理多种分割任务而受到瞩目。这些模型使用统一架构和一套权重，灵活适应全景、交互式和文本指代分割等任务。视频分割则更复杂，因为它不仅要处理图片分割，同时还要在连续的帧中准确追踪物体。解决不同视频分割任务的时序问题尤其棘手。例如，全景分割要求识别给定词表中包含的所有物体，而物体分割要在视频的后续帧中追踪第一帧中标记的特定目标。这些目标可能包含任何非常规物体，甚至是物体的一部分。我们的报告深入分析了当前最先进的视频分割通用模型框架，并比较了它们在不同任务和数据集上的效能及局限性。通过这些分析，我们旨在推动图片和视频分割技术的未来发展。

讲者简介：李明晗是香港理工大学计算系的博士生，由张磊教授指导。在攻读博士学位之前，她于 2019 年在西安交通大学获得数学与统计学硕士学位，导师为孟德宇教授。她的研究兴趣涵盖计算机视觉和机器学习领域，专注于多模态基础模型、图像/视频分割和视频恢复算法。在她的学术生涯中，李明晗在顶级会议和期刊上发表了 8 篇文章。作为第一作者，她在 CVPR 上发表了 4 篇文章，在 TIP 上发表了 1 篇文章。李明晗曾在阿里巴巴达摩院 IDST 及 OPPO 研究院担任算法工程师实习生，在 2018 年和 2020-2024 年积累了工业界经验。她曾担任计算机视觉领域顶级会议（包括 CVPR、ICCV 和 ECCV）以及《信号处理》期刊的审稿人。

王立元 清华大学

报告题目：生物智能启发的持续学习理论与方法

报告摘要：为了有效适应现实世界的动态变化，智能系统需要持续不断地获取、更新、积累和利用知识。这种能力被统称为持续学习或终身学习，是生物智能系统在“适者生存”的自然选择过程中形成的重要优势，但对于目前的人工智能系统来说却非常困难。如何充分挖掘和有效利用生物智能的持续学习机理，建立适用于生物智能与人工智能的通用计算模型，是我们长期以来一直在探索的重要课题。在本次报告中，我将简要介绍我们在生物智能启发持续学习方面的最新进展，包括对持续学习的优化目标在常规设定和预训练背景下的理论分析，以及对权重正则化、模块化结构、数据回放与生成回放、预训练模型微调等主流算法思路的前沿探索。

讲者简介：王立元先后在清华大学获得理学学士和理学博士学位，获得清华大学优秀毕业生、北京市优秀毕业生等荣誉，入选清华大学“水木学者”计划。他的研究兴趣为机器学习和神经科学的交叉方向，目前的研究方向主要包括持续学习、终身学习、脑启发 AI、以及 AI4Science，研究成果以第一作者发表在 Nature Machine Intelligence (被评为封面文章)、TPAMI、TNNLS、NeurIPS、ICLR、CVPR、ICCV、ECCV 等人工智能的顶级期刊和会议。

古祥 西安交通大学

报告题目：最优传输引导的条件得分扩散模型

报告摘要：条件得分扩散模型在文本引导图像生成、图像转换等诸多领域取得了显著成功，然而，现有条件得分模型的训练通常依赖于配对的源数据（条件）与目标数据，而现实世界的数据经常是非配对或者部分配对的。本报告中，我们针对非配对或部分配对设置的数据生成问题，提出了最优传输引导的条件得分扩散模型（OTCS），利用经典无监督最优传输以及我们建立的半监督最优传输数学模型分别构建非配对或部分配对数据的耦合关系，进一步将现有配对设置下的条件得分模型拓展到非配对和部分配对场景，拓展了条件得分模型的实用范围。非配对图像超分辨、半配对图像语义翻译等应用任务的实验证实了提出的 OTCS 的有效性。

讲者简介：古祥，西安交通大学数学与统计学院博士生，导师为徐宗本院士。主要研究人工智能（AI）的跨域泛化性数学方法、生成式 AI 模型的最优传输理论与算法。在人工智能领域顶级期刊（IEEE TPAMI 等）和顶级国际会议（NeurIPS, CVPR 等）发表论文 13 篇，其中包括第一作者论文 6 篇，共同一作论文 4 篇。担任 NeurIPS、ICLR、ICML、CVPR、ICCV、ECCV 等顶级会议程序委员会委员，以及 IEEE TPAMI、IEEE TIP 等顶级期刊审稿

人。

杨灵 北京大学

报告题目：扩散模型算法与应用

报告摘要：扩散模型以其良好的可控性和逼真的生成质量成为这两年来最火的生成模型之一，在各个领域都展现出优越的性能。然而，在一些复杂的条件生成场景下，如文生图、视频编辑以及药物发现等，扩散模型依然无法很好地应对。本次分享将介绍如何通过改进扩散模型算法、引入外部辅助（例如大模型）来提升扩散模型应对复杂条件生成的能力。

讲者简介：北京大学博士在读三年级，导师为崔斌教授，杨灵担任北大信息技术高等研究院人工智能与数据安全实验室专家，主编 AIGC 专著《扩散模型：生成式 AI 模型的理论、应用与代码实践》，一作在 CVPR/NeurIPS/ICML/ICLR/TKDE 等国际人工智能顶刊顶会上发表论文 10 余篇，长期担任 TPAMI, ICML, ICLR, NeurIPS, CVPR, KDD, AAAI 等顶刊顶会审稿人，主导并开源了文生图最强框架 RPG，性能超越 OpenAI 的 DALL-E 3，和 OpenAI、斯坦福大学等知名研究机构长期在文生图/视频扩散模型、大模型以及图模型等研究领域进行合作探索。

周尚辰 新加坡南洋理工大学

报告题目：Diffusion Models for Low-Level Vision

报告摘要：近来，随着视觉生成模型的日新月异，它们正逐渐成为推动底层视觉任务探索的重要引擎。生成模型通过其强大的生成先验极大地提升了底层视觉模型的性能，使输出画质得到质的提升。然而，强大的生成能力在底层视觉任务中也是一把双刃剑，模型输出常常伴随着明显的纹理改变，给底层视觉模型的设计带来新的问题和挑战。本报告主要探讨 Diffusion 生成模型如何驱动不同底层视觉任务的探索与进步。

讲者简介：新加坡南洋理工大学 MMLab@NTU 四年级博士生，师从 Chen Change Loy 教授。他在计算机视觉领域的顶级会议和期刊，如 CVPR、ICCV、ECCV、NeurIPS、TPAMI 等上发表了多篇论文。他曾获得 WAIC 青年优秀论文提名奖，ICIMCS 最佳论文奖，并在 NTIRE 视频复原比赛中获得三项冠军。他还与 ECCV2022、CVPR2023 和 CVPR2024 会议上参与组织了移动智能摄影和成像（MIPI）国际系列研讨会。目前主要研究兴趣包括图像/视频增强、生成模型。个人主页：https://shangchenzhou.com

夏俊 西湖大学

报告题目：利用预训练模型解读生物化学密码

报告摘要：近年来，预训练大模型的出现，如 ChatGPT，彻底改变了计算机理解和生成语言的方式。受到它们的启发，预训练模型在计算生物化学领域的应用有着巨大的潜力，可以解读生化密码并加速科学进步。在这次演讲中，我将介绍我们在化学小分子和蛋白质组学预训练模型研发方面的重要进展。首先，我将展示我们在化学小分子预训练模型方面的工作。我们开发了领域知识增强的化学小分子预训练模型 Mole-BERT，它在各种药物发现任务中表现出了良好的性能。这些模型利用了化学领域知识，实现

了准确的属性预测和潜在药物候选物的高效筛选。此外，我们还收集整理了目前世界上最大的蛋白质组学质谱数据集，并在此基础上研发 UniMass 蛋白质组质谱大模型。UniMass 在计算蛋白质组学的多种任务中取得了最先进的结果，包括肽段从头测序、谱图预测等。这些进展对于理解蛋白质功能及其在疾病机制中的作用具有深远意义。这次演讲将强调预训练模型在解读生化密码方面的变革性影响，并突出它们在加速计算生物化学领域科学突破中的潜力。

讲者简介：夏俊，西湖大学与浙江大学联合培养四年级直博生，师从人工智能讲席教授李子青(Stan Z. Li), 主要的研究方向是机器学习及其在生物化学中的应用。目前，在人工智能顶级会议期刊如 ICML, NeurIPS, ICLR 发表论文多篇，其中一作论文被 PaperDigest 选为 WWW 会议最有影响力论文之一，另有多篇论文入选 AI 顶会 Oral 或 Spotlight。曾入选 KAUST 评选的 Rising Stars in AI, 获得苹果奖学金提名（中国大陆 9 人），西湖大学校长奖章，浙江大学国家奖学金等荣誉和奖励。

韩晓光 香港中文大学（深圳）

讲者简介：韩晓光博士，现任香港中文大学（深圳）理工学院助理教授，校长青年学者。他于 2017 年获得香港大学计算机科学专业博士学位。其研究方向包括计算机视觉和计算机图形学等，在该方向著名国际期刊和会议已发表论文近 100 篇，包括顶级会议和期刊 SIGGRAPH(Asia), CVPR, ICCV, ECCV, NeurIPS, ACM TOG, IEEE TPAMI 等。他曾获得吴文俊人工智能优秀青年奖，广东省杰出青年基金资助，香港中文大学（深圳）青年科研奖。目前也担任 CVPR2023/2024，NeurIPS 2023 以及 ECCV2024 领域主席，同时也是 IEEE TVCG 的编委。他的工作曾两次获得 CCF 图形开源数据集奖（DeepFashion3D 和 MVImgNet），曾两次入选 CVPR 最佳论文列表。

胡建芳 中山大学

讲者简介：胡建芳博士，目前为中山大学副教授，博士生导师，针对不同应用场景下的视频解析问题进行了系统性研究，提出了一系列性能优秀的算法与模型解决视频行为识别、行为意图预测和视频-文本定位等问题。在 IEEE TPAMI、IEEE TIP 等国际顶级期刊和 CVPR、ICCV、NurIPS 等国际顶级会议发表多篇学术论文。主持多项国家自然科学基金、广东省科学基金和企业委托横向项目，获广东省杰出青年基金（2022 年）支持，广东省科学技术奖自然科学奖二等奖（2020 年）、中国图象图形学学会优秀博士学位论文奖（2017 年）。

任冬伟 哈尔滨工业大学

讲者简介：任冬伟，哈尔滨工业大学副教授，博士生导师。研究方向为计算机视觉，包括图像和视频复原以及物体检测等，在 TPAMI、TIP、CVPR、ICCV、ECCV、AAAI 等国际期刊和会议发表论文四十余篇，谷歌学术引用 6000 余次。主持国家自然科学基金面上项目、青年项目，主持黑龙江省优秀青年基金，获黑龙江省自然科学一等奖。

张鼎文 西北工业大学

讲者简介：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, IJCV, 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、中国图象图形学学会优秀博士论文奖等奖励。担任中国图象图形学学会青年工作委员会副秘书长、中国图象图形学学会优博俱乐部副主席，任 IEEE TMM、TCSVT、PR 等刊物（客座）编辑。

Workshop 18: 多模态大模型

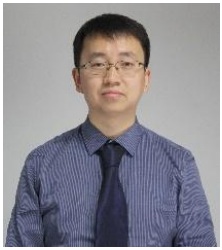
组织者：白翔（华中科技大学）、徐天阳（江南大学）、沈为（上海交通大学）、刘禹良（华中科技大学）

时间：5 月 7 日（周二）8:30-12:00 地点：欣悦厅

时间	主持人	内容
8:30-9:05	白翔	讲者：王士进（科大讯飞） 题目：通用人工智能技术进展和典型应用
9:05-9:40		讲者：王文海（香港中文大学） 题目：图文多模态大模型的研究与应用
9:40-10:15	徐天阳	讲者：丁二锐（百度） 题目：文心·CV 大模型 VIMER 闭环：数据视角
10:15-10:23		企业宣讲：联想研究院 联想端侧智能体解决方案（师忠超）
10:23-10:35	中场休息	
10:35-11:10	沈为	讲者：张博（浙江大学） 题目：面向实际场景体验的多模态大模型 DeepSeek VL
11:00-11:35		讲者：刘禹良（华中科技大学） 题目：通用多模态大模型细节描述能力提升方法
11:35-11:50	刘禹良	Panel 嘉宾： 白翔（华中科技大学）、王士进（科大讯飞）、王文海（香港中文大学）、丁二锐（百度）、张博（浙江大学）、沈为（上海交通大学）、徐天阳（江南大学）

王士进 科大讯飞

报告题目：通用人工智能技术进展和典型应用



报告摘要：本报告首先分析了人工智能的阶段，并提出当前以认知大模型为代表的通用人工智能技术引发全球广泛关注，然后还分析了从认知大模型到多模态大模型的技术特性、发展趋势及应用价值。其次，报告汇报了科大讯飞研发星火大模型的成果和研发经历，最后重点介绍了大模型服务多个行业的探索经验。

讲者简介：王士进博士，正高级工程师，现任科大讯飞副总裁、AI 研究院执行院长、认知智能全国重点实验室副主任。王士进博士主要从事自然语言处理、认知大模型等方向研究，承担多项国家重点研发计划，曾获安徽省科学技术进步奖一等奖、吴文俊人工智能科技进步奖一等奖、中国科协求是杰出青年成果转化奖。他主导了科大讯飞“认知智能大模型技术及应用”专项，发布的讯飞星火大模型达到国内领先水平。

王文海 香港中文大学

报告题目：图文多模态大模型的研究与应用

报告摘要：在当前人工智能领域，融合图像和文本信息的多模态大型模型正迅速成为关注焦点，催生出了 GPT-4V 和 Gemini 等创新性的产品。这些模型以大语言模型技术为基础，通过整合视觉和文本数据，在理解和生成复杂内容方面展现出了强大能力。本报告将探讨图文多模态大模型的基本原理和关键技术，涉及大规模视觉模型训练及大规模图文模型的对齐。其次，报告还将介绍图文多模态大模型在各个领域中的应用前景，并讨其在实际应用中的潜力与挑战。

讲者简介：南京大学博士，香港中文大学博士后。研究方向为视觉基础模型研究，上海人工智能实验室“书生”系列视觉基础模型核心开发者。主要成果发表在顶级期刊和会议 TPAMI、CVPR、ICCV、ECCV、ICLR、NeurIPS 等共 43 篇论文，其中 19 篇为一作/共一/通信。研究成果获得了总共超 1.6 万次引用，单篇最高引用超 3000 次。研究成果分别入选 CVPR 2023 最佳论文，世界人工智能大会青年优秀论文奖，CVMJ 2022 最佳论文提名奖，两次入选 ESI 高被引论文（前 1%）和热点论文（前 0.1%），6 次入选 Paper Digest CVPR、ICCV、NeurIPS、ECCV 年度十大最具影响力论文，入选 2023 年 CSIG 优博提名，斯坦福大学 2023 年度全球前 2%顶尖科学家。

丁二锐 百度

报告题目：文心・CV 大模型 VIMER 闭环：数据视角

报告摘要：面对视觉大模型在产业应用中存在的“数据少、模型多、落地难”鸿沟，VIMER 大模型闭环实践一年多来在效率及效果提升上取得了长足进步。当前，数据工作正逐渐从数据众包走向数据技术，逐步形成了数据挖掘、生成、标注、清洗四大特色。本报告，以多模态为线条，重点介绍四大特色技术近期的研究成果。从数据数量角度，着重介绍多模态大模型如何实现了开放域中的长尾稀缺数据挖掘，2D/3D AIGC 技术如何助力生成合理的、难以获取的数据。

从数据质量角度，着重介绍大小模型融合如何进行高效的全方位标注，尤其借助多模态提示下的“人机回环”密集标注，以及如何利用图文互补的复合表征提纯数据。

讲者简介：丁二锐，博士，正高级工程师，百度视觉技术部总监，2023 年被华中科技大学聘为“湖北产业教授”。长期从事计算机视觉及图形学的技术研发与产业应用工作。在图文跨模态、弱/半监督上的工作先后获国家技术发明奖二等奖（2020)、中国专利银奖（2022）、中国电子学会科技进步一等奖（2018）、地理信息科学进步一等奖（2023）。曾获 ICDAR 2019 最佳论文第二名奖、IEEE FG 2023 最佳学生论文奖。作为部门负责人，带领团队有力支撑了搜索、智能驾驶、智能云等核心业务的行业领先地位，其中，计算机视觉公有云连续 3 年被 IDC 评为市场第一，视觉检测技术还获得了北京市发明专利一等奖（2021）。

张博 浙江大学

报告题目： 面向实际场景体验的多模态大模型 DeepSeek VL

报告摘要：最近，开源多模态模型取得一系列重要进展，但在实际应用场景中仍面临诸多挑战。我们在设计 DeepSeek VL 时，以用户实际体验为中心，进行了一系列从数据、模型结构到训练策略的设计。首先，基于大模型无损压缩理论，我们通过大规模且多样化的数据预训练，赋予模型丰富的世界知识。在微调数据方面,我们针对实际使用场景构建了分类体系,并有针对性地收集微调样本,优化用户体验。模型结构上，我们的模型采用混合视觉编码器，实现对高分辨率输入的精细处理，捕捉全局语义信息及细节信息。同时，我们的训练同时权衡了语言、多模态能力，使得训练得到的通用模型在两种模态上保持竞争力。我们通过合理的构建评测体系，使得我们的实验可以在小模型上得到充分探究。DeepSeek-VL 系列多模态模型现实世界的应用中展示了优秀的用户体验，在相同模型大小的各种视觉语言基准测试中取得具有竞争力的性能，同时在以语言为中心的基准测试中保持了出色的性能。我们的实验报告相机介绍完整的实验训练细节，并同时开源模型参数。1.3B 的小模型可供端侧使用。我们希望该项目的完整开源，可为领域提供参照，并加速该领域的创新与发展。

讲者简介：浙江大学计算机学院 CAD&CG 国家重点实验室“百人计划”研究员、博士生导师。2019 年于香港科技大学获电子与计算机工程博士学位，此前 2013 年于浙江大学光电信息工程学院获学士学位。2019 年至 2023 年工作于微软亚洲研究院计算视觉组，任主管研究员。研究方向为深度内容生成技术（包括 2D、3D 内容生成）、虚拟人建模、多模态模型以及具身智能方向的研究。在内容生成领域有多项代表性工作。如高质量图像翻译 CoCosNet 系列工作（CoCosNet v2 获 CVPR 2021 最佳论文候选）、业界首个文生图生成扩散模型 VQ-Diffusion, 首个高质量 3D 扩散生成模型 Rodin，3D 生成技术 DreamCraft 3D，知名开源多模态大模型 DeepSeek-VL 等。此外，研究工作（Bringing old photos back to life）曾获被国外知名媒体 louisbouchard.ai 列为 2020 年三十大 AI 进展。近年来在 CVPR、ICCV、ECCV、ICLR、SIGGRAPH、TPAMI 等顶级刊物以核心作者身份发表文章 30 余篇。个人主页：bo-zhang.me

刘禹良 华中科技大学

报告题目：多模态大模型 Monkey 及其在文档智能中的应用

报告摘要：多模态大模型 Monkey 专为处理高分辨率图像设计，能有效识别和解析复杂视觉信息。该模型将图像切割成统一大小的小块逐块处理，每块大小与训练中使用的视觉编码器的尺寸相匹配，提升了对细节的捕捉能力，并支持处理高分辨率图像。Monkey 采用多层次描述生成方法，丰富场景与对象的关联描述，增强语言输出的详细性和准确性。在 18 个视觉语言数据集上的实验表明，Monkey 在图像标题生成和视觉问答任务中表现出较强的竞争力。基于 Monkey 模型，进一步针对文档任务引入移位窗口注意力机制和零初始化，提高了更高分辨率输入时的跨窗口连接性，优化了早期训练稳定性。进一步，通过筛选重复的图像标记简化了处理流程，提升了模型性能。同时该模型增强了文本定位和位置解释性，精确执行屏幕代理任务和文本识别。在多达 12 个的文本基准以及在 OCRBench 的评估测试中，该方法比过去的开源方法展现出更好的性能。

讲者简介：华中科技大学人工智能与自动化学院研究员，博士生导师。曾获评 CSIG 优秀博士论文，第九届中国科协青年托举人才、教育部海外高层次引才专项，湖北省百人计划，及华为学者。担任中国科学-信息科学专刊编委、中国图象图形学报编委等职务。主持国家自然科学基金青年项目及两项国家重点研发计划子课题。在一流国际期刊和会议如 PAMI、CVPR 等发表论文 40 余篇，其中第一作者 13 篇（8 篇引用过百，多篇获获口头报告邀请，国际顶会满分论文，及期刊封面论文），通讯 10 篇。曾获第六届及第八届互联网+ 金奖，获 CVPR、ICDAR 等国际权威学术竞赛冠军 6 项。9 项研发成果应用于金山办公、Adobe 等国内外知名企业。主持研发的 Monkey 多模态大模型（CVPR 2024）曾在 Meta AI 公认的 OpenCompass 多模态大模型排行榜中名列开源模型榜首。

Workshop 19: 端到端自动驾驶

主席：李弘扬（上海人工智能实验室）、张兆翔（中国科学院自动化研究所）、杨言超（香港大学）、马超(上海交通大学)

时间：5 月 7 日（周二）13:30-18:00　　地点：智悦厅

时间	主持人	内容
14:00-14:35	李弘扬	讲者：陈韫韬（中国科学院香港创新研究院） 题目：基于视频世界模型的闭环端到端自动驾驶
14:35-15:10		讲者：李镇（香港中文大学(深圳)） 题目：3D 车道线检测及多模态点云增强解析
15:10-15:45	张兆翔	讲者：李国法（重庆大学） 题目：全工况自动驾驶闭环工具链系统研究
15:45-15:55	中场休息	
15:55-16:30	杨言超	讲者：张力（复旦大学） 题目：大规模自动驾驶仿真系统研究
16:30-17:05	马超	讲者：冷佳旭（重庆邮电大学） 题目：面向自动驾驶的脑启发可信场景分析
17:05-17:40		Panel 嘉宾： 陈韫韬（中国科学院香港创新研究院）、李镇（香港中文大学(深圳)）、李国法（重庆大学）、张力（复旦大学）、冷佳旭（重庆邮电大学）、杨泽同（上海人工智能实验室）、金鑫（东方理工）、史少帅（滴滴）、李弘扬（上海人工智能实验室）、张兆翔（中国科学院自动化研究所）、杨言超（香港大学）

陈韫韬（中国科学院香港创新研究院）

报告题目：基于视频世界模型的闭环端到端自动驾驶

报告摘要：在自动驾驶中,提前预测未来事件并评估可预见的风险,可以使自动驾驶汽车更好地规划行动,提高道路安全性和效率。为此,我们提出了一种全新的闭环端到端方法，第一次提出现有端到端规划模型兼容的驾驶世界模型。通过视图分解促进的联合时空建模,我们的模型可以在驾驶场景中生成高保真度的多视图视频。基于其强大的生成能力,我们首次展示了将世界模型应用于安全驾驶规划的潜力。特别是,我们的方法能够根据不同的驾驶动作生成多个未来视频,并根据基于对未来图像的反馈确定最佳轨迹。在真实世界驾驶数据集上的评估验证了我们的方法可以生成高质量、一致且可控的多视图驾驶视频,为真实世界模拟和安全规划开辟了可能性。

讲者简介：陈韫韬，中国科学院香港创新研究院人工智能与机器人创新研究中心助理教授，2021 年毕业于中科院自动化研究所模式识别与智能系统专业。他有着丰富的自动驾驶从业经验，先后在图森未来和辉羲智能承担感知算法研究员与算法负责人的工作。他的主要研究方向为开放环境下的通用视觉感知与智能体的复杂行为决策。他在 TPAMI、IJCV、JMLR、R-AL、NeurIPS、CVPR、ICCV、ECCV、AAAI 等国际顶级期刊与会议上发表了多篇论文，Google Scholar 引用次数超过 2000 次。长期担任 TPAMI、IJCV、TIP、R-AL、NeurIPS、ICLR、ICML、CVPR、ICCV、ECCV 等国际顶级期刊会议的审稿人与 PRCV 领域主席。

李镇 香港中文大学(深圳)

报告题目：3D 车道线检测及多模态点云增强解析

报告摘要：在自动驾驶应用中，准确识别车道线是实现高级驾驶辅助系统(ADAS)和全自动驾驶的关键基础任务。由于相机价格低廉且能采集密集视觉信息等优势，基于视觉的 3D 车道线感知成为了研究的热点课题。我们首先介绍借助 Transformer 中的注意力机制及动态路面更新的创新 LATR 模型，取得了最优的单目 3D 车道线检测性能。另一方面，尽管激光雷达的应用成本较高，但它能够提供准确的 3D 位置结构信息，也在逐步被一些车企采用且在未来随着成本的降低也将会更广泛地部署到车端。为了探究拥有互补信息的点云数据与图像信息，我们设计了多模态车道线检测方案，即 DV-3DLane 模型，其通过在透视和俯视空间分别融合图像和激光雷达点云特征，大幅提升了 3D 车道线检测的定位精度，验证了端到端多模态模型设计在 3D 车道线检测下的可行性与有效性。同样地，我们进一步介绍多模态点云增强解析，并将其应用到室外自动驾驶的占据预测及开放场景的室内视觉定位任务。

讲者简介：李镇博士现任香港中文大学（深圳)理工学院助理教授， 未来智联网络研究院助理院长，校长青年学者。李镇博士获得香港大学计算机科学博士学位（2014-2018 年），他还于 2018 年在芝加哥大学担任访问学者。李镇博士荣获 2023 年吴文俊人工智能优秀青年，2021 年中国科协第七届青年托举人才，2023CVPR HOI4D 竞赛第一名，2022 年 SemanticKITTI 语义分割竞赛第一名，2023 年 IROS 最佳论文 Finalist，ICCV2021 Urban3D 竞赛第二名，CASP12 接触图预测全球冠军等。李镇博士还获得了来自于国家、省市级以及工业界的科研项目。李镇博士领导了港中深的 Deep Bit Lab（https://mypage.cuhk.edu.cn/academics/lizhen/ ），其主要的研究方向是 3D 视觉解析及应

用（包括但不限于点云解析，多模态联合解析），深度学习等基础理论算法研究，并致力于将 2D/ 3D 人工智能算法推广应用于交叉学科，自动驾驶，工业视觉等场景中，在该方向著名国际期刊和会议发表论文 60 余篇，包括顶级期刊 Cell Systems, Nature Communications, T-PAMI, TMI, TVCG，TNNLS 等，以及顶级会议 CVPR, ICCV, ECCV, NeurIPS, ICLR，IROS，ACM MM，AAAI, IJCAI, MICCAI 等。李镇博士担任 IEEE Transactions on Mobile Computing、IROS 副编以及众多顶刊顶会的审稿人，李镇博士还是广东院士联合会脑科学与类脑智能专委委员，VALSE、MICS、中国图象图形学学会机器视觉专委会，3DV 专委会等学术组织的委员。

李国法（重庆大学）

报告题目：全工况自动驾驶闭环工具链系统研究

报告摘要：长尾场景发生频次低但危险性大。围绕该问题，建立可编辑、可复用的自动驾驶仿真系统，高保真模拟全工况的自动驾驶环境，尤其关注高速公路匝道汇入汇出口、环岛、十字路口等强交互驾驶场景；提出基于合成数据或半监督学习的舱内/外多目标/状态识别技术，实现在多工况中的有效连续监测；建设全场景的实车路、虚危险虚实联动测试平台，实现强交互场景中的智能决策开发。

讲者简介：李国法，教授，博导，2023 年度小米青年学者、第四届中国科协青年托举人才、深圳市海外高层次人才。2010 年本科毕业于北京理工大学机械与车辆学院，2016 年博士毕业于清华大学汽车工程系（现车辆与运载学院），美国 University of Michigan 联合培养博士。研究方向为人工智能技术在智能网联汽车中的创新与应用。先后主持国家自然科学基金 2 项，中国科协青年人才托举项目 1 项，省部级项目 3 项，全国重点实验室开放基金重点项目 1 项等。以第一作者或通讯作者发表论文 70 余篇，其中，8 篇入选 ESI 高被引论文，1 篇入选热点论文，参编英文专著 1 部、中文专著 3 部，获国家发明专利授权 13 项，美国专利授权 1 项。获机械工程学会突出贡献团队（排第 3）、第六届中国科协优秀科技论文、Automotive Innovation 期刊 2020 年度 Best Paper 等荣誉。担任 IEEE Transactions on Affective Computing 和 IEEE Sensors Journal 的 Associate Editor，《机械工程学报》、《中国公路学报》和《Green Energy and Intelligent Transportation》青年编委。

张力 复旦大学

报告题目：大规模自动驾驶仿真系统研究

报告摘要：近年来，随着自动驾驶技术和仿真系统的快速进步，越来越多的研究致力于开发能够生成和模拟高度真实驾驶环境的系统。这些工作旨在基于复杂的交通场景和动态交通参与者行为，为自动驾驶算法提供丰富的训练和测试环境。然而，现有仿真系统在面对新的驾驶条件或交通场景时，往往受限于其对已有采集数据的依赖，难以实现泛化；虽然通过采用神经辐射场（NeRF）的三维重建技术，并将雷达点云数据作为重建过程的先验，一些研究在提高街景重建的三维一致性方面取得了进展，但在处理更广泛和复杂的动态场景时，这些方法仍面临成本和重建质量的挑战。此外，自动驾驶系统迫切需要解决数据多样性不足和处理复杂光照条件的问题，以满足对广泛和多样化数据的需求。本报告介绍从稀疏视角获取的数据

中生成连续时空场景的高精度仿真数据方法，能够准确捕捉和模拟对象的运动和环境变化，覆盖各种可能的环境条件、光照变化和动态场景，并保持时间和空间上的连续性和一致性，从而大幅提升下游应用的模型训练效率和预测准确性。

讲者简介：张力，复旦大学大数据学院研究员，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。获得国家级高层次青年人才计划；作为项目负责人先后主持国家自然科学基金面上项目、青年项目、临港实验室“求索杰出青年计划”、上海市自然科学基金面上项目；获得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 IEEE TPAMI、IJCV、NeurIPS 等人工智能国际期刊与会议论文 60 余篇；根据谷歌学术，近五年论文总被引 15000 余次。担任人工智能国际会议 NeurIPS 2023，CVPR 2023 与 CVPR 2024 的领域主席，国际期刊 Pattern Recognition 副主编，并组织 CVPR 2023 自动驾驶主题 Workshop。

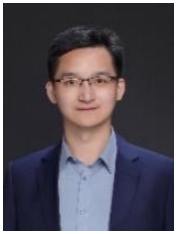
冷佳旭 重庆邮电大学

报告题目：面向自动驾驶的脑启发可信场景分析

报告摘要：在自动驾驶系统中，环境视觉感知的精准性是确保车辆能够自主导航并做出安全决策的关键。当前，尽管视觉感知技术迅速发展，其在处理复杂环境时仍面临可信度挑战，主要表现在视觉场景难以清晰识别、微小目标难以捕捉、以及场景内容难以理解等方面。本报告着重分析这些问题的根源，并探讨如何通过借鉴人脑的认知机理来提高自动驾驶系统视觉感知的可信性。报告中还将简要介绍团队在脑启发的可信场景分析方面的研究工作，并对脑机互鉴的未来发展趋势进行了展望。

讲者简介：冷佳旭，重庆邮电大学计算机科学与技术学院副教授，博士生导师。重庆市博士后创新人才计划入选者，中国计算机学会高级会员。主要从事可信视觉智能方面的研究工作，结合脑启发和类人学习机制的视觉计算模型，在以人为中心的可信场景分析上开展了系统的研究，在面向国家公共安全和智慧校园监管需求的系统平台上取得成功应用。近三年来，主持国家自然科学基金青年项目 1 项，中国博士后科学基金面向项目 1 项，主持省部级重点项目 1 项、一般项目 3 项，参与国自然区域联合重点项目和装备预研教育部联合基金重点项目各 1 项。目前，在国内外重要学术期刊或会议上发表论文 40 余篇，其中以第一作者或通讯作者在 SCI 一区 TOP/CCF A 类国际顶级期刊和会议上发表学术论文 20 余篇，授权国家发明专利共计 10 项，相关成果在甘肃省公安厅和重庆市南岸区公安局得到落地应用。担任 CVPR 和 IEEE TIP 等期刊/会议审稿人。

马超 上海交通大学



组织者简介：马超，上海交通大学人工智能研究院长聘轨副教授，博士生导师。国家优青、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。主要研究计算机视觉与机器学习。谷歌学术引用 1 万余次，连续入选爱思唯尔中国高被引学者 (2020-2023)。担任中国图象图形学学会优博俱乐部轮值主席。获中国图象图形学学会青年科学家奖、华为技术合作领域

2021 年度优秀技术成果奖，研究成果应用于华为达芬奇芯片及其无人驾驶 MDC 平台。

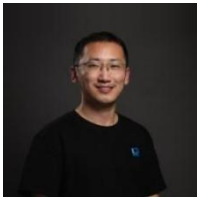
张兆翔 中科院自动化所

组织者简介：张兆翔，博士，研究员，博士生导师，任职于中国科学院自动化研究所模式识别国家重点实验室与智能感知与计算研究中心，是中国科学院大学岗位教授，中国科学院脑科学与智能技术卓越创新中心骨干，入选“教育部 CJ 学者奖励计划”、“国家万人计划青年拔尖人才”和“教育部新世纪优秀人才支持计划”，曾获得中国人工智能学会吴文俊技术发明奖。张兆翔博士的研究兴趣包括：模式识别、计算机视觉与深度学习，具体研究方向包括：视觉认知计算、类脑学习和面向开放环境的视觉感知与理解，在本领域国际主流期刊与会议上发表论文 200 余篇，近五年来在 IEEE T-PAMI、IJCV、JMLR、IEEE T-IP、IEEE T-NN 等顶级期刊与 CVPR、ICCV、ECCV、NIPS、ICLR、AAAI、IJCAI 等顶级会议发表论文 100 余篇，授权专利 20 余项，承担了国家自然科学基金重点项目、国家自然科学基金企业联合重点项目、国家重点研发项目等多项国家级科研项目和企业合作项目。张兆翔博士是 IEEE 高级会员，VALSE 常务 AC，中国计算机学会 CCF 杰出会员、中国人工智能学会 CAAI 杰出会员、中国人工智能学会 CAAI 副秘书长、中国人工智能学会 CAAI 理事、中国人工智能学会 CAAI 模式识别专委会秘书长。张兆翔博士是或者曾经是 IEEE T-CSVT、Patten Recognition、NeuroComputing 编委，担任 CVPR、ICCV、AAAI、IJCAI、ACM MM、ICPR、ACCV 等国际会议的领域主席（Area Chair）。

杨言超 香港大学

组织者简介：杨言超博士是香港大学电气与电子工程系和香港大学同心基金数据科学研究院（HKU IDS）的助理教授。他于 2019 年获得 UCLA 计算机科学博士学位，师从 IEEE/ACM Fellow、亚马逊 AI Lab & AWS 副总裁 Stefano Soatto 教授。此后，他在斯坦福大学担任博士后研究员，与美国三院院士、计算机几何、视觉、图形泰斗 Leonidas Guibas 教授共事。杨言超博士的研究方向包括机器视觉、统计学习、人机交互和具身人工智能，并深耕于自监督感知、推理和表征学习。例如，其工作开创了首个基于深度学习的自监督物体检测分割信息论模型，并建立了首个基于傅里叶空间的域适配框架，为实现通用人工智能、减少智能体对人工的依赖奠定了必要基础。杨言超博士的研究成果主要发表在 CV、ML & Robotics 领域的顶级会议或期刊，担任 CVPR、ECCV 等会议领域主席，并与业界在自动驾驶、智能医疗、虚拟、增强现实等方面有广泛合作。

李弘扬 上海人工智能实验室



组织者简介：李弘扬博士是上海人工智能实验室青年科学家，OpenDriveLab 团队负责人，研究方向聚焦于自动驾驶和具身智能。他于 2019 年在香港中文大学获得博士学位，并在工业界从事多年辅助驾驶研究工作。在 2021 年，他在上海人工智能实验室组建 OpenDriveLab 团队。他提出的鸟瞰图感知工作 BEVFormer，在 2022 年获得了百强影响力人工智能论文荣誉，并在 CES 2024 大会上获得了 Nvidia CEO 黄仁勋和 Mobile Eye CEO Shashua 教授的认可。他主导的端到端自动驾驶项目 UniAD 获得 IEEE CVPR 2023 最佳论文奖。UniAD 在工业界和学术界都产生了巨大的影响，包括特斯拉近期推出的 FSD V12。他还是 NeurIPS 2023 的 Notable Area Chair，IEEE 的高级成员。更多详情，请访问 https://opendrivelab.com

Poster 交流论文一览表

1	Xinke Li(李欣科), Junchi Lu(陆俊池), Henghui Ding(丁恒辉), Changsheng Sun(孙常晟), Joey Tianyi Zhou(周天异), Yeow Meng Chee	新加坡国立大学	AAAI 2024
	PointCVaR: Risk-optimized Outlier Removal for 3D Robust Point Cloud Classification		
2	Chong Yin, Siqi Liu, Kaiyang Zhou, Vincent Wai-Sun Wong, Pong C. Yuen	香港浸会大学	CVPR 2024
	Prompting Vision Foundation Models for Pathology Image Analysis		
3	林靖，曾爱玲，路舜麟，蔡元浩，张瑞茂，王好谦，张磊	IDEA 研究院	NeurIPS 2023
	Motion-X: A Large-scale 3D Expressive Whole-body Human Motion Dataset		
4	许悦聪，杨剑飞，周云娇，陈正华，吴敏，李晓黎	Institute for Infocomm Research, A*STAR, Singapore	ICCV 2023
	Augmenting and Aligning Snippets for Few-Shot Video Domain Adaptation		
5	黄怿豪, 徐觉非，郭青，张杰，吴与桐，胡铭，李恬霖，蒲戈光，刘杨	Nanyang Technological University	AAAI 2024
	Personalization as a shortcut for few-shot backdoor attack against text-to-image diffusion models		
6	李昊颖，Jixin Zhao，Shangchen Zhou，冯华君，李重仪，Chen Change Loy	Nanyang Technological University	ICLR 2024
	Adaptive Window Pruning for Efficient Local Motion Deblurring		
7	曹浩志，许悦聪，杨剑飞，尹鹏宇，袁圣海，谢立华	Nanyang Technological University, Singapore	ICCV 2023
	Multi-Modal Continual Test-Time Adaptation for 3D Semantic Segmentation		
8	杨剑飞，钱汉杰，许悦聪，王锴，谢立华	Nanyang Technological University, Singapore	ICLR 2024
	Can We Evaluate Domain Adaptation Models Without Target-Domain Labels?		
9	Jun Xiao, Zihang Lyu, Cong Zhang, Yakun Ju, Changjian Shui, Kin-man Lam	The Hong Kong Polytechnic University	CVPR 2024

	Towards Progressive Multi-frequency Representation for Image Warping		
10	邹阳，李星源，姜智颖，刘晋源	The University of Sydney	AAAI 2024
	Enhancing Neural Radiance Fields with Adaptive Multi-Exposure Fusion: A Bilevel Optimization Approach for Novel View Synthesis		
11	孙照东 李晓白	University of Oulu	IEEE TPAMI 2024
	Contrast-Phys+: Unsupervised and Weakly-supervised Video-based Remote Physiological Measurement via Spatiotemporal Contrast		
12	陈子恒，宋悦，刘高文, Ramana Rao Kompella, 吴小俊, Nicu Sebe	University of Trento	CVPR 2024
	Riemannian Multinomial Logistics Regression for SPD Neural Networks		
13	康霄阳，杨涛，欧阳雯琪，任沛然，李凌志，谢宣松	阿里巴巴	ICCV 2023
	DDColor: Towards Photo-Realistic Image Colorization via Dual Decoders		
14	叶晴昊，徐国海，严明，徐海洋，钱祺，张佶，黄非	阿里巴巴	ICCV 2023
	Hitea: Hierarchical temporal-aware video-language pre-training		
15	徐海洋，叶晴昊，叶加博，严明，胡安文，刘昊伟，钱祺，张佶，黄非，周靖人	阿里巴巴	CVPR 2024
	mPLUG-Owl2: Revolutionizing Multi-modal Large Language Model with Modality Collaboration		
16	王逍，王世傲，唐川明，朱林，江波，田永鸿，汤进	安徽大学	CVPR 2024
	Event Stream-based Visual Object Tracking: A High-Resolution Benchmark Dataset and A Novel Baseline		
17	洪雨辰，郑乾，赵泠然，Xudong Jiang，Alex C. Kot，施柏鑫	北京大学	IEEE TPAMI 2023
	PAR2Net: End-to-end Panoramic Image Reflection Removal		
18	Ang Li, Yifei Wang, Yiwen Guo, Yisen Wang	北京大学	NeurIPS 2023
	Adversarial Examples Are Not Real Features		
19	王珺，何红亮，魏鹏旭，刘畅，陈杰	北京大学	ICCV 2023
	Toposeg: Topology Aware Nuclear Instance Segmentation		
20	王宇，钟准，乔鹏冲，郑侠武，刘畅, Nicu Sebe, 纪荣嵘，陈杰	北京大学	NeurIPS 2023

	Discover and Align Taxonomic Context Priors for Open-world Semi-Supervised Learning		
21	吴睿海，铁宸睿，杜雨时，赵妍，董豪	北京大学	ICCV 2023
	Leveraging SE(3) Equivariance for Learning 3D Geometric Shape Assembly		
22	冶恩泽，王宇航，张宏，高毅勤，王欢，孙赫	北京大学	ICCV 2023
	Recovering a molecule's 3D dynamics from liquid-phase electron microscopy movies		
23	余济闻，张轩宇，许佑民，张健	北京大学	NeurIPS 2023
	CRoSS: Diffusion Model Makes Controllable, Robust and Secure Image Steganography		
24	朱磊；余琪；陈倩；孟祥溪；耿慕峰；金录嘉；张艺宝；任秋实；卢闫晔	北京大学	TPAMI 2023
	Background-aware Classification Activation Map for Weakly Supervised Object Localization		
25	王宇琪，何嘉伟，范略，李鸿鑫，陈韫韬，张兆翔	北京大学	AAAI 2024
	Driving into the Future: Multiview Visual Forecasting and Planning with World Model for Autonomous Driving		
26	程鑫华，杨田雨，王佳楠，李昱，张磊，张健，袁粒	北京大学	ICLR 2024
	Progressive3D: Progressively Local Editing for Text-to-3D Content Creation with Complex Semantic Prompts		
27	丁健豪，余肇飞，黄铁军, Jian K. Liu	北京大学	AAAI 2024
	Enhancing the Robustness of Spiking Neural Networks with Stochastic Gating Mechanisms		
28	吴虓杨，蒋理，王鹏帅，刘志健，刘希慧，乔宇，欧阳万里，贺通，赵恒爽	香港大学、上海人工智能实验室	CVPR 2024
	Point Transformer V3: Simpler, Faster, Stronger		
29	高志，杜云涛，张欣桐，马晓健，韩文娟、朱松纯、李庆	北京大学	CVPR 2024
	CLOVA: A Closed-Loop Visual Assistant with Tool Usage and Update		
30	蒋超亚，徐海洋，叶蔚，董孟帆，陈家兴，叶晴昊，严明，张佶，黄非，张世琨	北京大学	CVPR 2024
	Hallucination Augmented Contrastive Learning for Multimodal Large Language Model		
31	金录嘉，郭箐，赵时，朱磊，陈倩，任秋实，卢闫晔	北京大学	IJCV 2024

	One-Pot Multi-frame Denoising		
32	金鹏，Ryuichi Takanobu，张万才，操晓春，袁粒	北京大学	CVPR 2024
	Chat-UniVi: Unified Visual Representation Empowers Large Language Models with Image and Video Understanding		
33	雷廷，殷绍峰，刘洋	北京大学	CVPR 2024
	Towards Open-Vocabulary HOI Detection via Conditional Multi-level Decoding and Fine-grained Semantic Enhancement		
34	林志威，刘柘，夏仲禹，王新昊，王勇涛，漆晟翔，董洋，董楠，张乐，朱策	北京大学	CVPR 2024
	RCBEVDet: Radar-camera Fusion in Bird’s Eye View for 3D Object Detection		
35	刘昭辰，李志轩，蒋婷婷	北京大学	AAAI 2024
	BLADE: Box-Level Supervised Amodal Segmentation through Directed Expansion		
36	闵称，赵大伟，肖良，赵健，许新里，朱政，金磊，李健树，郭裕兰，兴军亮，景丽萍，聂一鸣，戴斌	北京大学	CVPR 2024
	DriveWorld: 4D Pre-trained Scene Understanding via World Models for Autonomous Driving		
37	田雨川、陈汉亭、王许涛、白哲源、张清华、李锐锋、许超、王云鹤	北京大学	ICLR 2024
	Multiscale Positive-Unlabeled Detection of AI-Generated Texts		
38	王茜，李玮琦，牟冲，程鑫华，张健	北京大学	CVPR 2024
	360DVD: Controllable Panorama Video Generation with 360-Degree Video Diffusion Model		
39	吴睿海，鲁浩然，王一言，王昱博，董豪	北京大学	CVPR 2024
	UniGarmentManip: A Unified Framework for Category-Level Garment Manipulation via Dense Visual Correspondence		
40	Jiyuan Zhang,Shiyan Chen,Yajing Zheng,Zhaofei Yu,Tingjun Huang	北京大学	CVPR 2024
	Spike-guided Motion Deblurring with Unknown Modal Spatiotemporal Alignment		
41	张心亮，朱磊，何航舟，金录嘉，卢闫晔	北京大学	AAAI 2024
	Scribble Hides Class: Promoting Scribble-based Semantic Segmentation with its Class Label		
42	张轩宇，李润一，余济闻，许佑民，李玮琦，张健	北京大学	CVPR 2024

	EditGuard: Versatile Image Watermarking for Tamper Localization and Copyright Protection		
43	赵睿，熊瑞勤，赵菁，张健，范晓鹏，余肇飞，黄铁军	北京大学	CVPR 2024
	Boosting Spike Camera Image Reconstruction from a Perspective of Dealing with Spike Fluctuations		
44	钟灏峰，洪雨辰，翁书晨，梁锦秀，施柏鑫	北京大学	CVPR 2024
	Language-guided Image Reflection Separation		
45	杨树州，丁莫轩，吴艳敏，李子晗，张健	北京大学深圳研究生院	ICCV 2023
	Implicit Neural Representation for Cooperative Low-light Image Enhancement		
46	乔鹏冲，商磊，刘畅，孙佰贵，季向阳，陈杰	北京大学深圳研究生院	CVPR 2024
	FaceChain-SuDe: Building Derived Class to Inherit Category Attributes for One-shot Subject-Driven Generation		
47	周啸宇，林志威，单骁骏，王勇涛, Deqing Sun, Ming-Hsuan Yang	北京大学王选计算机研究所	CVPR 2024
	DrivingGaussian: Composite Gaussian Splatting for Surrounding Dynamic Autonomous Driving Scenes		
48	郭效君，王一飞，魏泽明，王奕森	北京大学	NeurIPS 2023
	Architecture matters: Uncovering Implicit Mechanisms in Graph Contrastive Learning		
49	杜天祺，王一飞，王奕森	北京大学	ICLR 2024
	On the Role of Discrete Tokenization in Visual Representation Learning		
50	李帅伯，马伟，郭建伟，徐士彪，李本冲，张晓鹏	北京工业大学	CVPR 2024
	UnionFormer: Unified-Learning Transformer with Multi-View Representation for Image Manipulation Detection and Localization		
51	袁彤彤，张轩歌，刘鲲，刘波，陈宸，金键，焦臻帧	北京工业大学	CVPR 2024
	Towards Surveillance Video-and-Language Understanding: New Dataset, Baselines, and Challenges		
52	邓欣，许静怡，高方远，孙先成，徐迈	北京航空航天大学	IEEE 2023
	DeepM2CDL: Deep Multi-scale Multi-modal Convolutional Dictionary Learning Network		
53	李胜曦，张佳璐，李一飞，徐迈*，邓欣，李丽	北京航空航天大学	ICCV 2023

	Neural Characteristic Function Learning for Conditional Image Generation		
54	张晋卿，张延安，刘庆杰，王蕴红	北京航空航天大学	ICCV 2023
	SA-BEV: Generating Semantic-Aware Bird's-Eye-View Feature for Multi-view 3D Object Detection		
55	卓乐，王肇凯，王柏森，廖越，鲍晨曦，彭帅，韩松澔，张爱喜，方非，刘偲	北京航空航天大学	ICCV 2023
	Video background music generation: Dataset, method and evaluation		
56	方若桓，庞观松，百晓	北京航空航天大学	AAAI 2024
	Simple Image-Level Classification Improves Open-Vocabulary Object Detection		
57	高宸健，姜柏言，李星辉，张颖鹏，于茜	北京航空航天大学	CVPR 2024
	GenesisTex: Adapting Image Denoising Diffusion to Texture Space		
58	高渝路，孙奕帆，丁旭东，赵楚洋，刘偲	北京航空航天大学	CVPR 2024
	EASE-DETR: Easing the Competition among Object Queries		
59	樵明朗，徐迈，蒋铼，雷鹏，文师杰，陈运锦，Leonid Sigal	北京航空航天大学	IEEE 2024
	HyperSOR: Context-aware Graph Hypernetwork for Salient Object Ranking		
60	行习铭，周海涛，王闯，张晶，徐东，于茜	北京航空航天大学	CVPR 2024
	SVGDreamer: Text Guided SVG Generation with Diffusion Model		
61	张家伟，李嘉禾，黄雷，于笑寒，谷林，郑锦，百晓	北京航空航天大学	CVPR 2024
	Robust Synthetic-to-Real Transfer for Stereo Matching		
62	周楠、陈佳鑫、黄迪	北京航空航天大学	ICCV 2023
	DR-Tune: Improving Fine-tuning of Pretrained Visual Models by Distribution Regularization with Semantic Calibration		
63	韩坤洋，刘镛，Jun Hao Liew，丁恒辉，魏云超，刘佳军，王一童，唐彦嵩，杨余久，赵耀	北京交通大学	ICCV 2023
	Global Knowledge Calibration for Fast Open-Vocabulary Segmentation		
64	焦思宇，魏云超，王耀威，赵耀, Humphrey Shi	北京交通大学	NeurIPS 2023
	Learning Mask-aware CLIP Representations for Zero-Shot Segmentation		
65	廖康，聂浪，林春雨，郑子硕，赵耀	北京交通大学	ICCV 2023

	RecRecNet: Rectangling rectified wide-angle images by thin-plate spline model and DoF-based curriculum learning		
66	聂浪，林春雨， 廖康，刘帅成， 赵耀	北京交通大学	ICCV 2023
	Parallax-Tolerant Unsupervised Deep Image Stitching		
67	杨尚蓉，林春雨，廖康，赵耀	北京交通大学	ICCV 2023
	Innovating Real Fisheye Image Correction with Dual Diffusion Architecture		
68	朱泓光，魏云超，梁小丹，张淳杰，赵耀	北京交通大学	ICCV 2023
	CTP: Towards Vision-Language Continual Pretraining via Compatible Momentum Contrast and Topology Preservation		
69	方正伟，王睿，黄韬，景丽萍	北京交通大学	CVPR 2024
	Strong Transferable Adversarial Attacks via Ensembled Asymptotically Normal Distribution Learning		
70	回文君，朱振峰，郑帅，赵耀	北京交通大学	CVPR 2024
	Endow SAM with Keen Eyes: Temporal-spatial Prompt Learning for Video Camouflaged Object Detection		
71	刘欢，谭资昌，谭创创，魏云超，王井东，赵耀	北京交通大学	CVPR 2024
	Forgery-aware Adaptive Transformer for Generalizable Synthetic Image Detection		
72	瞿梦雪，武宇，刘武，梁小丹，宋井宽，赵耀，魏云超	北京交通大学	NeurIPS 2024
	RIO: A Benchmark for Reasoning Intention-Oriented Objects in Open Environments		
73	任中伟，黄至铖，魏云超，赵耀，付冬梅，冯佳时，靳潇杰	北京交通大学	CVPR 2024
	PixelLM:Pixel Reasoning with Large Multimodal Model		
74	谭创创，刘欢，赵耀，韦世奎，顾广华，刘平，魏云超	北京交通大学	CVPR 2024
	Rethinking the Up-Sampling Operations in CNN-based Generative Network for Generalizable Deepfake Detection		
75	许璟轩，陈武阳，赵耀，魏云超	北京交通大学	CVPR 2024
	Transferable and Principled Efficiency for Open-Vocabulary Segmentation		
76	宋昆，马惠敏，邹博超，张辉帅，黄维然	北京科技大学	NeurIPS 2023

	FD-Align: Feature Discrimination Alignment for Fine-tuning Pre-Trained Models in Few-Shot Learning		
77	丁鑫龙，陈健生，余宏伟，商宇，秦怡宁，马惠敏	北京科技大学	AAAI 2024
	Transferable Adversarial Attacks for Object Detection using Object-Aware Significant Feature Distortion		
78	徐婧林 郭奕杰 彭宇新	北京科技大学	CVPR 2024
	FinePOSE: Fine-Grained Prompt-Driven 3D Human Pose Estimation via Diffusion Models		
79	余宏伟，陈健生，丁鑫龙，张宇东，唐挺，马惠敏	北京科技大学	AAAI 2024
	Step Vulnerability Guided Mean Fluctuation Adversarial Attack against Conditional Diffusion Models		
80	马文轩，李爽，蔡林灿，康景炫	北京理工大学	NeurIPS 2023
	Language Semantic Graph Guided Data-Efficient Learning		
81	谢斌辉；李爽；郭庆举；刘驰；程新景	北京理工大学	NeurIPS 2023
	Annotator: A Generic Active Learning Baseline for LiDAR Semantic Segmentation		
82	谢蜜雪 李爽 袁龙辉 刘驰 戴泽辉	北京理工大学	NeurIPS 2023
	Evolving Standardization for Continual Domain Generalization over Temporal Drift		
83	冯汉森，王立志，汪彧之，范浩强，黄华	北京理工大学	TPAMI 2024
	Learnability Enhancement for Low-Light Raw Image Denoising: A Data Perspective		
84	齐雅昀 赵文天 吴心筱	北京理工大学	AAAI 2024
	Relational Distant Supervision for Image Captioning without Image-text Pairs		
85	汪启昕, 范超琼, 贾甜远, 韩宇阳, 邬霞	北京师范大学	AAAI 2024
	ND-MRM: Neuronal Diversity Inspired Multisensory Recognition Model		
86	陈梓宁、王伟秋、赵志诚、苏菲、门爱东、Hongying Meng	北京邮电大学	CVPR 2024
	PracticalDG: Perturbation Distillation on Vision-Language Models for Hybrid Domain Generalization		
87	吕昌盛，齐梦实，李夏，杨征元，马华东	北京邮电大学	AAAI 2024
	SGFormer: Semantic Graph Transformer for Point Cloud-based 3D Scene Graph Generation		

88	王涛，金磊，王正，李建树，李亮，赵放，程禹，袁粒，周礼，兴军亮，赵健	北京邮电大学	CVPR 2024
	SynSP: Synergy of Smoothness and Precision in Pose Sequences Refinement		
89	杜昀昊，雷程，赵志诚，苏菲	北京邮电大学	CVPR 2024
	iKUN: Speak to Trackers without Retraining		
90	康奔，陈鑫，王栋，彭厚文，卢湖川	大连理工大学	ICCV 2023
	Exploring Lightweight Hierarchical Vision Transformers for Efficient Visual Tracking		
91	刘晋源，刘铸，吴冠尧，马龙，刘日升，仲维，罗钟铉，樊鑫	大连理工大学	ICCV 2023
	Multi-interactive Feature Learning and a Full-time Multi-modality Benchmark for Image Fusion and Segmentation		
92	张吉庆；董博；傅应锴；王源琛；魏小鹏；尹宝才；杨鑫	大连理工大学	IJCV 2023
	A Universal Event-Based Plug-In Module for Visual Object Tracking in Degraded Conditions		
93	刁海文，万博，张莹，贾旭，卢湖川，陈隆	大连理工大学	CVPR 2024
	UniPT: Universal Parallel Tuning for Transfer Learning with Efficient Parameter and Memory		
94	董文，梅海洋，魏子麒，金傲，仇森，张强，杨鑫	大连理工大学	AAAI 2024
	Exploiting Polarized Material Cues for Robust Car Detection		
95	姜智颖，李星源，刘晋源，樊鑫，刘日升	大连理工大学	AAAI 2024
	Towards Robust Image Stitching: An Adaptive Resistance Learning against Compatible Attacks		
96	吴冠尧；傅泓铭；刘晋源；马龙；樊鑫；刘日升	大连理工大学	AAAI 2024
	Hybrid-Supervised Dual-Search: Leveraging Automatic Learning for Loss-free Multi-Exposure Image Fusion		
97	于佳左，诸葛云志，张璐，胡平，王栋，卢湖川，何友	大连理工大学	CVPR 2024
	Boosting Continual Learning of Vision-Language Models via Mixture-of-Experts Adapters		
98	余晨阳，刘雪虎，王英权，张平平，卢湖川	大连理工大学	AAAI 2024
	TF-CLIP: Learning Text-Free CLIP for Video-Based Person Re-identification		

99	赵骁骐，庞有伟，陈振宇，余谦，张立和，刘瀚淇，左嘉铭，卢湖川	大连理工大学	CVPR 2024
	Towards Automatic Power Battery Detection: New Challenge, Benchmark Dataset and Baseline		
100	王宇皓，刘雪虎，张平平，陆虎，涂铮铮，卢湖川	大连理工大学	AAAI 2024
	TOP-ReID: Multi-Spectral Object Re-identification with Token Permutation		
101	张平平，王宇皓，刘洋，涂铮铮，卢湖川	大连理工大学	CVPR 2024
	Magic Tokens: Select Diverse Tokens for Multi-modal Object Re-Identification		
102	罗旭，吴昊，张继，高联丽，徐菁，宋井宽	电子科技大学	ICML 2023
	A Closer Look at Few-shot Classification Again		
103	张继，吴世涵，高联丽，申恒涛，宋井宽	电子科技大学	CVPR 2024
	DePT: Decoupled Prompt Tuning		
104	谢昊宇；王昶棋；赵健；刘洋；但俊；付冲；孙佰贵	东北大学	IJCV 2024
	PRCL: Probabilistic Representation Contractive Learning for Semi-Supervised Semantic Segmentation		
105	卢雨洁，万龙，丁纳钰，王渝龙，申抒含，蔡梦*，高林*	东华大学	CVPR 2024
	Unsigned Orthogonal Distance Fields: An Accurate Neural Implicit Representation for Diverse 3D Shapes		
106	黄步真，李辰，许重阳，潘亮，王雁刚，Gim Hee Lee	东南大学	CVPR 2024
	Closely Interactive Human Reconstruction with Proxemics and Physics-Guided Adaption		
107	周思帆，李亮，张新雨，张勃，白世鹏，孙淼，赵子瑜，路小波，初祥祥	东南大学	ICLR 2024
	LIDAR-PTQ: Post-Training Quantization for Point Could 3D Object Detection		
108	丛晓峰，桂杰，张敬，后俊明，沈浩	东南大学	CVPR 2024
	A Semi-supervised Nighttime Dehazing Baseline with Spatial-Frequency Aware and Realistic Brightness Constraint		
109	陈斌，印佳丽，陈舒凯，陈柏豪，刘西蒙	福州大学	ICCV 2023

	An Adaptive Model Ensemble Adversarial Attack for Boosting Adversarial Transferability		
110	苏建楠，甘敏，陈光永，印佳丽，陈俊龙	福州大学	IEEE TPAMI 2023
	Global learnable attention for single image super-resolution		
111	Xiangyang Miao, Guobao Xiao, Shiping Wang, Jun Yu	福州大学	AAAI 2024
	BCLNet: Bilateral Consensus Learning for Two-View Correspondence Pruning		
112	丁宇晗，尹富坤，范佳缘，郦辉，陈欣，刘闻，陆崇善，俞刚，陈涛	复旦大学	NeurIPS 2023
	Point Diffusion Implicit Function for Large-scale Scene Neural Representation		
113	江彪，陈欣，刘闻，虞晶怡，俞刚，陈涛	复旦大学	NeurIPS 2023
	MotionGPT: Human Motion as a Foreign Language		
114	陆崇善，尹富坤，陈欣，陈涛，俞刚，范佳媛	复旦大学	ICCV 2023
	A Large-Scale Outdoor Multi-modal Dataset and Benchmark for Novel View Synthesis and Implicit Scene Reconstruction		
115	曹健健，叶鹏，李晟泽，余翀，唐彦嵩，鲁继文，陈涛	复旦大学	CVPR 2024
	MADTP: Multimodal Alignment-Guided Dynamic Token Pruning for Accelerating Vision-Language Transformer		
116	陈思锦，陈欣，张弛，李铭晟，俞刚，费豪，朱宏远，范佳媛，陈涛	复旦大学	CVPR 2024
	LL3DA: Visual Interactive Instruction Tuning for Omni-3D Understanding, Reasoning, and Planning		
117	黄新宇，张又才，马锦玓，田维维，冯瑞，张玥杰，李亚乾，郭彦东，张磊	复旦大学	(ICLR) 2024
	Tag2Text: Guiding Vision-Language Model via Image Tagging		
118	杨依颖，尹富坤，刘闻，范佳媛，陈欣， 于刚，陈涛	复旦大学	AAAI 2024
	PM-INR: Prior-Rich Multi-Modal Implicit Large-Scale Scene Neural Representation		
119	屈德林，严驰，王栋，殷杰，陈其志，张艺亭，徐旦，赵斌，李学龙	复旦大学、上海人工智能实验室	CVPR 2024

	Implicit Event-RGBD Neural SLAM		
120	杨铠成, 邓健康, 安翔, 李嘉伟, 冯子勇, 郭佳, 杨静, 刘同亮	格灵深瞳	ICCV 2023
	ALIP: Adaptive Language-Image Pre-training with Synthetic Caption		
121	李博扬，王应谦，王龙光，张菲，刘婷，林再平，安玮，郭裕兰	国防科技大学	ICCV 2023
	Monte Carlo Linear Clustering with Single-Point Supervision is Enough for Infrared Small Target Detection		
122	梁政宇，王应谦，王龙光，杨俊刚，周石琳，郭裕兰	国防科技大学	ICCV 2023
	Learning Non-Local Spatial-Angular Correlation for Light Field Image Super-Resolution		
123	刘猛，刘悦，梁科，涂文轩，王思为，周思航，刘新旺	国防科技大学	ICLR 2024
	Deep Temporal Graph Clustering		
124	陇盛，陶蔚，李硕豪，雷军，张军	国防科技大学	AAAI 2024
	On the Convergence of an Adaptive Momentum Method for Adversarial Attacks		
125	杨志雄，夏靖远，李胜曦，黄星棹，张双辉，刘振，付耀文，刘永祥	国防科技大学	CVPR 2024
	A Dynamic Kernel Prior Model for Unsupervised Blind Image Super-Resolution		
126	李奇璋，郭怡文，陈浩，左旺孟	哈尔滨工业大学	NeurIPS 2023
	Towards Evaluating Transfer-based Attacks Systematically, Practically, and Fairly		
127	刘晓禹，刘铭，李俊义，Shuai Liu，Xiaotan Wang，LeiLei，左旺孟	哈尔滨工业大学	ICCV 2023
	Beyond Image Borders: Learning Feature Extrapolation for Unbounded Image Composition		
128	尚玮，任冬伟， 冯超宇，汪晓涛， 雷磊， 左旺孟	哈尔滨工业大学	ICCV 2023
	Self-supervised Learning to Bring Dual Reversed Rolling Shutter Images Alive		
129	尹志存，刘铭，李晓明，杨辉，肖龙安，左旺孟	哈尔滨工业大学	ICCV 2023
	MetaF2N: Blind Image Super-Resolution by Learning Efficient Model Adaptation from Faces		

130	潘昱辰，江俊君，江奎，吴志昊，余科垣，刘贤明	哈尔滨工业大学	CVPR 2024
	OpticalDR: A Deep Optical Imaging Model for Privacy-Protective Depression Recognition		
131	王晨阳，江俊君，江奎，刘贤明	哈尔滨工业大学	AAAI 2024
	Low-Light Face Super-resolution via Illumination, Structure, and Texture Associated Representation		
132	吴刚，江俊君，江奎，刘贤明	哈尔滨工业大学	AAAI 2024
	Learning from History: Task-agnostic Model Contrastive Learning for Image Restoration		
133	张志路，王浩宇，Shuai Liu，Xiaotao Wang，Lei Lei，左旺孟	哈尔滨工业大学	ICLR 2024
	Self-Supervised High Dynamic Range Imaging with Multi-Exposure Images in Dynamic Scenes		
134	魏于翔，张亚博，冀志龙，白锦峰，张磊，左旺孟	哈尔滨工业大学、香港理工大学	ICCV 2023
	Elite: Encoding visual concepts into textual embeddings for customized text-to-image generation		
135	黄域青，李鑫，周子坤，王耀威，何震宇，Ming-Hsuan Yang	哈尔滨工业大学(深圳)	CVPR 2024
	RTracker: Recoverable Tracking via PN Tree Structured Memory		
136	徐杭，刘鑫源，徐皓楠，马宜科，朱尊杰，颜成钢，代锋	杭州电子科技大学	CVPR 2024
	Rethinking Boundary Discontinuity Problem for Oriented Object Detection		
137	刘缨杰，刘璇，俞辉，汤璇，魏宪	华东师范大学	NeurIPS 2023
	Learning Dictionary for Visual Attention		
138	李晓凡，张志忠，谭鑫，陈成伟，曲延云，谢源，马利庄	华东师范大学	CVPR 2024
	PromptAD: Learning Prompts with Only Normal Sample for Few-Shot Anomaly Detection		
139	温凌峰，汤璇，杨剑，陈铭松，魏宪	华东师范大学	AAAI 2024
	Hyperbolic Graph Diffusion Model		
140	诸嘉逸，郭青，Felix-徐觉非，黄怿豪，刘杨，蒲戈光	华东师范大学	CVPR 2024
	CosalPure: Learning Concept from Group Images for Robust Co-Saliency Detection		
141	黄思铨，黎弈江，陈冲，时乐宇，高英	华南理工大学	ICCV 2023
	Multi-Metrics Adaptively Identifies Backdoors in Federated Learning		
142	蒋擎，汪嘉鹏，刘崇宇，彭德智，金连文	华南理工大学	ICCV 2023
	Revisiting Scene Text Recognition: A Data Perspective		

143	陈祎旸，段伦豪，赵杉杉，丁长兴，陶大程	华南理工大学	CVPR 2024
	Local-consistent Transformation Learning for Rotation-invariant Point Cloud Analysis		
144	李卓龙，李星奥，丁长兴，徐向民	华南理工大学	CVPR 2024
	Disentangled Pre-training for Human-Object Interaction Detection		
145	谭文韬，丁长兴，江佳瑜，王飞，詹忆冰，陶大鹏	华南理工大学	CVPR 2024
	Harnessing the Power of MLLMs for Transferable Text-to-Image Person ReID		
146	邹文斌，高红霞，叶田，陈亮，杨伟朋，黄莎莎，陈洪盛，陈思翔	华南理工大学	AAAI 2024
	VQCNIR: Clearer Night Image Restoration with Vector-Quantized Codebook		
147	蔡世铝，陈立群，钟胜，颜露新，周嘉欢，邹旭	华中科技大学	AAAI 2024
	Make Lossy Compression Meaningful for Low-Light Images		
148	郭赫 叶紫璇 曹治国 陆昊	华中科技大学	CVPR 2024
	In-Context Matting		
149	李煦蕤、黄子鸣、薛峰、周瑜	华中科技大学	ICLR 2024
	MuSc: Zero-Shot Industrial Anomaly Classification and Segmentation with Mutual Scoring of the Unlabeled Images		
150	刘昊岳，彭仕涵，朱林，昌毅，周寒宇，颜露新	华中科技大学	CVPR 2024
	Seeing motion during nighttime with event camera		
151	汪先奇，许刚伟，贾浩，杨欣	华中科技大学	CVPR 2024
	Selective-Stereo: Adaptive Frequency Information Selection for Stereo Matching		
152	王岩杰，邹旭，颜露新，钟胜，周嘉欢	华中科技大学	CVPR 2024
	SNIDA: Unlocking Few-Shot Object Detection with Non-linear Semantic Decoupling Augmentation		
153	许刚伟，王韵，程俊达，唐金辉，杨欣	华中科技大学	TPAMI 2024
	Accurate and Efficient Stereo Matching via Attention Concatenation Volume		
154	张化鑫，王翔，徐晓豪，卿志武，高常鑫，桑农	华中科技大学	AAAI 2024
	HR-Pro: Point Supervised Temporal Action Localization via Hierarchical Reliability Propagation		

155	周寒宇，昌毅，石志伟，颜露新	华中科技大学	CVPR 2024
	Bring Event into RGB and LiDAR: Hierarchical Visual-Motion Fusion for Scene Flow		
156	左嘉龙，周寒宇，聂迎，张锋，郭天宇，桑农，王云鹤，高常鑫	华中科技大学	CVPR 2024
	UFineBench: Towards Text-based Person Retrieval with Ultra-fine Granularity		
157	孙昊，周明瑶，陈文静，谢伟	华中师范大学	AAAI 2024
	TR-DETR: Task-Reciprocal Transformer for Joint Moment Retrieval and Highlight Detection		
158	林舒源，黄斐然，赖桃桃，赖剑煌，王菡子，翁健	暨南大学	IJCV 2024
	Robust heterogeneous model fitting for multi-source image correspondences		
159	赵少川；徐天阳；吴小俊；Josef Kittler	江南大学	IJCV 2023
	A Spatio-Temporal Robust Tracker with Spatial-Channel Transformer and Jitter Suppression		
160	Cong Wu, Xiao-Jun Wu, Josef Kittler, Tianyang Xu, Sara Atito, Muhammad Awais, Zhenhua Feng	江南大学	AAAI 2024
	SCD-Net: Spatiotemporal Clues Disentanglement Network for Self-Supervised Skeleton-Based Action Recognition		
161	朱学峰，徐天阳，刘宗陶，汤张泳，吴小俊，Josef Kittler	江南大学	IJCV 2024
	UniMod1K: Towards a More Universal Large-Scale Dataset and Benchmark for Multi-modal Learning		
162	胡青松，户战选	昆明理工大学	AAAI 2024
	Catalyst for Clustering-Based Unsupervised Object Re-identification: Feature Calibration		
163	廖馨婷，刘伟明，陈超超，周芃旸，于丰源，朱华彬，姚滨辉	浙江大学	CVPR 2024
	Rethinking the Representation in Federated Unsupervised Learning with Non-IID Data		
164	郭颖，甄成，闫鹏飞	美团	ICCV 2023
	Controllable Guide-Space for Generalizable Face Forgery Detection		
165	段晨，付培，郭山，姜仟艺，魏晓明	美团	CVPR 2024
	ODM: A Text-Image Further Alignment Pre-training Approach for Scene Text Detection and Spotting		
166	刘茜，郭颖，甄成，李通，敖莹莹，闫鹏飞	美团	CVPR 2024
	CustomListener: Text-guided Responsive Interaction for User-friendly Listening Head Generation		
167	费俊杰，王腾，张金锐，何震宇，汪铖杰，郑锋	南方科技大学	ICCV 2023
	Transferable Decoding with Visual Entities for Zero-Shot Image Captioning		

168	邱忠喜，胡衍，陈晓珊，曾丹，胡清勇，刘江	南方科技大学	IEEE TPAMI 2024
	Rethinking Dual-Stream Super-Resolution Semantic Learning in Medical Image Segmentation		
169	叶顶强；樊超；马靖哲；Xiaoming Liu；于仕琪	南方科技大学	CVPR 2024
	BigGait: Learning Gait Representation You Want by Large Vision Models		
170	姜伟，Jiayu Qin，吴凌宇, Changyou Chen, Tianbao Yang，张利军	南京大学	ICML 2023
	Learning Unnormalized Statistical Models via Compositional Optimization		
171	邱梓豪，Quanqi Hu, Zhuoning Yuan, Denny Zhou，张利军, Tianbao Yang	南京大学	ICML 2023
	Not All Semantics are Created Equal: Contrastive Self-supervised Learning with Automatic Temperature Individualization		
172	侍蒋鑫，魏通，项于珂，李宇峰	南京大学	NeurIPS 2023
	How Re-sampling Helps for Long-Tail Learning?		
173	付明浩，朱可，吴建鑫	南京大学	AAAI 2024
	DTL: Disentangled Transfer Learning for Visual Recognition		
174	李志琦，Zhiding Yu, Shiyi Lan，李嘉涵，Jan Kautz， 路通，Jose M. Alvarez	南京大学	CVPR2024
	Is Ego Status All You Need for Open-Loop End-to-End Autonomous Driving?		
175	周大蔚，孙海龙，叶翰嘉，詹德川	南京大学	CVPR 2024
	Expandable Subspace Ensemble for Pre-Trained Model-Based Class-Incremental Learning		
176	唐嘉良，陈硕，牛罡，Masashi Sugiyama，宫辰	南京理工大学	ICCV 2023
	Distribution Shift Matters for Knowledge Distillation with Webly Collected Images		
177	杨凌风，王岳泽，李翔，王鑫龙，杨健	南京理工大学	NeurIPS 2023
	Fine-grained visual prompting		
178	陈翔，潘金山，董姜鑫	南京理工大学	CVPR 2024
	Bidirectional Multi-Scale Implicit Neural Representations for Image Deraining		
179	孙彦鹏，陈家辉，张珊，张欣彧，谌强，张刚，丁二锐，王井东，李泽超	南京理工大学	CVPR 2024
	VRP-SAM: SAM with Visual Reference Prompt		
180	王凤云，孙倩如，张冬，唐金辉	南京理工大学	CVPR 2024
	Unleashing Network Potentials for Semantic Scene Completion		
181	Zhiqiang Yan, Yuankai Lin, Kun Wang, Yupeng Zheng, Yufei Wang, Zhenyu Zhang, Jun Li, Jian Yang	南京理工大学	CVPR 2024

	Tri-Perspective View Decomposition for Depth Completion		
182	Ziqi Gu, Chunyan Xu, Jian Yang, Zhen Cui	南京理工大学	ICCV 2023
	Few-shot Continual Informax Learning		
183	余昊，程旭，彭伟，刘伟昊，赵国英（芬兰）	南京信息工程大学	ICCV 2023
	MDTv2: Masked Diffusion Transformer is a Strong Image Synthesizer		
184	高尚华，周攀，程明明，颜水成	南开大学	ICCV 2023
	MDTv2: Masked Diffusion Transformer is a Strong Image SynthesizerMDTv2: Masked Diffusion Transformer is a Strong Image Synthesizer		
185	靳鑫，肖嘉文，韩凌昊，郭春乐，张瑞勋，刘夏雷，李重仪	南开大学	ICCV 2023
	Lighting Every Darkness in Two Pairs: A Calibration-Free Pipeline for RAW Denoising		
186	李宇轩，侯淇彬，郑兆晖，程明明，杨健，李翔	南开大学	ICCV 2023
	Large Selective Kernel Network for Remote Sensing Object Detection		
187	曹续生，卢浩日，黄林澜，刘夏雷，程明明	南开大学	ICVPR 2024
	Generative Multi-modal Models are Good Class Incremental Learners		
188	李森茂，Joost van de Weijer，胡泰航，Fahad Shahbaz Khan，候淇彬，王亚星，杨健	南开大学	ICLR 2024
	Get What You Want, Not What You Don't: Image Content Suppression for Text-to-Image Diffusion Models		
189	李震，曹铭登，王鑫涛，祁仲昂，程明明，单瀛	南开大学	ICVPR 2024
	PhotoMaker: Customizing Realistic Human Photos via Stacked ID Embedding		
190	李政，李翔，傅欣艺，张鑫，王维强，陈硕，杨健	南开大学	ICVPR 2024
	PromptKD: Unsupervised Prompt Distillation for Vision-Language Models		
191	刘云，吴宇寰，张世辰，Li Liu，Min Wu，程明明	南开大学	IEEE TPAMI 2024
	Revisiting Computer-Aided Tuberculosis Diagnosis		
192	孙博远，杨雨奇，张乐，程明明，侯淇彬	南开大学	ICVPR 2024
	CorrMatch: Label Propagation via Correlation Matching for Semi-Supervised Semantic Segmentation		
193	尹博文， 张旭迎， 李钟毓， 刘丽， 程明明，侯淇彬	南开大学	ICLR 2024
	DFormer: Rethinking RGBD Representation Learning for Semantic Segmentation		
194	张旭迎，尹博文，陈宇铭，林铮，李运恒，侯淇彬，程明明	南开大学	ICVPR 2024

	TeMO: Towards Text-Driven 3D Stylization for Multi-Object Meshes		
195	张知诚，胡钧耀，程文韬， Danda Paudel，杨巨峰	南开大学	ICVPR 2024
	ExtDM: Distribution Extrapolation Diffusion Model for Video Prediction		
196	赵攀诚，徐鹏，秦鹏达，范登平，张知诚， 贾国力，周伯文，杨巨峰	南开大学	ICVPR 2024
	LAKE-RED: Camouflaged Images Generation by Latent Background Knowledge Retrieval-Augmented Diffusion		
197	杨雨奇 姜鹏涛 侯淇彬 张昊 陈锦伟 李博	南开大学、vivo	ICVPR 2024
	Multi-Task Dense Prediction via Mixture of Low-Rank Experts		
198	姜鹏涛 杨雨奇 曹洋 侯淇彬 程明明 沈春华	南开大学、浙江大学	ICVPR 2024
	Traffic Scene Parsing through the TSP6K dataset		
199	岳宗胜，雍宏巍，赵谦，张磊，孟德宇，Kenneth K.Y. Wong	南洋理工大学	IEEE TPAMI 2024
	Deep Variational Network Toward Blind Image Restoration		
200			
201	崔鹏，张丹，邓志杰，董胤蓬，朱军	清华大学	NeurIPS 2023
	learning sample difficulty from pre-trained models for reliable prediction		
202	王语霖，乐阳，卢睿，刘天娇， 钟钊，宋士吉，黄高	清华大学	ICCV 2023
	EfficientTrain: Exploring Generalized Curriculum Learning for Training Visual Backbones		
203	何春明，李凯，张亚超，张宇伦， 郭振华，李秀，Martin Danelljan	清华大学	ICLR 2024
	Strategic preys make acute predators: Enhancing camouflaged object detectors by generating camouflaged objects		
204	胡锦毅，姚远，王崇屹，王珊，潘寅旭，陈乾瑜，余天予，吴航昊，赵越，张皓烨，韩旭， 林衍凯，薛娇，李大海，刘知远，孙茂松	清华大学	ICLR 2024
	Large Multilingual Models Pivot Zero-Shot Multimodal Learning across Languages		
205	林嘉琪，李志豪，唐逍， 刘健庄，吴小飞，许松岑	清华大学	CVPR 2024
	VastGaussian: Vast 3D Gauussians for Large Scene Reconstruction		
206	吴凌轩，杨啸，董胤蓬，谢留葳，苏航，朱军	清华大学	ICLR 2024
	Embodied Active Defense: Leveraging Recurrent Feedback to Counter Adversarial Patches		
207	Han Yu, Xingxuan Zhang, Renzhe Xu, Jiashuo Liu, Yue He, Peng Cui	清华大学	CVPR 2024

	Rethinking the Evaluation Protocol of Domain Generalization		
208	Yachao Zhang, Runze Hu, Ronghui Li, Yanyun Qu, Yuan Xie, Xiu Li	清华大学	AAAI 2024
	Cross-Modal Match for Language Conditioned 3D Object Grounding		
209	Youze Wang, Wenbo Hu, Yinpeng Dong, Hanwang Zhang, Richang Hong	清华大学	CVPR 2024
	Exploring the Transferability of Visual Prompting for Multimodal Large Language Models		
210	Lin Zhao, Tianchen Zhao, Zinan Lin, Xuefei Ning, Guohao Dai, Huazhong Yang, Yu Wang	清华大学	NeurIPS 2022
	FlashEval: Towards Fast and Accurate Evaluation of Text-to-image Diffusion Generative Models		
211	Yifan Pu, Weicong Liang, Yiduo Hao, Yuhui Yuan, Yukang Yang, Chao Zhang, Han Hu, Gao Huang	清华大学、北京大学、剑桥大学	NeurIPS 2023
	Rank-DETR for High Quality Object Detection		
212	Yaohua Zha, Huizhen Ji, Jinmin Li, Rongsheng Li, Tao Dai, Bin Chen, Zhi Wang, Shu-Tao Xia	清华大学深圳国际研究生院	AAAI 2024
	Towards Compact 3D Representations via Point Feature Enhancement Masked Autoencoders		
213	Ronghui Li, YuXiang Zhang, Yachao Zhang, Hongwen Zhang, Jie Guo, Yan Zhang, Yebin Liu, Xiu Li	清华大学深圳国际研究生院	CVPR 2024
	Lodge: A Coarse to Fine Diffusion Network for Long Dance Generation Guided by the Characteristic Dance Primitives		
214	Wang Jinpeng, Wang Yuting, Chen Bin, Zeng Ziyun, Xia Shutao	清华大学深圳国际研究生院	AAAI 2023
	GMMFormer: Gaussian-Mixture-Model Based Transformer for Efficient Partially Relevant Video Retrieval		
215	Wang Wenbin, Wang Ruiping, Shan Shiguang, Chen Xilin	中国科学院计算技术研究所	IJCV 2023
	Importance First: Generating Scene Graph of Human Interest		
216	Wang Liqiong, Yang Jinyu, Zhang Yanfu, Wang Fangyi, Zheng Feng	三峡大学、南方科技大学	CVPR 2024
	Depth-Aware Concealed Crop Detection in Dense Agricultural Scenes		
217	Hao Xiaoshuai, Zhang Wanqian	三星中国研究院	NeurIPS 2023
	Uncertainty-Aware Alignment Network for Cross-Domain Video-Text Retrieval		
218	Chen Manzhao, Shao Wenqi, Xu Peng, Lin Mingbao, Zhang Kaipeng, Chao Fei, Ji Rongrong, Qiao Yu, Luo Ping	厦门大学	ICCV 2023
	DiffRate: Differentiable Compression Rate for Efficient Vision Transformers		
219	Gen Luo, Yiyi Zhou, Tianhe Ren, Shengxin Chen, Xiaoshuai Sun, Rongrong Ji	厦门大学	NeurIPS 2023

	Cheap and quick: Efficient vision-language instruction tuning for large language models		
220	Luo Xiaotong, Ai Zekun, Liang Qiuyuan, Liu Ding, Xie Yuan, Qu Yanyun, Fu Yun	厦门大学	AAAI 2024
	AdaFormer: Efficient Transformer with Adaptive Token Sparsification for Image Super-resolution		
221	Qiming Xia, Wei Ye, Hai Wu, Shijia Zhao, Leyuan Xing, Xun Huang, Jinhao Deng, Xin Li, Chenglu Wen, Cheng Wang	厦门大学	CVPR 2024
	HINTED: Hard Instance Enhanced Detector with Mixed-Density Feature Fusion for Sparsely-Supervised 3D Object Detection		
222	Yang Zhou, Hao Shao, Letian Wang, Steven L. Waslander, Hongsheng Li, Yu Liu	商汤科技研究院	CVPR 2024
	SmartRefine: A Scenario-Adaptive Refinement Framework for Efficient Motion Prediction		
223	Jiazhong Cen, Zanwei Zhou, Jieming Fang, Chen Yang, Wei Shen, Lingxi Xie, Dongsheng Jiang, Xiaopeng Zhang, Qi Tian	上海交通大学	NeurIPS 2023
	Segment Anything in 3D with NeRFs		
224	Tongkun Guan, Wei Shen, Xue Yang, Qi Feng, Zekun Jiang, Xiaokang Yang	上海交通大学	ICCV 2023
	Self-supervised Character-to-Character Distillation for Text Recognition		
225	Zhiwei Zhang, Zhizhong Zhang, Qian Yu, Ran Yi, Yuan Xie, Lizhuang Ma	上海交通大学	ICCV 2023
	LiDAR-Camera Panoptic Segmentation via Geometry-Consistent and Semantic-Aware Alignment		
226	Shangzhe Di, Weidi Xie	上海交通大学	CVPR 2024
	Grounded Question-Answering in Long Egocentric Videos		
227	Chengyang Hu, Keyue Zhang, Taiping Yao, Shouhong Ding, Lizhuang Ma	上海交通大学	CVPR 2024
	Rethinking Generalizable Face Anti-spoofing via Hierarchical Prototype-guided Distribution Refinement in Hyperbolic Space		
228	Yonglu Li, Xiaoqian Wu, Xinpeng Liu, Yiming Dou, Yikun Ji, Junyi Zhang, Yixing Li, Jingru Tan, Xudong Lu, Cewu Lu	上海交通大学	CVPR 2024
	From Isolated Islands to Pangea: Unifying Semantic Space for Human Action Understanding		
229	Zelin Peng, Zhengqin Xu, Zhilin Zeng, Lingxi Xie, Qi Tian, Wei Shen	上海交通大学	CVPR 2024
	Parameter Efficient Fine-tuning via Cross Block Orchestration for Segment Anything Model		
230	Chongjie Si, Zekun Jiang, Xuehui Wang, Yan Wang, Xiaokang Yang, Wei Shen	上海交通大学	AAAI 2024
	Partial Label Learning with a Partner		
231	Fan Wu, Jinling Gao, Lanqing Hong, Xinbing Wang, Chenghu Zhou, Nanyang Ye	上海交通大学	AAAI 2024
	G-NAS: Generalizable Neural Architecture Search for Single Domain Generalization Object Detection		

232	Shaofeng Zhang, Jinfa Huang, Qiang Zhou, Zhibin Wang, Fan Wang, Jiebo Luo, Junchi Yan	上海交通大学	ICLR 2024
	Continuous-Multiple Image Outpainting in One-Step via Positional Query and A Diffusion-based Approach		
233	Shaobin Zhuang, Kunchang Li, Xinyuan Chen, Yaohui Wang, Ziwei Liu, Yu Qiao, Yali Wang	上海交通大学	CVPR 2024
	Vlogger: Make Your Dream A Vlog		
234	Wanxing Chang, Ye Shi, Jingya Wang	上海科技大学	NeurIPS 2024
	CSOT: Curriculum and Structure-Aware Optimal Transport for Learning with Noisy Labels		
235	Bin Li, Ye Shi, Qian Yu, Jingya Wang	上海科技大学	AAAI 2024
	Unsupervised Cross-Domain Image Retrieval via Prototypical OPtimal Transport		
236	Rongjie Li, Yu Wu, Xuming He	上海科技大学	CVPR 2024
	Learning by Correction: Efficient Tuning Task for Zero-Shot Generative Vision-Language Reasoning		
237	Xuekun Jiang, Anyi Rao, Jingbo Wang, Dahua Lin, Bo Dai	上海人工智能创新中心	CVPR 2024
	Cinematic Behavior Transfer via NeRF-based Differentiable Filming		
238	Yihao Liu, Hengyuan Zhao, Jinjin Gu, Yu Qiao, Chao Dong	上海人工智能实验室	TPAMI 2023
	Evaluating the Generalization Ability of Super-Resolution Networks		
239	Zeyu Lu, Di Huang, Lei Bai, Jingjing Qu, Chengyue Wu, Xihui Liu, Wanli Ouyang	上海人工智能实验室	NeurIPS 2024
	Seeing is not always believing: Benchmarking Human and Model Perception of AI-Generated Images		
240	Yao Mu, Qinglong Zhang, Mengkang Hu, Wenhai Wang, Mingyu Ding, Jun Jin, Bin Wang, Jifeng Dai, Yu Qiao, Ping Luo	上海人工智能实验室	NeurIPS 2024
	EmbodiedGPT: Vision-Language Pre-Training via Embodied Chain of Thought		
241	Zhe Chen, Jiannan Wu, Wenhai Wang, Weijie Su, Guo Chen, Sen Xing, Muyan Zhong, Qinglong Zhang, Xizhou Zhu, Lewei Lu, Bin Li, Ping Luo, Tong Lu, Yu Qiao, Jifeng Dai	上海人工智能实验室	CVPR 2024
	InternVL: Scaling up vision foundation models and aligning for generic visual-linguistic tasks		
242	Hao Li, Xue Yang, Zhaokai Wang, Xizhou Zhu, Jay Zhou, Yu Qiao, Xiaogang Wang, Hongsheng Li, Lewei Lu, Jifeng Dai	上海人工智能实验室	CVPR 2024
	Auto MC-Reward: Automated Dense Reward Design with Large Language Models for Minecraft		
243	Wenqi Shao, Mengzhao Chen, Zhaoyang Zhang, Peng Xu, Lirui Zhao, Zhiqian Li, Kaipeng Zhang, Peng Gao, Yu Qiao, Ping Luo	上海人工智能实验室	ICLR 2024
	Omniquant: Omnidirectionally calibrated quantization for large language models		

244	Chi Yan, Delin Qu, Dan Xu, Zhigang Wang, Bin Zhao, Dong Wang, Xuelong Li	上海人工智能实验室	CVPR 2024
	GS-SLAM: Dense Visual SLAM with 3D Gaussian Splatting		
245	Chao Yang, Bo Zou, Yu Qiao, Chengbin Quan, Youjian Zhao	上海人工智能实验室	CVPR 2024
	LLaMA-Excitor: General Instruction Tuning via Indirect Feature Interaction		
246	Yi Yu, Xue Yang, Qingyun Li, Feipeng Da, Jifeng Dai, Yu Qiao, Junchi Yan	上海人工智能实验室	CVPR 2024
	Point2RBox: Combine Knowledge from Synthetic Visual Patterns for End-to-end Oriented Object Detection with Single Point Supervision		
247	Wenlong Zhang, Xiaohui Li, Xiangyu Chen, Xiaoyun Zhang, Yu Qiao, Xiaoming Wu, Chao Dong	上海人工智能实验室、香港理工大学	ICLR 2024
	SEAL: A FRAMEWORK FOR SYSTEMATIC EVALUATION OF REAL-WORLD SUPER-RESOLUTION（spotlight)		
248	Weiwen Wang, Min Shi, Qingyun Li, Wenhai Wang, Zhenhang Huang, Linjie Xing, Zhe Chen, Hao Li, Xizhou Zhu, Zhiguo Cao, Yushi Chen, Tong Lu, Jifeng Dai, Yu Qiao	上海人工智能实验室、复旦大学	ICLR 2024
	The All-Seeing Project: Towards Panoptic Visual Recognition and Understanding of the Open World		
249	Runhao Zeng, Xiaoyong Chen, Jiaming Liang, Huisi Wu, Guangzhong Cao, Yong Guo	深圳北理莫斯科大学	CVPR 2024
	Benchmarking the Robustness of Temporal Action Detection Models Against Temporal Corruptions		
250	Shuo Yang, Yongqi Wang, Xiaofeng Ji, Xinxiao Wu	深圳北理莫斯科大学、北京理工大学	AAAI 2024
	Multi-Modal Prompting for Open-Vocabulary Video Visual Relationship Detection		
251	Changsheng Chen, Liangwei Lin, Yongqi Chen, Bin Li, Jishen Zeng, Jiwu Huang	深圳大学	CVPR 2024
	CMA: A Chromaticity Map Adapter for Robust Detection of Screen-Recapture Document Images		
252	Qinliang Lin, Cheng Luo, Zenghao Niu, Xilin He, Weicheng Xie, Yuanbo Hou, Linlin Shen, Siyang Song	深圳大学	AAAI 2024
	Boosting Adversarial Transferability across Model Genus by Deformation-Constrained Warping		
253	Qinliang Lin, Cheng Luo, Zenghao Niu, Xilin He, Weicheng Xie, Yuanbo Hou, Linlin Shen, Siyang Song	深圳大学	AAAI 2024
	HACDR-Net: Heterogeneous-Aware Convolutional Network for Diabetic Retinopathy Multi-Lesion Segmentation		
254	Yijie Lin, Jie Zhang, Zhenyu Huang, Jia Liu, Zujie Wen, Xi Peng	四川大学	ICLR 2024
	Multi-granularity Correspondence Learning from Long-term Noisy Videos		
255	Yiding Lu, Yijie Lin, Mouxing Yang, Dezong	四川大学	AAAI 2024

	Peng, Peng Hu, Xi Peng		
	Decoupled Contrastive Multi-View Clustering with High-Order Random Walks		
256	Yang Qin, Yingke Chen, Dezhong Peng, Xi Peng, Joey Tianyi Zhou, Peng Hu	四川大学	CVPR 2024
	Noisy-Correspondence Learning for Text-to-Image Person Re-identification		
257	Mouxing Yang, Yunfan Li, Changqing Zhang, Peng Hu, Xi Peng	四川大学	ICLR 2024
	Test-time Adaption against Multi-modal Reliability Bias		
258			
259	Tingliang Feng, Hao Shi, Xueyang Liu, Wei Feng, Liang Wan, Yanlin Zhou, Di Lin	天津大学	NeurIPS 2023
	Open Compound Domain Adaptation with Object Style Compensation for Semantic Segmentation		
260	Anan Liu, Chenxi Huang, Ning Xu, Hongshuo Tian, Jing Liu, Yongdong Zhang	天津大学	TMM 2024
	Counterfactual Visual Dialog: Robust Commonsense Knowledge Learning from Unbiased Training		
261	黄淑英，陈格，杨勇，王晓争，梁晨滨	天津工业大学	AAAI 2024
	MFTN: Multi-Level Feature Transfer Network Based on MRI-Transformer for MR Image Super-Resolution		
262	谢驰，张钊，吴逸璇，朱烽，赵瑞，梁爽	同济大学	NeurIPS 2023
	Described Object Detection: Liberating Object Detection with Flexible Expressions		
263	王一飞，张吉哲，王奕森	北京大学数学科学学院、西安交通大学人工智能与机器人研究所	ICLR 2024
	Do Generated Data Always Help Contrastive Learning?		
264	李博，肖浩科，唐律	维沃移动通信有限公司（IQ）	CVPR 2024
	ASAM: Boosting Segment Anything Model with Adversarial Tuning		
265	段伦豪，赵杉杉，薛楠，Mingming Gong，夏桂松，陶大程	武汉大学	NeurIPS 2023
	ConDaFormer: Disassembled Transformer with Local Structure Enhancement for 3D Point Cloud Understanding		
266	范颖颖、武宇、杜博、林雨恬	武汉大学	NeurIPS 2023
	Revisit Weakly-Supervised Audio-Visual Video Parsing from the Language Perspective		
267	武辰，杜博，张良培	武汉大学	IEEE TPAMI 2023

	Fully convolutional change detection framework with generative adversarial network for unsupervised, weakly supervised and regional supervised change detection		
268	张韬，田兴业，武宇，季顺平，王学博，张渊，万鹏飞	武汉大学	ICCV 2023
	DVIS: Decoupled Video Instance Segmentation Framework		
269	赵恒伟, 王心宇，李静涛，钟燕飞	武汉大学	ICCV 2023
	Class Prior-Free Positive-Unlabeled Learning with Taylor Variational Loss for Hyperspectral Remote Sensing Imagery		
270	黄文柯，叶茫，石泽昆，杜博	武汉大学	IEEE TPAMI 2024
	Generalizable Heterogeneous Federated Cross-Correlation and Instance Similarity Learning		
271	李贺，叶茫，张明，杜博	武汉大学	CVPR 2024
	All in One Framework for Multimodal Re-identification in the Wild		
272	李卓鸿，贺威，李杰潘，卢方骁，张洪艳	武汉大学	CVPR 2024
	Learning without Exact Guidance: Updating Large-scale High-resolution Land Cover Maps from Low-resolution Historical Labels		
273	罗俊伟，杨学，俞毅，李青云，严骏驰，李彦胜	武汉大学	CVPR 2024
	PointOBB: Learning Oriented Object Detection via Single Point Supervision		
274	王俊珏，郑卓，陈梓航，马爱龙，钟燕飞	武汉大学	AAAI 2024
	EarthVQA: Towards Queryable Earth via Relational Reasoning-Based Remote Sensing Visual Question Answering		
275	王赛；林雨恬；武宇	武汉大学	CVPR 2024
	Omni-Q: Omni-Directional Scene Understanding for Unsupervised Visual Grounding		
276	张嘉旭，黄少立，涂志刚，陈欣，詹晓航，俞刚，单瀛	武汉大学	ICLR 2024
	TapMo: Shape-aware Motion Generation of Skeleton-free Characters		
277	赵天浩，陈永灿，武宇，刘天阳，杜博，肖培伦，邱石，杨宏达，李国镇，杨毅，林雨恬	武汉大学	CVPR 2024
	Improving Bird’s Eye View Semantic Segmentation by Task Decomposition		
278	王若瑜，杨永祺，钱志浩，竺烨，武宇	武汉大学	ICLR 2024
	Diffusion in Diffusion: Cyclic One-Way Diffusion for Text-Vision-Conditioned Generation		
279	段志斌，吕之一，王超杰，陈渤，安波，周明远	西安电子科技大学	NeurIPS 2023
	Few-shot Generation via Recalling the Episodic-Semantic Memory like Human Being		
280	郑皓天；王启舟；方震；夏晓波；刘峰；刘同亮；韩波	西安电子科技大学	NeurIPS 2023

	Out-of-distribution Detection Learning with Unreliable Out-of-distribution Sources		
281	曾泽群，谢岩，张昊，陈驰宇，王正珏，陈渤	西安电子科技大学	CVPR 2024
	MeaCap: Memory-Augmented Zero-shot Image Captioning		
282	陈志文；朱智宇；张异凡；侯军辉；石光明；吴金建	西安电子科技大学	CVPR 2024
	Segment Any Events via Weighted Adaptation of Pivotal Tokens		
283	李柯，王笛，胡张源，朱文瑄，李少峰，王泉	西安电子科技大学	CVPR 2024
	Unleashing Channel Potential: Space-Frequency Selection Convolution for SAR Object Detection		
284	刘德成，王西军，彭春蕾，王楠楠，胡瑞敏，高新波	西安电子科技大学	AAAI 2024
	Adv-Diffusion: Imperceptible Adversarial Face Identity Attack via Latent Diffusion Model		
285	刘永旭，全英汇，肖国尧，李奥博，吴金建	西安电子科技大学	AAAI 2024
	Scaling and Masking: A New Paradigm of Data Sampling for Image and Video Quality Assessment		
286	马彦彪，焦李成，刘芳，杨淑媛，刘旭，陈璞花	西安电子科技大学	(IJCV) 2024
	Geometric Prior Guided Feature Representation Learning for Long-Tailed Classification		
287	王昭阳，李东阳，张明阳，罗浩，公茂果	西安电子科技大学	AAAI 2024
	Enhancing Hyperspectral Images via Diffusion Model and Group-Autoencoder Super-resolution Network		
288	徐偲，司佳俊，管子玉，赵伟，武越，高熙越	西安电子科技大学	AAAI 2024
	Reliable Conflictive Multi-view Learning		
289	杨文，吴金建，马居坡，李雷达，石光明	西安电子科技大学	AAAI 2024
	Motion Deblurring via Spatial-Temporal Collaboration of Frames and Events		
290	吕奕龙、李敏、何玉杰、李少朋、贺鬯祯、杨爱涛	西安高新技术研究所	ICCV 2023
	Anchor-Intermediate Detector: Decoupling and Coupling Bounding Boxes for Accurate Object Detection		
291	罗倚斯，赵熙乐，李哲民，Michael K. Ng.，孟德宇	西安交通大学	IEEE TPAMI 2023
	low-rank tensor function representation for multi-dimensional data recovery		
292	束俊，袁翔，孟德宇*，徐宗本	西安交通大学	IEEE TPAMI 2023
	CMW-Net: Learning a class-aware sample weighting mapping for robust deep learning		
293	王海林，彭江军，覃文金，王建军，孟德宇	西安交通大学	IEEE TPAMI 2023

	Guaranteed Tensor Recovery fused Low-rankness and Smoothness		
294	王全子昂，王仁振，吴一尘，贾西西，孟德宇	西安交通大学	ICCV 2023
	CBA: Improving Online Continual Learning via Continual Bias Adaptor		
295	谢琦，赵谦，徐宗本，孟德宇	西安交通大学	IEEE TPAMI 2023
	Fourier Series Expansion Based Filter Parametrization for Equivariant Convolutions		
296	于曦，古祥，刘浩志，孙剑	西安交通大学	NeurIPS 2023
	Constructing Non-isotropic Gaussian Diffusion Model Using Isotropic Gaussian Diffusion Model for Image Editing		
297	庞立，芮翔宇，崔龙，王洪中，孟德宇，曹相湧	西安交通大学	CVPR 2024
	HIR-Diff: Unsupervised Hyperspectral Image Restoration Via Improved Diffusion Models		
298	丘琮培，Tong Zhang，伍彦豪，柯炜， Mathieu Salzmann，Sabine Süsstrunk	西安交通大学	(ICLR) 2024
	Mind Your Augmentation: The Key to Decoupling Dense Self-Supervised Learning		
299	王家浩，闫彩霞，张未展，刘欢，孙皓，郑庆华	西安交通大学	AAAI 2024
	SAUI: Scale-Aware Unseen Imagineer for Zero-Shot Object Detection		
300	伍彦豪、Tong Zhang、柯炜、丘琮培、Sabine Susstrunk、Mathieu Salzmann	西安交通大学	CVPR 2024
	Mitigating Object Dependencies: Improving Point Cloud Self-Supervised Learning through Object Exchange		
301	Hengfei Cui, Yifan Wang, Yan Li, Lei Jiang, Yong Xia, Yanning Zhang	西北工业大学	IEEE JBHI 2023
	An Improved Combination of Faster R-CNN and U-Net Network for Accurate Multi-Modality Whole Heart Segmentation		
302	Binglu Wang, Lei Zhang, Zhaozhong Wang, Yongqiang Zhao, Tianfei Zhou	西北工业大学	ICCV 2023
	CORE: Cooperative Reconstruction for Multi-Agent Perception		
303	Hao Li, Dingwen Zhang, Yalun Dai, Nian Liu, Lechao Cheng, Jingfeng Li, Jingdong Wang, Junwei Han	西北工业大学	CVPR 2024
	GP-NeRF: Generalized Perception NeRF for Context-Aware 3D Scene Understanding		
304	Peng Wu, Xuerong Zhou, Guansong Pang, Yujia Sun, Jing Liu, Peng Wang, Yanning Zhang	西北工业大学	CVPR 2024
	Open-Vocabulary Video Anomaly Detection		
305	Wenqi Zhong, Linzhi Yu, Chen Xia, Junwei Han, Dingwen Zhang	西北工业大学	AAAI 2024
	SpFormer: Spatio-Temporal Modeling for Scanpaths with Transformer		
306	Gaoge Han, Shaoli Huang, Mingming Gong, Jinglei Tang	西北农林科技大学	AAAI 2024

	HuTuMotion: Human-Tuned Navigation of Latent Motion Diffusion Models with Minimal		
307	Mingjia Li, Binhui Xie, Shuang Li, Chi Harold Liu, Xinjing Cheng	西南交通大学	CVPR 2024
	SVDinsTN: A Tensor Network Paradigm for Efficient Structure Search from Regularized Modeling Perspective		
308	Jie Liu, Yixiao Zhang, Jieneng Chen, Yongyi Lu, Bennett A. Landman, Yixuan Yuan, Alan Yuille, Yucheng Tang, Zongwei Zhou	香港城市大学	ICCV 2023
	clip-driven universal model for organ segmentation and tumor detection		
309	Yichen Wu, Longkai Huang, Renzhen Wang, Deyu Meng, Ying Wei	香港城市大学	ICLR 2024
	Meta Continual Learning Revisited：Implicitly Enhancing Online Hessian Approximation via Variance Reduction （Oral presentation）		
310	Zhixuan Liang, Yao Mu, Hengbo Ma, Masayoshi Tomizuka, Mingyu Ding, Ping Luo	香港大学	CVPR 2024
	SkillDiffuser: Interpretable Hierarchical Planning via Skill Abstractions in Diffusion-Based Task Execution		
311	Xiaoyang Wu, Zhuotao Tian, Xin Wen, Bohao Peng, Xihui Liu, Kaicheng Yu, Hengshuang Zhao	香港大学	CVPR 2024
	Towards Large-scale 3D Representation Learning with Multi-dataset Point Prompt Training		
312	Yutao Hu, Tianbin Li, Quanfeng Lu, Wenqi Shao, Junjun He, Yu Qiao, Ping Luo	香港大学、上海人工智能实验室	CVPR 2024
	OmniMedVQA: A New Large-Scale Comprehensive Evaluation Benchmark for Medical LVLM		
313	Kai Chen, Enze Xie, Zhe Chen, Yibo Wang, Lanqing Hong, Zhenguo Li, Dit-Yan Yeung	香港科技大学	ICLR 2024
	GeoDiffusion: Text-Prompted Geometric Control for Object Detection Data Generation		
314	Feng Li, Qin Jiang, Hao Zhang, Tianhe Ren, Shilong Liu, Xueyan Zou, Huaizhe Xu, Hongyang Li, Chunyuan Li, Jianwei Yang, Lei Zhang, Jianfeng Gao	香港科技大学	CVPR 2024
	Visual In-context Prompting		
315	Minghan Li, Shuai Li, Xindong Zhang, Lei Zhang	香港理工大学	CVPR 2024
	UniVS: Unified and Universal Video Segmentation with Prompts as Queries		
316	Yabin Zhang, Wenjie Zhu, Hui Tang, Zhiyuan Ma, Kaiyang Zhou, Lei Zhang	香港理工大学	CVPR 2024
	Dual Memory Networks: A Versatile Adaptation Approach for Vision-Language Models		
317			
318	Tianyu Huang, Liangzu Peng, Rene Vidal, Yunhui Liu	香港中文大学	CVPR 2024
	Scalable 3D registration via Truncated Entry-wise Absolute Residuals		

319	Yiyuan Zhang; Xiaohan Ding; Kaixiong Gong; Yixiao Ge; Ying Shan; Xiangyu Yue	香港中文大学	CVPR 2024
	Multimodal Pathway: Improve Transformers with Irrelevant Data from Other Modalities		
320	Mingli Zhu, Shaokui Wei, Hongyuan Cha, Baoyuan Wu	香港中文大学（深圳）	NeurIPS 2023
	Neural Polarizer: A Lightweight and Effective Backdoor Defense via Purifying Poisoned Features		
321	秦逸然，周恩申，刘齐畅，尹臻菲，盛律，张瑞茂，乔宇，邵婧	香港中文大学（深圳）	CVPR 2024
	MP5: A Multi-modal Open-ended Embodied System in Minecraft via Active Perception		
322	王炯，杨丰羽，李炳良，苟文博，闫丹琪，曾爱玲，高一君，王君乐，荆彦青，张瑞茂	香港中文大学（深圳）	CVPR 2024
	FreeMan: Towards benchmarking 3D human pose estimation under Real-World Conditions		
323	朱梓豪，张明达，魏少魁，吴秉哲，吴保元	香港中文大学（深圳）	ICLR 2024
	VDC: Versatile Data Cleanser for Detecting Dirty Samples via Visual-Linguistic Inconsistency		
324	刘垦坤，金德荣，曾爱玲，韩晓光，张磊	香港中文大学（深圳）、大湾区数字经济研究院（IDEA）	NeurIPS 2023
	HMNeRFBench: A Comprehensive Benchmark for Neural Human Radiance Fields		
325	李博豪，葛玉莹，葛艺潇，王广智，王瑞，张瑞茂，shan ying	香港中文大学（深圳）	CVPR 2024
	SEED-Bench: Benchmarking Multimodal Large Language Models		
326	杨杰，李炳良，曾爱玲，张磊，张瑞茂	香港中文大学（深圳）	CVPR 2024
	Open-World Human-Object Interaction Detection via Multi-modal Prompts		
327	Yuzhou Huang(黄聿洲), Liangbin Xie(谢良彬), Xintao Wang(王鑫涛), Ziyang Yuan(袁梓洋), Xiaodong Cun(寸晓东), Yixiao Ge(葛艺潇), Jiantao Zhou(周建涛), Chao Dong(董超), Rui Huang(黄锐), Ruimao Zhang(张瑞茂), Ying Shan	香港中文大学（深圳）	CVPR 2024
	SmartEdit: Exploring Complex Instruction-based Image Editing with Multimodal Large Language Models		
328	郭子尧, 王锴，George Cazenavette,李晖, 张凯鹏, 尤洋	新加坡国立大学	ICLR 2024
	Towards Lossless Dataset Distillation via Difficulty-Aligned Trajectory Matching		
329	金晔莹， Wei Ye, Wenhan Yang, Yuan Yuan, Robby T. Tan	新加坡国立大学	AAAI 2024
	DeS3: Adaptive Attention-Driven Self and Soft Shadow Removal using ViT Similarity		

330	梁哲昕，李重仪，周尚辰，冯锐成，Chen Change Loy	新加坡南洋理工大学	ICCV 2023
	Iterative Prompt Learning for Unsupervised Backlit Image Enhancement		
331	Lin Li，Haoyan Guan，Jianing Qiu，Michael Spratling	英国伦敦大学国王学院（King's College London）	CVPR 2024
	One Prompt Word is Enough to Boost Adversarial Robustness for Pre-trained Vision-Language Models		
332	Daiwei Yu, Zhuorong Li, Lina Wei, Canghong Jin, Yun Zhang, Sixian Chan	浙大城市学院	CVPR 2024
	Soften to Defend: Towards Adversarial Robustness via Self-Guided Label Refinement		
333	方涛，郑乾，潘纲	浙江大学	NeurIPS 2023
	Alleviating the Semantic Gap for Generalized fMRI-to-Image Reconstruction		
334	陈佳炫，祁玉，王跃明，潘纲	浙江大学	CVPR 2024
	Mind Artist: Creating Artistic Snapshots with Human Thought		
335	陈秭达 张自然 李昊颖 李孟灏 陈跃庭 李奇 冯华君 徐之海 陈世镐	浙江大学	AAAI 2024
	Deep Linear Array Pushbroom Image Restoration: A Degradation Pipeline and Jitter-Aware Restoration Network		
336	何威振，邓以恒，唐诗翔，陈祺浩，谢青松，王逸舟，白磊，朱烽，赵睿，欧阳万里，齐冬莲，闫云凤	浙江大学	CVPR 2024
	Instruct-ReID: A Multi-purpose Person Re-identification Task with Instructions		
337	黄思腾，龚镖，冯玉彤，张敏，吕逸良，王东林	浙江大学	CVPR 2024
	Troika: Multi-Path Cross-Modal Traction for Compositional Zero-Shot Learning		
338	景宸琛，李煜堃，陈昊，沈春华	浙江大学	AAAI 2024
	Retrieval-Augmented Primitive Representations for Compositional Zero-Shot Learning		
339	廖展锋，郑乾，刘彦，潘纲	浙江大学	AAAI 2024
	Spiking NeRF: Representing the Real-World Geometry by a Discontinuous Representation		
340	闵致远 罗亚威 杨卫 王跃嵩 杨易	浙江大学	CVPR 2024
	Entangled View-Epipolar Information Aggregation for Generalizable Neural Radiance Fields		
341	申江荣，倪文遥，徐齐，唐华锦	浙江大学	AAAI 2024
	Efficient spiking neural networks with sparse activations for continual learning		
342	徐晓刚，孔庶，胡涛，刘哲，鲍虎军	浙江大学	CVPR 2024
	Boosting Image Restoration via Priors from Pre-trained Models		

343	杨鸿辉，张莎，黄迪，吴虓杨，朱皓怡，贺通，唐诗翔，赵恒爽，邱奇波，林彬彬 ，何晓飞，欧阳万里	浙江大学	CVPR 2024
	UniPAD: A Universal Pre-training Paradigm for Autonomous Driving		
344	叶振辉；钟添芸；任意；杨嘉琪；李伟创；黄家伟；江子越；何金铮；黄融杰；刘静林；章晨；殷翔；马泽君；赵洲	浙江大学	ICLR 2024
	Real3D-Portrait: One-shot Realistic 3D Talking Portrait Synthesis		
345			
346	王闻箫；陈蔚；邱奇波；陈隆；武伯熹；何晓飞；刘威	浙江大学、 浙江大学 CAD&CG 实验室	IEEE TPAMI 2023
	CrossFormer++: A Versatile Vision Transformer Hinging on Cross-Scale Attention		
347	胡世哲，娄铮铮，闫小强，叶阳东*	郑州大学	IEEE TPAMI 2024
	A Survey on Information Bottleneck		
348	邵帅，白雨，王焱，刘宝弟，刘斌	之江实验室	AAAI 2024
	Collaborative Consortium of Foundation Models for Open-World Few-Shot Learning		
349	李越，张越一，叶俊天，徐飞虎，熊志伟	中国科学技术大学	NeurIPS 2023
	Deep Non-line-of-sight Imaging from Under-scanning Measurements		
350	徐光锴，尹炜，陈昊，沈春华，程恺，赵峰	中国科学技术大学	ICCV 2023
	FrozenRecon: Pose-free 3D Scene Reconstruction with Frozen Depth Models		
351	姚明德 黄杰 金鑫 周满 徐瑞康 周生龙 熊志伟	中国科学技术大学	ICCV 2023
	Generalized Lightness Adaptation via Channel Selective Normalization		
352	曹成志，张瑞茂，李爽	中国科学技术大学	ICLR 2024
	Enhancing Human-AI Collaboration Through Logic-Guided Reasoning		
353	李和倍，汪金，袁嘉辉，李越，翁文明，彭岩松，张越一，熊志伟，孙晓艳	中国科学技术大学	CVPR 2024
	Event-assisted low light video object segmentation		
354	彭岩松，李和倍，张越一，孙晓艳，吴枫	中国科学技术大学	CVPR 2024
	Scene Adaptive Sparse Transformer for Event-based Object Detection		
355	唐传波，盛锡华，李卓元，张昊田，李礼，刘东	中国科学技术大学	AAAI 2024
	Offline and Online Optical Flow Enhancement for Deep Video Compression		

356	肖杰，冯睿鑫，张晗，刘志恒，阳展韬，朱禹睿，傅雪阳，朱凯，刘宇，查正军	中国科学技术大学	ICLR 2024
	DreamClean: Restoring Clean Image Using Deep Diffusion Prior		
357	杨子瑾，曾凯，陈可江，方涵，张卫明，俞能海	中国科学技术大学	CVPR 2024
	Gaussian Shading: Provable Performance-Lossless Image Watermarking for Diffusion Models		
358	张雅琪, 黄迪, 刘斌, 唐诗翔, 陆岩, 陈路, 白磊, 储琪, 俞能海, 欧阳万里	中国科学技术大学	AAAI 2024
	MotionGPT: Finetuned LLMs Are General-Purpose Motion Generators		
359	朱佳莹，李东，傅雪阳，阳港，黄杰，刘爱萍，查正军	中国科学技术大学	AAAI 2024
	Learning Discriminative Noise Guidance for Image Forgery Detection and Localization		
360	朱禹睿，傅雪阳，姜鹏涛，张昊，孙启彬，陈锦伟，查正军，李博	中国科学技术大学	CVPR 2024
	Revisiting Single Image Reflection Removal In the Wild		
361	李鑫，廉东泽，卢治合，白家旺，陈志波，王鑫超	中国科学技术大学、新加坡国立大学	NeurIPS 2023
	GraphAdapter: Tuning Vision-Language Models With Dual Knowledge Graph		
362	齐天浩，方山城，邬彦泽，谢洪涛，刘佳伟，陈朗，何茜，张勇东	中国科学技术大学、字节跳动	CVPR 2024
	DEADiff: An Efficient Stylization Diffusion Model with Disentangled Representations		
363	赖浩然 姚青松 蒋子航 王嵘晟 贺志阳 陶晓东 周少华	中国科学技术大学 苏州高等研究院	CVPR 2024
	CARZero: Cross-Attention Alignment for Radiology Zero-Shot Classification		
364	胡世宇，赵鑫，黄梁华，黄凯奇	中国科学院大学	IEEE TPAMI 2023
	Global Instance Tracking: Locating Target More Like Humans		
365	杨禹雪；范略；张兆翔	中国科学院自动化研究所	ICLR 2024
	MixSup: Mixed-grained Supervision for Label-efficient LiDAR-based 3D Object Detection		
366	吴优，刘克安，弭晓月，唐帆，曹娟，李锦涛	中国科学院计算技术研究所	CVPR 2024
	U-VAP: User-specified Visual Appearance Personalization via Decoupled Self Augmentation		
367	刘俐君 王蕊 王源 荆丽桦 王川	中国科学院大学信息工程研究所	AAAI 2024
	Frequency Shuffling and Enhancement for Open Set Recognition		
368	张一帆，文青松，王雪，陈维奇，孙亮，张彰，王亮，金榕，谭铁牛	中国科学院大学自动化研究所	NeurIPS 2024

	OneNet: Enhancing Time Series Forecasting Models under Concept Drift by Online Ensembling		
369	焦培淇 闵越聪 李雅楠 王晓涛 雷磊 陈熙霖	中国科学院计算技术研究所	ICCV 2023
	CoSign: Exploring Co-occurrence Signals in Skeleton-based Continuous Sign Language Recognition		
370	胡民阳，常虹，马丙鹏，山世光，陈熙霖	中国科学院计算技术研究所	ICLR 2024
	Scalable Modular Network: A Framework for Adaptive Learning via Agreement Routing		
371	王思博、张杰、袁峥、山世光	中国科学院计算技术研究所	CVPR 2024
	Pre-trained Model Guided Fine-Tuning for Zero-Shot Adversarial Robustness		
372	王子涵，黎向阳，杨嘉豪，刘烨琦，胡俊杰，蒋明，蒋树强	中国科学院计算技术研究所	CVPR 2024
	Lookahead Exploration with Neural Radiance Representation for Continuous Vision-Language Navigation		
373	杨传广，安竹林，黄礼泊，毕君郁，于新强，杨晗，刁博宇，徐勇军	中国科学院计算技术研究所	CVPR 2024
	CLIP-KD: An Empirical Study of CLIP Model Distillation		
374	袁峥，张杰，蒋昭炎，李亮亮，山世光	中国科学院计算技术研究所	IEEE TPAMI 2024
	Adaptive Perturbation for Adversarial Attack		
375	孙隽姝，王树徽，韩歆哲，薛哲，黄庆明	中国科学院计算技术研究所、中国科学院大学	ICML 2023
	All in a Row: Compressed Convolution Networks for Graphs		
376	黎昆昌，王亚立，张钧皓，高鹏，宋广录，刘宇，李鸿升，乔宇	中国科学院深圳先进院	IEEE TPAMI 2023
	UniFormer: Unifying Convolution and Self-attention for Visual Recognition		
377	孙干，梁文奇，董家华，李俊，Zhengming Ding，丛杨	中国科学院沈阳自动化研究所	IEEE TPAMI 2024
	Create your world: Lifelong text-to-image diffusion		
378	刘佳伟，王强，范慧杰，王一秾，唐延东，屈靓琼	中国科学院沈阳自动化研究所	CVPR 2024
	Residual Denoising Diffusion Models		
379	朱子璇，王蕊，邹聪，荆丽桦	中国科学院信息工程研究所	ICCV 2023

	The Victim and The Beneficiary: Exploiting a Poisoned Model to Train a Clean Model on Poisoned Data		
380	何润泽，黄少飞，聂学成，惠天瑞，刘洛麒，戴娇，韩冀中，李冠彬，刘偲	中国科学院信息工程研究所	CVPR 2024
	Customize your NeRF: Adaptive Source Driven 3D Scene Editing via Local-Global Iterative Training		
381	荆丽桦，王蕊，任文琦，董鑫，邹聪	中国科学院信息工程研究所	CVPR 2024
	PAD: Patch-Agnostic Defense against Adversarial Patch Attacks		
382	刘洋，王峰，王乃岩，张兆翔	中国科学院自动化所	NeurIPS 2023
	Echoes Beyond Points: Unleashing the Power of Raw Radar Data in Multi-modality Fusion		
383	王宇琪，陈韫韬，廖星宇，范略，张兆翔	中国科学院自动化所	CVPR 2024
	PanoOcc: Unified Occupancy Representation for Camera-based 3D Panoptic Segmentation		
384	黄中昱，王赢珩，李朝卓，何晖光	中国科学院自动化研究所	IEEE TPAMI 2023
	Growing Like a Tree: Finding Trunks From Graph Skeleton Trees		
385	谭淏，李俊，周亦庄，万军，雷震，张祥雨	中国科学院自动化研究所	AAAI 2024
	Compound Text-Guided Prompt Tuning via Image-Adaptive Cues		
386	游泽彬，钟勇，鲍凡，孙嘉城，李崇轩，朱军	中国人民大学	NeurIPS 2023
	Diffusion Models and Semi-Supervised Learners Benefit Mutually with Few Labels		
387	郑晨宇，吴国强，李崇轩	中国人民大学	NeurIPS 2023
	Toward Understanding Generative Data Augmentation		
388	彭子乔、胡文韬、史悦、朱翔昱、张小梅、赵昊、何军、刘红岩、范肇心	中国人民大学	CVPR 2024
	SyncTalk: The Devil is in the Synchronization for Talking Head Synthesis		
389	卫雅珂，冯若轩，王子贺，胡迪	中国人民大学	CVPR 2024
	Enhancing Multi-modal Cooperation via Sample-level Modality Valuation		
390	杨泽群, 卫雅珂, 梁策, 胡迪	中国人民大学	ICLR 2024
	Quantifying and Enhancing Multi-modal Robustness with Modality Preference		
391	吕凡，孙晴，尚凡华，万亮，冯伟	中国科学院自动化研究所、天津大学	ICCV 2023
	Measuring asymmetric gradient discrepancy in parallel continual learning		

392	杨泽娴，吴大衍，吴陈铭，林政，谷井子，王伟平	中国科学院信息工程研究所	CVPR 2024
	A Pedestrian is Worth One Prompt: Towards Language Guidance Person Re-identification		
393	李鸿鑫，苏靖然，陈韫韬，李青，张兆翔	中国科学院自动化研究所	NeurIPS 2023
	SheetCopilot: Bringing Software Productivity to the Next Level through Large Language Models		
394	苏修，游山，谢吉洋，王飞，钱晨，张长水，徐畅	中南大学	IEEE TPAMI 2023
	Searching for Network Width With Bilaterally Coupled Network		
395	何宗耀、金枝	中山大学	CVPR 2024
	Latent Modulated Function for Computational Optimal Continuous Image Representation		
396	钟珊珊，黄中展，高尚华，温武少，林倞，Marinka Zitnik，周攀	中山大学	CVPR 2024
	Let's Think Outside the Box: Exploring Leap-of-Thought in Multimodal Large Language Models with Creative Humor Generation		
397	周昱臣，刘林凯，苟超	中山大学	CVPR 2024
	Learning from Observer Gaze: Zero-Shot Attention Prediction Oriented by Human-Object Interaction Recognition		
398	黄中展，周攀，颜水成，林倞	中山大学	NeurIPS 2023
	ScaleLong: Towards More Stable Training of Diffusion Model via Scaling Network Long Skip Connection		
399	黄福香，张磊，符小伟，宋苏琪	重庆大学	AAAI 2024
	Dynamic Weighted Combiner for Mixed-Modal Image Retrieval		
400	任开洁，张磊	重庆大学	CVPR 2024
	Implicit Discriminative Knowledge Learning for Visible-Infrared Person Re-Identification		
401	张磊，唐林渝	重庆大学	CVPR 2024
	Robust Overfitting Does Matter: Test-Time Adversarial Purification With FGSM		
402	张磊，符小伟，黄福香，杨易，高新波	重庆大学	IJCV 2024
	An Open-World, Diverse, Cross-Spatial-Temporal Benchmark for Dynamic Wild Person Re-Identification		
403			

404	陈松灿，Weikai Li	重庆交通大学	IEEE TPAMI 2023
	Partial Domain Adaptation without Domain Alignment		
405	徐宗懿，袁波，赵杉杉，Qianni Zhang，高新波	重庆邮电大学	ICCV 2023
	Hierarchical Point-based Active Learning for Semi-supervised Point Cloud Semantic Segmentation		
406	范略、杨禹雪、毛一鸣、王峰、陈韫韬、王乃岩、张兆翔	中国科学院自动化研究所	ICCV 2023
	Once Detected, Never Lost: Surpassing Human Performance in Offline LiDAR based 3D Object Detection		
407	刘波，胡斌，毕秀丽，李伟生，肖斌	重庆邮电大学	AAAI 2024
	Focus Stacking with High Fidelity and Superior Visual Effects		

合作单位简介

铂金合作单位

视拓云

视拓云团队来自于山世光老师创建的中科视拓，专注云计算和 AI 开发者社区两个细分领域，面向“大 AI 圈”内的科研工作者运营 AutoDL.com 和 CodeWithGPU.com 两个产品。AutoDL 提供弹性、好用、省钱的 GPU 云算力；CodeWithGPU 提供算法复现服务和内容交流社区。有算法就有复现，能复现才是好算法，复现算法就上 CodeWithGPU.com.

铂金合作单位

OPPO

OPPO 是全球五大智能手机品牌之一，也是全球领先的智能设备制造商和创新者。作为技术驱动型的科技公司，OPPO 在全球建立了 9 大智能制造中心、6 大研究所、5 大研发中心，以人工智能、云服务、大数据等前沿技术驱动、软件产品和互联网服务的开发，引领 5G、AI、影像处理、新材料新工艺等技术在智能终端上的发展应用。万物互融时代，OPPO 坚持“3+N+X”的科技跃迁战略。其中“3”指的是三项基础技术，包括硬件基础技术、软件基础技术和服务基础技术；“N”指的是能力中心，包括互联互通、多媒体、人工智能和安全隐私等；“X”指的是差异化技术，其中包含了 OPPO 的强项闪充、影像、新形态和 AR 增强现实技术等。

截至 2021 年 3 月，OPPO 全球专利申请数量超过 61,000 件，专利授权数量超过 26,000 件，并仍在持续增长。据世界知识产权组织(WIPO)发布 2020 年国际专利条约(PCT)申请数量排行榜，OPPO 全球排名前十。

专利搭载在产品上，就变成了功能与用户体验。除了持续引领 SUPERVOOC 闪充生态，推进 5G 普及之外，OPPO 近期迭代了一系列凝结科技的功能：凭借更准确的语义分析与用户实现情感交互冰冷的小布助手，带动机器学习与语音语义突破；借助全链路色彩管理系统、感知人像与画质增强引擎、视频防抖，背后是计算机视觉与多媒体等技术在硬件及软件上的创新。

OPPO 在智能终端布局日趋完善，以高端旗舰 Find X3 系列手机、Reno5 系列新品手机等智能手机为中心，持续完善 IoT 生态场景，推出了智能电视、OPPO Watch、OPPO 手环、Enco X 真无线降噪耳机等产品，为用户提供智能美好的完整体验。

在前沿科技层面，OPPO 秉承对美的一贯追求。OPPO X 2021 卷轴屏概念机是 OPPO 手机形态探索的最新成果，屏幕可如画卷般伸展，呈现几乎“零折痕”的屏幕效果。OPPO AR Glass 2021 则结合了全时空间计划 AR 应用，探索“虚实融合”数字世界进一步升级。

铂金合作单位

华为

华为是全球领先的 ICT 基础设施和智能终端提供商，致力于把数字世界带入每个人、每个家庭、每个组织，构建万物互联的智能世界。我们在通信网络、IT、智能终端和云服务等领域为客户提供有竞争力、安全可信的产品、解决方案与服务，持续为客户创造价值。

华为云 EI 是企业智能的使能者，通过云服务的方式（公有云、专属云等模式），提供一个开放、可信、智能的平台，结合产业场景，使能企业应用系统能看、能听、能说，让更多的企业便捷地使用 AI 和大数据服务，加速业务发展，造福社会。

华为消费者 BG AI 技术应用部是华为面向全场景智慧硬件的 AI 应用技术研发和能力中心。聚焦基于 1+8 硬件的计算机视觉、听觉、多传感器融合感知和识别技术、算力提升和数据分析预测技术，致力于面向消费者全场景打造 1+8 产品的硬件智慧体验。

中央媒体技术院是华为公司媒体技术创新与工程能力中心，肩负着公司手机拍照、ARVR、视频、音频及音视频标准、媒体体验与测评等领域的技术研究、创新和突破任务，确保华为公司媒体产品技术竞争力业界持续领先。

诺亚方舟实验室是华为的人工智能研究中心，立足于人工智能基础算法研究，聚焦打造数据高效和能耗高效的 AI 引擎，推动计算机视觉、语音和自然语言处理、决策推理等 AI 领域发展，助力华为公司主航道业务 AI 使能。

铂金合作单位

茶思屋

科学与技术上的学习争辩、切磋讨论是没有物理边界的。黄大年茶思屋就是这样一个不受物理位置的约束，线上线下无缝的、全天候的、开放的科学与技术交流平台。在这里，用户可以：通过学术热点和精选论文，共享学术前沿趋势、分享学术成功，呈现全球科研智慧；通过咖啡茶话，共享原汁原味的学术交流活动和专家观点，联接全球科研思想；通过高端论坛，参与全球重量级技术峰会，紧跟重大技术变革动态；通过科技赛事，观看全球顶级的科技热门赛事赛题，了解产学研合作的最新成果；此外，我们还将通过这一平台，发布技术难题，聚合最强大脑，展开技术合作。线上黄大年茶思屋聚焦学术领域的探索、牵引、开放、思辨，不触碰与科学和技术无关的话题，只要您走进黄大年茶思屋，都可以在这里就科学与技术话题展开思想碰撞、学术交流，畅所欲言，协作共进。

铂金合作单位

马上消费

马上消费金融股份有限公司（简称“马上消费”）是一家经原中国银监会批准，持有消费金融 牌照的科技驱动型数字金融机构。公司始终秉承“科技让生活更轻松”

的使命，以用户为中心，聚焦普惠金融，通过科技赋能创新， 致力于打造成为全球最被信赖的金融服务商，为有金融服务需求的社会各阶层和群体提供小额分散的消费金融服务。公司积极助力消费升级、拉动内需和服务实体经济发展，不断为地方税收、就业等经济社会发展作出积极贡献。截至 2023 年底，累计纳税超 78.1 亿元，直接 创造就业 3100 余人。

作为科技驱动的数字金融机构，马上消费在全球范围内选拔金融、大数据和人工智能等领域 高精尖人才，组建了超 2000 人的技术团队，以及大数据风控团队，先后成立了人工智能研 究院、国家级博士后科研工作站、大数据智能与隐私计算重庆市重点实验室、认知智能和数 智金融重庆市重点实验室，还与中国科学院大学、重庆师范大学等科研院校共建国家应用数 学中心，开展横向课题研究，不断提升产学研协作能力。基于上述科研基础，公司自主研发 了 1000 余套涵盖消费金融全业务流程、全生命周期的核心技术系统，累计提交发明专利申 请 1600 余项，稳居行业首位，获得软件著作权登记证书近 100 项。

截至 2023 年底，公司已获主体 AAA 信用评级和国家高新技术企业认定，入围工业和信息 化部“人工智能产业创新揭榜优胜单位”名单，当选全国信息技术标准化技术委员会人工智能 分技术委员会（SAC/TC28/SC 42）首批单位委员，两项科研课题获中国银保监会银行业信 息科技风险管理课题非银机构组一类成果奖，多次获评央行征信中心“个人征信系统数据质 量工作优秀机构”，连续五年入选“中国服务业企业 500 强”等荣誉。

腾讯优图

金牌合作单位

腾讯优图

腾讯优图实验室成立于 2012 年，是腾讯公司旗下顶级人工智能实验室。优图聚焦计算机视觉，专注人脸识别、图像识别、OCR 等领域开展技术研发和行业落地，在推动产业数字化升级过程中，优图始终专注基础研究、产业落地两条腿走路的发展战略，与腾讯云与智慧产业深度融合，挖掘客户痛点，切实为行业降本增效。

与此同时，优图关注科技的社会价值，践行科技向善理念，致力于通过视觉 AI 技术解决社会问题，帮助弱势群体。

INTSIG
合合信息

金牌合作单位

合合信息

合合信息是一家人工智能及大数据科技企业，基于自主研发的领先的智能文字识别及商业大数据核心技术，为全球 C 端用户和多元行业 B 端客户提供数字化、智能化的产品及服务。

公司 C 端业务主要为面向全球个人用户的 APP 产品，包括扫描全能王（智能文字扫描及识别 APP）、名片全能王（智能名片及人脉管理 APP）、启信宝（企业商业信息查询 APP）3 款核心产品；公司 B 端业务为面向企业客户提供以智能文字识别、商业大数据为核心的服务，形成了包括基础技术服务、标准化服务和场景化解决方案的业务矩阵，满足客户降本增效、风险管理、智能营销等多元需求，助力客户实现数字

化与智能化的转型升级。凭借领先的自主研发技术、成熟的产品落地能力、优质的用户体验及服务质量，公司的 C 端产品覆盖了全球百余个国家和地区的亿级用户，B 端服务覆盖了近 30 个行业的企业客户。在 B 端业务方面，公司智能文字识别与商业大数据服务已覆盖了银行、证券、保险、政府、物流、制造、地产、零售等近 30 个行业的众多头部客户，《财富》杂志 2020 年发布的世界 500 强公司名单中，公司客户已覆盖超过 80 家。

金牌合作单位

阿里妈妈

阿里妈妈成立于 2007 年，是淘天集团商业数智营销中台，管理并运营淘天集团搜索、展示、信息流等多种消费场景下的营销产品及技术解决方案、覆盖互联网主流 APP 全域营销资源。

阿里妈妈拥有淘天集团独家核心商业数智能力，致力于深研 AI 前沿技术，引领了 AI 在互联网营销领域的探索和大规模应用，并通过技术创新驱动淘天集团数智营销业务高速成长。

2021 年阿里妈妈发布全新的品牌主张“让每一份经营都算数”及全新的经营方法论-“深链经营”， 全面实现从全域营销到全域经营的升级。

金牌合作单位

美团

美团是一家科技零售公司。美团以“零售+科技”的战略践行“帮大家吃得更好，生活更好”的公司使命。自 2010 年 3 月成立以来，美团持续推动服务零售和商品零售在需求侧和供给侧的数字化升级，和广大合作伙伴一起努力为消费者提供品质服务。2018 年 9 月 20 日，美团在港交所挂牌上市。

美团始终以客户为中心，不断加大在新技术上的研发投入。美团会和大家一起努力，更好承担社会责任，更多创造社会价值。

金牌合作单位

百度

百度是拥有强大互联网基础的领先 AI 公司。是全球为数不多的提供 AI 芯片、软件架构和应用程序等全栈 AI 技术的公司之一，被国际机构评为全球四大 AI 公司之一。百度以“用科技让复杂的世界更简单”为使命，坚持技术创新，致力于“成为最懂用户，并能帮助人们成长的全球顶级高科技公司”。

百度公司 2000 年 1 月 1 日创立于中关村，创始人李彦宏拥有“超链分析”技术专

利，也使中国成为美国、俄罗斯、和韩国之外，全球仅有的 4 个拥有搜索引擎核心技术的国家之一。百度每天响应来自 100 余个国家和地区的数十亿次搜索请求，是网民获取中文信息和服务的最主要入口，服务 10 亿互联网用户。

基于搜索引擎，百度演化出语音、图像、知识图谱、自然语言处理等人工智能技术；最近 10 年，百度在深度学习、对话式人工智能操作系统、自动驾驶、AI 芯片等前沿领域投资，使得百度成为一个拥有强大互联网基础的领先 AI 公司。

金牌合作单位

极视角科技

极市(Extreme Mar)是极视角科技旗下的 Al 开发者生态，为开发者提供一站式线上便
捷算法开发平台，同时提供大咖技术分享及直播、社区交流与线下沙龙、以及一系列的算法竞赛等丰富的内容与服务。

极市开发者生态自 2015 年起，迄今已经积累超 240,000 名海内外专业算法开发者，影响力覆盖 300,000+Al 从业者/学生群体，极市希望与开发者们一起打造计算机视觉行业的生态圈，携手用算法改变世界。

金牌合作单位

联想研究院

联想研究院创立于 1999 年，是联想集团的公司级技术研发机构。从 PC 互联网时代，到移动互联网时代，再到智能互联网时代，联想研究院一直致力于推动 IT、计算机领域和智能设备和服务的技术发展，为联想的众多高科技产品和服务注入了最前沿的科技成果和理念，为提升联想全球用户的用户体验、打造引领时代潮流的生活方式而不断奋斗。

银牌合作单位

趋动科技

趋动科技作为软件定义 AI 算力技术的领导厂商，专注于为全球用户提供国际领先的数据中心级 AI 算力虚拟化和资源池化软件及解决方案，已完成中关村高新、国高新、“专精特新”等企业认证。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。趋动科技的双子座 GEMINI AI 训练平台，为客户提供强大的 AI 算力管理服务以及高效的算法开发和训练支持，能够化繁为简，帮助企业建好 AI 平台、管好 GPU、用好 AI 服务。

依托全球领先的 AI 算力池化技术，趋动科技重磅推出趋动云 VirtAI Cloud，为万千企业和 AI 开发者带来又便宜、又好用的 AI 算力池化云服务。凭借标准化、可复制的产品架构，趋动科技得到了包括互联网、金融、电信运营商、自动驾驶、能源、科研机构和高校等大量行业头部客户的认可。
资本市场对于趋动科技的发展充满信心——趋动科技成立两年多已经完成近亿美元的融资，顶级的投资机构持续支持趋动科技的发展，包括国开装备基金、沙特阿美旗下多元化风投基金 Prosperity7 Ventures、元禾重元、招银国际、顺为、高瓴、嘉御、戈壁、讯飞和涌铧在内的国内外顶级 VC 正在见证趋动科技锐意进取的脚步。

银牌合作单位

融科联创（天津）

融科联创（天津）信息技术有限公司，总部位于天津市武清区京津科技谷，是一家集服务器研发、生产、营销为一体的高新技术企业。公司建有专业的服务器定制化生产基地，可实现年产量 20 万台，并在北京、成都、杭州、长沙、深圳、武汉、西安、新加坡等地设立分支业务机构。

目前，融科联创业务已覆盖人工智能、云计算、互联网、教育、科研院所、广电、政府、工业企业、物联网等众多行业和应用领域。先后在云计算数据中心、高校 AI 教育、冷冻电镜生物分子研究、基因测序、广播电视录播系统、铁路智能安检、人脸识别、高性能计算平台、自动驾驶、无人配送车等行业应用方向成功落地。

多年来，融科联创获得了众多自主知识产权和荣誉资质。包括国家高新技术企业、天津市雏鹰企业、天津市瞪羚企业、天津市创新型中小企业、ISO 三体系认证、中国强制性产品 3C 认证，以及 80 余项专利和软件著作权认证等。

同时，融科联创还与超微（Supermicro STAP 成员）、英特尔（Intel 钛金级合作伙伴）、英伟达（NVIDIA 合作伙伴）等国际行业巨头紧密合作，整合产业优势资源，为客户提供优质的产品和解决方案，致力于成为人工智能、云计算、大数据、物联网领域卓越的解决方案提供商！

银牌合作单位

并行科技

北京并行科技股份有限公司（简称并行科技，股票代码：BJ839493）成立于 2007 年，总部位于北京市海淀区，注册资本 5823 万元。公司于 2023 年 11 月 1 日在北交所上市，是国内领先的超算云和智算云算力服务商，主要业务包括通用云、行业云、AI 云、设计仿真云等。

公司发展 17 年以来，旗下拥有 4 个研发中心、26 项专利和近百项计算机软件著作权，在北京、天津、广州、长沙、宁夏等 5 地设有子公司，在上海、南京、厦门、西安、武汉等城市设立 20 余个办事处和服务站。并行科技始终积极参与国家“算力网络”建设，打造了跨地域全网智能调度的算力网络 PaaS 平台，建立了业内领先的工业软件

SaaS 化服务体系，已发展成为国家高新技术企业、北京市“专精特新”中小企业、中关村高新技术企业。凭借领先的行业地位和高效的算力服务，公司拥有大批忠实用户，涵盖 300 多所知名科研机构、400 多所重点高校及 500 多家头部企业。

并行科技始终坚持以“助力科技强国，让计算更简单”为使命，构建国内领先的云上超算科研环境，形成集算力资源、应用资源、服务资源和人才资源于一体的超算云服务平台。公司聚焦各类用户在不同场景下的业务需求，持续推出满足各行业、各领域科研需求的综合性超算云服务解决方案，为众多来自科研高校、石油勘探、智能制造、地球环境、生命科学、人工智能等各应用领域的用户提供中国超算算力、应用、用户一体化云上科研工作环境。

并行科技基于算力网络服务模式、灵活的超算云平台架构，实现了全网超大规模算力调度技术突破，为用户提供最符合业务需求的算力资源，专业的技术服务团队，为用户提供 7×24 小时实时保障服务。2022 年，Frost&Sullivan（沙利文）发布《2021 年中国超算云服务市场报告》称：在中国通用超算云服务市场，并行科技排名第一。同年，公司与中山大学国家超级计算广州中心、国防科技大学共同完成的“工程计算与智能计算融合超算应用支撑平台”，荣获 2022 年度中国电子学会科技进步一等奖和广东省人民政府颁发的科技进步特等奖。

并行科技致力于成为世界领先的算力服务供应商，为人工智能、科研教育、智能制造、生命科学等领域的用户，持续提供高质量、高性能、高性价比的算力服务。助力科技强国，让计算更简单。

金山办公 KINGSOFT OFFICE

银牌合作单位

金山办公

北京金山办公软件股份有限公司（688111.SH）（以下简称“金山办公”），是中国领先的办公软件产品和服务提供商。作为一家源自中国的科技公司，秉持“让智慧绽放”的品牌理念，金山办公在过去 34 年持续深耕办公赛道，不断打磨技术和产品服务，始终致力于把最简单高效的办公体验带给众多个人、家庭和组织，帮助个人用户更轻松快乐的创作和生活，帮助组织客户更高效安全的运行与发展。

凭借以 WPS、金山文档、稻壳儿等为代表的办公产品和服务，金山办公为来自全球 220 多个国家和地区的用户提供办公服务，截至 2021 年 12 月，公司主要产品月度活跃设备数为 5.44 亿，其中 WPS office PC 版月度活跃设备数 2.19 亿，移动版月度活跃设备数 3.21 亿，持续领先其他国产办公软件。

思腾合力 SITONHOLY 加速AI梦想

银牌合作单位

思腾合力

思腾合力（SITONHOLY）成立于 2009 年，总部位于天津滨海区逸仙科学工业园，是 AI 服务器与 HPC 基础架构解决方案商，作为 NVIDIA 精英级别合作伙伴，专注于人工智能领域，提供云计算、AI 服务器、AI 工作站、系统集成、产品定制、软件开发、

边缘计算等产品和整体解决方案，拥有自主品牌 AI 服务器及通用 X86 服务器，适用于深度学习训练及推理等场景，覆盖服务器、静音工作站等多种产品形态，已经打造出了一套完全自主的软硬件结合的产品生态，致力于成为行业领先的人工智能基础架构解决方案商。

思腾合力拥有完善的研发、生产、制造基地，2021 年收购包头市易慧信息科技有限公司开启云计算业务，形成以天津为生产及研发基地，北京为营销中心，南京、深圳、成都、武汉、西安、内蒙古覆盖全国主要区域的营销和售后服务机构，为更全面地服务客户提供了有力保障。

思腾合力成立十多年来深耕教育及科研行业，从业 AI 领域研究、高性能计算的重点高校百分之八十都采用了思腾产品及解决方案，为各专业的科学实验研究提供了完备的 AI 加速解决方案。目前合作客户包括清华大学、北京大学、北京理工大学、中科院计算所、中科院自动化所、中科院半导体所、中科院信息工程所，以及国内知名人工智能公司等各企事业单位。

meitu | MT Lab 美图影像研究院

银牌合作单位

美图

美图公司成立于 2008 年，是一家以美为内核，以人工智能为驱动的科技公司。秉承着“让艺术与科技美好交汇”的使命，美图公司致力于打造优秀的影像与设计产品，让图像、视频、设计的制作变得更简单，并通过美业解决方案助力产业数字化升级。

2010 年，美图成立了核心研发部门——美图影像研究院（MT Lab），致力于计算机视觉、深度学习、计算机图形学等人工智能（AI）相关领域的研发，以核心技术创新推动公司业务发展。截至 2023 年 12 月，美图公司月活跃用户总数为 2.5 亿，其中海外月活跃用户数约 7768 万。

美图影像研究院（MT Lab, Meitu Imaging & Vision Lab）是美图公司致力于计算机视觉、机器学习、增强现实、云计算等领域的算法研究、工程开发和产品化落地的团队。MT Lab 位于北京、厦门、深圳三地，成员来自国内外知名高校和企业，期待和您一起让世界变得更美！

爱诗 AIsphere

银牌合作单位

爱诗科技

爱诗科技有限公司，简称“爱诗科技”，英文名为 AISphere。成立于 2023 年，总部位于北京。 爱诗科技是一家 AIGC 视觉多模态算法开发商，致力于探索 AI 视觉生成的无限潜力，通过 AIGC 技术赋能不同行业和场景，为视觉创意和智能生成开辟崭新前景，让用户拓展想象、释放潜力。 团队成员由一群拥有深厚学术背景和行业经验的专家组成，在计算机视觉、机器学习、计算机工程和算法设计等领域具备丰富的经验与技术储备。

银牌合作单位

快手

快手是一款由中国大陆的科技公司快手科技开发的短视频分享平台，用户可以在平台上创作、编辑和分享各种类型的短视频。快手于 2011 年正式上线，迅速成为中国最受欢迎的短视频应用之一。快手以其独特的算法和社交属性，吸引了大量年轻用户，形成了一个庞大的内容创作者和观众群体。快手的内容涵盖了生活、娱乐、教育、科技等多个领域，为用户提供了丰富多样的短视频体验。

银牌合作单位

虹软科技

虹软科技股份有限公司是计算机视觉行业领先的算法服务提供商及解决方案供应商，服务于世界各地的客户，将领先的计算机视觉技术与人工智能技术商业化应用在智能手机、智能汽车、智能家居、智能零售、互联网视频等领域，并且仍在不断探索新的领域与方向。公司在杭州、上海、南京、深圳、台北、硅谷、东京、都柏林等地设有商业与研发基地。

基于拥有自主知识产权的世界先进计算机视觉技术与人工智能技术，虹软为各领域提供一站式视觉解决方案，为全球各类知名的设备制造商提供个性化具有市场竞争力优势的行业解决方案与产品。在保持技术领先的同时，虹软率先发布了提供支持离线式图像技术的虹软视觉开放平台，与广大合作伙伴携手推动各类视觉技术应用深入到旅游、教育、政务、出行、社区楼宇、互联网应用等各个领域，引领和推动着视觉技术赋能和落地。

在超过 20 年的发展过程中，虹软成功聚集了众多的视觉领域专家，并吸纳和培养了来自国内外一流高校的优秀人才作为生力军。虹软将坚持聚焦在技术，注重技术与行业结合的应用经验，融合先进的学术科研力量，为全球的客户和消费者带来真正价值的视觉享受与体验。

银牌合作单位

生数科技

北京生数科技有限公司（简称“生数科技”）成立于 2023 年 3 月，核心团队成员来自清华大学人工智能研究院，此外汇集了来自阿里、腾讯、字节等知名科技公司的顶尖人才，是全球范围内领先的深度生成式算法研究团队，拥有扩散概率模型底层创新研发能力。公司致力打造世界领先的多模态大模型，融合文本、图像、视频、3D 等多模态信息，探索生成式 AI 在艺术设计、游戏制作、影视后期、内容社交等场景的商业赋能，通过 AI 提升人类的创造力和生产力。

银牌合作单位
思谋科技

思谋科技（SmartMore），智能制造的持续创新者，基于先进视觉技术，以深度学习和机器视觉作为技术引擎，应用领先的光学、机械、工业自动化、大模型等技术，助力加快发展新质生产力，扎实推进高质量发展。 经过二十多年的技术积累与沉淀，结合独特的商业思考，持续打造更具拓展性和普惠价值的智能工业和数智创新平台、工业大模型，以大模型和工业 AI-AOI 为核心，推动探索工业的智能化升级和数字化转型。

目前，思谋已通过自研的智能工业平台、智能传感器产品以及智能一体化设备，服务了卡尔蔡司、空客、博世、佳能、大陆集团、舍弗勒等来自全球超过 200 家行业头部企业，以技术促进更高效、更灵活、更先进智造的发展；此外，思谋还不断拓宽智造外延，基于“智造+”平台与数智化解决方案，自主研发了数字化制造管理系统，覆盖了从产线到工厂的应用场景，为客户在全球范围内提供全面而优质的产品与方案服务。

思谋由计算机视觉国际顶尖专家创立，自成立以来一直坚持国际化发展路线，成为了一家服务全球的全场景智能公司，吸引了近千名来自世界知名学府和行业领先企业的国际化人才。公司已在香港、深圳、上海、北京、苏州、杭州、重庆、新加坡和日本东京等多地设有前沿技术研发与商务中心，在东南亚和欧洲等地建立了代表处及合作伙伴网络，商业布局遍布全球

银牌合作单位
重庆长安

长安汽车是中国汽车四大集团阵营企业，拥有 162 年历史底蕴、40 年造车积累，全球有 12 个制造基地、22 个工厂。作为中国汽车品牌的典型代表之一，长安汽车旗下包括长安启源、深蓝、阿维塔、长安引力、长安凯程、长安福特、长安马自达、江铃等品牌。2017 年，长安汽车发起“第三次创业——创新创业计划”，将文化、效率和软件能力打造成为核心竞争力，向智能低碳出行科技公司转型。在新能源领域，2017 年发布“香格里拉”计划，2022 年发布数字纯电品牌“深蓝汽车”，目前已掌握 400 余项核心技术。在智能化领域，2018 年发布“北斗天枢”计划，2022 年发布智能品牌“诸葛智能”，目前已掌握 200 余项核心技术。在海外领域，2023 年发布“海纳百川”计划，坚持“长期主义、绿色低碳、本地运营、发展共赢”的原则，确立“海外市场投资突破 100 亿美元，海外市场年销量突破 120 万辆，海外业务从业人员突破 10000 人，将长安汽车打造成世界一流汽车品牌”的“四个一”发展目标。长安汽车始终以“引领汽车文明，造福人类生活”为使命，以科技创新为驱动，重塑能力、升级产业，以更快的速度、更大的强度，坚定坚决向智能低碳出行科技公司转型，向社会作出源源不断的贡献，不断满足人们更加美好的生活需要，奋力推进第三次创业——创新创业计划，为打造世界一流汽车品牌努力奋斗。

MRDVS
Mobile Robot Vision Expert

银牌合作单位
迈尔微视

浙江迈尔微视科技有限公司，自 2016 年组建核心技术团队，汇聚了中科院、复旦、浙大及国内顶尖机器视觉企业的精英。公司秉承“让机器人‘看懂’世界，服务世界”的使命，专注于为移动机器人和人形机器人开发专用的 3D 视觉传感器，并提供 3D+AI 算法一体化的视觉解决方案。迈尔微视已成功推出 S、M、V、H 四大产品系列，累计交付超过万台 3D 视觉传感器，显著提升了移动机器人在视觉避障、对接、导航和抓取等关键功能上的性能。

S 系列是专为避障应用设计的 3D 工业相机，结合内置 AI 算法，能智能识别并分类低矮、悬空障碍物，实现精准避障；V 系列深度相机为定位导航而生，1-12m 的测量范围配合顶视导航方案，通过天花板特征进行稳定定位，内置 IMU 和定位算法，实现±1cm 的重复定位精度；M 系列相机专注对接应用，提供 0.3m-5m 的深度数据，搭载自研托盘识别、多层料笼堆叠、库位检测、体积测量等 AI 算法，为无人叉车和智能物流提供高效视觉识别整体解决方案；H 系列是是专为高精度抓取设计的结构光 RGB-D 相机，结合迈尔微视自主研发的 AI 算法，实现了±0.1mm 的抓取精度，适用于机器人的抓取应用。

迈尔微视以技术创新为驱动，不断推动机器人视觉技术的发展，致力于让机器人更好地服务世界。

西云算力
算 力 改 变 未 来

银牌合作单位
西云算力

西云算力科技有限公司，原宁夏天云网络科技有限公司，致力于成为全球最有效率的绿色安全算力运营商。西云算力于 2016 年成立，由亚信科技联合创始人、宽带资本董事长田溯宁博士创立，是宁夏中卫市第一家云计算大数据企业。

随着 2022 年国家“东数西算”战略的提出，西云算力依托其高可靠、高效能、绿色数据中心，启动了宁夏“十四五”重点项目--总投资数十亿元的“西部云基地算力自主可控服务平台”建设，为大数据、人工智能、高性能计算等行业提供专用算力服务。


博汇巴渝
智创未来
博士后管理服务“掌中宝”
招收计划
识别二维码
成果集
解读
快闪VCR
直达全服务
政策
服务
重庆市博士后管理办公室
cqpostdoctoral@163.com
CHONGQING

部分组织单位简介

中国人工智能学会 Chinese Association for Artificial Intelligence

主办单位：中国人工智能学会

中国人工智能学会（Chinese Association for Artificial Intelligence，CAAI）成立于 1981 年，是经国家民政部正式注册的我国智能科学技术领域唯一的国家级学会，是全国性 4A 级社会组织，挂靠单位为北京邮电大学；是中国科学技术协会的正式团体会员，具有推荐“两院院士”的资格。

目前拥有 51 个分支机构，包括 43 个专业委员会和 8 个工作委员会，覆盖了智能科学与技术领域。学会活动的学术领域是智能科学技术，活动地域是中华人民共和国全境，基本任务是团结全国智能科学技术工作者和积极分子通过学术研究、国内外学术交流、科学普及、学术教育、科技会展、学术出版、人才推荐、学术评价、学术咨询、技术评审与奖励等活动促进我国智能科学技术的发展，为国家的经济发展、社会进步、文明提升、安全保障提供智能化的科学技术服务。

学会自主创办全球人工智能技术大会、中国人工智能大会、中国智能产业高峰论坛、中国 AI+创新创业大赛、“华为杯”全国大学生智能设计竞赛、国际人工智能会议、IEEE 云计算与智能系统国际会议等规模化、系列化学术活动。拥有“吴文俊人工智能科学技术奖”、“中国人工智能学会优秀博士学位论文评选”、“学会会士评选”、“学会先进个人”等多个奖项与评选。

欢迎广大科技工作者踊跃加入中国人工智能学会！

中国图象图形学学会 CHINA SOCIETY OF IMAGE AND GRAPHICS

主办单位：中国图象图形学学会

中国图象图形学学会（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成员。由从事图像图形基础理论与应用研究、软硬件技术开发及应用推广的专家学者和相关科技工作者组成。

中国图象图形学学会的宗旨是团结广大图像图形领域的科技工作者，积极开展图像图形基础理论和高新技术的研究，促进该学科技术的发展和在国民经济各个领域的推广应用。本学会专业领域涵盖了数字图像处理、图像理解、计算机视觉、图像压缩与传输、体视技术、科学计算可视化、虚拟现实、多媒体技术、模式识别、计算机图像图形学、医学影像处理、计算机动画、空间信息系统等。

中国图象图形学学会的主要任务是：开展科学研究与学术交流，活跃学术思想，促进学科发展，推广先进技术，发现、培养和举荐人才，提供技术咨询与服务，普及图像图形科技知识，传播科学思想和方法，编辑出版学术和科普书刊，加强同国内外学术团体和科技工作者的友好交往。

VALSE——学术华尔兹

VALSE发起于2011年，是Vision And Learning SEminar的简写，取“华尔兹舞”之意。旨在为全球计算机视觉、模式识别、机器学习、多媒体技术等相关领域的华人青年学者提供一个平等、自由的学术交流舞台。发起VALSE的主要动机是我们深感中国计算机视觉与机器学习领域缺少一个以华人青年学者为主体的常态化学术交流舞台。有鉴于此，2011 年初，山世光、潘纲、刘青山和颜水成共同讨论了发起一个视觉与学习领域华人青年学者研讨会的想法，之后该想法得到了李学龙、徐东、周志华、马毅等青年学者的大力支持。为此，首届视觉与学习青年学者研讨会于2011年4月8日-9日在杭州成功举行。此后，主要发起人山世光、潘纲、刘青山、颜水成、李学龙等共同讨论确定了VALSE这一名称。之后由山世光牵头起草、逐步完善了VALSE作为一个学术社区的组织原则和发展规划，特别是大会讲者由指导委员会按照“诺贝尔奖”模式推荐和选举产生的原则。

截至目前，VALSE已成功举办13届，分别为VALSE2011（杭州），VALSE2012（西安），VALSE2013（南京），VALSE2014（青岛），VALSE2015（成都），VALSE2016（武汉），VALSE2017（厦门），VALSE2018（大连），VALSE2019（合肥），VALSE2020（线上），VALSE2021（杭州），VALSE2022（天津），VALSE 2023（无锡）。VALSE2024将于2024年5月5-7日在重庆举行，由中国人工智能学会、中国图象图形学学会主办，重庆邮电大学承办，重庆大学微电子与通信工程学院、中国图象图形学学会青年工作委员会协办。除大会演讲之外，VALSE年度大会还不断推陈出新，逐渐增加了 Poster/Spotlight、Demo、Tutorial、年度进展评述、Workshops和工业界技术分享等环节，参会人数也逐渐增长到了5000人以上。在此过程中，逐渐形成了由山世光、潘纲、刘青山、颜水成、李学龙、周志华、徐东、马毅、周昆、高新波、何晓飞、余凯、杨健、黄华、白翔等15名青年学者组成的指导委员会。2018年，VALSE确立了指导委员会委员45岁退休的原则，故高新波、马毅、杨健、周志华、汤进五位老师进入顾问委员会，同时吸纳了华刚、汪萌、虞晶怡和张敏灵四位老师进入指导委员会。此外，VALSE大会也吸引了越来越多的企业参加，已成为计算机视觉与机器学习领域“产-学-研”合作交流的重要平台。

为配合VALSE系列年度研讨会，VALSE主要发起人之一山世光于2014年6月18日创建了VALSE专业学术交流QQ群，即VALSE-A群（2000人满）。此后，逐渐开通了VALSE-B-R群（群号：137634472）。从而形成了一个近两万两千人的视觉与学习青年学者在线社区。自2014年9月开始VALSE每周或隔周定期举办VALSE Webinar学术报告会，在已有品牌活动的基础上，我们最新推出了《VALSE短教程》、《VALSE论文速览》以经济、便捷的在线形式，将众多青年学者的最新工作和学术思想呈现给散落在世界各地的青年学者和研究生。自2020年4月以来，活动迁移至B站直播。VALSE Online 迄今已组织288期的在线学术报告。特别是众多知名青年学者的亲临报告（如：颜水成、王晓刚、屠卓文、凌海滨、沈春华、张磊、朱军、李纯明、印卧涛、熊红凯、刘利刚、齐国君、刘烨斌、毕彦超、Philip Torr、华刚、刘小明等），更大大激励了VALSE Online的发展，目前VALSE B站有3.9万粉丝，B站所存放的视频都是VALSE每周的Webinar录制的视频，视频的累计播放量71.6万，单个视频的最高播放量在4.9万次。逐渐形成了一个独具特色、经济高效、便捷实用的在线学术交流舞台。VALSE历史视频都会更新在B站空间，欢迎在B站搜索VALSE_Webinar关注我们！也可以通过链接直接观看：https://space.bilibili.com/562085182/。

VALSE Online 是青年学者自组织、自管理的舞台。其兴起不仅得益于VALSE指导委员会成员的大力支持，更离不开逐渐形成的VALSE Online组织团队。除发起人山世光之外，越来越多的青年学者参与了进来。特别是白翔（华中科技大学）、程明明（南开大学）、孟德宇（西安交通大学）、彭玺（四川大学）、贾伟（合肥工业大学）、郑海永（中国海

洋大学）、纪荣嵘（厦门大学）、姬艳丽（电子科技大学）、张利军（南京大学）、章国锋（浙江大学）、左旺孟（哈尔滨工业大学）、张兆翔（自动化所）、何晖光（自动化所）、禹之鼎（CMU）、王乃岩（图森未来）、苏航（清华大学）、欧阳万里（悉尼大学）等，都为VALSE Online的发展付出了大量心血。为了更好地组织VALSE年度及在线活动，VALSE成立了常务AC委员会(LACC)，资深AC委员会(SACC)，执行AC委员会(EACC)（名单参见后面的委员会名单），220名青年学者积极参与到了相关活动的组织中。

关于VALSE的更多信息，请访问VALSE总主页：http://valser.org（特别鸣谢中国海洋大学郑海永教授搭建该平台）。欢迎大家扫码关注之后页面的VALSE微信公众号，查看VALSE的最新消息。

借此机会，我们要诚挚感谢本届VALSE大会的合作单位，包括：视拓云、OPPO、华为、茶思屋、马上消费、腾讯优图、合合信息、阿里妈妈、美团、百度、极视角科技、联想研究院、趋动科技、融科联创（天津）、并行科技、金山办公、思腾合力、美图、爱诗科技、快手、虹软科技、生数科技、思谋科技、重庆长安、迈尔微视、西云算力。感谢这些公司的负责人和联系人为赞助VALSE而做出的努力，谢谢您们！

上述成绩的取得更离不开众多 VALSER 们的支持和鼓励，尤其是众多常态化参与 VALSE Webinar 报告会的老师和同学们，我们深表谢意！今后，我们将继续集思广益、创新学术交流和合作模式，更好地搭建视觉与学习领域华人学术交流大舞台，为本领域的产学研发展起到更好的促进作用。

VALSE 在线活动参与方法介绍

1、VALSE 每周举行的 Webinar 活动依托 B 站直播平台进行，欢迎在 B 站搜索 **VALSE_Webinar** 关注我们！

直播地址：

https://live.bilibili.com/22300737；

历史视频观看地址：

https://space.bilibili.com/562085182/

2、VALSE Webinar 活动通常**每周三**晚上 20:00 进行，但偶尔会因为讲者时区问题略有调整，为方便您参加活动，请关注 VALSE 微信公众号：**valse_wechat** 或加入 VALSE QQ **R 群**，群号：**137634472**）；

***注：**申请加入 VALSE QQ 群时需验证**姓名、单位和身份，**缺一不可。入群后，请实名，姓名身份单位。身份：学校及科研单位人员 T；企业研发 I；博士 D；硕士 M。

3、VALSE 微信公众号一般会在**每周四**发布下一周 Webinar 报告的通知。

4、您也可以通过访问 VALSE 主页：**http://valser.org/**直接查看 Webinar 活动信息。Webinar 报告的 PPT（经讲者允许后），会在 VALSE 官网每期报告通知的最下方更新。

VALSE 各委员会

荣誉委员会

高新波	西安电子科技大学	马　毅	UC Berkeley
周志华	南京大学	杨　健	南京理工大学
汤　进	安徽大学	吴小俊	江南大学

指导委员会

白　翔	华中科技大学	何晓飞	浙江大学
华　刚	Wormpex AI Research	黄　华	北京师范大学
李学龙	中国电信集团	刘青山	南京邮电大学
潘　纲	浙江大学	山世光	中国科学院计算技术研究所
汪　萌	合肥工业大学	徐　东	香港大学
颜水成	天工智能	虞晶怡	上海科技大学
余　凯	地平线机器人	张敏灵	东南大学
周　昆	浙江大学		

常务 AC 委员会（LACC）

主席：	白　翔	华中科技大学	
副主席：	程明明	南开大学	纪荣嵘　厦门大学
常务 AC：	程明明	南开大学	白　翔　华中科技大学
	纪荣嵘	厦门大学	姬艳丽　电子科技大学
	韩　琥	中国科学院计算技术研究所	刘日升　大连理工大学
	贾　伟	合肥工业大学	欧阳万里　上海人工智能实验室
	孟德宇	西安交通大学	王楠楠　西安电子科技大学
	彭　玺	四川大学	章国锋　浙江大学
	王　琦	西北工业大学	张兆翔　中国科学院自动化所
	张利军	南京大学	左旺孟　哈尔滨工业大学
	郑海永	中国海洋大学	郭裕兰　国防科技大学
	戴玉超	西北工业大学	王兴刚　华中科技大学
	卢策吾	上海交通大学	魏秀参　东南大学
	苏　航	清华大学	周晓巍　浙江大学
	张姗姗	南京理工大学	

资深 AC 委员会（SACC）

主席：	姬艳丽	电子科技大学	
副主席：	严骏驰	上海交通大学	谢凌曦　华为技术有限公司
	张姗姗	南京理工大学	
资深 AC：	陈　涛	复旦大学	樊　彬　北京科技大学
	丛润民	北京交通大学	高常鑫　华中科技大学
	洪晓鹏	哈尔滨工业大学	高陈强　重庆邮电大学

胡　鹏	四川大学	连宙辉	北京大学
林巍峣	上海交通大学	刘　偲	北京航空航天大学
李冠彬	中山大学	明　悦	北京邮电大学
潘金山	南京理工大学	任传贤	中山大学
马　超	上海交通大学	王利民	南京大学
王文冠	浙江大学	王兴刚	华中科技大学
王云鹤	华为诺亚方舟实验室	夏　勇	西北工业大学
魏云超	北京交通大学	徐　畅	悉尼大学
许永超	武汉大学	严骏驰	上海交通大学
杨　猛	中山大学	张　磊	重庆大学
张　林	同济大学	张姗姗	南京理工大学
赵　健	中国电信&西北工业大学	郑伟诗	中山大学
谢凌曦	华为技术有限公司	江　波	安徽大学

执行 AC 委员会（EACC）

主席： 郭裕兰　国防科技大学

副主席： 胡　鹏	四川大学	雷柏英	深圳大学
李冠彬	中山大学	郑　乾	浙江大学
执行 AC： 白亚龙	京东 AI 研究院	贲晛烨	山东大学
曹　越	微软亚洲研究院	常晓军	蒙纳士大学
陈　涛	复旦大学	丛润民	北京交通大学
代登信	ETH Zurich	邓　欣	北京航空航天大学
丁长兴	华南理工大学	董　超	中国科学院深圳先进技术研究院
董伟生	西安电子科技大学	董宣毅	悉尼科技大学
窦　琪	香港中文大学	段立新	电子科技大学
冯尊磊	浙江大学	宫　辰	南京理工大学
宫明明	墨尔本大学	顾舒航	悉尼大学
郭晓杰	天津大学	韩　波	香港浸会大学
韩晓光	香港中文大学（深圳）	洪晓鹏	西安交通大学
胡　迪	中国人民大学	胡建方	中山大学
胡　鹏	四川大学	胡　玮	北京大学
黄　高	清华大学	黄　雷	北京航空航天大学
贾　旭	大连理工大学	江　波	安徽大学
焦剑波	牛津大学	柯秋红	墨尔本大学
雷柏英	深圳大学	李　策	兰州理工大学
李冠彬	中山大学	李皓亮	香港城市大学
李雷达	西安电子科技大学	李　爽	北京理工大学
李　文	电子科技大学	李　镇	香港中文大学（深圳）
林　迪	天津大学	林绍辉	华东师范大学
刘　峰	密歇根州立大学	刘　昊	宁夏大学
刘　俊	新加坡科技设计大学	刘同亮	悉尼大学

刘　洋	北京大学
马　超	上海交通大学
彭春蕾	西安电子科技大学
任文琦	中国科学院信息工程研究所
盛　律	北京航空航天大学
宋　杰	浙江大学
隋　尧	哈佛大学
涂志刚	武汉大学
汪婧雅	上海科技大学
王　鹤	北京大学
王旗龙	天津大学
王奕森	北京大学
王智慧	大连理工大学
文碧汉	新加坡南洋理工大学
吴　琦	阿德莱德大学
夏　勇	西北工业大学
徐　迈	北京航空航天大学
徐天阳	江南大学
杨　恒	深圳爱莫科技有限公司
杨文瀚	南洋理工大学
杨　旭	西安电子科技大学
姚权铭	第四范式
叶　茫	武汉大学
张鼎文	西北工业大学
张　健	北京大学
张瑞茂	香港中文大学（深圳）
张　正	哈尔滨工业大学（深圳）
赵文达	大连理工大学
郑　乾	新加坡南洋理工大学
周　毅	东南大学
朱霖潮	浙江大学
朱盈盈	华中科技大学
董胤蓬	清华大学
陈冠英	香港中文大学（深圳）
陈使明	美国卡耐基梅隆大学、阿联酋人工智能大学
程光亮	英国利物浦大学
代季峰	清华大学
冯　婕	西安电子科技大学
高广谓	南京邮电大学
郭　青	新加坡科技研究局
刘　宇	大连理工大学
牛玉磊	新加坡南洋理工大学
秦　杰	阿联酋起源人工智能研究院
沈　为	上海交通大学
舒祥波	南京理工大学
眭亚楠	清华大学
谭明奎	华南理工大学
万人杰	新加坡南洋理工大学
王　栋	大连理工大学
王立君	大连理工大学
王　鑫	清华大学
王　正	东京大学
韦星星	北京航空航天大学
吴金建	西安电子科技大学
吴庆波	电子科技大学
谢凌曦	华为数字技术有限公司
徐　易	阿里巴巴（美国）集团
杨二昆	北卡罗来纳大学教堂山分校
杨　曦	西安电子科技大学
杨　旭	中国科学院自动化所
杨　欣	华中科技大学
于乐全	香港大学
张弘扬	滑铁卢大学
张平平	大连理工大学
张长青	天津大学
赵恒爽	牛津大学
赵　洋	合肥工业大学
周天飞	ETH Zurich
朱　磊	University of Cambridge
朱鹏飞	天津大学
邹常青	华为加拿大研究院
曹相湧	西安交通大学
陈　浩	香港科技大学
陈威华	阿里巴巴达摩院
崔兆鹏	浙江大学
丁明宇	加州大学伯克利分校
傅雪阳	中国科学技术大学
顾　实	电子科技大学
郭宗辉	中国科学院计算技术研究所
黄文炳	中国人民大学高瓴人工智能学院

	胡庆拥	University of Oxford	况　琨	浙江大学
	霍　静	南京大学	李　响	西安交通大学
	李崇轩	中国人民大学	刘夏雷	南开大学计算机学院
	刘　晗	大连理工大学	刘　勇	中国人民大学高瓴人工智能学院
	刘瑶瑶	约翰霍普金斯大学	倪张凯	同济大学
	陆　昊	华中科技大学	任冬伟	哈尔滨工业大学
	屈靓琼	香港大学	唐彦嵩	清华大学
	孙奕帆	百度	王　啸	北京航空航天大学
	田亚鹏	德克萨斯大学达拉斯分校	武　宇	武汉大学计算机学院
	王一帆	大连理工大学	杨　帅	MMLab@南洋理工大学
	谢雨彤	澳大利亚阿德莱德大学	弋　力	清华大学
	杨　旭	东南大学计算机学院	于　茜	北京航空航天大学
	易　冉	上海交通大学	张　力	复旦大学
	元玉慧	微软亚洲研究院		
秘书处：	朱盈盈	华中科技大学	程　一	中国科学院计算技术研究所
	班瀚文	中国科学院计算技术研究所		

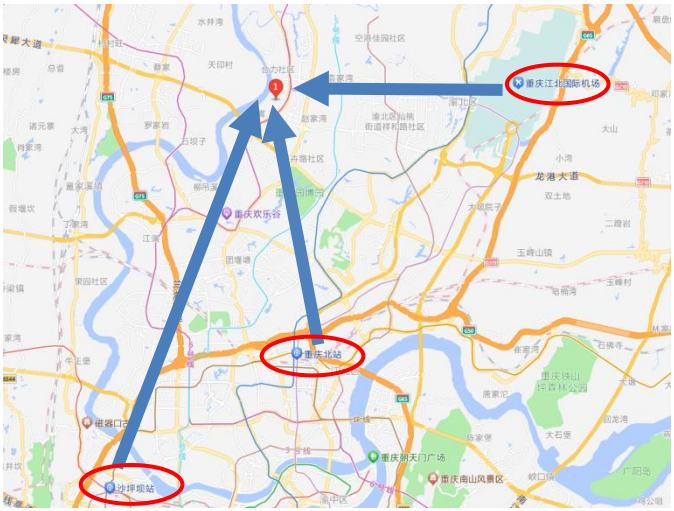
会场及酒店

会场位置

会议地点：重庆悦来国际会议中心
交通地址：重庆市渝北区滨江大道 86 号悦来会展城悦来国际博览中心内

交通信息

交通情况概览

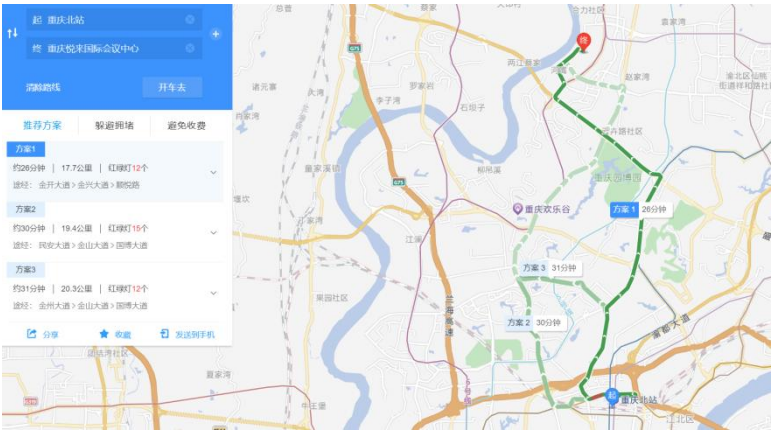


重庆江北国际机场

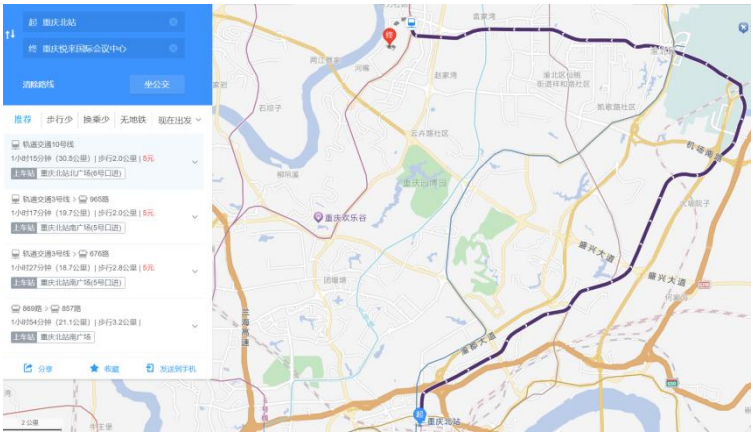
打车路线【全程估计 20 分钟】
距离会场约 11.4 公里，出租车约 20 元。

重庆北站

打车路线【全程约 26 分钟】
距会场约 17.7 公里，出租车约 27 元。

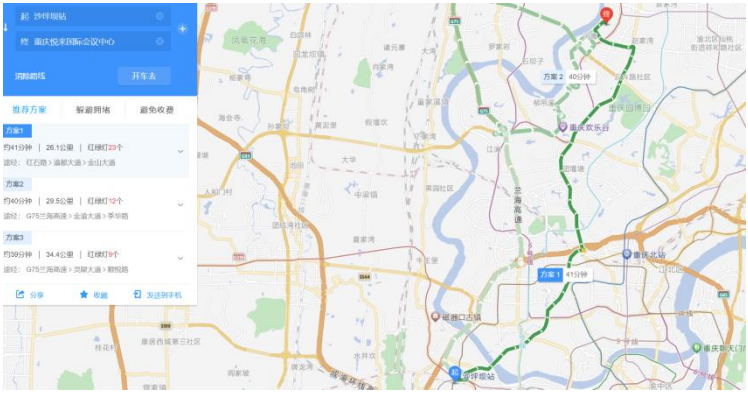


地铁路线【全程 1 小时 15 分钟】
乘坐轨道交通 10 号线（王家庄方向）：重庆北站北广场上车，悦来站下车。



沙坪坝站

打车路线【全程约 40 分钟】
距会场约 26.1 公里，出租车约 52-76 元



酒店信息

大会协议酒店信息

为方便各位老师和同学参会，组委会前期调研了会场周边的酒店，重庆悦来国际会议中心周围酒店列表如下。同时也非常建议参会者自行通过携程等公共平台自行搜索会场附近酒店。

组委会将在大会网站上公开电子版会议通知，方便各位参会者报销。酒店特惠订房协议价如下：

重庆金陵大饭店	单标间	370	重庆市渝北区仙桃街道春华大道99号	陈灏	17338333391	预定需要说出优惠口令“VALSE2024”，尽快预定，距离会场6.1公里/车程11分钟。
重庆熙美酒店	贵宾大/双床房	350	重庆市渝北区睦portal路160号13幢	周总	18696533241	预定需要说出优惠口令“VALSE2024”，尽快预定，距离会场6公里/车程11分钟。
重庆华辰国际大酒店	豪华大/双床房	350	重庆市渝北区空港新城百果路33号	肖娟	13883153410	预定需要说出优惠口令“VALSE2024”，尽快预定，距离会场8公里/车程15分钟。
维也纳酒店（国博店）	愉梦大/双床房	339	重庆市渝北区国博城悦城路53号	何经理	15922931042	预定需要说出优惠口令“VALSE2024”，尽快预定，距离会场3.7公里/车程8分钟。
禧满鸿福酒店	豪华大/双床房	509	重庆市渝北区腾芳大道8号	总机	023-61962999	预定需要说出优惠口令“VALSE2024”，尽快预定，距离会场7.4公里/车程14分钟。
重庆尚高丽呈酒店	优享轻奢大/双床房	268	重庆市渝北区空港工业园区长凯路315号	彭经理	17265160109	预定需要说出优惠口令“VALSE2024”，尽快预定，距离会场13公里/车程27分钟。
重庆银鑫世纪酒店	豪华大床房	459	重庆市渝北区回兴宝桐路9号	张经理	17708336410	预定需要说出优惠口令“VALSE2024”，尽快预定，距离会场13公里/车程15分钟。
	豪华双床房	489				
新华酒店	精致大/双床房	240	重庆市渝北区双龙大道138号	吴经理	13527354959	预定需要说出优惠口令“VALSE2024”，尽快预定，距离会场12公里/车程26分钟。
	商务大/双床房	260				
	豪华大/双床房	280				
重庆澜庭悦酒店	豪华景观大床房	540	重庆市渝北区滨江路84号附1-3号	总机	023-67807676	预定需要说出优惠口令“VALSE2024”，尽快预定，距离会场步行5分钟/500米以内。
	豪华景观双床房	620				
青风客栈	江景单人间	430	重庆市渝北区滨江路84号附1-4号	总机	023-67169955	预定需要说出优惠口令“VALSE2024”，尽快预定，距离会场步行5分钟/500米以内。
	商务标准间	430				
观江酒店	江景雅致大床房	530	重庆市渝北区悦来滨江路94号附1号	总机	023-81925966	预定需要说出优惠口令“VALSE2024”，尽快预定，距离会场步行5分钟/500米以内。
	舒适商务双床房	530				
朗逸酒店	商务单间	440	重庆市渝北区悦来滨江路94号附9号	总机	023-81925968	预定需要说出优惠口令“VALSE2024”，尽快预定，距离会场步行5分钟/500米以内。
	江景标间	540				
重庆曼德姆酒店	豪华大床房	530	重庆市渝北区悦来滨江大道94号附13-22	总机	023-67270777	预定需要说出优惠口令“VALSE2024”，尽快预定，距离会场步行5分钟/500米以内。
	豪华双床房	550				
铂尔斯酒店	舒适商务双床房	500	重庆市渝北区悦来滨江大道44号附6-30	总机	023-67252188	预定需要说出优惠口令“VALSE2024”，尽快预定，距离会场步行5分钟/500米以内。
	豪华商务大床房	540				
馨乐庭桃源居公寓酒店	行政单房公寓（大/双床）	440	重庆市渝北区长凯路2号	总机	023-67172555	预定需要说出优惠口令“VALSE2024”，尽快预定，距离会场13公里/车程18分钟。
奥蓝酒店	舒适商务大/双床房	230	重庆市渝北区双龙大道218号	张经理	18423198502	预定需要说出优惠口令“VALSE2024”，尽快预定，距离会场8公里/车程25分钟。


VALSE