

CSIG 图像图形学科前沿讲习班——“大模型理论方法及 技术实战”



讲习班现场

2023 年 12 月 3 日，由中国图象图形学学会（CSIG）主办，联合 CSIG 文档图像分析与识别专委会、中国科学技术大学以及科大讯飞股份有限公司共同承办的第 27 期 CSIG 图像图形学科前沿讲习班（IGAL27）在安徽合肥圆满闭幕。本次讲习班主题聚焦于“大模型理论方法及技术实战”，由北京科技大学的殷绪成教授和中国科学技术大学的杜俊副教授共同担任学术主任。来自全国各地的教师、学生以及企业工程师参与讲习班，现场座无虚席，气氛热烈。

开班仪式于 12 月 2 日上午举行，北京科技大学殷绪成教授代表主办方发表致辞。他对学术主任和报告嘉宾分享领域核心观点和技术表示感谢，并衷心感谢来自全国各地的参会者对活动的支持。殷教授希望与会的科技工作者能在为期两天的学习和交流中有所收获，并欢迎他们加入中国图象图形学学会。



吕岳教授作报告

讲习班的首场报告由华东师范大学通信与工程学院院长吕岳教授主讲，题为“多模态模型和生成式方法在文本识别中的应用”。吕岳教授分享了针对文本识别领域的创新方法，开拓了多模态模型的发展方向。他提出了图像-文本联合引导的学习方法，通过图像-文本对比学习建立了文本与图像之间紧密联系的方式。这种方法不仅将图像和文本视为同等重要的模态，还在训练过程中自动学习多模态特征表达，为提升识别性能带来了全新的思路。



张娅教授作报告

随后，上海交通大学首席研究员张娅教授分享了关于“医疗领域基础模型的探索与实践”的演讲。张娅教授就医疗领域基础模型构建方面的问题，深入探讨了隐私挑战下数据汇聚的难题，并专注于多方协同计算的解决方案；同时，她还深入研究了知识与数据融合的代表学习方法，以提高模型在医学诊断中的准确性和可解释性。



王士进博士作报告

科大讯飞副总裁王士进博士分享了题为“通用人工智能的技术进展和典型应用”的演讲。在报告中，王士进博士深入探讨了人工智能的发展阶段，着重突显了认知大模型作为通用人工智能代表所引发的全球关注和科技革命浪潮。他还详细分析了从认知大模型到多模态大模型的技术特性、发展趋势以及应用价值，并指出了大模型面临的三大挑战和未来技术发展的方向。最后，王士进博士总结了科大讯飞在星火大模型研发中所取得的成果，特别强调了在教育、医疗、办公和工业领域的应用探索经验，以及大模型在各行业应用中扮演的关键角色。



殷绪成教授给吕岳教授颁发证书



殷绪成教授给张娅教授颁发证书



殷绪成教授给王士进博士颁发证书

最后，殷绪成教授亲自颁发了证书给上午做报告的三位嘉宾，表达了对他们精彩报告的感激之情。这个简短的仪式不仅是对他们卓越贡献的认可，也是对他们会议上所分享见解的深深感谢。



乔宇教授作报告

下午的首场报告由上海人工智能实验室主任助理、领军科学家乔宇教授分享，主题是“书生通用大模型体系”。乔宇教授在报告中介绍了通用多模态大模型的最新进展，重点聚焦于上海人工智能实验室打造的“书生”通用大模型体系，并深入分析了未来的发展趋势。



刘静研究员作报告

接着，中科院自动化所，国家优青获得者刘静研究员分享题为“多模态大模型的研究与应用”。刘静研究员详细阐述了从单模态到多模态预训练模型的发展历程，指出了多模态预训练在各领域如多模态理解、搜索、推荐、问答等应用中的巨大潜力。报告包括对多模态预训练模型重要性与必要性的分析、关键技术的概要描述、当前前沿进展的回顾以及对未来发展的思考和展望。



杜俊副教授给乔宇教授颁发证书



杜俊副教授给刘静研究员颁发证书

最后，由中国科学技术大学的杜俊副教授颁发证书给下午报告的两位嘉宾，以表示对他们出色演讲的诚挚感谢。



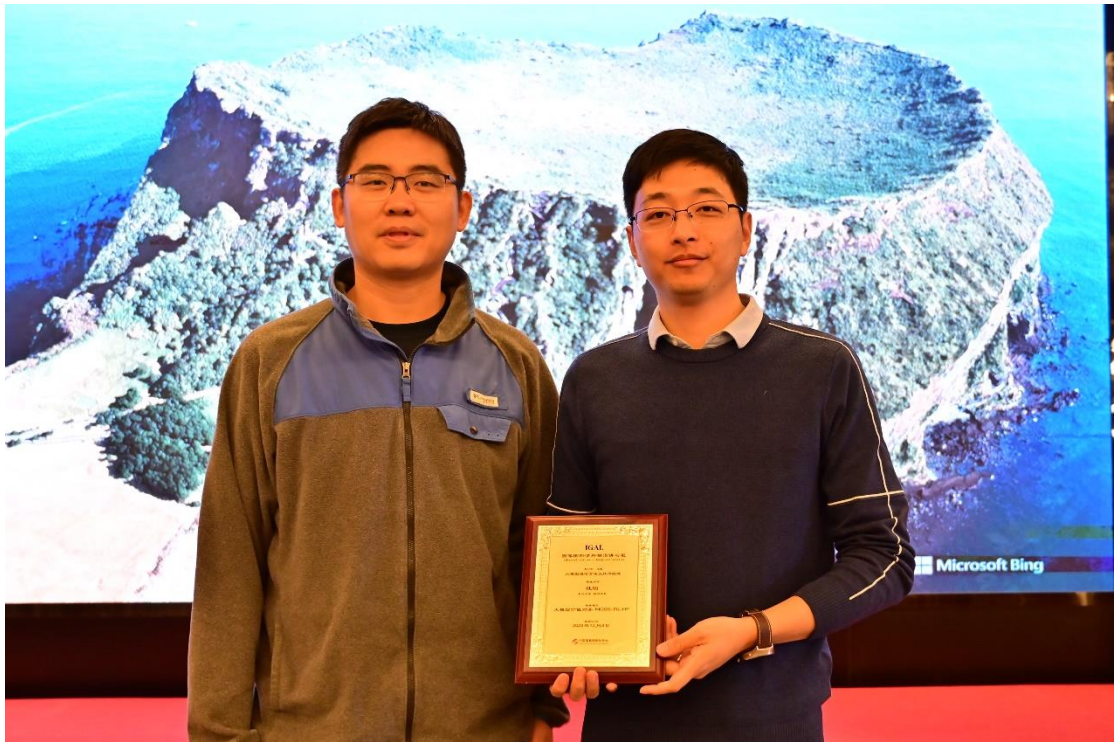
桂韬博士作报告

第二天上午，复旦大学青年副研究员桂韬博士分享了题为“大模型能力对其 MOSS-RLHF”的报告。他聚焦于基于 Transformer 的大型模型在自然语言处理领域的崭露，并对 NLP 及其他领域中模型同质化趋势提出了关切。重点强调了大模型同质化可能带来的安全、公平和隐私等问题，并探讨了解决这些挑战的方法。报告着重涵盖了大模型在安全伦理、社会影响方面的风险，并探讨了通过强化学习算法实现价值观对齐的方法。



白帅高级算法工程师作报告

随后，阿里巴巴通义实验室的高级算法工程师白帅研究员分享了题为“Qwen-VL：探索通用多模态大模型”的报告。这份报告围绕着多模态大模型的发展，着重介绍了 Qwen-VL 通用多模态大模型的构建、训练方法以及评估标准。白帅研究员特别强调了 Qwen-VL 在实际应用中的情况和实践经验，并对未来多模态大模型领域的发展趋势进行了展望，为进一步探索提供了基础和方向。



杜俊副教授给桂韬博士颁发证书



杜俊副教授给白帅研究员颁发证书

最后，由中国科学技术大学的杜俊副教授颁发证书给上午报告的两位嘉宾，以表彰他们出色的演讲表现。



王井东博士作报告

下午，百度计算机视觉首席科学家王井东博士分享了题为“视觉大模型：算法、应用与思考”的报告。报告首先介绍了文心·CV 大模型中的自监督表征学习算法 Context Autoencoder (CAE) 和基于预训练的目标检测算法 Group DETR。随后，从学习物体部件的角度阐述了几个自监督预训练算法的特点，并详细探讨了工业视觉大模型、人体图像大模型等的应用场景。此外，报告还涉及了 OCR 领域中大模型的研究和应用，同时对视觉大模型的未来方向进行了深入探讨，包括视觉模型预训练的重点、语言模型与视觉模型的融合，以及大模型时代下视觉研究课题的新趋势，如数据、优化和网络结构。



杜俊副教授给王井东博士颁发证书

随后，由中国科学技术大学杜俊副教授给王井东博士颁发证书，以表彰其所作的精彩报告。



Panel 环节

最后进行了大模型研讨环节，由中国科学技术大学的杜俊副教授担任主持人，邀请了来自学术界和工业界的专家学者参与讨论，包括华东师范大学的吕岳教授，中国科学技术大学的凌震华教授，上海合合信息技术的总监郭丰俊，以及科大讯飞的高级算法工程师张建树博士。在会上，他们就大模型对 OCR 领域的影响、如何优化 OCR 领域中的大模型，以及大模型未来发展的趋势等议题展开了深入的讨论。



合影环节

本次讲习班专家报告精彩、内容丰富，学员们踊跃提问、积极互动、收获颇丰，不仅加深了对专业领域的认知，还拓宽了视野和思维，更促进了同行之间的交流与合作。报告结束后，学术主任杜俊副教授在结营仪式上作了总结致辞，并对报告嘉宾的分享和参会者的支持表示感谢。第 27 期大模型理论方法及技术实战讲习班在各位报告嘉宾、广大同行以及学会的各级领导的大力支持下，取得了圆满成功。