

附件：

紫东太初多模态大模型研究集体

中国科学院自动化研究所

1、推荐单位推荐意见

该研究集体围绕多模态智能领域的国际前沿和国家战略需求，提出了多模态通用智能基础理论，基于全栈国产化基础软硬件研发了全球首个千亿参数“紫东太初”多模态人工智能大模型，开拓性提出跨模态多任务自监督学习等关键技术，实现了图像、文本、语音、视频等多种模态数据间的“统一表示”与“相互生成”，打通了模态间语义鸿沟，实现了人工智能模型从“预先定义的模式分类”到“开放式学习推理”的技术突破，形成了完整的智能表示、推理和生成能力，迈出了人工智能通用化的重要一步。建立了全栈自主可控多模态大模型开放服务平台，并在手语教考、交通检测、广电监管、手术导航等场景实现示范应用。2022年9月获得了世界人工智能大会最高奖项SAIL奖。

2、代表性论文专著和核心知识产权列表

代表性论文专著列表：

序号	论文（专著）名称	刊名	年卷页码	发表时间	全部作者
1	UniVIP: A Unified Framework for Self-Supervised Visual Pre-training	IEEE/CVF Conference on Computer Vision and Pattern Recognition	2022 年 14627-14636 页	2022 年 6 月 18 日	Zhaowen Li, Yousong Zhu, Fan Yang, Wei Li, Chaoyang Zhao, Yingying Chen, Zhiyang Chen, Jiahao Xie, Liwei Wu, Rui Zhao, Ming Tang, Jinqiao Wang
2	Dual Attention Network for Scene Segmentation	IEEE/CVF Conference on Computer Vision and Pattern Recognition	2019 年 3146- 3154 页	2019 年 6 月 16 日	Jun Fu, Jing Liu, Haijie Tian, Yong Li, Yongjun Bao, Zhiwei Fang, Hanqing Lu
3	Synchronous Bidirectional Neural Machine Translation	Transactions of the Association for Computational Linguistics	2019 年 7 卷 91 - 105 页	2019 年 4 月 1 日	Long Zhou, Jiajun Zhang, Chengqing Zong

4	MST: Masked Self-Supervised Transformer for Visual Representation	Advances in Neural Information Processing Systems	2021 年 34 卷 13165-13176 页	2021 年 12 月 6 日	Zhaowen Li, Zhiyang Chen, Fan Yang, Wei Li, Yousong Zhu, Chaoyang Zhao, Rui Deng, Liwei Wu, Rui Zhao, Ming Tang, Jinqiao Wang
5	Sequence to Multi-Sequence Learning via Conditional Chain Mapping for Mixture Signals	Advances in Neural Information Processing Systems	2020 年 33 卷 3735-3747 页	2020 年 10 月 20 日	Jing Shi, Xuankai Chang, Pengcheng Guo, Shinji Watanabe, Yusuke Fujita, Jiaming Xu, Bo Xu, Lei Xie

核心知识产权列表:

序号	知识产权 (标准) 类别	知识产权(标准)具体名称	国家 (地区)	授权号(标准编 号)	授权(标准发 布)日期	发明人(标准起草 人)	发明专利(标 准)有效 状态
1	发明专利	端到端的在线语音检测与识别方法、系统及设备	中国	CN112951213B	2022.05.24	周世玉, 徐波, 李蒙	有效

2	发明专利	掩码图像模型训练方法、掩码图像内容预测方法和设备	中国	CN114842307B	2022.10.28	朱优松, 李朝闻, 赵朝阳, 陈志扬, 王金桥	有效
3	发明专利	基于多模态预训练模型的跨模态理解与生成方法和装置	中国	CN113591902B	2022.09.27	刘静, 朱欣鑫, 刘飞、郭龙腾	有效
4	发明专利	多模态机器翻译方法、装置、电子设备和存储介质	中国	CN112800785B	2021.07.27	宗成庆, 黄鑫, 张家俊, 周玉	有效
5	发明专利	低资源多语言的语音识别模型、语音识别方法	中国	CN110428818B	2021.09.28	周世玉, 徐波	有效
6	发明专利	少样本目标检测方法、装置和电子设备	中国	CN114861842B	2022.10.28	朱贵波, 李宗树, 葛国敬, 赵旭, 王金桥	有效
7	发明专利	图像生成方法、装置、电子设备和介质	中国	CN115018954B	2022.10.28	吴凌翔, 王金桥, 牛蕴方	有效
8	发明专利	一种视频生成方法、装置、电子设备及存储介质	中国	CN114598926B	2023.01.03	王卫宁, 朱欣鑫, 刘静, 孙铭真, 刘佳伟	有效
9	发明专利	视觉声纹辅助的语音分离方法及装置	中国	CN113035225B	2023.02.28	许家铭, 张鹏, 石晶, 徐波	有效
10	发明专利	篇章因果关系判断方法、系统、装置	中国	CN109918646B	2020.08.11	向露, 刘洋, 张家俊, 周玉, 宗成庆	有效

3、研究集体成员贡献情况

姓名	人员类型	工作单位	主要贡献
徐波	突出贡献者	中国科学院自动化研究所	研发听觉预训练大模型、研究设计了多模态总体方案、领导大模型团队集体攻关
王金桥	突出贡献者	中国科学院自动化研究所	提出了通用视觉预训练模型设计及训练技术、高效自监督预训练等系列算法，搭建通用人工智能开放平台
刘静	突出贡献者	中国科学院自动化研究所	提出了图文音多模态大模型理解与生成理论框架
张家俊	主要完成者	中国科学院自动化研究所	提出以中文为核心的多语言大模型设计方法、自然语言生成新范式和超大规模大模型高效利用方法
周世玉	主要完成者	中国科学院自动化研究所	提出了语音大模型预训练训练方法和语音多任务统一建模技术
朱贵波	主要完成者	中国科学院自动化研究所	负责大模型国产化适配优化，构建了当时全国最大的中文多模态训练数据集进行多模态大模型训练

吴凌翔	主要完成者	中国科学院自动化研究所	负责国产化适配、多模态系统集成、大模型行业应用，负责与华为等公司合作，推动产业化落地
朱欣鑫	主要完成者	中国科学院自动化研究所	负责图文音三模态大模型的设计和实现，基于国产昇腾平台实现千亿参数模型，参与构建“小初”虚拟人
石晶	主要完成者	中国科学院自动化研究所	深入参与并研发了紫东太初多模态大模型语音部分的底层架构
朱优松	主要完成者	中国科学院自动化研究所	负责紫东太初多模态大模型中视觉基础预训练模型，提出了多粒度视觉掩码自监督预训练技术和大规模目标检测与识别技术

说明：公示内容必须与推荐书相关部分一致。