

大会简介

人工智能与健康医疗行业的深度融合迎来了最好的时代，中国也已成为全球数字医疗的领军者。为进一步引领学术发展，加强我国在此领域的影响力，上海交通大学决定主办 2024 年世界生物医药与人工智能大会。2024 年世界生物医药与人工智能大会主要主题为元宇宙医学与人工智能药物开发，设六个分会场：1、人工智能与药物开发；2、人工智能与计算生物学；3、人工智能与合成生物学、生物工程；4、人工智能与医学影像、数字病；5、企业家论坛；6、生物信息学方法。会议主要包含三个版块：报告交流、论文征集以及会员招募。我们成立了国际生物信息学会并招募会员，希望本次会议能提供一个平台，促进会员之间的交流与合作。大会最终将评选出优秀青年报告，以及优秀墙报。

The deep integration of artificial intelligence and the healthcare industry has entered its best era, with China now becoming a global leader in digital healthcare. To further lead academic development and strengthen China's influence in this field, Shanghai Jiao Tong University has decided to host the 2024 International Conference on Biomedicine and AI. The main themes of the 2024 conference are Metaverse Medicine and AI Drug Development, with five sub-forums: 1. AI and Drug Development; 2. AI and Computational Biology; 3. AI, Synthetic Biology and Bioengineering; 4. AI and Medical Imaging, Digital Diseases; 5. Entrepreneur Forum; 6. Computational Biology. The conference will include three key sections: report exchanges, paper submission, and membership recruitment. We have established the International Society for Bioinformatics and are recruiting members. We hope that this conference will provide a platform to promote communication and collaboration among members. The conference will recognize outstanding young presentations and excellent posters to honor the remarkable contributions and innovative thinking of young researchers.

大会主席（排列顺序不分先后）

魏冬青 彭绍亮

大会名誉主席（排列顺序不分先后）

陈润生 张春霆 李明 周如鸿 马剑鹏

Dennis Salahub Gren Patey

程序委员会主席（排列顺序不分先后）

马步勇 熊毅 顾若虚

大会秘书长

毛雪莹

副秘书长

孙鹤淇

会务组成员

袁雅静 安丹丹 申京冉 贾贵华 彭镇南

组织委员会成员（按姓氏拼音排序）

Abbas Khan Faez Iqbal Khan Gurudeeban Selvaraj

Kashif Iqbal Sahibzada Mtarhe Rui Satyavani Kaliamurthi

安丹丹 白晨 毕夏安 蔡宏民 曾坚阳 曾湘祥 陈海峰 陈铭 陈琦
陈图南 陈伟 陈禹保 陈增喜 陈宗青 程伟 褚晏伊 丁彦蕊 方艺田
高峰 高洁 龚新奇 关晓青 管仁初 郭祖奉 韩葳葳 胡广 胡黄丹
胡秋媛 胡小华 黄健 黄强 姬广富 蒋兴鹏 赖英杰 雷秀娟 李代禧
李国辉 李慧玉 李甲乙 李健 李金艳 李婧 李钧涛 李敏 李明
李书艳 李文金 李华 林志国 刘桂霞 刘娟 刘俊彦 刘笑 刘元宁刘
振栋 刘振明 吕晖 马梦玥 马婷 毛雪莹 孟浩 欧阳德方 潘建波庞
朝阳 邱育珊 曲强 石铁流 石婷 孙鹤淇 宋震 唐旭清 唐贇 王恒王
珊珊 王涛 王鲜芳 王艳菁 王燕 王宇 韦朝春 魏乐义 魏彦杰
吴红艳 吴辉 夏必胜 夏克林 谢鹭 谢宛青 谢志 徐峻 徐沁 徐云刚
许兆斌 薛志东 杨帆 杨风雷 杨广宇 杨晓慧 杨跃东 叶凯 余国先
俞章盛 员丽丽 袁雅静 张法 张贵军 张浩 张红燕 张雷 张永红
张友威 张玉娟 张岳 章文 赵建元 赵瑞丽 郑明月 周梅 周猛
周耀旗 邹权

上海交通大学生命科学技术学院（简介）

生命科学技术学院由上海交通大学与中国科学院上海分院联合于 1997 年 2 月建立。学院全体师生秉承交大“敢为人先”的精神和气魄，以“修德厚爱，健己惠人，强队谐群，凝特聚优”的学院文化为引领，二十年来，历经从无到有、由弱攀强的发展过程，2016 年第四轮全国一级学科评估中生物学跻身第一方阵，2017 年生物学被列入“双一流”建设学科名单，生物工程、生物与医药一级学科博士点新获批建设，学院走出了一条具有“交大”特色的生命学科跨越式发展之路。

学院目前设有生物化学与分子生物学系、遗传与发育科学系、微生物科学系、生物信息学与生物统计学系和生物工程系 5 个系，生命科学实验教学中心以及上海交通大学 Bio-X 研究院。学院拥有微生物代谢国家重点实验室、教育部遗传发育与精神神经疾病重点实验室、教育部代谢与发育科学国际合作联合实验室、农业部国家转基因生物分子特征验证测试中心、上海市转基因生物和食品安全专业技术服务平台等多个重点实验室和研究平台。

学院现拥有生物学、生物工程、生物与医药三个一级学科博士点，下设生物化学与分子生物学、微生物学、遗传学、发育生物学、细胞生物学、生物信息学、生物工程、生物与医药等二级学科；设有生物科学、生物技术、生物工程、生物信息学等 4 个本科专业及方向，生物技术专业、生物科学专业、生物工程专业入选国家级一流本科专业，生物科学专业入选教育部首批强基计划招生专业；另外还有生物工程领域工程硕士点、生物学博士后流动站。学院目前共拥有 2 个国家重点学科，3 个上海市重点学科，建成了具有交大特色且较为完善的现代生物学学科体系，生物学与生物化学、分子生物学与遗传学进入 ESI 全球前 1‰，微生物学、植物学与动物学、免疫学、精神病学与心理学等保持 ESI 全球前 1%。

大会议程

2024 年 10 月 25 日/下午		地点
14:00-20:00	报道注册	上海交通大学徐汇校区文治堂

2024 年 10 月 26 日/上午			
会议开幕式			
地点：文治堂 主持人：魏冬青教授			
8:20-8:30	刘卫东副校长致辞		
8:30-8:40	吕晖系主任致辞		
	合照		
大会报告			
地点：文治堂 主持人：熊毅研究员 顾若虚教授			
时间	报告人	单位	报告题目
8:40 - 9:00	李明院士	滑铁卢大学	人工智能解码人类免疫肽段组
9:00 - 9:20	周如鸿教授	浙江大学	Intelligent Screening and Design of Neoantigens for Cancer Immunotherapy
9:20 - 9:40	沈定刚教授	上海科技大学	多模态医疗大模型 - 从数据采集到疾病诊疗
9:40 - 10:00	黄宪达教授	香港中文大学	多组学与计算生物学解析中药分子机制及调控网络：以 microRNA 为关键调控分子

2024 年世界生物医药与人工智能大会 · 会议手册

10:00 – 10:20	彭绍亮教授	湖南大学	多模态生物医药大数据和人工智能 大模型
10:20 – 10:30	茶歇		
10:30 – 10:50	Diwakar Shukla 教授 (线上)	University of Illinois at Urbana- Champaign	Predicting substrate selectivity of proteins using Artificial Intelligence based methods
10:50 – 11:10	蒋太交教授	广州实验室	Predicting Influenza Virus Antigenic Variation and Vaccine Strains with PREDAC
11:10 – 11:30	李帅成教授	香港城市大学	Coding genomes with gapped pattern graph convolutional network
11:30 – 11:50	马步勇教授	上海交通大学	Protein conformation dynamics and AI prediction based on experimental structures in PDB
11:50 – 12:10	许东教授 (线上)	密苏里大学	Large protein language models and their prompt-based learning

2024 年世界生物医药与人工智能大会 · 会议手册

2024 年 10 月 26 日/下午 (分会场一)				
主题: 人工智能与药物开发				
地点: 新上院 N312 主席: 张玉娟、李惠玉、王鲜芳				
时间	报告人	所属机构	报告题目	报告类型
13:30 - 13:50	马剑竹	清华大学智能产业 研究院	基于靶点 3D 结构的小分子及短肽药物生成模型	特邀报告
13:50 - 14:10	何松	军事医学研究院	面向药物多终点毒性预测的智能算法	特邀报告
14:10 - 14:30	李代禧	上海理工大学	Intelligent Identification of Foodborne Pathogenic Bacteria by Self-Transfer Deep Learning and Ensemble Prediction Based on Single-Cell Raman Spectrum	特邀报告
14:30 - 14:50	张文艺	西湖大学	Towards structure-based AI drug design	特邀报告
14:50 - 15:05	雷秀娟	陕西师范大学	A Multi-modal Drug Target Affinity Prediction Based on Graph Features and Pre-trained Sequence Embeddings	口头报告
15:05 - 15:20	王鲜芳	河南工学院	Research on the Mechanism of Ziyin-Jianghuo Decoction in Treating Diabetes Based on PPI	口头报告
15:20 - 15:35	李慧玉	上海电力大学	Investigate the Potential Inhibitors of SphK1 with Molecular Dynamics and Artificial Intelligence Drug Design methods	口头报告
15:35 - 15:50	茶歇/墙报展示			
15:50 - 16:05	李瑞芳	河南工业大学	Deep learning combined with quantitative structure -	口头报告

2024 年世界生物医药与人工智能大会 · 会议手册

			activity relationship accelerates <i>de novo</i> design of antifungal peptides	
16:05 – 16:20	王小奇	西北工业大学	基于图深度学习的药物发现	口头报告
16:20 – 16:35	张佩宇	晶泰科技 CSO	AI+机器人赋能药物发现	口头报告
16:35 – 16:50	王旭辉	上海临科智华董事长	生成式 AI 对生物信息数据的分析的几点思考	口头报告
16:50 – 17:05	蚁佳才	国防科技大学	基于贝叶斯元学习的小样本分子性质预测	口头报告
17:05 – 17:20	吴锦洲	重庆师范大学	多模态药物分子性质预测中的图对比学习研究	口头报告
17:20 – 17:35	Muhammad Imran	上海交通大学	Molecular dynamics assisted in-silico analysis of DNA-binding role of myocyte enhancer factor-2 to explore potential inhibitors	口头报告
17:35 – 17:50	舒鑫	南昌大学公共卫生学院	Artificial Intelligence-based Digital Pathology for Early Tumor Metastasis Prediction and Prevention	口头报告
17:50 – 18:05	于筱艺	浙江师范大学	DASNet: A Convolutional Neural Network with SE Attention Mechanism for ccRCC Tumor Grading	口头报告

2024 年世界生物医药与人工智能大会 · 会议手册

2024 年 10 月 26 日/下午 (分会场二)				
主题: 人工智能与计算生物学				
地点: 新上院 N314 主席: 薛志东、刘振栋、李代禧				
时间	报告人	所属机构	报告题目	报告类型
13:30 - 13:50	肖传乐	中山大学	Development and application of multi-dimensional genome information analysis technology for three-generation sequencing	特邀报告
13:50 - 14:10	姜昊	中国人民大学	Heterogeneity analysis and modelling in single-cell data	特邀报告
14:10 - 14:30	王红	山东师范大学	Multi-Level Knowledge-Inspired Molecular Property Prediction	特邀报告
14:30 - 14:50	彭昱忠	南宁师范大学	Molecular properties prediction based on multi-view molecular representation learning	特邀报告
14:50 - 15:10	郑伟	南开大学统计与数据科学学院	Protein Complex Structure Prediction Enhanced by Deep Learning and Huge Metagenomics Data	特邀报告 (线上)
15:10 - 15:30	罗军伟	河南理工大学	Genomic structural variants detection based on deep learning	特邀报告
15:30 - 15:45	茶歇/墙报展示			
15:45 - 16:05	刘振栋	上海第二工业大学	The Algorithm of DNA Binding Site with Combined Feature Encoding	特邀报告
16:05 - 16:25	卢萌	北京大学临床医学高等研究院	ERnet: Transformer-Based Precise Segmentation and Multi-Parametric Analysis of Endoplasmic Reticulum from Super-Resolution Images	特邀报告
16:25 - 16:45	王雷	华中科技大学	Ultra-fast and Accurate Prediction of Protein-protein	特邀报告

2024 年世界生物医药与人工智能大会 · 会议手册

			Interaction Sites with ProtFormer-Site	
16:45 - 17:00	崔英博	国防科技大学	Deep learning-based structural variant filtering method	口头报告
17:00 - 17:15	许兆斌	德州学院	Dynamic Modeling of Antibody Repertoire Reshaping in Response to Viral Infections	口头报告
17:15 - 17:30	吕昊	电子科技大学	人工智能助力三维基因组结构解析	口头报告
17:30 - 17:45	李杰	北京全球健康药物研发中心	Leveraging AI Technology to Accelerate Preclinical Drug Pipelines: Insights from Academic to Industrial Transition	口头报告
17:45 - 18:00	姜自杰	上海交通大学	A deep learning-based method enables the automatic and accurate assembly of chromosome-level genomes	口头报告
18:00 - 18:15	刀福英	南洋理工大学	An Artificial Intelligence Platform for Identifying 3D Genome Organization for Cancer Treatment	口头报告

2024 年世界生物医药与人工智能大会 · 会议手册

2024 年 10 月 26 日/下午 (分会场三)				
主题: 人工智能与合成生物学、生物工程				
地点: 新上院 S411 主席: 杨广宇、张晨				
时间	报告人	所属机构	报告题目	报告类型
13:30 - 13:50	谢震	清华大学自动化系	ERNIE-RNA: An RNA Language Model with Structure-enhanced Representations	特邀报告
13:50 - 14:10	王宇光	上海交通大学自然科学学院	Generative AI for Protein Design	特邀报告
14:10 - 14:30	胡黔楠	中科院系统	AI+合成生物创新技术驱动高附加值原料合成	特邀报告
14:30 - 14:50	韩彦强	上海交通大学	人工智能赋能合成生物学	特邀报告
14:50 - 15:10	潘小勇	上海交通大学	蛋白绑定 RNA 位点预测及设计	特邀报告
15:10 - 15:30	王晓雷	上海交通大学	基于序列和结构信息的蛋白质多模态语言模型	特邀报告
15:30 - 15:50	茶歇/墙报展示			
15:50 - 16:10	卢元	清华大学	Toward intelligent artificial life	特邀报告
16:10 - 16:30	崔海洋	南京师范大学	The integration of protein engineering and computational science: a complementary path forward	特邀报告
16:30 - 16:50	周冰心	上海交通大学	Understanding and Engineering Proteins with Geometric Deep Learning	特邀报告
16:50 - 17:05	聂挺	上海交通大学	General Protein Language Model-assisted Multi-	口头报告

2024 年世界生物医药与人工智能大会 · 会议手册

			objective Optimization of CDase-mediated Transglycosylation for the Synthesis of α -O-oligosaccharides	
17:05 - 17:20	王瑞	重庆师范大学	蛋白质语言模型特征表示策略改进及泛素化位点预测	口头报告
17:20 - 17:35	关家浩	上海交通大学	基于蛋白质预训练语言模型的细菌 II 型毒素-抗毒素系统预测	口头报告
17:35 - 17:50	卞佳豪	上海交通大学	Efficient Enzyme Stabilization by Combining Multiple Mutations Using Protein Language Model	口头报告
17:50 - 18:05	Kashif Iqbal Sahibzada	河南工业大学	Enzyme Evo Predictor: An Ensemble-Classifiers Based Approach for Making Synergistic Prediction of Biodegradable Enzymes	口头报告

2024 年世界生物医药与人工智能大会 · 会议手册

2024 年 10 月 26 日/下午 (分会场四)				
主题: 人工智能与医学影像、数字病理				
地点: 新上院 S413 主席: 李书艳、杨旸、柯晶				
时间	报告人	所属机构	报告题目	报告类型
13:30 - 13:50	徐军	南京信息工程大学 信息学院	数据与知识驱动的图像计 算及其对精准医学的价值	特邀报告
13:50 - 14:10	艾丹妮	北京理工大学	多模态图像融合导航肝癌 精准诊疗研究	特邀报告
14:10 - 14:30	庄吓海	复旦大学	医学影像可解释分析	特邀报告
14:30 - 14:50	高伟	加州理工学院 (Caltech)	Skin-Interfaced Wearable Biosensors	特邀报告
14:50 - 15:10	蔡云鹏	中国科学院深圳先 进技术研究院	面向医学影像 AI 模型的可 解释知识提取方法及应用	特邀报告
15:10 - 15:30	胡振华	中国科学院自动化 研究所	近红外二区光学分子影像 技术及临床探索	特邀报告
15:30 - 15:45	茶歇/墙报展示			
15:45 - 16:05	金成	上海交通大学	基于多模态、多组学数据 融合的复杂疾病智能诊断 与归因分析 —— 大模型 时代的新范式	特邀报告

2024 年世界生物医药与人工智能大会 · 会议手册

16:05 - 16:25	陈宇飞	同济大学	证据深度学习与可信医学 辅助诊断决策	特邀报告
16:25 - 16:45	洪义	上海交通大学	Medical Multimodal Fusion for Computer-Aided Diagnosis of Alzheimer's Disease	特邀报告
16:45 - 17:00	魏红江	上海交通大学	人工智能驱动的精准确医学 成像	口头报告
17:00 - 17:15	沈逸卿	Johns Hopkins University	医学病理影像中的基础模 型	口头报告（线 上）
17:15 - 17:30	刘森昊	医流巴巴智慧科技 重庆有限公司	多模态一站式 AI 心理健康 诊疗评估系统	口头报告
17:30 - 17:45	Du Haoze	北卡罗来纳州立大 学	CT-Semi-Net: Segmentation of infected areas in lung CT images based on attention mechanism and semi- supervised learning	口头报告（线 上）
17:45 - 18:00	Mahrukh Latif	The University of Lahore	AI-Driven Fusion of Medical Imaging Modalities: Integrating PET, CT, and MRI	口头报告（线 上）

2024 年世界生物医药与人工智能大会 · 会议手册

2024 年 10 月 26 日/下午 (分会场五)			
主题：企业家论坛			
地点：新上院 S301			
时间	报告人	所属机构	报告题目
13:50 - 14:20	魏冬青 教授	上海交通大学生命科学 技术学院	人工智能超级计算精准药物发现 一花素抗衰老，癌症免疫与新冠
14:20 - 14:50	林菀骐 博士	中国康之莱健康管理创 办人	预见方能遇见-智能体在健康行业 的应用
14:50 - 15:20	郑超 博士	上海拜文生物医药股份 有限公司总经理	从生物医学科研成果到产业转化的 个人实践
15:20 - 15:50	益喜卓玛	扎西班智雅品牌负责人	融合传统智慧与 AI 科技:中藏药养 生产品 的创新研发
15:50 - 16:10	茶歇		
16:10 - 16:50	下半场圆桌对话		

2024 年 10 月 26 日/晚上 (晚宴)	
时间	地点
18:30 - 20:30	教师活动中心-聚贤阁

2024 年世界生物医药与人工智能大会 · 会议手册

2024 年 10 月 27 日/上午 (分会场六)				
主题：生物信息学方法				
地点：新上院 N100 主席：邵文广、张晨				
时间	报告人	所属机构	报告题目	报告类型
8:30 – 8:50	魏东辉	郑州大学	Predicting chemoselectivity of enzyme-catalyzed reactions based on Pywfn programme and artificial intelligence	特邀报告
8:50 – 9:10	黄琳	上海交通大学	基于人工智能的组学分析及临床诊断应用	特邀报告
9:10 – 9:30	孙端辰	山东大学	Identifying phenotype-associated cell subpopulations based on multi-omics data	特邀报告
9:30 – 9:50	邵文广	上海交通大学	基于生物质谱的免疫肽组学分析与研究	特邀报告
9:50 – 10:10	茶歇			
10:10 – 10:25	李甲乙	上海交通大学	Maneuver Interaction and quantum calculation of enzyme design: From molecular simulations to machine learning	口头报告
10:25 – 10:40	张育芳	上海交通大学	Prediction of drug-target interactions based on hypergraph neural networks with multimodal feature fusion	口头报告
10:40 – 10:55	林圣庚	上海交通大学	RNASTOP: Prediction and optimization of mRNA stability by integrating deep learning and thermodynamic	口头报告

2024 年世界生物医药与人工智能大会 · 会议手册

			property-guided heuristic search	
10:55 – 11:10	张晨	河南淇河实验室	基于自适性图卷积与膨胀卷积的空间转录组学空间域划分算法 STGIC	口头报告
11:10 – 11:40	闭幕式/颁奖			

2024 年 10 月 27 日/下午	
观赏交大校园	
时间	行程
14:00 – 15:00	参观徐汇校区新中院董浩云航运博物馆（注：请带实体身份证）
15:00 - 17:00	参观交大校园

主会场 · 报告嘉宾

李明 院士

滑铁卢大学校级教授 (University Professor)，加拿大皇家科学院院士，鹏城实验室前沿部智能数据压缩基础研究室责任院士，2010 年获得 Killam Prize。在 Nature, Nature Methods, Nature Machine Intelligence, Nature Communications, PNAS, Scientific American, J.ACM, CACM, FOCS, STOC 等杂志会议发表了具有影响力的文章。李明院士是研究 Kolmogorov 复杂性的世界权威专家，在研究机器学习，自然语言处理，算法平均复杂度，信息距离和生物信息学等方面解决了多个几十年未解难题，开创了新的研究领域，作出了重大贡献。他的工作开创并系统发展了 Kolmogorov 复杂性和计算机科学结合的交叉领域，这是计算机科学的基本支柱之一。



报告题目：

人工智能解码人类免疫肽段组

报告摘要：

人类基因组和人类蛋白组已经为我们带来了 21 世纪生物技术的革命。但是他们离大部分疾病或者疾病治疗有一定距离。距离大部分疾病（包括癌症，自免疫系统疾病，和各种感染）最近的是我们的免疫肽段组。今天，我们讲如何用人工智能结合质谱技术全面解码人类免疫肽段组。

周如鸿 教授

AAAS Fellow, APS Fellow, 浙江大学求是讲席教授, 生命科学学院院长, 上海高等研究院院长, 科技部十四五“生物大分子与微生物”专家组成员, 哥伦比亚大学化学系兼职教授。授权 30 余项国际专利, 在国际权威学术刊物发表论文 300 余篇 (其中 2 篇 Nature, 2 篇 Science, 1 篇 Cell, 23 篇 Nature 子刊, 17 篇 PNAS, 14 篇 JACS), 文章总引用超过 30000 余次, H 因子 88。曾获全美化学学会计算化学奖 (DEC Award), IBM 杰出科技贡献奖 (IBM 技术类最高奖), IBM 杰出创新奖, 并于 2009 年以 IBM 蓝色基因计划主要成员之一荣获奥巴马总统颁发的美国国家技术奖章 (总统奖)。



报告题目:

Intelligent Screening and Design of Neoantigens for Cancer Immunotherapy

报告摘要:

Cancer immunotherapy has been among the most promising breakthroughs in oncology, particularly in the case of immune check point inhibitors, however, the effective response rate remains quite low, only about 20-30%. In this presentation, I will talk about our recent work with molecular modeling and machine learning which solves one mystery behind this low response rate. We found that patients with certain HLA genotype (HLA-B44) have consistently higher survival rate, while patients with some other type (HLA-B15) have much poorer survival rates. Large scale molecular dynamics simulations further reveal that HLA-B15 proteins with poorer therapeutic outcomes had structural appendages that closed over the cancer neoantigens with much less flexibility. HLA typing thus might serve as a useful biomarker in future immunotherapy. Two effective neoantigens Neil3 and Myole are further identified for bladder cancer. The same techniques have also been applied to the design and development of HIV and T1D vaccines, which has been of great interest as well in recent years. With a combined in silico and in vivo approach, we studied the HLA-peptide-TCR interactions from multiple clonotypes specific for a well-defined HIV-1 epitope, and found that effective and ineffective clonotypes bind to the terminal portions of the HLA-peptide through similar salt bridges, but their hydrophobic side-chain packings can be very different, which accounts for the major part of the differences among these clonotypes. Meanwhile, a new x-autoantigen from a dual expressor (X-cell) has been identified for T1D patients, which shows a super potent binding affinity to HLA-DQ8, the main risk allele for T1D. Newly designed autoantigen mutants might serve as HLA-DQ8 blockers or vaccine candidates.

沈定刚 教授

上海科技大学教授、生物医学工程学院创始院长，IEEE/AIMBE/IAPR/MICCAI/ISMRM/IAMBE Fellow，联影智能联席 CEO，美国 The Academy for Radiology & Biomedical Imaging Research 杰出研究者奖。曾任美国 UNC-Chapel Hill 终身教授、冠名杰出教授，宾夕法尼亚大学 (UPenn) 助理教授，约翰霍普金斯大学 (Johns Hopkins University) 讲师。世界上最早开展医学影像人工智能研究的科学家之一，并最先将深度学习应用于医学影像。



报告题目：

多模态医疗大模型——从数据采集到疾病诊疗

报告摘要：

探讨多模态医疗大模型（特别医疗专业性大模型）的构建，为复杂医疗场景提升诊疗和决策，实现手术室的全景建模、实时术中导航、术后报告生成等应用。

黄宪达 教授

黄宪达教授曾于 1997 年台湾中央大学资讯工程学系获学士学位、2003 年台湾中央大学获得博士学位，曾任台湾交通大学生物科技学系终身教授、特聘教授、讲座教授、系主任、副院长、台湾生物资讯学会理事长。现任香港中文大学(深圳)生命与健康科学学院校长讲座教授、副院长(教学)和瓦谢尔计算生物研究院执行院长。



报告题目：

多组学与计算生物学解析中药分子机制及调控网络：以 microRNA 为关键调控分子

报告摘要：

中药是中国数千年医疗实践的宝贵经验结晶。然而，由于其配方复杂、治疗效果多样以及多成分多靶点的特性，使得在分子层面系统解析中药的机制充满挑战。近年来，我们结合计算与实验，从成分、靶标到调控网络多维度解析中药的作用机理。在成分与靶标层面，利用基于人工智能的药物靶标预测模型以及热漂移蛋白组等高通量实验技术，宏观推断中药的关键成分和潜在靶标。随后，通过分子动力学模拟和多种分子实验技术进一步锁定关键成分及其靶标。在调控网络层面，我们构建了以 miRNA 和氧化 miRNA 为核心的人类分子调控知识图谱，结合中药作用的多组学数据，使用多种算法对中药的调控网络进行量化和情景化分析，从而推断中药的作用机制及关键节点。这些推断的作用网络可以通过分子实验进行验证，揭示中药新的作用原理。我们认为结合人工智能与系统生物学，可以更全面地解析中药在分子层面的治疗机制，使得全球科学家更好地理解中药，并推动中药的现代化，提供新的见解和治疗选择。

彭绍亮 教授

长江学者，主持湖南省“创新群体”和“杰出青年”基金，国家超级计算长沙中心副主任，湖南大学“岳麓学者”二级教授/博导，2023 年获评湘江榜样（10 位），2023 湖南省数字化十大杰出人物。从事大数据、生物信息、人工智能、区块链等技术研究。出版学术专著 8 部，在 Nature Machine Intelligence 等期刊发表学术论文 200 余篇，论文引用 9000 余次，2024 年入选全球前 2% 顶尖科学家榜单。获 2019 年国家科技进步二等奖，2019 年湖南省技术发明一等奖（排名 1），2013 年军队科技进步一等奖 1 项，2021 年 CCF 技术发明二等奖（排名 1），2018 年 CCF 自然科学二等奖（排名 1），2016 年立三等功。



报告题目：

多模态生物医药大数据和人工智能大模型

报告摘要：

生命科学在大数据和人工智能的技术驱动下全面进入数智时代，从单细胞组学数据到群体队列数据，从电子病历到医学影像，从中心法则到精准医疗和手术机器人，从生物医药到人工智能疾病诊疗和数智大健康，生命科学全面进入了一个多模态大数据驱动的人工智能时代。2022 年当通用人工智能和生成式大模型的出现和全速发展，无论是人类三大疾病，还是地球物种繁衍，人工智能会帮助我们重新认识、解释和生成生命，打破时间和空间对生命的束缚，揭示人类疾病、健康和智能本原问题。国际国内目前也展开了对生物医药大数据和大模型的全面研究，报告会重点介绍目前生物医药领域的多模态数据类型和面临的存储、计算等问题和挑战，包括团队和国际国内最新的生物医药大数据、大模型研究进展。2023 年自主研发国产药物大模型-神农 GPT（开源），可以辅助医生精准处方和患者精准服药。通过建立多任务协同预训练大模型，深度融合分子结构和相互作用的多视图药物表征，进行大规模精准药物发现（包括疾病分型、敏感性、不良反应、毒性等）。实现了高通量药物大数据和人工智能双驱动下的干湿实验闭环研究。神农 GPT 药物大模型目前正在湘雅医院、广东省人民医院、北海市第二人民医院等单位得到应用。2021-2024 年，还持续研发了空间转录组单细胞大模型、单细胞多组学联合预训练大模型、多组学融合的肿瘤免疫表型计算平台、DNA 结合蛋白质大语言模型 ESM-DBP。

Diwakar Shukla 教授

Diwakar Shukla is an associate professor in the Department of Chemical and Biomolecular Engineering at University of Illinois at Urbana-Champaign. He is also an affiliate faculty in the Center for Biophysics and Quantitative biology, National Center for Supercomputing Applications, Department of Plant Biology and Department of Bioengineering. His research work is focused on understanding the complex biological processes using novel physics-based models and techniques. Shukla has received numerous early-career faculty awards from the American Chemical Society, American Institute of Chemical Engineers, Foundation for Food & Agriculture Research, Alfred P. Sloan Foundation, National Institutes of Health and the National Science Foundation.



报告题目:

Predicting substrate selectivity of proteins using Artificial Intelligence based methods

报告摘要:

In this talk, we will present recent work on developing a tailored language model for lasso peptides, which is one of the classes of Ribosomally synthesized and post-translationally modified peptides (RiPPs). Lasso peptides display a characteristic rotaxane conformation formed by a lasso cyclase. This unique, threaded conformation endows lasso peptides with diverse biological activities and remarkable thermal and proteolytic stability. The prediction of lasso peptide properties, such as substrate compatibility with a particular lasso cyclase or desired biological activity, remains challenging due to limited experimental data and the intricate nature of the substrate fitness landscapes. Protein language models (PLMs) have demonstrated impressive performance in predicting protein structure and function. However, general-purpose PLMs perform poorly in lasso peptide-related predictive tasks. Therefore, there is a need to provide effective representations for lasso peptides to enable enhanced property prediction. Furthermore, we also demonstrate a cross-attention-empowered SE(3)-equivariant graph neural network architecture named EZSpecificity for predicting enzyme substrate specificity, which was trained on a comprehensive tailor-made database of enzyme-substrate interactions at sequence and structural levels.

蒋太交 研究员

蒋太交，研究员，博士生导师，获中国科学院上海生物化学研究所生物化学与分子生物学博士学位以及耶鲁大学计算机科学硕士学位，并在耶鲁大学和麻省理工学院从事博士后研究工作；现任广州实验室研究员，当选第十四届全国人大代表。曾入选中国科学院人才计划，获得国家自然科学基金委杰出青年基金支持，被授予国家百千万人才工程“有突出贡献中青年专家”，获得中华医学科技奖一等奖、二等奖等荣誉。目前兼任中国生物物理学会常务理事及生物医学信息分会会长和中国预防医学会生物信息分会副主任委员。



报告题目：

Predicting Influenza Virus Antigenic Variation and Vaccine Strains with PREDAC

报告摘要：

Due to the development of high throughput sequencing technologies, large-scale sequencing of pathogens, such as influenza viruses, SARS-Cov-2 and HIV, has become a routine work in surveillance of infectious diseases, significantly enhancing our understanding towards the evolution of diseases, promoting new strategies for prevention and control of these viruses. In my laboratory, we are interested in developing computational methods to model the antigenic changes of influenza viruses from the viral surface protein haemagglutinin (HA) genes. In our previous work, we developed a computational method, denoted as PREDAC, to predict antigenic clusters of influenza A (H3N2) viruses with high accuracy from viral HA sequences. Later on, we extended this method to other influenza viruses including A(H1N1), A(H5N1) and type B viruses. In collaboration with China CDC, these computational methods have now been in use in China CDC for assisting flu vaccine strains selection.

李帅成 教授

香港城市大学教授，博导。研究重点是开发算法和计算工具来分析大规模基因组和蛋白质组数据，目的是了解复杂生物过程的潜在机制。在生物信息学方面的显著贡献包括开发用于基因组组装和注释的新算法、分析基因表达数据以识别疾病相关基因，以及开发用于预测蛋白质。在 *Nature Evolution and Ecology*、*Nature Communications*、*NAR*、*Bioinformatics* 等期刊发表学术论文多篇。



报告题目：

Coding genomes with gapped pattern graph convolutional network

报告摘要：

Genome sequencing technologies reveal a huge amount of genomic sequences. Neural network-based methods can be prime candidates for retrieving insights from these sequences because of their applicability to large and diverse datasets. However, the highly variable lengths of genome sequences severely impair the presentation of sequences as input to the neural network. Genetic variations further complicate tasks that involve sequence comparison or alignment. Inspired by the theory and applications of “spaced seeds,” we propose a graph representation of genome sequences called “gapped pattern graph.” These graphs can be transformed through a Graph Convolutional Network to form lower-dimensional embeddings for downstream tasks. On the basis of the gapped pattern graphs, we implemented a neural network model and demonstrated its performance on diverse tasks involving microbe and mammalian genome data. The method consistently outperformed all the other state-of-the-art methods across various metrics on all tasks, especially for the sequences with limited homology to the training data. In addition, our model was able to identify distinct gapped pattern signatures from the sequences. The framework is available at <https://github.com/deepomicslab/GCNFrame>.

马步勇 教授

马步勇，上海交通大学长聘教授，博士生导师。1995 年获美国佐治亚大学博士学位；1995-1998 佐治亚大学博士后，1998 年加入美国癌症研究所任资深科学家，2020 年回国加入上海交通大学药学院，致力于生物药物设计，结合大规模超级计算和人工智能推动新型药物开发。他在生物大分子动态结构、相互作用和功能方面颇有研究，总结提出了生物分子识别的构象选择（conformational selection）理论，成为与经典的‘Lock-Key’和‘Induced fit’并行的模型。在世界上首次揭示了阿尔兹海默病 A 淀粉体结构。研究结果在蛋白质结构的动态特性表征，酶催化，蛋白质-蛋白质相互作用和药物设计，神经退行疾病等方面被大量引用，促进了这些领域的发展。



报告题目：

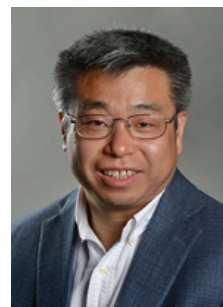
Protein conformation dynamics and AI prediction based on experimental structures in PDB

报告摘要：

Protein structure prediction has reached revolutionary levels of accuracy, implying biophysical energy function can be learned from known protein structures. However apart from single static structure, conformational distributions and dynamics often control protein biological functions. Towards this goal, we develop DeepConformer, a diffusion generative model for sampling protein conformation distributions from a given amino acid sequence. Despite the lack of molecular dynamics (MD) simulation data in training process, DeepConformer captured conformational flexibility similar to MD simulation and reproduced experimentally observed conformational variations. Our study demonstrated that DeepConformer learned energy landscape can be used to efficiently explore protein conformational distribution and dynamics.

许东 教授

许东教授是美国密苏里大学电子工程和计算机科学系的校董、讲座教授，同时担任 Christopher S. Bond 生命科学中心研究员。许东教授在北京大学获得本科与硕士学位，于 1995 年在伊利诺伊大学获得博士学位。他在美国国家癌症研究所做了两年博士后工作，此后在橡树岭国家实验室做了六年研究工作，一直到 2003 年加入密苏里大学，从 2007-2016 年，他担任该校计算机科学系的系主任。许教授的研究领域包括计算生物学和生物信息学，涵盖单细胞数据分析、蛋白质结构预测、蛋白质翻译后修饰识别等。他是美国科学促进会（AAAS）会士和美国医学与生物工程院（AIMBE）会士。



报告题目：

Large protein language models and their prompt-based learning

报告摘要：

Protein language models (PLMs) offer powerful representations of protein sequences and their evolutionary patterns. We introduced S-PLM, a structure-aware PLM that enhances protein prediction by integrating both sequence and structural information. This model employs a multi-view contrastive learning strategy to align protein sequences with their structures at both the protein and residue levels within a shared embedding space. Equipped with an extensive set of fine-tuning tools, S-PLM achieves superior prediction performance compared to other PLMs. In addition, we developed Prot2Seq to extend PLMs' capabilities for multitasking in protein prediction and design. Prot2Seq utilizes an autoregressive language modeling approach, where task-specific tokens are added to the decoder to enhance simultaneous multitasking training within a single model. Furthermore, we implemented a Parameter-Efficient Fine-Tuning framework on the ESM2 model, employing various prompting methods such as Prompt Tuning, LoRA, and Adapter Tuning for tasks like predicting protein signal peptides and localization signals.

分会场一 · 特邀嘉宾

马剑竹 副教授

马剑竹，现任清华大学智能产业研究院副研究员/副教授，博导，海外优青项目入选者。2016 年于芝加哥丰田工业研究院获得博士学位，随后在加州大学圣地亚哥分校医学院担任项目科学家。2020 年开始担任美国普渡大学计算机系 Walther 助理教授，兼任普渡大学生物化学系助理教授，2021 年担任北京大学人工智能研究院副教授，北京大学公共卫生学院副教授，主要研究领域为人工智能、系统生物学、生物制药和智慧医疗。曾获得生物计算学国际顶级会议 RECOMB 的最佳论文奖，国际计算生物学会 ISMB 的 Warren DeLano 奖，国际 RNA 和蛋白质折叠大会的最佳海报奖，论文曾被 Nature Methods, Nature Machine Intelligence 评为封面论文。



报告题目：

基于靶点 3D 结构的小分子及短肽药物生成模型

报告摘要：

现有的 AI-based Drug Design (AIDD) 大多采用虚拟筛选和分子对接的方式进行辅助药物开发。本次报告中，我们将探索另一种 AIDD 的方式 - 根据靶点的 3D 结构采用生成模型直接生成 3D 的小分子和短肽药物。我们将讨论三种不同的机器学习算法并分别从不同角度来解决这一问题。这些模型包括了自回归模型、扩散模型和基于先验知识的生成模型。我们将首先分析生成模型和传统的虚拟筛选模型的区别，然后探讨这三种模型生成的小分子和多肽在 2D、3D 和药物靶点对接上的各项指标，最后分析他们各自的优势和劣势以及不同的应用场景。

何松 副研究员

军事科学院军事医学研究院副研究员、硕士生导师。主要从事基于人工智能和大数据的重大疾病机理解析和防治药物研发、多组学数据智能整合分析等领域的科研与教学工作。近 5 年来以第一作者和通讯作者身份在 *Nucleic Acids Research*、*Briefings in Bioinformatics*、*Genomics, Proteomics & Bioinformatics* 等 SCI 期刊和 *IEEE International Conference on Bioinformatics and Biomedicine* 等学术会议上发表研究论文 20 余篇。近 2 年来以第一发明人身份获得国家发明专利授权 2 项，以第二发明人身份获得国家发明专利授权 1 项。担任 *Briefings in Bioinformatics*、*BMC Bioinformatics* 等 SCI 期刊审稿人。兼任中国自动化学会智能健康与生物信息专委会委员。



报告题目：

面向药物多终点毒性预测的智能算法

报告摘要：

基于人工智能的毒理学研究在新药研发、环境监测与风险评估、化学品管理与法规等方面有重要的价值。通过大数据和人工智能方法在识别某些毒性方面能够媲美甚至优于动物实验。本次报告主要介绍课题组开发的一套化合物毒性综合数据库 TOXRIC (<http://toxric.bioinformai.tech/>)，提供了 11 万余种化合物 1400 多个毒性终点的记录，数据可直接用于 AI 计算，并介绍课题组在此基础上开发的多终点毒性预测算法。

李代禧 教授

Dr. Daixi Li(李代禧) is distinguished professor of Hujiang Scholars at University of Shanghai for Science and Technology, researcher at Peng Cheng National Laboratory, and associate editor of three international journals including “Interdisciplinary Sciences--Computational Life Sciences”, “Current Computer-Aided Drug Design” and “Computational Biomedicine”. He mainly engages in research in computational biology, network pharmacology, drug design, foodborne pathogens, and artificial intelligence. He has published over 150 international papers, including one paper recommended from the World Academic Organization of the “Faculty Opinions” and one highly cited ESI paper. He won three provincial and ministerial level awards.



报告题目:

Intelligent Identification of Foodborne Pathogenic Bacteria by Self-Transfer Deep Learning and Ensemble Prediction Based on Single-Cell Raman Spectrum

报告摘要:

Foodborne pathogenic infections pose a significant threat to human health. Accurate detection of foodborne diseases is essential in preventing disease transmission. This study proposed an AI model for precisely identifying foodborne pathogenic bacteria based on single-cell Raman spectrum. Self-transfer deep learning and ensemble prediction algorithms had been incorporated into the model framework to improve training efficiency and predictive performance, significantly improving prediction results. Our model can identify simultaneously gram-negative and positive, genus, species of foodborne pathogenic bacteria with an accuracy over 99.99%, as well as recognized strain with over 99.81%. At all four classification levels, unprecedented excellent predictive performance had been achieved. This advancement holds practical significance for medical detection and diagnosis of foodborne diseases by reducing false negatives.

张文艺 助理研究员

张文艺，西湖大学助理研究员，2020 年博士毕业于东北师范大学信息技术学院。2016 年至 2019 年在密歇根大学计算医学与生物信息学系张阳实验室进行博士联合培养。2020 年 8 月，加入西湖大学黄晶实验室从事博士后研究，专注于计算机辅助药物设计相关算法的开发，研究领域包括分子对接、虚拟筛选和药物设计等。相关研究成果已发表在《J. Cheminf.》、



《JCIM》、《Science》和《ACS Central Science》等学术期刊上。此外，参与了新冠病毒 3CL 抑制剂艾普司韦的研发，目前已完成三期临床试验。

报告题目：

Towards structure-based AI drug design

报告摘要：

Computational drug design methods are broadly classified into ligand-based and structure-based approaches. Structure-based drug design utilizes the three-dimensional (3D) structures of target proteins to identify and optimize active compounds. One widely used approach in this domain is molecular docking, which predicts the binding affinity of candidate molecules to known receptor structures from large chemical libraries containing millions of compounds. In this presentation, I will highlight advancements in structure-based virtual screening and molecular design through our methods, EViS and SWIT. EViS integrates ligand docking, protein pocket template searching, and ligand template shape similarity calculations. It introduces the PL-score to efficiently assess pocket-ligand template similarity, leading to improved enrichment factors (EF) on the DUDE dataset compared to traditional docking-based virtual screening methods. SWIT, on the other hand, incorporates 3D target structural information into molecular generative models for structure-based drug design. This method combines a message-passing neural network, which predicts docking scores, with a generative neural network as its reward function, enabling more efficient exploration of chemical space. By constructing target-specific molecular sets for training, SWIT addresses the transferability challenges often faced by conventional docking models, enabling the generation of novel compounds that bind favorably to specific targets without prior knowledge of active compounds.

分会场二 · 特邀嘉宾

肖传乐 研究员

肖传乐，生物信息学博士，广东省杰出青年基金获得者，中山大学中山眼科中心研究员。他长期致力于开发和应用基于第三代测序的基因组和表观遗传分析方法。近年来，他通过建立一系列关键算法和支撑软件，解决了第三代测序基础和应用基因组学研究的瓶颈。作为第一作者和通讯作者在 Nature Methods、Nature Structural & Molecular Biology、Molecular Cell、Nature Communications 等 CNS 子刊上发表论文十余篇。主持国家级项目 6 项，并担任 Nature Methods、Nature Communications、Genome Biology 等期刊审稿人。



报告题目：

Development and application of multi-dimensional genome information analysis technology for three-generation sequencing

报告摘要：

The interplay among genomic sequences and epigenetic modifications plays an essential role in shaping phenotypes and disease development. Decoding such high-dimensional genomic information using third-generation sequencing presents challenges. We leveraged interdisciplinary knowledge and techniques from bioinformatics, biomedicine, and other fields to focus on the analysis and application of high-dimensional genomic information using third-generation sequencing. We have developed fundamental tools for efficient processing of third-generation sequencing data, significantly reducing reliance on high-performance computing. Additionally, we have developed methods for long-read based DNA modification detection and 3D genomic structure analysis to reveal epigenetic regulatory patterns at the single-molecule and haplotype levels. Applying these methods to disease research, we have discovered new pathogenic mechanisms and drug targets.

姜昊 副教授

姜昊，中国人民大学数学学院副教授、博士生导师，担任中国运筹学会女性工作委员会副秘书长、中国生物信息学(筹)生物信息学算法研究专业委员会副秘书长、中国工业与应用数学会数学与生命科学专业委员会委员，主要从事机器学习、数据挖掘、计算生物信息学、基于学习的建模、优化和控制等方面的研究工作，主持、完成国家自然科学基金项目 3 项，并以核心成员身份参与国家自然科学基金重大研究计划集成项目。在 *Information Sciences*, *IEEE Transactions on Neural Networks and Learning Systems*, *Bioinformatics*, *Applied Mathematical Modeling*, *Applied Soft Computing* 等国际权威期刊和会议发表论文 50 余篇。



报告题目：

Heterogeneity analysis and modelling in single-cell data

报告摘要：

Single-cell transcriptomics has transformed our ability to characterize cell states. New methods for simultaneous profiling of multi-omics single cell data enable a better understanding of the cellular states and functions. Cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq), allowed for parallel quantification of cell-surface protein expression and transcriptome profiling in the same cells; Methylome and transcriptome sequencing from single-cells (scM&T-Seq) allows for analysis of transcriptomic and epigenomic profiling in the same individual cells. However, effective integration method for mining the heterogeneity of cells over the noisy, sparse and complex multi-modal data is in growing need. In this talk, we will address the problem of heterogeneity analysis and representation learning in single cell data, for analyzing the optimal embedding representation and identifying cell clusters in a robust manner.

王红 教授

王红教授是山东师范大学的教授和博士生导师。她获得了中国科学院计算技术研究所的计算机应用博士学位，并在山东大学完成了博士后研究工作。她的研究兴趣包括人工智能、机器学习、社会计算、个性化推荐、知识图谱、医疗大数据和生物信息学等领域。作为第一作者和通讯作者，王红教授在国际知名期刊上发表了 100 多篇学术论文。她已获得 30 项专利，并拥有 16 项软件著作权。她曾荣获山东省科学技术奖、山东省计算机应用优秀成果奖以及山东省计算机学会奖等多项荣誉。



报告题目：

Multi-Level Knowledge-Inspired Molecular Property Prediction

报告摘要：

Molecular property prediction is pivotal in the realm of artificial intelligence-driven drug design. Deep learning strategies have become highly effective tools for this purpose. Yet, a common shortcoming among current methods is the underutilization of domain-specific knowledge, such as physical and chemical insights. To surmount these obstacles, our team has developed innovative computational models, including CasANGCLE (BIB2023), MPPNet (ESWA2024), MGPhyche (ESWA2023), and MolLLMKD (accepted). These models leverage multi-granularity physicochemically-inspired representation learning and the power of Large Language Models (LLMs) to significantly enhance the accuracy of molecular property predictions. Our cutting-edge methods have consistently achieved top-tier performance on various benchmark datasets, underscoring their exceptional potential and superiority in molecular property prediction. This advancement not only streamlines the drug discovery process but also promises to revolutionize the way we approach molecular analysis in pharmaceutical research.

彭昱忠 教授

彭昱忠博士，毕业于复旦大学，获博士学位。博士，现任南宁师范大学教授，广西科学院人工智能与生物医学信息研究中心兼职主任。主持国家自然科学基金项目 2 项，省部级项目 4 项。以第一作者或通讯作者在 IEEE TMM、Briefings in Bioinformatics、BIBM、information sciences、Knowledge-Based Systems 等著名期刊和学术会议上发表论文 40 余篇。曾以第一完成者身份获得 BIBM2019 最佳学生论文奖、巴盖人工智能科学技术奖。



报告题目：

Molecular properties prediction based on multi-view molecular representation learning

报告摘要：

Accurate prediction of the relevant properties of drug candidate molecules is crucial in the drug discovery process. With the advancement of technology, various machine learning methods and different forms (e.g., text sequences or graphs) of molecular representation methods have been developed for molecular modelling, which have contributed to the development of areas such as molecular property prediction. Molecular representation and its learning affect the accuracy and efficiency of molecular modelling related tasks. However, previous molecular representation learning methods often have limitations, therefore, we have carried out research on issues related to how to effectively learn and integrate multiple molecular representations. In this report, we introduce two molecular property prediction models based on multi-view molecular representation learning, which break through the limitations of a single molecular representation learning method and realize more accurate molecular property prediction, which is expected to facilitate the process of drug discovery and material development.

郑伟 教授

郑伟博士，南开大学统计与数据科学学院教授。他是 ModelCIF 蛋白质结构预测标准国际委员会的成员。郑博士主导开发了多种蛋白质结构和功能预测算法，包括 C-I-TASSER、D-I-TASSER、DMFold 等，并多次获得 CASP 竞赛冠军。在 NatureMethods、NatureCommunications、PNAS 等高影响力期刊上发表多篇论文，引用近 3000 次。他的算法已为 100 多个国家的 90,000 多名用户提供服务。



报告题目：

Protein Complex Structure Prediction Enhanced by Deep Learning and Huge Metagenomics Data

报告摘要：

Protein-protein complexes are crucial to biological processes, making them key targets for disease research, therapeutics, biomarker discovery, and drug development. Traditional methods like X-ray crystallography and Cryo-EM for protein complex structure determination have limitations. To overcome these challenges, we developed DMFold, a deep learning-based tool that enhances protein-protein complex structure prediction through accurate multiple sequence alignment (MSA) construction using large metagenomics databases. In the CASP15 world-wide competition, DMFold outperformed all 80 participated methods, including AlphaFold2, particularly excelling in large protein complex and antibody-antigen/nanobody-antigen predictions. Recent benchmarks also show DMFold outperforming AlphaFold3. Moreover, DMFold generated high-quality global folds (pLDDT ≥ 0.7) for 1,934 of 5,042 difficult human proteins that AlphaFold2 failed to model. The latest DMFold server integrates functional annotations such as Gene Ontology, Enzyme Commission, and Ligand Binding Information, significantly enhancing its capacity for structural biology studies of protein complexes. DMFold is available at <https://zhanggroup.org/DMFold>.

罗军伟 副教授

罗军伟博士，副教授，中原英才-中原科技创新青年拔尖人才、河南省教育厅学术技术带头人、河南省青年骨干教师、河南理工大学“元培学者”。近年来主要从事深度学习、大数据技术应用、序列组装、结构变异检测、基因组三维结构检测等研究。



报告题目：

Genomic structural variants detection based on deep learning

报告摘要：

Structural variations (SVs) play important roles in human genetic diversity, accurately detecting and genotyping SVs is significant for disease research. To further improve the accuracy of SV detection, our team utilizes multiple deep learning networks to extract complex features and has developed several SV detection methods based on deep learning networks. These proposed methods are benchmarked against some other structural variation detection methods on multiple different datasets, our work shows that deep learning network can be combined with short reads and long reads to call SVs more effectively and sensitively.

刘振栋 教授

刘振栋，上海第二工业大学 计算机与信息工程学院 三级教授，滨海贤人领军人才获得者，揽蓄人才 B 类获得者。清华大学优秀访问学者，哈佛大学访问学者，获山东大学博士学位。山东省生物信息学会副理事长，上海市生物信息学会监事。中国人工智能学会生物信息学与人工生命专委会常务委员，中国生物信息学会算法专委会常务委员，中国计算机学会生物信息学专委会执行委员，山东省人工智能学会常务理事。研究方向为生物信息学、算法分析与设计，发表 SCI/EI 论文 120 余篇，荣获中国人工智能学会吴文俊科学技术奖（排名第一）。



报告题目：

The Algorithm of DNA Binding Site with Combined Feature Encoding

报告摘要：

Due to the richness of biological and drug data, identifying hidden DNA binding sites has been a rewarding but difficult task. For the final prediction of DNA binding sites, we propose the OptimDase algorithm, which combines an optimal decision idea, combined feature encoding and weighted multi-granularity scans. The algorithm OptimDase is very robust, which is not only suitable for classification tasks but also regression tasks. Experiments based on two different datasets show that OptimDase achieved an accuracy of 0.8943 in classification tasks and an RMSE of 0.0054 in regression tasks, achieving excellent high accuracy and low error rates, respectively. In general, the OptimDase algorithm is excellent to the current algorithms in all evaluation indicators and has better portability and higher prediction accuracy. The research in this paper can not only detect hidden DNA binding sites but also provide ideas for the study of efficient recombination systems for drug design.

卢萌 助理教授

卢萌，2015 年博士毕业于剑桥大学，2022 年晋升为剑桥大学 Senior Research Associate, 2024 年入职北京大学，任临床医学高等研究院助理教授、研究员、独立 PI, 建立人工智能与动态结构实验室。卢萌从事超分辨率成像、人工智能图像分析与神经退行性疾病机制的交叉研究多年，在基于人工智能技术的生物图像分析领域取得了一系列成果，1) 研发了整合超分辨率成像、电生理记录和人工智能模型的双模态神经元结构与活动的记录与分析方法，检测了单突触的结构变化和活动；2) 构建了基于人工智能的视频图像分析工具 ERnet, 解决了内质网结构精确识别和定量分析的难题，能够对内质网在不同条件下的结构形态变化进行准确测量；3) 并应用以上研发的图像识别与分析技术，结合先进分子工具和疾病细胞模型，揭示了溶酶体对内质网形态和胞内分布的因果性调控机制，阐明了溶酶体在修复神经元断裂内质网中的关键作用；以上研究为深入阐明神经元亚结构动态变化在神经元生长和退行性疾病中的作用机制提供了新的研究方向和技术基础。



报告题目:

ERnet: Transformer-Based Precise Segmentation and Multi-Parametric Analysis of Endoplasmic Reticulum from Super-Resolution Images

报告摘要:

Quantifying structural changes in the endoplasmic reticulum (ER) is essential for understanding its architecture and function. However, the dynamic behavior and complex topology of ER networks present significant challenges to this task. To address these challenges, we developed a state-of-the-art semantic segmentation tool, ERnet, for the automatic classification of ER sheet and tubular domains within individual cells. The data are processed through skeletonization and represented as connectivity graphs, allowing for precise and efficient quantification of network connectivity. ERnet provides metrics on the topology and structural integrity of ER networks, enabling the assessment of changes in response to genetic or metabolic manipulations. We validated ERnet using data from multiple ER imaging techniques across various cell types and synthetic ER structure images. This tool supports automatic, high-throughput, and unbiased analysis, detecting subtle changes in ER phenotypes that could provide insights into disease progression and therapeutic responses.

王雷 博士

王雷博士，华中科技大学软件工程学院博士后研究员。他的研究兴趣集中于大型蛋白质语言模型（ProtFlash）的开发和设计。迄今为止，他已以第一作者身份在领先的生物信息学期刊和会议上发表了七篇论文，包括《Cell Reports Physical Science》、《Briefings in Bioinformatics》、《Journal of Molecular Biology》以及《Computational and Structural Biotechnology Journal》。此外，王博士还在分子生物学方法系列中贡献了题为“基于法学硕士的蛋白质生物信息学方法”的章节。还受邀担任 NPJ Digital Medicine、Journal of Cheminformatics、BMC Genomics 等多个国际学术期刊的审稿人。



报告题目：

Ultra-fast and Accurate Prediction of Protein-protein Interaction Sites with ProtFormer-Site

报告摘要：

Proteins play a critical role in various biological processes such as signal transduction, metabolic regulation, and gene expression control. These processes largely rely on the physical contact and molecular docking between proteins within cells or organisms, a phenomenon known as protein-protein interactions. Identifying PPI sites is essential for predicting protein functions, revealing the molecular mechanisms of diseases, and designing therapeutic drugs. In this study, we developed two models, ProtFormer-Site-Single and ProtFormer-Site-3D, using advanced deep learning techniques to predict protein-protein interaction sites. ProtFormer-Site-Single relies solely on protein sequence information, utilizing Protein language model to generate rich feature embeddings. These embeddings are fine-tuned using low-rank adaptation to enhance feature representation for PPI site prediction. ProtFormer-Site-3D incorporates additional 3D structural information. This integrates both sequence and structural information to improve prediction accuracy. ProtFormer-Site demonstrated outstanding performance across all evaluation metrics on three benchmark datasets, with Matthew's correlation coefficient improvements ranging from 22.4% to 61.5% across different datasets. These results demonstrate that ProtFormer-Site offers significant advantages in PPI site prediction, providing a more accurate and efficient solution.

分会场三 · 特邀嘉宾

谢震 研究员

谢震是清华大学信息科学与技术国家实验室的研究员和博士生导师。主要研究兴趣包括建立合成生物学使能技术平台，利用合成基因元件和基因线路研究转录、后转录调控网络的设计原理和基因相互作用网络的调控机制，构建以非破坏性方式监控细胞状态变化的合成生物系统，致力于生物基因功能网络研究和基因治疗应用领域中的关键技术研究。在包括 Science、Nature 子刊、PNAS、NAR 等杂志发表 SCI 论文 19 篇。现任《Quantitative Biology》助理主编。2011 年 11 月入选第一批国家“青年千人”计划。



报告题目：

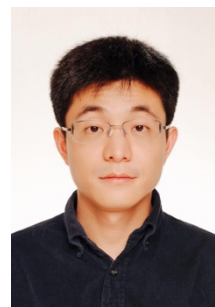
ERNIE-RNA: An RNA Language Model with Structure-enhanced Representations

报告摘要：

Ribonucleic acids (RNAs) act as essential regulators and structural components influencing a wide range of life processes. Nowadays, the advancements in high throughput sequencing technology have produced a wealth of unlabeled data, which contain rich information about RNA structures and functions. Many BERT-style RNA language models trained on abundant RNA sequences have been reported. However, existing RNA language models overlook the impact of structure when modeling the RNA semantic space, resulting in incomplete feature extraction and suboptimal performance across various downstream tasks. The self-attention mechanism used in the transformer-based models such as Alphafold2 and Uni-mol plays a pivotal role in the feature extraction. In this study, we developed a RNA pre-trained language model named ERNIE-RNA (Enhanced Representations with base-pairing restriction for RNA modeling) based on a modified BERT (Bidirectional Encoder Representations from Transformers) by incorporating base-pairing restriction with no MSA (Multiple Sequence Alignment) information. We found that the attention maps from ERNIE-RNA with no fine-tuning are able to capture RNA structure in the zero-shot experiment more precisely than conventional methods such as fine-tuned RNAfold and RNAstructure, suggesting that the ERNIE-RNA can provide comprehensive RNA structural representations. ERNIE-RNA achieved SOTA (state-of-the-art) performance after fine-tuning for various downstream tasks, including RNA structural and functional predictions. In summary, our ERNIE-RNA model provides general features which can be widely and effectively applied in various subsequent research tasks.

王宇光 副教授

王宇光现任上海交通大学的自然科学研究院、数学科学学院、计算机科学与工程系和张江高等研究院 AI 生物医学中心的副教授，并在教育部科学与工程计算重点实验室（MOE-LSC）担任研究员。他也是上海人工智能实验室的首席研究员，并兼任新南威尔士大学（UNSW）副教授。



王宇光教授的研究领域包括人工智能、计算数学、统计学和数据科学，特别关注几何深度学习、图神经网络、应用谐波分析、贝叶斯推断、信息几何和数值分析，以及这些技术在生物医学和蛋白质设计中的应用。

此前，王宇光曾在马普数学科学研究所担任深度学习理论组的研究科学家，并获得了新南威尔士大学的应用数学博士学位，师从 Ian Sloan 教授和 Rob Womersley 教授。他曾获得布朗大学 ICERM 学期博士后奖学金（2018），并在 2019 年访问了加州大学洛杉矶分校（UCLA）的 IPAM，2022 年长期访问了剑桥大学的 AI 研究组。

报告题目：

Generative AI for Protein Design

报告摘要：

Generative AI plays a crucial role in molecular design and can be applied to protein and drug design. We have developed a large protein model based on 50 billion tokens of high-quality protein data and connected it with over 20 downstream synthetic biology design modules through an agent system, achieving "Protein Design All in One." This enables the design of new molecules and proteins with superior functionality. Our model uses two generative models: a diffusion model based on geometric deep learning and a sequence-based conversational protein large model (ChatGPT-like). The basic modules of the model, such as Transformer and deep equivariant graph neural networks, combine Bayesian statistics, particle equations, harmonic analysis, and learning theory. They are highly interpretable and demonstrate stronger generalization and representation capabilities for molecule designs.

胡黔楠 教授

Prof. Hu received Ph.D. degree in applied chemistry from Central South University in 2004. He joined Computer Chemistry Center at University of Erlangen Nurnberg as a Postdoctoral Research Associate in 2004. Dr. Hu joined School of Informatics and Computer Sciences University of California Irvine in 2007 as a Postdoctoral Research Associate, and then Bioinformatics Center at Kyoto University as a staff Scientist in 2008. He joined Wuhan University as Associate Professor of School of Pharmacy in 2010. Then, Dr. Hu joined in Chinese Sciences of Academy as Full Professor since 2014, and his research interest focuses on data-driven synthetic biology for solving global health challenges.



报告题目：

AI+合成生物创新技术驱动高附加值原料合成

报告摘要：

高附加值原料有三个主要的制造方式，天然产物提取、化学合成、以及生物合成，本次报告主要讲述如何利用大数据和人工智能技术进行高附加值原料合成的前沿科技动态。本次主要讲述：（1）医药中间体行业疾病-原料生物合成信息学知识库平台 RDBridge（发表于 *Bioinformatics*, 2023），以及全球最大的高附加值原料有关的合成生物学合成反应/途径数据库。（2）全球首个基于微生物代谢网络的高附加值原料生物合成途径从头智能设计技术工具（*novoPathFinder*, 发表于 *NAR.*, 2020）；（3）全球领先的基于酶序列和高附加值原料分子结构的催化酶挖掘工具（发表于生物催化顶级杂志 *ACS Catalysis*, 2024 封面文章；*Briefings in Bioinformatics*, 2023）。胡博士团队成功搭建了从大数据、AI 生物计算设计、到菌株构建、实验室规模验证直至工艺优化、中试放大及工业生产的高附加值原料合成生物制造的全链条技术能力。

韩彦强 助理研究员

韩彦强博士， 现任上海交通大学微纳电子学系助理研究员。2017 年本科毕业于南开大学物理系，之后进入上海交通大学电子信息与电气工程学院攻读博士学位。博士毕业后加入上海交通大学微纳电子系人工智能与微结构实验室，从事科研工作。目前从事微纳材料的第一性原理与人工智能算法以及交叉领域的研究工作。研究内容主要包括高性能量子分块、第一性原理和人工智能等算法框架的开发，并重点探索了相关算法在微米/纳米结构、生物体系以及晶体结构中的交叉应用，致力于实现大尺度体系的量子力学精度计算和材料、器件功能导向设计。先后在 *Physics Reports* (IF=30.5)、*Briefings in Bioinformatics* (IF=13.9)、*Journal of Physical Chemistry Letters* (IF=9.4) 等高水平学术期刊发表 20 余篇学术论文，授权 10 余项发明专利和软件著作权。



报告题目：

人工智能赋能合成生物学

报告摘要：

人工智能在数据处理、模式识别、实验优化等方面的强大能力，将帮助合成生物学加速研究进展，并为医疗、农业、环境保护等多个领域带来突破性的创新。团队以人工智能、第一性原理计算等技术为核心，在 AI 大模型的基础上，专注发展针对特定领域或任务设计的 AI for Science 专业模型，赋能合成生物学上下游垂直领域研究，包括机器学习势能面何动力学模拟、生物体系性能和功能预测、酶蛋白设计和优化、生物合成途径设计、发酵工艺设计和优化等。开发了蛋白质 AI 建模技术、蛋白质双杂化泛函精度力场、基于 transformer 的 AIMD 加速算法、蛋白质折叠自由能和催化活性预测等 AI 模型，基于此完成了 PET 塑料降解、亚精胺和褪黑素等生物合成关键酶的优化改造和合成效率的提升，并实现了生物合成生产工艺的改造和优化。

潘小勇 长聘副教授

潘小勇，现为上海交通大学的长聘副教授。他曾于 2019 年 2 月至 7 月在比利时根特大学和 BASF 担任博士后研究员。此前，他还在荷兰鹿特丹伊拉斯姆斯大学医学中心的医学信息学系从事博士后研究工作。潘小勇于 2017 年在哥本哈根大学获得博士学位，导师为 Jan Gorodkin 教授和 Lars Juhl Jensen 教授；2011 年在上海交通大学获得硕士学位，导师为沈红斌教授。2011 年至 2013 年期间，他曾在 IBM 公司担任软件工程师两年半，随后开启了在哥本哈根大学的博士研究，并于 2013 至 2016 年期间获得该校博士奖学金资助。此外，他的硕士论文曾获得上海市优秀硕士论文奖（2013 年）。潘小勇已在多本顶级期刊上发表了 20 余篇学术论文，现任 BMC Bioinformatics 的副主编。



报告题目：

蛋白绑定 RNA 位点预测及设计

报告摘要：

蛋白与 RNA、化合物分子相互作用形成的网络是认识生命系统复杂现象的重要内容。DNA 元件百科全书(Encyclopedia of DNA Elements, ENCODE)项目绘制了 150 多种 RNA 绑定蛋白的结合与功能图谱，以及大量实验验证的蛋白-分子绑定数据。这些数据使得我们能够开发机器学习模型挖掘 RNA 和蛋白绑定的序列与结构特征。我们考虑绑定点的序列和结构上下文、全局和局部特性，开发了一系列的基于深度网络模型的蛋白在 RNA 上的片段和碱基水平绑定位点预测，从而实现基于生成模型的蛋白绑定 RNA 序列设计。

王晓雷 博士

王晓雷，博士，毕业于北京大学，现任上海交通大学生命科学技术学院助理研究员，主要从事生物信息学、计算生物学及人工智能等方面研究，致力于人工智能等前沿方法在蛋白质功能预测及设计等领域的应用。在 JACS、ACS Catal.、npj Digit. Med.、Nat. Comm. 等期刊发表 SCI 论文 20 余篇，被引百余次。主持国家自然科学基金青年基金、博士后基金，以及国家自然科学基金重点项目子课题、上海市智慧医疗专项子课题等。



报告题目：

基于序列和结构信息的蛋白质多模态语言模型

报告摘要：

蛋白质语言模型（如 AlphaFold、ESM 系列等）仅依靠序列信息就能够可靠的解决蛋白质折叠问题，目前已经被广泛地用于蛋白质工程的各项常见任务。由于蛋白质功能取决于其结构，因此我们尝试进一步将结构信息有效地融合到蛋白质语言模型中，以提高其在蛋白质工程任务中的能力。我们提出了一种新的蛋白质序列结构多模态架构 ProBEiT，在百万级的高质量蛋白质序列和结构数据上进行了预训练，并进一步在蛋白质适应度和注释等下游任务中进行评估，结果表明这一模型在六项不同的任务中表现均优于成熟的更大规模的基线模型，证明结构信息的引入能够有效提升模型在蛋白质功能理解方面的能力；而且在不同参数规模下，我们构建的序列-结构多模态模型都显示出明确的任务表现提升，证明在有效引入结构模态后，通过相对较少的训练量就能够提升模型在蛋白质功能预测方面的能力，这为蛋白质语言模型在蛋白质设计任务中的应用提供一种新的策略。

卢元 副教授

Yuan Lu is an Associate Professor at Tsinghua University. After receiving B.S. (2004) and Ph.D. (2009) from Tsinghua University, China, he received postdoctoral training from Johns Hopkins University (2009-2010) and Stanford University (2010-2014), USA. He was a project researcher from 2014 to 2016 at the University of Tokyo, Japan. Dr. Lu joined the faculty of Chemical Engineering at Tsinghua University in 2016. His central research paradigm is to exert extreme interdisciplinary approaches, intelligently manipulate atoms, molecules, complex networks, and interface fusion at different levels, constantly break through human cognitive boundaries, and trigger a chain of technological innovation changes. His main research focus is to aim at the commanding heights of future technology, develop the world's most cutting-edge technologies, break through the limitations of natural life, accelerate technical innovation in fields such as human health, material manufacturing, energy and power, information technology, and environmental protection, and cultivate emerging and future industries.



报告题目:

Toward intelligent artificial life

报告摘要:

Innovatively creating artificial life with intelligent features is a persistent dream and pursuit of scientists, striving to achieve precise control of life systems from different physical scales to meet the diverse needs of basic science and applied research. Our research objectives include three aspects: (1) Assemble artificial cells or living systems to understand basic life and employ a new biosynthesis philosophy to overcome problems in living cells; (2) Break through the limitations of life systems and build unnatural systems that surpass natural cells for infinite possibilities; (3) Create innovative applications for flexible manufacturing, human health, and environment governance. We first copied nature in artificial systems to realize biological central dogma, compartment, metabolism, and communication features. Then we reconstruct the basic life features from bottom to top and then integrate them with cutting-edge technologies, such as intelligent materials, optogenetics, electronic engineering, microfluidics, swarm robotics, etc., to achieve artificial living systems with smart responses to physical and chemical signals in space and time. Furthermore, by fusing with artificial intelligence technologies, we expand its applications in biomolecular synthesis, diagnostics, and drug delivery fields.

崔海洋 教授

崔海洋教授，博士生导师，南京师范大学生命科学学院，国家级青年人才，江苏省特聘教授。2016 年硕士毕业于中科院天津工业生物技术研究所，师从结构生物学领域专家郭瑞庭教授；2020 年，博士毕业于德国亚琛工业大学（导师：Ulrich Schwaneberg 教授，蛋白质工程领域专家）；同年，进入德国 DWI-莱布尼茨互动材料研究所进行博士后研究；2021-2023 年于美国伊利诺伊大学厄巴纳-香槟分校继续进行博士后研究（导师：赵惠民教授，合成生物学领域先驱）。2024 年初回国组建“人工智能与工程生物学”研究团队，搭建载有世界顶尖 GPU-A100 芯片的超算平台 Infinity（无穷生物超算集群）。近五年，以第一作者/通讯作者（含共同）在 *Science*, *Angew. Chem.*, *ACS Catal.*, *Nat. Commun.* 等期刊发表 SCI 论文 22 篇，并应邀在国际权威丛书 *Methods in Molecular Biology* 分别撰写独立章节 3 章。申请发明专利 4 项，授权 1 项。



报告题目：

The integration of protein engineering and computational science: a complementary path forward

报告摘要：

Enzyme is an indispensable building block in synthetic biology. Synthetic biology often exploits the efficiency and selectivity of enzymes to transform natural substrates under mild conditions. However, natural enzymes are often not suitable for actual production in harsh industrial environments (organic solvents, high temperatures, disorder, etc.). Herein, we developed a series of significant and broad protein engineering strategies by integrating directed evolution and computational biology. Meanwhile, our molecular understanding has provided effective explanation for achieving high robustness and targeted immobilization of enzymes. Furthermore, deep learning has opened a novel route for developing an enzyme function prediction model, which is then verified through experimental characterization. This approach overcomes the challenges of traditional enzyme function learning, which is often difficult and imprecise. In short, by leveraging protein engineering and computational science, biocatalysis can advance a range of solutions in synthetic biology.

周冰心 助理研究员

Bingxin Zhou is currently a Research Scientist at Shanghai Jiao Tong University. She obtained her Ph.D. degree at the University of Sydney, Australia, in 2022 and later became a visiting scholar at the University of Cambridge. Her research primarily focuses on the development of deep learning techniques, especially geometric deep learning, to address challenges in biology such as enzyme engineering, metabolic gene networks, and proteome-wide evolutionary analysis. For model development, she has published useful Graph Neural Network models for static, dynamic, heterophilic, and noisy graphs in IEEE TPAMI, JMLR, ICML, NeurIPS, etc. In the field of protein analysis and application, she has established general deep learning frameworks for protein engineering and sequence design, with promising experimental evaluations. Some results have been published in Cell Discovery, eLife, Chemical Science, JCI, etc.



报告题目:

Understanding and Engineering Proteins with Geometric Deep Learning

报告摘要:

Protein engineering plays a pivotal role in addressing global challenges, from healthcare to sustainability. Recent research leverages deep learning methods, such as language models and graph neural networks, to analyze protein sequences, structures, and functions. This emerging biotechnology significantly reduces the cost and complexity of studying and modifying proteins. This talk introduces our recent deep learning solutions for protein engineering, aimed at enhancing protein functionality and properties to meet practical needs. We address a range of challenges faced by biologists, including zero-shot mutations, deep mutations, few-shot dry-wet experimental iterations, and patent blockades. The reliability and generalizability of our solutions have been validated through wet lab experiments on a variety of proteins and protein assays.

分会场四 · 特邀嘉宾

徐军 教授

徐军，现任南京信息工程大学未来技术学院副院长，二级教授；智能医学图像计算江苏省高校重点实验室主任。2007 年于浙江大学控制科学与工程系获博士学位，先后在美国 Rutgers 大学、凯斯西储大学生物医学工程系任博士后研究员和访问教授。兼任江苏省人工智能学会医学图像分析专委会副主任委员；中华医学会病理学分会数字病理与人工智能委员会委员；中国生物医学工程学会医学图像信息与控制分会常务委员；第四届医学图像计算青年研讨会轮值主席等。研究成果主要发表在 Nature Communications, Radiology, IEEE Trans. on Medical Imaging, Medical Image Analysis 等期刊。主要研究方向是：医学图像计算；计算病理；数字病理；基于影像和常规病理切片定量分析的疾病辅助诊疗和预后。他所领导的智能医学图像计算重点实验室现有全职教师 20 余名，90%以上毕业于中国、美国、英国、新加坡等国家世界前 200 位一流顶尖高校。实验室主要致力于运用先进的人工智能、信息科学的手段，探索和解决医学与健康学科前沿领域的重大科学问题。经过多年发展，实验室凝练了如下四个主要研究：医学影像成像与分析，脑科学与神经影像，病理智能辅助诊断与干预，以及智能感知与医学大数据。



报告题目：

数据与知识驱动的图像计算及其对精准医学的价值

报告摘要：

精准医学的目标是针对患者在恰当的时间制定有效的治疗方案。影像和常规组织学病理切片包含许多有价值的肿瘤预防、诊断、治疗和预后信息。医学图像数据占医学数据的 80%-90%，随着机器学习及医学图像分析技术的快速发展，我们可以运用数据与知识驱动的机器学习模型，分别从影像和组织病理学切片中提取疾病的多尺度图像生物标志物，例如毫米级和微米级较为宏观的图像表型，从而进一步补充纳米尺度如基因，蛋白、分子等数据在疾病表达方面的不足。在南京信息工程大学智能医学图像计算重点实验室，我们团队致力于探索面向影像和组织病理学图像的数据与知识驱动机器学习方法，通过对影像图像和数字化组织样本的“图像生物标记物”的提取和分析，构建多尺度图像生物标记物，以帮助医生从影像和组织学图像中更准确地检测和诊断疾病，并尝试预测复发风险、疾病侵袭性、患者生存率以及患者对治疗的响应，从而辅助医生为肿瘤患者制定最优的治疗方案和策略。在本次报告中，我将介绍我们团队以上领域的相关工作进展。

艾丹妮 研究员

艾丹妮，北京理工大学特别研究员，博士生导师，结合国家临床医学重大需求，围绕医学图像处理、手术导航、虚拟现实与增强现实进行理论及应用研究，先后主持/参与国家 973、863、十二五支撑计划项目、十三五重点研发计划项目、国家自然科学基金重点/仪器专项等项目近 10 项。近三年，以第一/通讯作者身份发表 SCI 顶级期刊论文 1 篇、重要期刊论文 6 篇；作为第一发明人申报国家发明专利 5 项、授权发明专利 4 项。作为第五发明人获中国图象图形学学会科学技术一等奖。作为导师指导博士生获 2019 年校级优秀博士学位论文，作为第一指导教师获第五届“互联网+”大学生创新创业大赛北京市赛区总决赛一等奖，作为通讯作者指导硕士生获国际知名会议 ICIG 2019 最佳学生论文奖。



报告题目：

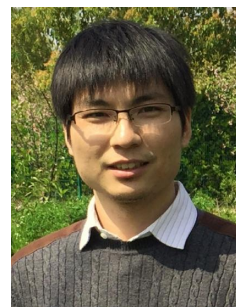
多模态图像融合导航肝癌精准诊疗研究

报告摘要：

肝脏是易发肿瘤的主要内脏器官，其恶性肿瘤发生率列全球和我国前十位。超声因实时动态、无辐射等优势，成为肝脏肿瘤经皮穿刺诊疗最常用的影像引导方式。然而，超声引导穿刺仍面临以下难题：1) 肝脏和前列腺解剖结构复杂，穿刺因缺少空间信息，需在二维影像反复调试，精度低、耗时长且高度依赖医生经验；2) 呼吸、体位和应力变化导致柔性脏器形变大且缺少实时跟踪反馈，使穿刺精度受到严重干扰。针对上述临床问题，本研究团队突破了多模影像智能建模与动态穿刺规划、组织形变跟踪及移位补偿的关键问题，提出了超声三维重建与多模态图像配准方法，突破了组织弹性形变及运动影响下精确跟踪难题，显著提升了手术穿刺的精度和效率，有效降低了手术操作难度，实现了超声主导多模影像穿刺由经验穿刺向智能、精准和可控的方向发展。

庄吓海 教授

庄吓海，复旦大学大数据学院教授，主要研究方向是可解释医学影像人工智能等。近五年以第一或通讯作者发表中科院一区 24 篇，包括 6 篇 IEEE TPAMI；工作入选顶刊医学影像分析高引论文和 ESI 高引论文、被 Nature 子刊作为心脏影像人工智能分析的成功案例引用。多项第一/通讯作者工作被国际组织或顶会提名最佳论文奖和青年科学家奖（如 IPMI 和 MICCAI 等），其中 2023 年的可解释医学影像分析工作获得爱思唯尔出版社-顶刊 Med Imag Ana 和国际 MICCAI 学会联合颁发最佳论文奖（1/1865+），为国内第一次以第一单位获此奖项。获 2023 年上海市自然科学二等奖（排名第一）；自 2020 年以来连续入选爱思唯尔-斯坦福大学发布全球顶尖科学家“终身”和“年度”科学影响力榜单；目前担任国际 MICCAI 学会常务理事，担任 IEEE TMI、Med Imag Anal、Neurual Networks 等多中科院一区期刊的编委。



报告题目：

医学影像可解释分析

报告摘要：

医学影像智能分析在计算机辅助诊断和治疗等现代医学中发挥着重要的作用；其中，人工智能方法的模型泛化能力和可解释性对临床应用至关重要。然而，医学影像数据的异质性和人工智能方法的黑盒性对医学影像智能分析的安全可靠性提出了挑战；因此在涉及多模态、跨中心图像和弱监督无监督学习的场景中，模型的泛化能力和可推广性不足。本次讲座将介绍我们近期从弱监督无监督学习总结出的经验，提出主动构建可解释深度神经网络架构方法，基于显式建模和概念学习的架构可解释和决策可解释方法；实验证明模型自身可解释的人工智能方法可以提高模型的泛化能力和可理解交互性。报告将介绍课题组近期初步结果，更多关于可解释框架、可泛化方法及应用研究请参考实验室主页，其中部分数据和代码也通过该主页公布（<https://zmiclabor.github.io/projects.html>）。

高伟 教授

高伟是加州理工学院的医学工程教授、Ronald and JoAnne Willens Scholar、Heritage Medical Research Institute Investigator。他目前担任 Science Advances、npj Flexible Electronics、Biosensors and Bioelectronics 和 Sensors & Diagnostics 的副主编。他的成就获得了众多奖项和荣誉，例如 NSF Career Award, ONR Young Investigator Award, IAMBE Early Career Award, Sloan Research Fellowship, Pittcon Achievement Award, IEEE EMBS Early Career Achievement Award, IEEE Sensor Council Technical Achievement Award, Falling Walls Breakthrough of the Year 2023 in Engineering and Technology, 3M Non-Tenured Faculty Award, MIT Technology Review 35 Innovators Under 35, ACS DIC Young Investigator Award, 以及 Materials Today Rising Star Award。他的研究兴趣涵盖可穿戴传感器、生物电子学、柔性电子学和微/纳米机器人技术等多个领域。



报告题目：

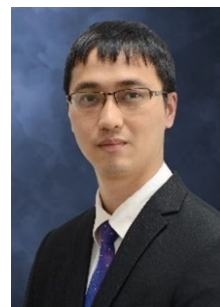
Skin-Interfaced Wearable Biosensors

报告摘要：

The growing interest in personalized medicine is set to transform conventional healthcare, offering new avenues for predictive analytics and tailored treatment approaches. In this talk, I will present our advancements in developing wearable biosensors for non-invasive molecular analysis. These wearables autonomously access and sample body fluids, such as sweat and exhaled breath condensate, continuously monitoring a wide array of analytes—including metabolites, nutrients, hormones, proteins, and drugs—during various daily activities. To enable large-scale, cost-effective manufacturing of these high-performance nanomaterial-based sensors, we leverage techniques such as laser engraving, inkjet printing, and 3D printing. The clinical utility of our wearable systems is assessed through a series of human trials, focusing on precision nutrition, stress response and mental health assessment, chronic disease management, and drug personalization. Our wearable systems' clinical applications are evaluated through human trials in areas like precision nutrition, stress response monitoring, mental health assessment, chronic disease management, and drug personalization. Furthermore, I will explore our efforts in energy harvesting from both the human body and the environment, paving the way for battery-free, wireless biosensing devices. This integration of wearable technologies has the potential to revolutionize personalized healthcare, spanning diagnostics, real-time monitoring, and therapeutic innovations.

蔡云鹏 研究员

蔡云鹏，中国科学院深圳先进技术研究院研究员，生物医学信息技术研究中心执行主任，广东省健康大数据分析技术工程中心主任，深圳市健康大数据分析技术工程实验室主任。已发表学术论文一百余篇，包括作为主要作者在 IEEE TNNLS、IEEE TOC、IEEE ICDE、Nucleic Acids Research 等机器学习和生物信息学顶级期刊、会议上的多篇论文，获得授权发明专利 40 多项。担任 SCI 一区期刊 Health Information Science and Systems 主编，获广东省科学技术奖二等奖 2 项，深圳市技术发明一等奖、中华医学会卫生管理奖各一项。



报告题目：

面向医学影像 AI 模型的可解释知识提取方法及应用

报告摘要：

医疗领域对可解释 AI 模型的需求日益迫切，现有针对深度学习的可解释性方法难以有效获取全局性知识规律，导致产生错误和含糊的解释结果。本报告介绍课题组提出的基于类关联子流形学习的深度学习模型解释技术，通过子流形引导的反事实生成模型同时提升全局规律解释和个体归因解释的准确性，并在病灶定位、亚型发现、诊断规律挖掘等典型医学影像分析问题上进行了应用。

胡振华 研究员

胡振华，中科院自动化所，研究员、博导，国家杰青、优青、北京市杰青。聚焦信息技术与生物医学的交叉融合，长期从事光学分子影像及相关信息技术研究，深入开展了光学分子影像新技术研发与临床转化应用的全链条工作。近五年以一作和通讯（含并列）发表高质量论文 20 余篇，包括在综合类 Nature 子刊 Nature Biomedical Engineering 2 篇、信息类 IEEE Trans、光学类 OSA、临床及综述类 Nature Reviews Clinical Oncology 等期刊发表了系列高水平论文。相关工作获 Nature 子刊专题评述及多位国际权威专家的“近红外二区荧光成像国际首次临床应用”评价。相关工作被国内临床教科书章节收录。同时，报告人牵头制订了近红外二区荧光成像团体标准，并与临床专家合作制定肝癌诊疗专家共识，推动技术标准化和临床应用规范化。主持基金委重点项目、科技部、北京市杰青等国家省部级项目。任国家一级学会中国图学学会理事及多个医学专委会常务委员、国际期刊编委及多个国际学术期刊论文审稿人。



报告题目：

近红外二区光学分子影像技术及临床探索

报告摘要：

报告人围绕肿瘤精准成像难的临床挑战，提出临床高灵敏近红外二区荧光分子成像新技术。研发成像方法，构建成像系统，通过基础研究验证，将新技术应用于肝癌临床成像，将肝癌检出率提高 10%。相关工作发表于 Nature Biomedical Engineering，并获 Nature 子刊专题评述及多位国际权威专家“近红外二区荧光成像国际首次临床应用”评价。在此基础上，将新技术推广至多癌种及血管、神经的术中荧光成像中，提高肿瘤完整切除率，延长患者生存时间，同时，保护患者正常生理功能，提高患者术后生活质量。

金成 副教授

上海交通大学数字医疗研究所发起人、副教授、博士生导师。上海国家 AI 实验室高级研究顾问。斯坦福大学高级研究科学家。国家级海外高层次人才，上海领军人才。美国国家人工智能咨询委员会生成式 AI 工作组成员。本人致力于多模态数据融合，临床级别的 AI 辅助诊疗系统。目前在 Nature



Medicine、TPAMI、Science Translational Medicine、Nature communications、Cell Reports Medicine、Patterns 等期刊发表论文 40 余篇，部分研究成果已落地转化为临床应用、被人民日报、路透社等知名媒体报道。最新代表成果：①构建联邦学习平台，开展直肠癌新辅助治疗的疗效评估项目荣获第九届国际互联网+大赛季军；②利用美国 5 家保险公司和中国 5 家体检中心的 60 万人数据开发的个性化康复大模型项目在 2023 年中国国际大学生创新创业大赛荣获金奖；③ 2024 年 6 月，在第十六届中国精神心理卫生学术年会上，发布全球首款面向医生的精神心理专精大模型。

报告题目：

基于多模态、多组学数据融合的复杂疾病智能诊断与归因分析 —— 大模型时代的新范式

报告摘要：

复杂疾病往往由多种遗传、环境和生活方式因素共同作用引起，其诊断和归因分析具有较高难度。随着检测技术的不断更新和信息记录的多元化，复杂疾病的多模态信息和多组学信息越来越丰富，包括多组学数据（基因组、转录组、蛋白质组、代谢组等数据），表型数据，影像数据，数字病理图像，临床记录数据，电子健康记录，行为数据，环境和生活方式数据等。多组学数据提供了关于基因变异、基因表达、蛋白质功能和代谢产物的详细信息，而临床和影像等多模态数据则提供了患者的具体表现和疾病进展情况。多模态、多组学数据融合的难点在于不同数据源之间的异质性和尺度差异，以及如何有效处理高维度和稀疏数据。为了解决这些问题，确保多模态数据能够在同一框架内进行综合分析，探讨将不同模态的数据映射到同一向量空间中，从而实现数据的有机融合，Transformer 结构的多头自注意力机制，则可以在高维空间中捕捉不同数据模态之间的相互关系，提升模型的解释力和预测能力。此外，研究还注重复杂疾病的归因分析，旨在揭示疾病发生发展的机制。通过模型解释技术，可以理解各特征对模型预测的贡献，从而识别出关键的致病因素和潜在的治疗靶点。这些发现不仅有助于个性化医疗的实施，还为新药开发和预防策略提供了科学依据。以阿尔兹海默症等疾病智能诊断与归因分析为例，利用 embedding 和 transformer 结构，显著提升其诊断精度和机制研究深度，为精准医学的发展提供了重要支持。

陈宇飞 副教授

同济大学电子与信息工程学院副教授，博士生导师。曾赴德国弗劳恩霍夫(Fraunhofer)图像数据处理研究所医学影像中心任访问研究员。主要研究方向为医学影像精准分割、多源影像配准与融合、影像不确定性分类等。在领域内知名学术期刊及会议累计发表论文 100 余篇，其中包括 TMI、TNNLS、TIP、NeurIPS、AAAI、ACMMM、MICCAI、CVPR 等高水平期刊会议 50 余篇，授权发明专利 14 项。为 MICCAI-CLIP 主席、MICCAI-SIG-xMedIA 教育主席、MICS 委员会委员、全国仿生学标委会委员等。近年来，主持国家自然科学基金项目 4 项，国家重点研发计划课题 1 项，子课题 1 项，省部级项目 2 项等，作为主要成员获上海市技术发明一等奖、科技进步三等奖、宁波市科技进步一等奖。具有丰富的项目开发经验。



研究方向：医学影像分析，计算机辅助诊断

报告题目：

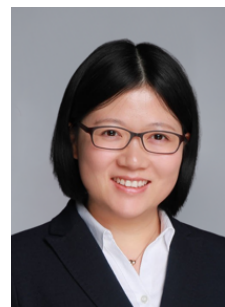
证据深度学习与可信医学辅助诊断决策

报告摘要：

多源数据智能分析为降低疾病辅助决策风险提供了途径，但由于多源数据普遍存在的不完备与不一致情况，会造成疾病诊断具有不确定性，从而提升决策风险。本报告从肿瘤临床诊断需求出发，聚焦深度神经网络在影像数据分析应用中的局限与挑战，包括多源影像数据稀缺与认知不充分、数据融合的不确定性与不一致性、模型缺乏对分布外数据的检测能力等难题，研究融合深度学习不确定性的智能影像分析技术，并应用于多源影像数据分割、疾病分类等多个场景，从而生成安全可信的医学辅助决策。

洪义 副教授

洪义，上海交通大学计算机科学与工程系长聘教轨副教授，博士生导师。主要研究医学图像分析、人工智能与智慧医疗交叉方向，以第一作者/唯一通讯发表 30 余篇论文，包括 TPAMI、Medical Image Analysis、MICCAI、IPMI 等国际著名期刊和会议论文。主持国家自然科学基金和美国自然科学基金项目，曾获得医学图像分析顶会 MICCAI 2014 青年学者奖、美国 NSF CISE CRII Award、北卡罗莱纳大学教堂山分校博士完成奖学金、上海市海外领军青年人才称号、2023 年度上海开源创新卓越成果奖二等奖（排名第一）、2023 年度上海市计算机学会教学成果奖一等奖（排名第二）、2024 年上海交通大学教学成果奖二等奖（排名第二）等荣誉。



报告题目：

Medical Multimodal Fusion for Computer-Aided Diagnosis of Alzheimer's Disease

报告摘要：

Alzheimer's Disease (AD) is a complex neurodegenerative disorder that requires comprehensive diagnostic approaches, integrating both imaging and clinical data. Multimodal fusion, which combines diverse data sources such as brain scans, genetics, and clinical assessments, holds great potential for improving diagnostic accuracy. However, challenges arise in aligning multimodal features, managing redundant information, and ensuring interpretability in clinical contexts. This talk addresses these key challenges, presenting innovative solutions that enhance the effectiveness and interpretability of computational tools for AD diagnosis. Our approach shows promising results in advancing the field of medical multimodal fusion for AD diagnosis.

分会场六 · 特邀嘉宾

魏东辉 教授

Dr. Donghui Wei is a Professor at College of Chemistry, Zhengzhou University. He performed a series of theoretical studies on the mechanisms of enzyme reactions (such as the reactions between proteasome and different peptides, cofactor-free oxidases and oxygenases catalyzed reactions, and so on) and organocatalytic reactions (such as N-heterocyclic carbene (NHC) catalyzed reactions). Recently, he proposed the concept of directional reactivity index, and developed wave function analysis programme Pywfn and AI models for predicting the chemoselectivity of homogeneous catalytic reaction.



报告题目：

Predicting chemoselectivity of enzyme-catalyzed reactions based on Pywfn programme and artificial intelligence

报告摘要：

Understanding and even predicating chemoselectivity of the reactions is a challenging question in both enzyme catalysis and organocatalysis field. To reduce the experimental cost, many computational approaches for predicating reaction chemoselectivity have been developed by our team, for example, Projection of Orbital Coefficient Vector (POCV) method, wave function analysis programme Pywfn, and artificial intelligence tools. Recently, we have disclosed the origin of chemoselectivity of halohydrin dehalogenase (HHDH)-catalyzed ring-opening reactions of epoxide with the nucleophilic reagent NO_2^- (J. Chem. Inf. Model. 2024, 64, 4530), and have planned to update and develop the wave function analysis programme Pywfn and AI tools to predicate the chemoselectivity of the enzyme and organocatalytic reactions.

黄琳 博士

黄琳，上海市胸科医院 PI，国家优青、上海市巾帼创新新秀、上海“高峰高原”学科建设双百人、上海市扬帆人才、上海交通大学医学院九龙青年人才。担任 *Materials Today Bio* (IF=8.2)等 6 个期刊编委及 20 多个国际学术期刊审稿人；服务国产期刊建设，获 VIEW 期刊突出贡献奖；工作受国际认可，受邀评审美国-以色列两国科学基金等国内外基金。近年来主持国家、上海市自然科学基金项目 5 项，获上海市技术发明一等奖(排名第 3)、中国实验医学优秀论文奖、Springer Nature 纳米女科学家、Royal Society of Chemistry 新兴研究员等奖项。

报告题目：

基于人工智能的组学分析及临床诊断应用

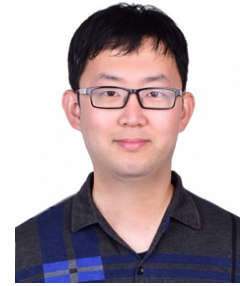
报告摘要：

肺癌是全球范围内癌症相关死亡的首要风险，其中重要原因是癌细胞转移。外科手术以及术后放化疗可有效控制原发肿瘤，但针对远端转移肿瘤预后不佳。以往研究表明，代谢重编程是恶性肿瘤的重要特征，可重塑肿瘤细胞以增强其生长和存活。然而，代谢重塑与肺癌发生发展进程间关系仍鲜有报道。此外，面向临床开展代谢重塑研究面临三个难点：（一）代谢组技术工具不理想；（二）数据分析算法不适用；（三）临床数据库构建困难等。

为解决上述科学问题，我们研发了适于肺癌早期诊断、疗效监测和预后判断的新型代谢组技术平台。该平台主要由三个关键部分构成，即新型代谢组质谱技术开发、AI 代谢重塑分析及临床应用转化验证。

孙端辰 博士

Dr. Duanchen Sun graduated from the Academy of Mathematics and Systems Science (AMSS), Chinese Academy of Sciences, in 2017, majoring in operations research and bioinformatics. His research interest mainly involved cross-research on mathematics, computers, and cutting-edge biomedical issues. Published articles as the first or co-first author in journals such as Nature Biotechnology, Molecular Cancer, Nature Communications, Briefings in Bioinformatics, and Clinical Cancer Research.



报告题目：

Identifying phenotype-associated cell subpopulations based on multi-omics data

报告摘要：

Single-cell RNA sequencing (scRNA-seq) distinguishes cell types, states and lineages within the context of heterogeneous tissues. However, current single-cell data cannot directly link cell clusters with specific phenotypes. Here we present Scissor, a method that identifies cell subpopulations from single-cell data that are associated with a given phenotype. Scissor integrates phenotype-associated bulk expression data and single-cell data by first quantifying the similarity between each single cell and each bulk sample. Applied to a lung cancer scRNA-seq dataset, Scissor identified subsets of cells associated with worse survival and with TP53 mutations. In melanoma, Scissor discerned a T cell subpopulation with low PDCD1/CTLA4 and high TCF7 expression associated with an immunotherapy response. Beyond cancer, Scissor was effective in interpreting facioscapulohumeral muscular dystrophy and Alzheimer's disease datasets. Scissor identifies biologically and clinically relevant cell subpopulations from single-cell assays by leveraging phenotype and bulk-omics datasets.

上海交通大学（徐汇校区）简介

上海交通大学徐汇校区位于中国上海市徐汇区华山路 1954 号，占地面积为占地面积 236490 平方米。该校区是上海交大整体搬迁至闵行校区前的主校区，也是承载老交大历史的校区，1896 年南洋公学始建于此。

初建时的校区四面环河，位置在现今校区的东部，徐汇校区的多数历史建筑亦分布于此。现今东侧草坪原为露天体育场，建筑环绕四周分布，建于 1899 年的中院位于草坪北侧，是徐汇校区现存最早的教学楼。后经多次购地扩建，逐步形成今日淮海西路、华山路、广元西路、番禺路间的校区格局。其中，南洋公学时期建筑及交通大学工程馆、新上院分别位列第二、四批上海市优秀历史建筑。2019 年 10 月 16 日，徐汇校区的历史建筑由中华人民共和国国务院以“上海交通大学早期建筑”之名公布为第八批全国重点文物保护单位。

目前上海交通大学安泰经济与管理学院、国际与公共事务学院、凯原法学院、高级金融学院设在此校区。

会议地点

开幕式/大会报告：文治堂（10 月 26 日上午）

分会场报告：新上院（10 月 26 日下午及 27 日上午）

晚宴：教师活动中心-聚贤阁（10 月 26 日晚上）

参观：新中院董浩云航运博物馆（10 月 27 日下午）

