

范文睿

✉ wenrui.fan@sheffield.ac.uk ☎ +44 (0) 7572 857388 🌐 wenruifan.github.io 👤 wenruifan
🔍 Google Scholar 📖 Web of Science 🔗 LinkedIn ☁ Hugging Face 🆔 ORCID: 0009-0007-1394-0092

教育经历

- **博士研究生：计算机科学**
– 英国谢菲尔德大学 01.2024 – 至今
- **理学硕士：机器人学 (Robotics)**
– 英国谢菲尔德大学；优等学位 (Distinction) 09.2021 – 09.2022
- **工学学士：飞行器设计与制造工程（卓越工程师计划）**
– 北京理工大学；优秀毕业论文 09.2017 – 07.2021

工作经历

- **AI Research Engineer – Centre for Machine Intelligence, University of Sheffield** 09.2023 – 至今
– 牵头并参与神经影像、精准医疗、癌症、蛋白质科学及分子建模等方向的多模态人工智能研究。
– 负责模型、数据集、评估方案、可解释性方法及可复现科研软件的研发与落地。
– 协调临床、生物医学及其他跨学科合作者的技术工作，推动研究任务按计划交付。
– 参与跨机构基金申请及开放科研基础设施建设。
- **智能制造人工智能算法工程师（实习）– 哈尔滨工业大学重庆研究院** 11.2022 – 07.2023
– 牵头开发面向半导体晶圆异常检测，注塑件外观异常检测，软性材料印刷质量检测等的机器学习方法。
– 优化模型检测性能与计算效率，提升方案的实际应用能力。
– 参与面向智能制造场景的基础模型架构研发。
– 参与工业异常检测上下游流水线的嵌入设计。

专业技能

- **人工智能与机器学习**：多模态学习、科学智能 (AI for Science)、基础模型适配、视觉 – 语言预训练、图学习、智能体人工智能及生成式建模。
- **模型评估与分析**：基线方案设计、独立队列验证、亚组分析、模型解释、因果分析及基准评测。
- **科研软件工程**：Python、PyTorch 生态、scikit-learn、Hugging Face、Git/GitHub、Linux、Docker、自动化测试、持续集成、Python 软件包发布及 FAIR 科研流程。

精选科研项目

- **Neuropathic Pain Drug-response Prediction – 代码**
– 牵头设计并基于 PyTorch 实现多模态学习框架，融合静息态 fMRI 时间序列与功能连接信息。
– 将基于 UK Biobank 数据预训练的基础模型适配至小样本临床队列。
– 针对两个临床数据集设计基线方案与评估流程，并开展亚组层面的模型解释分析。
– 将模型学习到的模式与医学、生物学研究发现及相关理论进行对照验证。
- **MeDSLIP: Medical Vision-Language Pre-training – 代码 | 模型**
– 牵头面向细粒度医学图像 – 文本对齐的模型设计、实现、实验与发布。
– 系统评估零样本学习、迁移学习与微调性能。
– 开源代码及预训练模型，支持外部复现与二次开发。
- **Parkinson's Disease Risk-factor Analysis**
– 牵头 UK Biobank 潜在帕金森病风险因素研究中的机器学习工作。
– 综合采用预测建模、年龄与性别分层分析、模型解释及因果推断方法。
– 区分预测贡献与因果证据，为形成审慎、可信的研究结论提供依据。
- **AI-driven Multistate Protein Design**
– 构建整合蛋白质序列、结构、状态、状态转换及功能信息的数据基础设施。
– 面向多状态结构预测与逆向折叠，设计基准任务、数据集划分、基线方法及评估指标。
– 研发状态感知的注释、表征学习与生成式建模方法，服务多状态蛋白质设计。

• Selective Anticancer Peptide Design

- 参与从头肽生成、选择性抗癌筛选及癌症类型预测。
- 设计补充计算实验，并支持生成候选分子的评估与验证。

开源软件与数据基础设施

• KaleProtein (v0.1.0a1, PyPI) – 创始人、首席开发者

- 全面负责软件架构、功能实现、测试、文档、打包、版本管理及发布流程。
- 提供标准化的模型与数据集注册、加载接口，服务蛋白质科学研究。
- 集成可直接运行的 DrugBAN 与 MapDiff 工作流及其配套数据集。
- 基于 GitHub Actions 维护自动化测试与持续集成流程。

• PyKale (GitHub 480+ Stars; 累计下载 3.4 万+ 次) – 核心贡献者、维护者

- 参与开发已纳入 PyTorch Ecosystem 的多模态机器学习库。
- 开发元学习模块、测试与文档，并构建端到端癌症多组学教程。
- 参与仓库维护；基于问卷调查与半结构化访谈，推动以用户需求为导向的项目重构。
- 按提交次数计，为三位联合创始人之后的第四大贡献者。

• MuSProt – 首席开发者

- 构建开放的百万级多状态蛋白质数据库。
- 开发并部署托管于 Hugging Face 的网站，提供检索、可视化及数据下载功能。
- 开源可复现的数据集构建流程，涵盖数据采集与自动化注释方法。

Publications

*Equal contribution.

Conference Papers

- [C1] **W. Fan**, L.M.R. Rizky, J. Zhang, C. Chen, H. Lu, K. Teh, D. Selvarajah, and S. Zhou, “Foundation-Model-boosted multimodal learning for fMRI-based neuropathic pain drug response prediction”, in *the 28th Int. Conf. on Medical Image Computing and Computer Assisted Intervention (MICCAI, CORE A)*, pp. 238–248, Sep. 2025. DOI: 10.1007/978-3-032-05182-0_24.
- [C2] J. Zhang, X. Liu, W. Wu, S. Tabakhi, **W. Fan**, S. Zhou, K.L. Tee, T.S. Wong, and H. Lu, “Classifying the Stoichiometry of Virus-like Particles with Interpretable Machine Learning”, in *47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 1–7, Jul. 2025. DOI: 10.1109/EMBC58623.2025.11254123.
- [C3] M.N.I. Suvon, P.C. Tripathi, **W. Fan**, S. Zhou, X. Liu, S. Alabed, V. Osmani, A. Swift, C. Chen, and H. Lu, “Multimodal Variational Autoencoder for Low-cost Cardiac Hemodynamics Instability Detection”, in *the 27th Int. Conf. on Medical Image Computing and Computer Assisted Intervention (MICCAI, CORE A)*, pp. 296–306, Oct. 2024. DOI: 10.1007/978-3-031-72378-0_28.
- [C4] J. Wang, B. Chen, **W. Fan**, and Y. Liu, “Co-salient object detection via discriminative prototypes contrast”, in *the 2024 IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 7390–7394, Apr. 2024. DOI: 10.1109/ICASSP48485.2024.10446685.
- [C5] **W. Fan**, “Path planning for mobile robot in 3D environment based on ant colony algorithm”, in *the 2nd Int. Conf. on Artificial Intelligence and Information Systems (ICAIS)*, p. 012095, May. 2021. DOI: 10.1088/1742-6596/1982/1/012095.

Journal Papers

- [J1] P. Bai, X. Liu, **W. Fan**, T. Jiang, S. Alabed, W.K. Cheung, and H. Lu, “Geometry-aware Line Graph Transformer Pre-training for Molecular Property Prediction”, *IEEE Transactions on Neural Networks and Learning Systems (IF: 8.9)*, May. 2026.
- [J2] T. Payne*, **W. Fan***, C. Chen, S. Zhou, V. Osmani, H. Lu, and O. Bandmann, “Risk of Parkinson’s disease following cholecystectomy: a prospective cohort study in the UK Biobank,” *npj Parkinson’s Disease (IF: 8.2)*. Under revision.
- [J3] M. N. I. Suvon, S. Zhou, P. C. Tripathi, **W. Fan**, S. Alabed, B. Khanal, V. Osmani, A. J. Swift, C. Chen and H. Lu, “Multimodal Latent Fusion of ECG Leads for Early Assessment of Pulmonary Hypertension,” *IEEE Journal of Biomedical and Health Informatics (IF: 6.7)*. arXiv: 2503.13470, Under review after major revision.

Preprints & Abstracts

- [P1] **W. Fan**, M. N. I. Suvon, S. Zhou, X. Liu, S. Alabed, V. Osmani, A. Swift, C. Chen, & H. Lu, “MeDSLIP: Medical Dual-Stream Language-Image Pre-training with Pathology-Anatomy Semantic Alignment”, *arXiv: 2403.10635*, 2024.

项目领导力

- **技术与科研项目管理**
 - 将生物医学问题转化为可执行的机器学习任务，并在神经病理性疼痛与帕金森病项目中负责关键方法决策。
 - 统筹技术交付、合作方沟通、行动项跟踪及后续推进。
 - 独立管理 KaleProtein 从研发到发布的全流程。
- **Bio-lab Research Assistant Hackathon Team – Team Lead; Second Place**
 - 搭建核心智能体系统，覆盖实验室管理、实验规划、参数计算及实时语音问答。
 - 统筹并集成团队成员开发的配套工具，完成原型系统交付。
- **Self-driving Protein and Materials Design Lab Hackathon Team – Team Lead; Second Place**
 - 带领由生物学、化学、数学、主动学习及人工智能研究人员组成的跨学科团队。
 - 设计并实现需求驱动的智能体工作流，用于金属结合蛋白的设计、评估、筛选与优化。
 - 牵头原型集成、演示方案设计及技术汇报。
- **参与组织活动**
 - Fourth Workshop on Multimodal AI，英国伦敦，2026 年 6 月。
 - Second Mini-Hackathon on Multimodal AI，英国伦敦，2026 年 6 月。
 - Second Multimodal AI Community Forum，英国谢菲尔德，2026 年 3 月。
 - Third Workshop on Multimodal AI，英国伦敦，2025 年 9 月。
 - First Mini-Hackathon on Multimodal AI，英国伦敦，2025 年 9 月。
 - PyKale tutorial at IEEE EMBC 2025，丹麦哥本哈根，2025 年 7 月。
 - Second Workshop on Multimodal AI，Sheffield，英国伦敦，2024 年 9 月。

代表性基金申请

- **UK Sovereign AI Strategic Assets – 共同研究者；约 £1.27M**
 - 提出核心构想：构建覆盖蛋白质与有机材料、可直接用于人工智能研究的基准数据集。
 - 牵头完成大部分意向书撰写，并协调学术界与产业界合作伙伴的后续技术工作。
- **ARIA Activation Partner – 共同申请人；£6.39M**
 - 共同设计项目构想、技术路线及交付计划。
- **UKRI Generative AI Hub – 共同申请人；两项申请，各 £100K**
 - 牵头其中一项申请的项目构想、技术规划及文本起草。
 - 参与另一项申请的项目设计。
- **Innovate UK Frontier AI Benchmarking Datasets – 共同申请人；两项申请，各 £750K**
 - 贡献技术构想与基准评测方法。
- **ARIA x Encode AI Fellowship – 申请人；£115K**
 - 独立完成以人工智能驱动多状态蛋白质设计为主题的申请方案。

学术交流、培训与专业服务

- **代表性学术报告**
 - “MuSProt: A Large-scale Multistate Protein Database” (Poster)，*Prague Protein Spring*，捷克布拉格，2026 年 5 月。
 - “State-Aware Functional Annotations for Multi-State Design of Dynamic Proteins” (Poster/Pitch)，*AI, Engineering Biology and Beyond*，英国布里斯托，2026 年 1 月。
 - “Neuropathic Pain Drug-response Prediction” (Poster)，*MICCAI*，韩国大田，2025 年 9 月。
 - “AI for Parkinson’s Disease” (Talk)，*Sheffield AI Community Meeting*，英国谢菲尔德，2024 年 3 月。
- **期刊审稿人**
 - IEEE Transactions on Neural Networks and Learning Systems.
 - IEEE Journal of Biomedical and Health Informatics.
 - Computational and Structural Biotechnology Journal.
- **会议审稿人**
 - International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), 2026.
 - MICCAI Workshop on Advancing Data Solutions in Medical Imaging AI (ADSMI), 2024.