

孙 浩

+86-17863107076 | sunhaoxx@zju.edu.cn | kiva12138.github.io (个人网页)

Hao Sun | Kiva12138 | 0000-0001-8094-1991 | Kiva12138

个人简介

孙浩，博士研究生，毕业于浙江大学计算机科学与技术学院，并曾在日本立命馆大学担任高级研究员与访问学者。研究方向涵盖多模态学习、大规模语言模型（LLM）、视觉语言动作模型（VLA）、强化学习及情感计算，致力于推动通用人工智能（AGI）及多模态技术的发展。以第一作者在 ACM Multimedia、Information Fusion、IEEE Transactions on Affective Computing、Pattern Recognition 等国际顶级期刊与会议发表多篇论文，总引用量超过 480 次，并拥有多项中美专利。参与国家重点研发计划、省级自然科学基金及产学研合作项目，具备扎实的科研与工程实践能力。曾获浙江大学优秀研究生、华为奖学金等多项荣誉，兼具学术创新力与项目组织协调能力。

教育经历

- | | |
|--|-------------------------|
| • 浙江大学 [1] | 2020 年 9 月 - 2025 年 6 月 |
| 博士研究生，计算机科学与技术专业；获“优秀毕业生”荣誉，学业成绩位列前 10%。 | 中国，杭州 |
| • 哈尔滨工业大学 [1] | 2016 年 9 月 - 2020 年 6 月 |
| 本科，软件工程专业；获“优秀毕业生”荣誉，成绩位列前 8%。 | 中国，哈尔滨 |

科研与工作经历

- | | |
|---|-------------------------|
| • 立命馆大学 - 计算机科学与工程学院 [1] | 2025 年 6 月 - 至今 |
| 高级研究员（受日本工程院院士陈延伟教授邀请，日本文部省资助） | 日本，大阪 |
| 负责所在实验室的 LLM、多模态、VLA 和强化学习方向研究工作。 | |
| 主导以多模态大模型、强化学习及仿生学为核心的 VLA 与通用 AGI 研究项目。 | |
| • 立命馆大学 - 计算机科学与工程学院 [1] | 2023 年 8 月 - 2024 年 8 月 |
| 访问学者（受日本工程院院士陈延伟教授邀请） | 日本，大阪 & 大津 |
| 获浙江大学“博士学术之星”项目资助（该奖项授予前 100 名优秀博士研究生）。 | |
| 主导基于大语言模型的统一多模态、多任务框架研究。 | |
| 主导探索通过参数高效微调（PEFT）提升大模型多模态处理能力的研究。 | |
| 研究成果发表于 IEEE Transactions on Affective Computing、Pattern Recognition 等国际权威期刊。 | |

科研项目经历

- | | |
|--|-------------|
| • 类风湿性关节炎智能化综合分析平台建设 | 2022 - 2025 |
| 国家重点研发计划项目 (2022YFC2504605)，科技部资助 | 核心成员 |
| 致力于开发基于人工智能的综合分析平台，以提升类风湿性关节炎的诊断与治疗水平。 | |
| 构建融合多模态临床数据的新方法，使诊断准确率提升约 10%。 | |
| 负责多模态方法的开发、实现与验证。 | |
| • 基于联邦学习的肝细胞癌术前早期复发检测与预测 | 2022 - 2024 |
| 浙江省自然科学基金重点项目 (LZ22F020012) | 核心成员 |
| 开发联邦学习解决方案，实现术前肝癌复发的早期预测并兼顾隐私保护。 | |
| 在保证数据隐私的前提下，将复发预测准确率提升约 13%。 | |
| 负责项目方案撰写、多模态方法研究与最终验证。 | |
| • 基于计算机视觉的智慧工地管理平台关键技术研究 | 2022 - 2024 |
| 杭州新中大科技股份有限公司产学研合作项目 (2022AIZD0147-02) | 核心成员 |
| 开发智慧工地监测与管理平台，以提升施工安全与管理效率。 | |
| 成功搭建实时监测系统，使违规率下降约 15%，并简化现场管理流程。 | |
| 负责项目方案设计、研究方法论、成果验收与部分管理工作。 | |

- [C.1] [第一作者]. **CubeMLP: An MLP-based Model for Multimodal Sentiment Analysis and Depression Estimation.** In *ACM Multimedia Proceedings*. 2022. pp.3722–3729. Association for Computing Machinery. Lisboa, Portugal. DOI: 10.1145/3503161.3548025. 引用超过 150.
- [J.1] [第一作者]. **Modality-Invariant Temporal Representation Learning for Multimodal Sentiment Classification.** *Information Fusion*. 2023. Vol.91. pp.504-514. Elsevier. DOI: 10.1016/j.inffus.2022.10.031. Cited by 40 (as of June 2025). 影响因子: 18.1.
- [J.2] [第一作者]. **Tensorformer: A tensor-based multimodal transformer for multimodal sentiment analysis and depression detection.** *IEEE Transactions on Affective Computing*. 2023. Vol.14(4). pp.2776-2786. IEEE. DOI: 10.1109/TAFFC.2022.3233070. Cited by 45 (as of June 2025). 影响因子: 13.9.
- [J.3] [第一作者]. **Multimodal Sentiment Analysis with Mutual Information-based Disentangled Representation Learning.** *IEEE Transactions on Affective Computing*. 2025. Vol.16(3). pp.1606-1617. IEEE. DOI: 10.1109/TAFFC.2025.3529732. 影响因子: 13.9.
- [J.4] [第一作者]. **One Framework to Rule Them All: Unifying Multimodal Tasks with LLM Neural-Tuning.** *Pattern Recognition*. 2025. Early Access. Elsevier. DOI: 10.1016/j.patcog.2025.112275. 影响因子: 8.5.
- [J.5] [第一作者]. **Multi-Modal Adaptive Fusion Transformer Network for The Estimation of Depression Level.** *SENSORS*. 2021. Vol.21(14). pp.4764. MDPI. DOI: 10.3390/s21144764. Cited by 80 (as of June 2025).
- [S.1] [第一作者]. **Multimodal Infusion Tuning for Large Models.** *ArXiv.2403.05060*, 2024.
- [S.2] [第一作者]. **Robust Latent Representation Tuning for Image-text Classification.** *ArXiv.2406.06048*, 2024.
- [S.3] [第一作者]. **Modality-invariant and Specific Prompting for Multimodal Human Perception Understanding.** *ArXiv.2311.10791*, 2023.
- [C.2] [共同第一作者]. **EPIC: Efficient Prompt Interaction for Text-Image Classification.** In *IEEE International Conference on Multimedia & Expo (ICME)*. 2025. IEEE.
- [C.3] [第二作者]. **IRLSG: Invariant Representation Learning for Single-Domain Generalization in Medical Image Segmentation.** In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2024. pp.5585-5589. IEEE. DOI: 10.1109/ICASSP48485.2024.10446700.
- [C.4] [第三作者]. **LGA: A Language Guide Adapter for Advancing the SAM Model's Capabilities in Medical Image Segmentation.** In *International Conference on Medical Image Computing and Computer-Assisted Intervention Proceedings (MICCAI)*. 2024. pp.610-620. Springer. DOI: 10.1007/978-3-031-72390-2_57.
- [C.5] [第三作者]. **MCKD: Mutually Collaborative Knowledge Distillation for Federated Domain Adaptation And Generalization.** In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2023. pp.1-5. IEEE. DOI: 10.1109/ICASSP49357.2023.10095699.
- [J.6] [第三作者]. **CoSTHR: A Heart Rate Estimating Network with Adaptive Color Space Transformation.** *IEEE Transactions on Instrumentation and Measurement*. 2022. Vol.71. pp.1-10. IEEE. DOI: 10.1109/TIM.2022.3170976. Cited by 25 (as of June 2025).
- [C.6] [第三作者]. **Enhanced Multimodal Depression Detection With Emotion Prompts.** In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2025. pp.1-5. IEEE. DOI: 10.1109/ICASSP49660.2025.10889035.
- [P.1] **Engineering Progress Determining Method and Device Based on Multi-Mode Time Sequence Information Fusion.** 专利申请号: US20250005475A1, 中国/美国, 公开日期: 2023 年 7 月。
- [P.2] 基于多模态时序信息融合的工程进度确定方法及装置。专利授权号: CN116502882B, 中国, 授权日期: 2023 年 10 月。
- [P.3] 基于多模态数据的类风湿关节炎活动度分级装置。专利授权号: CN116797572B, 中国, 授权日期: 2025 年 9 月。
- [P.4] 一种面向医学图像分割的单一域泛化方法。专利授权号: CN116596832B, 中国, 授权日期: 2025 年 7 月。

荣誉与奖励

• 优秀毕业研究生	浙江大学；授予学业成绩突出的博士研究生，获奖比例约前 10%。	2025 年 6 月 中国，杭州
• 优秀研究生（四次）	浙江大学；授予学业成绩优异的博士研究生，获奖比例约前 15%。	2024 年 12 月 / 2023 年 12 月 / 2022 年 12 月 / 2021 年 12 月 中国，杭州
• 学业奖学金	浙江大学；用于资助科研表现突出的优秀博士研究生。	2022 年 12 月 中国，杭州
• 优秀研究生干部（两次）	浙江大学；授予在科研及社会服务中展现突出领导力与贡献的研究生。	2024 年 12 月 / 2023 年 12 月 中国，杭州
• 五好优秀学生	浙江大学；授予兼具优异学业成绩与积极社会活动表现的研究生。	2023 年 12 月 中国，杭州
• 华为奖学金	浙江大学；授予在计算机科学与人工智能领域学业成绩优异、科研创新突出的优秀学生。	2023 年 12 月 中国，杭州
• 优秀毕业生	哈尔滨工业大学；授予成绩优异、综合表现突出的本科毕业生，获奖比例约前 15%。	2020 年 6 月 中国，哈尔滨
• 国家励志奖学金	哈尔滨工业大学；授予表现优异的本科生（前 5%），以表彰其卓越成绩。	2018 年 12 月 中国，哈尔滨

技能与专长

- 学术能力：学术论文写作、学术出版、同行评审、学术会议报告、科研项目申报
- 人工智能研究：算法研发、模型训练与微调、数据处理、模型评估
- 大语言模型：模型定制化、知识集成、多模态调优、可扩展性与效率优化
- 多模态研究：框架设计、任务适配、多模态系统部署
- 软件工程：可行性与需求分析、系统及详细设计、实现与软件维护
- 编程语言：Python、PyTorch、NumPy、TensorFlow、Java、C++、C、HTML、Go 等
- 全栈开发：前端开发、后端设计与实现、数据库系统
- 语言能力：中文（母语）、英文（CET-6: 620）、日语

其他信息

2020–2021：在慕课网（IMOOC）发布 TensorFlow 教程。

2016 - 2025：独立开发超过 30 个项目，部分成果可见 GitHub 主页 (<https://github.com/kiva12138>)。

推荐人

1. 陈延伟，教授

日本立命馆大学 - 信息理工学院

日本工程院院士

邮箱：chen@is.ritsumei.ac.jp

推荐关系：博士/博后研究合作导师

2. 林兰芬，教授

浙江大学 - 计算机科学与技术学院

浙江大学人工智能研究所副所长

邮箱：llf@zju.edu.cn

推荐关系：博士生导师