

李浩勋

(+86) 13601156976 | lihaoxun23@mails.ucas.ac.cn

男 25岁 (2000年11月)



教育背景

本科: 北京工业大学 **专业:** 计算机科学与技术(荣誉学院) **CET-6:** 473 **GPA:** 3.78/4.0(推免) **2019.9-2023.6**

硕士: 中国科学院大学 (杭高院, 导师: 李太豪研究员) **专业:** 人工智能 **GPA:** 3.78/4.0 **2023.9 至今**

研究兴趣: 情感语音生成、识别, 多模态情感计算

荣誉奖项

本科期间: 第七届中国国际“互联网+”大学生创新创业大赛校级特等奖、北京市二等奖; 第十五届中国大学生计算机设计大赛北京市二等奖、全国三等奖; 优秀毕业设计。

硕士期间: 中国科学院大学新生奖学金、一等学业奖学金、三好学生; 国家奖学金。

项目经历

浙江省尖兵计划: 基于大语言模型的数字化治疗系统研发及在重大脑疾病的应用 <核心成员> 2024.9.2 至今

· **项目描述:** 该项目由本团队与浙大一院精神卫生科、西湖大学、浙江大学以及三所公司共同承担, 最终目标是获得医疗器械注册证, 并在超过 10 所以上学校或医院进行部署。

· **负责内容:** 本团队负责其中的“抑郁症数字靶点挖掘”子任务, 从文本、视频、语音、多模态信息、生理信号的角度分别对数字靶点进行挖掘。本人负责其中语音部分的撰写任务和实现, 通过一种无监督方式解耦语音中的韵律表征, 实现说话人无关的语音韵律数字靶点自动挖掘。其间被任命为技术实现负责人, 主要工作是设计一套 AI 自动问诊多模态抑郁症检测系统的技术路线并协助与医院方的沟通, 带领团队推进进度。

组内核心项目: 多模态情感交互系统构建与应用 <负责人> 2024.11.28 至今

· **项目描述:** 该项目构建一个具有多模态情感识别、共情对话、图片生成、诗词生成、音乐生成等功能的综合性系统, 旨在集成组内以情感为主线的各项研究成果, 作为组内招牌。

· **负责内容:** 在关键节点主动承担责任, 被任命为项目负责人, 负责整体规划、工作调度、系统设计、系统部署, 并在三周时间内完成部署。后续参与联通、省建投、卓迈汽车等公司的商业交涉, 并建立良好合作关系。被导师评为组内能力最突出、最刻苦、贡献最大成员, 并颁发奖金。目前正在迭代系统的 2.0 版本, 带领全组成员使用多智能体革新整个架构, 实现研究和工程化的共同推进, 并在毕业前与 25 级学生完成交接工作。

横向课题: 基于大语言模型的专利价值数智化评估 <负责人> 2024.9.9 至今

· **项目描述:** 传统的专利价值评估方式往往依赖于专家的主观判断, 缺乏统一标准, 且耗时长、成本高。本项目旨在探索大语言模型在专利价值评估中的应用, 结合法律、技术和经济等多维度信息, 打造一个公正、可信、智能化的专利价值评估系统。

· **负责内容:** 作为项目负责人, 从项目初期开始负责管理项目相关全部事宜。设计了基于数据与知识双驱动的多源信息融合方案, 将模型训练分为数据准备、信息提取、信息融合三个部分。实现过程中主要负责信息提取部分, 即利用领域大模型从爬取到的专利数据中挖掘多维信息、构建专利与市场信息知识库。目前在我们爬取的 4000+ 数据上进行五分类的训练、测试, 准确率已达到 93%, 并取得省领导、专利局的进一步支持, 预计在实验室成立试点, 同时提供大量真实交易数据以供模型优化。

短期项目: 基于视觉行为特征的疲劳与压力智能检测系统 <负责人> 2025.9.23 至 2025.10.22

· **项目描述:** 青少年使用平板学习过程中容易因疲劳与压力导致注意力下降、学习效率降低。本项目面向真实学习场景, 构建一个疲劳与压力智能检测系统, 实现学习状态的实时评估与反馈, 并支持关键证据留存与可追溯分析。

· **负责内容:** 作为项目负责人, 首先, 完成需求拆解与指标体系设计, 明确疲劳/压力/注意力等状态的定义、判定规则与验收标准。其次, 主导核心算法模块落地, 统一集成人脸检测与表情识别, 并结合眼部开闭、哈欠、微睡眠等疲劳事件检测逻辑实现实时判定, 同时设计“疲劳触发—关键帧保存—会话级记录”的证据链闭环。最后搭建并交付“在线实时 + 离线验证”的系统流程与展示页面, 组织测试与问题复盘, 确保在周期内完成可用版本交付。

研究成果

* denotes Co-first Author, † denotes Advisor

· **Haoxun Li, Yu Liu, et al.** EMORL-TTS: Reinforcement Learning for Fine-Grained Emotion Control in LLM-based TTS[J]. arXiv preprint arXiv:2510.05758, 2025. (Accepted by ICASSP 2026)

· **Haoxun Li, Yuqing Sun, et al.** MSF-SER: Enriching Acoustic Modeling with Multi-Granularity Semantics for Speech Emotion

Recognition[J]. arXiv preprint arXiv:2510.05749, 2025. (Accepted by ICASSP 2026)

- Yu Liu*, Hanlei Shi*, **Haoxun Li**, et al. Centering Emotion Hotspots: Multimodal Local-Global Fusion and Cross-Modal Alignment for Emotion Recognition in Conversations[J]. arXiv preprint arXiv:2510.08606, 2025.
- **Hanlei Shi***, **Yu Liu***, **Haoxun Li***, **Yuxuan Ding***, et al. HOPE: Hierarchical Fusion for Optimized and Personality-Aware Estimation of Depression[C]//Proceedings of the 33rd ACM International Conference on Multimedia. 2025: 13937-13943.
- **Haoxun Li**, Leyuan Qu†, et al. EME-TTS: Unlocking the Emphasis and Emotion Link in Speech Synthesis. [C]//Proc. Interspeech 2025. 2025: 4368-4372.
- Jiaxi Hu, Leyuan Qu, **Haoxun Li**, et al. Label Semantic-Driven Contrastive Learning for Speech Emotion Recognition[C]//Proc. Interspeech 2025. 2025: 4348-4352.
- 《一种细粒度强调可控的情感语音合成方法》，发明专利，CN202411669784.7，学生第一作者
- 《一种基于有监督对比学习的语音情感识别方法》，发明专利，CN202411669786.6，学生第二作者
- 《一种动态表情识别方法、电子设备及计算机可读存储介质》，发明专利，CN202411669783.2，学生第三作者
- 《一种基于和弦情感的动机发展多声部音乐生成算法》，发明专利，CN202410175359.6，第一作者
- 软件著作权一篇（2024SR0123464），实用新型专利一篇（CN202220719665.8）

学术报告

- Interspeech 2025 Poster 汇报，荷兰鹿特丹；
- 第五届中国情感计算大会 (Chinese Conference on Affective Computing, CCAC) Poster 汇报，四川成都。

技能摘要

- 掌握 Python, Java, C++ 等编程语言；
- 掌握 Conda, PyTorch, HuggingFace Transformers, Sox, MySQL 等工具的使用；
- 熟练使用 Cursor 进行 AI 辅助编码与代码补全；
- 具有使用框架 (LlamaFactory, Axolotl) 进行 LLM 微调 (全量, LoRA, adapter tuning) 及强化学习的经验；
- 长期担任组内项目组组长 (研二至今)，具有一定布局、策划、沟通、协调、管理项目的能力；
- 长期担任组内服务器管理员 (研二至今)，具有服务器相关管理、维护等经验。

个人特色

- 自小学二年级学棋开始，仅三年升至围棋业余最高段位（五段），国家二级运动员；
- 古典吉他十级，指弹吉他演奏级。本科阶段担任民谣吉他社副社长，成立乐队并创作院歌带领全员进行录制，组织教学、小规模演出数场。研究生阶段成立器乐社，创立杭高院第一支乐队并出席各项演出；
- 羽毛球、乒乓球、网球、足球、台球、排球均为院队主力，自读研以来共取得 5 枚奖牌。