序号	年级	姓名	本科院校	硕士/博士论文题目（中英文）	主要研究内容	科研成果说明	就业单位	地点
1	2022级博士生	杨绪升	哈尔滨工业大学/郑州大学	低资源神经语音编解码及多任务协同方法研究	该论文研究低资源场景下的神经语音编解码及多任务协同方法。论文完成的主要创新性工作如下: (1)针对动态带宽多码率条件下语音编解码感知质量不稳定的问题，提出了一种基于多码率标量量化联合优化的神经语音编解码方法，改善了低码率语音重构质量下降、离散表示鲁棒性不足的问题; (2)针对个性化实时通信中的级联系统复杂度高与协同缺失的问题，提出了一种将变声功能直接嵌入神经语音编解码器的多功能协同框架，增强了神经语音编解码系统在个性化实时通信下的多功能协同能力; (3)针对低帧率下离散语音表示的信息保持与生成效率权衡的问题，提出了一种面向超低帧率的神经语音编解码方法，提高了极端压缩条件下的重构质量。	**发表论文:** 发表1篇ICASSP（CCF-B）国际会议论文（2026，一作）； 发表1篇CIKM（CCF-B）国际会议论文（2024，一作） 发表1篇APSIPA ASC（EI）国际会议论文（2024，一作）； 发表1篇ICASSP（CCF-B）国际会议论文（2025，二作）； 发表1篇INTERSPEECH（CCF-B）国际会议论文（2026，一作）； 发表1篇NCMMSC（CCF-C）国际会议论文（2026，一作）； 发表1篇Neural Networks（CCF-B、JCR Q1）国际期刊论文（2026，一作）； **在投论文:** 投稿1篇 ACM TOMM（CCF-B、JCR Q1）国际期刊论文（一作，大修） **参与项目:** 深圳市高等院校稳定支持计划：基于表征学习和知识辅助的口语对话关键技术研究	中国电信	北京
2	2023级硕士生	陈展鹏	中山大学	基于先验调控机制的视觉-语言幻觉缓解方法研究	论文研究基于先验调控机制的视觉-语言幻觉缓解方法，选题具有重要的理论意义和应用价值。论文完成的主要研究工作如下: 1. 提出了一种基于视觉先验重构的动态多粒度位置编码方法，提升了模型细粒度视觉特征感知能力与视觉词元准确性; 2. 提出了一种基于模态先验解耦的跨模态因果推断方法，提升了复杂场景下模型跨模态对齐忠实度与输出可靠性。	**发表论文:** 发表1篇ACL Findings（CCF-A Findings）国际会议论文（2025，一作）； 发表2篇EMNLP（CCF-B）国际会议论文（2024，一作）； **参与项目:** 1）AI4S交叉研究专项——“AI驱动的深蓝光OLED材料寿命因子发现与分子设计” **获奖情况:** 1）2024-2025学年北京大学“三好学生”	美团	深圳
3	2023级硕士生	林义凯	同济大学	脑卒中多模态数据集构建与诊断模型研究	该论文研究脑卒中多模态数据集构建与诊断方法，选题具有重要的理论意义与应用价值。 论文完成的主要工作如下: 1、构建了一个脑卒中多模态视觉问答数据集，提高了数据构建效率并验证了数据集可用性; 2、提出了一种面向病例级多模态信息的后验评估与重排序方法，提升了诊断结果的稳定性与可靠性; 3、提出了一种一致性约束下的迭代优化策略，提升了模型性能。	**发表论文:** 发表1篇ICASSP（CCF-B）国际会议论文（2026，一作）；	待定	待定
4	2022级硕士生	聂宸涞	吉林大学	弱监督数据构建与风格可控语音合成方法	论文研究弱监督数据构建与风格可控语音合成方法，选题具有重要的理论意义和应用价值。论文完成的主要研究工作如下: 1、提出了一种基于双轨特征和分层量化的语音半自动标注生成方法，构建了多维度语音风格标签数据集; 2、提出了一种基于cT5中文编码器和情感强度连续建模的风格语义增强方法，提升了合成语音的表现力与可控性。	**构建数据库:** 构建1个中文语音风格数据集CN-StyleCaps（规模约 15000 小时， 覆盖约 8000 位说话人， 包含 5000 余种自然语言风格描述） **申请国内发明专利:** 申请国内发明专利一篇 **参与项目:** 校企合作项目，“语音合成大模型”，2024/10-2025/07	比亚迪	深圳
5	2023级硕士生	谭凤	上海财经大学	知识增强的皮肤病医学报告生成方法研究	该论文研究知识增强的皮肤病医学报告生成方法，选题具有重要的理论意义与应用价值。 论文完成的主要研究工作如下: 1. 构建了一个面向可解释性的皮肤病病灶特征增强数据集，为提升模型的可解释性提供了数据支撑。 2. 提出了一种基于知识增强的结构化多模态推理框架，提升了皮肤病医学报告生成的性能;	**发表论文** [1]发表一篇IJCNN（CCF-C）国际会议论文（2026，一作）； **申请国内发明专利** [1]邹月娴，谭凤，张军斌，冯林.一种基于动态检索和链式推理的皮肤病辅助诊断方法、装置及介质：202610271796.7，2026-03-06. **参与项目** [1]重庆市中医院，数智中医首批开放课题，“基于中西医多模态信息的皮肤病诊疗专业大模型研究”2025/11-2026/03	比亚迪上海有限公司	上海
6	2023级硕士生	王之畅	东南大学	面向大语言模型的结构化思维链蒸馏方法研究	论文研究面向大语言模型的结构化思维链蒸馏方法，选题具有重要的理论意义与应用价值。论文完成的主要工作如下: 1、提出了一种基于多步递归的层次化链式思维迁移方法，提升了复杂推理任务中的结构化表达能力; 2、提出了一种基于层次一致性学习与层次思维对比学习的思维链蒸馏方法，提高了推理知识迁移的准确性; 3、提出了一种结构化思维链蒸馏框架，实现了对链式、树式与图式推理结构的统一建模与泛化。	**发表论文:** 发表1篇ICASSP（CCF-B）国际会议论文（2025，一作）； 发表1篇ICLR（CCF-A）国际会议论文（2025，一作）；	凌影科技（杭州）有限公司	杭州
7	2023级硕士生	谢宇昕	电子科技大学	面向音频-文本检索的细粒度对齐和知识解耦方法研究	该论文研究面向音频-文本检索的细粒度对齐与知识解耦方法，选题具有重要的理论意义与应用价值。 论文完成的主要工作如下: 1、提出了一种基于原型的细粒度特征对齐方法，显式建模了音频与文本间的局部语义对应关系; 2、提出了一种基于Sinkhorn-Knopp算法的检索分数校正机制，缓解了A2T检索中的实例不平衡问题; 3、提出了一种基于知识解耦的动态注入增强框架DIKE，提升了模型的鲁棒性与泛化能力。	**发表论文:** 发表 1 篇 Interspeech（CCF-B）国际会议论文（2024，一作，Oral）； 发表 1 篇 ACL（CCF-A）国际会议论文（2025，二作）； 发表 1 篇 CVPR（CCF-A）国际会议论文（2025，三作）。 **参与项目:** 1）深圳市高校稳定支持计划项目"基于表征学习和知识辅助的口语对话关键技术研究"（2022.09.01 - 2024.03.20）。	阿里巴巴	杭州
8	2023级硕士生	徐晚诗	大连理工大学	面向知识对话的多粒度信息感知检索增强生成方法研究	该论文研究面向知识对话的检索增强生成方法，选题具有重要的理论意义与应用价值。论文完成的主要工作如下: 1、提出了一种基于动态偏好对齐的端到端任务型对话生成方法，提升了对话实体预测准确率与回复流畅度; 2、提出了一种基于对抗式课程学习的细粒度检索增强生成方法，抑制了大模型幻觉。	**发表论文:** 发表2篇EMNLP国际会议论文（2024，一作）；	小红书	上海

9	2023级 硕士生	庄先炜	浙江大学	视觉语言大模型幻觉成因分析与缓解方法研究	该论文研究视觉语言大模型幻觉成因与缓解方法，选题具有重要的理论意义与应用价值。 论文完成的主要工作如下： 1. 提出了一种基于视觉感知词元稀疏化的视觉幻觉缓解方法，平衡了解码效率与幻觉缓解性能； 2. 提出了一种基于视觉感知增强和博弈论的视觉幻觉缓解方法，增强了对多粒度视觉语义的幻觉缓解能力； 3. 提出了一种基于强化学习的幻觉缓解方法，提升了模型在视觉理解与生成任务中的视觉忠实度。	**发表论文：** 发表1篇CVPR（CCF-A）国际会议论文（2025，一作）； 发表1篇ACL（CCF-A）国际会议论文（2024，一作）； 发表1篇AAAI（CCF-A）国际会议论文（2024，一作）； 发表1篇EMNLP（CCF-B）国际会议论文（2024，一作）； 发表1篇ACM MM（CCF-A）国际会议论文（2024，一作）； **参与项目：** 1）公司委托项目“北大-腾讯犀牛鸟项目” **获奖情况：** 1）北京大学优秀毕业生 2）北京大学2023-2024学年度北京大学国家奖学金； 3）北京大学2023-2024学年三好学生标兵； 4）北京大学2024-2025学年三好学生；	阿里通义千问	北京