

刘佳璇



性别：男

籍贯：河北衡水

政治面貌：中共党员

Lab：语音及语言信息处理

国家工程研究中心

本科专业：计算机科学与技术

本科综合成绩排名：1/234

研究生专业：信息与通信工程

研究方向：语音合成、语音大模型

邮箱：jxliu@mail.ustc.edu.cn



教育经历

- 2017年9月至2020年7月：河北衡水中学 化奥生
- 2020年9月至2024年6月：西北工业大学计算机学院 本科生
- 2024年9月至今：中国科学技术大学信息科学技术学院 研究生



获奖情况

奖学金

- 2021年国家奖学金
- 2022年国家励志奖学金
- 2022年吴亚军专项奖学金
- 2023年国家励志奖学金
- 2023年小米优秀奖学金
- 2021年西北工业大学一等奖学金
- 2022年西北工业大学一等奖学金
- 2023年西北工业大学一等奖学金
- 2024年中国科学技术大学一等奖学金

学科竞赛（部分）

- 国际数学建模竞赛 MCM 特等奖
- 全国大学生数学建模竞赛陕西省一等奖
- 全国大学生物理学术竞赛陕西省个人一等奖
- 中国大学生物理学术竞赛陕西省团队一等奖
- 百度文心大模型创意应用项目 A-科技奖励
- “挑战杯”陕西省大学生创业计划竞赛铜奖
- 西北工业大学程序设计竞赛一等奖
- 西北工业大学大学生数学竞赛一等奖
- 陕西省语音与图像信息处理重点实验室实习优秀个人

其他荣誉

- 西北工业大学优秀毕业生
- 2021西北工业大学优秀大学生
- 2022西北工业大学优秀大学生
- 2023西北工业大学优秀大学生
- 2021年五四表彰“优秀共青团员”
- 2022年五四表彰“优秀共青团员”
- 2025年优秀共产党员
- 授权一项国家发明专利



科研及实习经历

语音方向科研项目

- 2022年6月至2022年7月，我在“陕西省语音与图像信息处理重点实验室”实习，基于 Opencpop 数据集和 ByteSing 完成了歌唱合成项目，基于 BLSTM 优化了 Duration&LF0&UV Predictor 来捕获长期的时间依赖性，显著降低了 F0 误差，提高了音频合成质量，荣获优秀实习个人。
- 2023年6月至2023年7月，我参与“华为有声书高表现力连续情感空间建模”实习，解决有声书听觉疲劳问题。我参与实验建立了多模态的情感空间解决了显性离散情感类别局限的问题，基于 T5 大模型和对比学习，完成了跨模态监督的文本情感感知模块，用 text embedding 监督 text emotion encoder 的学习，增加了相似度 loss 加速收敛。“Text-aware and Context-aware Expressive Audiobook Speech Synthesis”。
- 2023年12月至2024年4月，我在科大讯飞股份有限公司实习，在核心研发平台担任助理研究算法工程师，负责语音合成方向的科研工作，开展基于扩散模型的韵律多样可控语音合成声学模型研究，完成本科毕业设计。
- 2024年6月至2024年10月期间，我与科大讯飞和中国移动研究院继续合作完成基于扩散模型的韵律多样可控语音合成声学模型研究，提出了 DiffStyleTTS，将语音韵律分层建模，基于条件扩散模型和改进的无分类引导策略，实现了语音韵律的多样性和可控性，在高自然度、韵律迁移和韵律控制方面均取得优越效果。本人为第一作者在 COLING 2025 上发表长论文并作了 Oral 学术报告。“DiffStyleTTS: Diffusion-based Hierarchical Prosody Modeling for Text-to-Speech with Diverse and Controllable Styles”
- 2024年11月至2025年5月，我独立开展一项情感语音合成的研究，致力于在 LLM 背景下帮助开发者基于大规模情感数据集构建可控的情感 TTS 系统，最终提升人机交互中情感表达的自然度。我以 CosyVoice 作为可扩展 baseline 提出了一种融合离散情感和维度情感的 TTS 统一框架 UDDETTS，构建了新颖的 Codec，实现了与多种情感语音数据集的兼容。UDDETTS 引入 ADV 情感维度空间，克服了情感类别的局限性，建模了细粒度情感表征；并引入半监督训练和 ADV 量化器，缓解了 ADV 空间中的稀疏性和不平衡问题。UDDETTS 在自然度、情感控制准确性等方面均显著优于基线，在三个可解释的维度上实现了情感的线性控制，并证明了 ADV 在端到端的自适应情感 TTS 中的可行性。本人为第一作者预发表长论文“UDDETTS: Unifying Discrete and Dimensional Emotions for Controllable Emotional Text-to-Speech”。

其他科研经历

- 2021年5月至2022年5月，我参与“神经网络算法在偏微分方程中的应用研究”大创项目，我运用物理神经网络，对初始和边界条件约束，对抛物线型方程、椭圆型方程、双曲型方程等基本二阶偏微分方程进行参数拟合，工作以学术论文优秀结题。
- 2022年5月至2023年5月，我参与“基于目标检测的全维度智慧交通巡检机器人”大创项目，我提出一种轻量化的端到端道路损伤检测目标分割方法，基于骨干特征提取网络 Shuffle-ECANet 将拍摄的实时图像转化为多通道特征图，基于 FPN 的多尺度特征融合网络引入了自适应注意力模块 AAM 和特征增强模块 FEM，来对不同尺度的特征融合。嵌入式轻量级注意力模块 ECA-Net 能比以往研究用更少的参数来提高检测精度，实现了精准识别路面损伤的类型。最终授权一项国家发明专利。
- 2022年8月至2022年10月，我参加百度文心大模型峡谷社区创意应用项目，获得百度 A-评级奖励，有三千多名用户使用量。
- 2022年10月至2022年12月，我完成了“药物相互作用关系链接预测”项目，应用图卷积神经网络进行链路预测，解决了不同种药物之间是否有相互作用关系的预测。工作以学术论文优秀结题。
- 2023年4月至2023年6月，我完成了“医学 X 线影像关键点检测”课程项目，基于人体姿态模型实现了病理判断。



其他工作实践

- 2020年9月至2024年6月：任学习委员，两年带领班级平均成绩名列专业第一。
- 2020年10月至2021年7月：任西北工业大学学生会干事和学习辅导中心朋辈导师。
- 2021年8月至2021年10月：参加十四运开闭幕式，获中华人民共和国第十四届运动会开闭幕式表彰。
- 2020年9月至今：参与社会实践十余项，总计志愿服务时长917小时。
- 2024年9月至今：任中国科学技术大学 6 系 24 级硕士第二党支部书记。