

任新蕾

手机: 18740451049 | 邮箱: 18740451049@163.com | 主页: <https://xinleiren.github.io/blog/>

研究方向

回声消除、单/多通道语音降噪、音质评估、歌唱评价、音频事件检测、音频内容分类、音频编解码、音效等

教育背景

2015.09 ~ 2018.03	西北工业大学	计算机学院	硕士
2011.09 ~ 2015.07	西北工业大学	软件与微电子学院	本科

竞赛经历

- ConferencingSpeech2021: INTERSPEECH2021 举办的多通道语音增强挑战赛 (负责人)
主要职责: 1. 仿真多通道数据; 2. 设计、训练多通道语音增强模型;
竞赛成果: 1. 提出的多通道因果 U-Net 波束形成网络获得**两个赛道的冠军**; 2. 发表学术论文一篇;
- L3DAS21: IEEE MLSP2021 举办的 3D 语音挑战赛 (负责人)
主要职责: 1. 训练、优化基于多通道 U-Net 波束形成网络的 3D 语音增强模型;
竞赛成果: 1. 获得语音增强赛道的**冠军**; 2. 发表学术论文一篇;
- ConferencingSpeech2022: INTERSPEECH2022 举办的语音质量评估挑战赛 (负责人之一)
主要职责: 1. 设计、训练基于 GCNN + BiGRU + Dense 网络结构的非侵入式音质评估模型;
竞赛成果: 1. 获得比赛的**亚军**; 2. 发表学术论文一篇;
- L3DAS22: ICASSP2022 举办的 3D 语音挑战赛 (国内负责人)
主要职责: 1. 和意大利罗马大学一起举办比赛; 2. 搭建语音增强赛道的基线模型;
竞赛成果: 1. 比赛成功举办, 吸引全球近二十支队伍参赛; 2. 发表学术论文一篇;

项目经历

- | 2023.04 ~ 至今 | StarMaker | 音频算法工程师 |
|--|-----------|-------------------|
| ● 气口和重音检测算法 (负责人) | | 2024.03 ~ 至今 |
| 主要职责: 1. 设计、研发基于信号处理 (基频、能量、过零率和谱质心等) 和神经网络的气口检测算法; | | |
| 2. 设计、研发基于信号处理 (基频、beats/downbeats 等) 的重音检测算法; | | |
| 3. 研究开源软件 UVR5, 并抽取其中的去混响、去和声和降噪模块, 以对数据进行批量处理; | | |
| 项目成果: 1. 气口检测精确率 92%, 召回率 93%; 2. 重音检测精确率 85%, 召回率 83%; | | |
| ● 歌曲的主歌部分提取算法 (负责人) | | 2024.06 ~ 2024.06 |
| 主要职责: 1. 设计、研发主歌提取算法, 提取出时长 30 秒左右的主歌, 为 AI 音乐生成提供训练数据; | | |
| 项目成果: 1. 算法准确率达 99%, 已交付 AI 音乐项目组使用; | | |
| ● “嘶~”音检测算法和基于 PYIN 的实时音域检测算法 (负责人) | | 2023.12 ~ 2024.02 |
| 主要职责: 1. 设计、研发基于谱质心特征的“嘶~”音检测算法; | | |
| 2. 研究开源算法 PYIN, 在保证效果的前提下, 降低其计算复杂度, 以满足实时性要求; | | |
| 项目成果: 1. 效果通过认证, 均已上线 EasySing, 用户反馈检测结果准确; | | |
| 2. 总结 YIN 及 PYIN 算法, 并输出文档【 基频提取算法-PYIN 】; | | |

- 多维度歌唱评价 (负责人) 2023.04 ~ 2023.11
主要职责: 1. 调研、设计并实现五个维度 (音准、节奏、技巧、情感和音域) 来评价用户演唱的作品;
项目成果: 1. 评价结果准确率为 90%;
2. 效果通过认证, 已上线 StarMaker, 用户反馈很好, 认可该评价系统;

2020.07 ~ 2023.01 快手 音频算法工程师

- 基于神经网络的实时 K 歌喷麦噪声抑制 (负责人) 2020.07 ~ 2020.10
主要职责: 1. 从 0 到 1 构建喷麦噪声数据集; 2. 设计、训练喷麦噪声抑制模型;
项目成果: 1. 作为业内首发的 K 歌喷麦抑制方案, 效果通过认证, 已上线快手和回森 (K 歌 APP);
- 基于神经网络的实时通用噪声抑制 (负责人之一) 2020.11 ~ 2021.11
主要职责: 1. 把控、提升语音和噪声数据集的质量; 2. 设计、训练基于 CNN + GRU + DENSE 的模型;
项目成果: 1. 效果通过认证, 已上线快手 APP、快影 (剪辑 APP) 和 KIM (快手内部办公 APP);
- 基于 SN-Net 的非实时噪声抑制 (负责人) 2021.12 ~ 2022.04
主要职责: 1. 实现、训练并优化基于时-频 Self-Attention 的 SN-Net 网络;
项目成果: 1. 复现的 SN-Net 的 PESQ 指标比原始论文高 0.2; 2. 总结并输出文档【[SN-Net](#)】;
- 快手 APP 直播音频码率优化 (负责人) 2022.05 ~ 2022.07
主要职责: 1. 设计、训练语音、音乐、噪声三分类模型; 2. 在保证音质不降的前提下, 定制音频码率;
项目成果: 1. 在 QoE 持平的同时, 音频直播码率下降 18.2%, 大幅节省了音频带宽, 已上线快手;
- 基于神经网络的 K 歌假声技巧检测 (负责人) 2022.08 ~ 2022.12
主要职责: 1. 构建假声数据集; 2. 设计、训练假声检测模型;
项目成果: 1. 假声检出率 83%, 准确率 80%; 2. 效果通过认证, 已上线回森;

2018.04 ~ 2020.07 百度 音频算法工程师

- 基于 RNNoise 的非平稳噪声抑制 (负责人之一) 2020.01 ~ 2020.06
主要职责: 1. 采集、清洗噪声数据集; 2. 训练、优化 RNNoise 模型;
项目成果: 1. 抑制常见非平稳噪声 20dB 以上; 2. 效果通过认证, 已上线百度 Hi;
- 百度 Hi 音频 3A 算法优化 (负责人之一) 2018.04 ~ 2019.12
主要职责: 1. 优化延迟估计算法; 2. 优化 AEC 线性滤波和后处理算法; 3. 优化降噪算法;
项目成果: 1. 延迟估计算法检出率提升 20%; 2. AEC 线性部分消除量提升 6dB, 双讲剪切率降低 10%;
3. 噪声抑制量提升 6dB; 4. 效果通过认证, 已上线百度 Hi;
- 基于 G.722.1 音频编解码器的识别率优化 (负责人) 2019.01 ~ 2019.03
主要职责: 1. 优化 G.722.1 音频编解码器, 以提升识别率;
项目成果: 1. 错字率由 4.8% 下降到 4.3%; 2. 总结并输出文档【[音频编解码之 G.722.1](#)】;

发表论文 (共发表论文 8 篇)

-
- [1] Ren, X., Zhang, X., Chen, L., Zheng, X., Zhang, C., Guo, L., & Yu, B. (2021, August). A Causal U-Net Based Neural Beamforming Network for Real-Time Multi-Channel Speech Enhancement. In *Interspeech* (pp. 1832-1836).
- [2] Ren, X., Chen, L., Zheng, X., Xu, C., Zhang, X., Zhang, C., ... & Yu, B. (2021, October). A neural beamforming network for b-format 3d speech enhancement and recognition. In *2021 IEEE 31st International Workshop on Machine Learning for Signal Processing (MLSP)* (pp. 1-6). IEEE.

- [3] Chen, L., **Ren, X.**, Zhang, X., Zheng, X., Zhang, C., Guo, L., & Yu, B. (2022). Impairment Representation Learning for Speech Quality Assessment. *Proc. Interspeech 2022*, 3323-3327.
- [4] Zhang, X., **Ren, X.**, Zheng, X., Chen, L., Zhang, C., Guo, L., & Yu, B. (2021). Low-Delay Speech Enhancement Using Perceptually Motivated Target and Loss. In *Interspeech* (pp. 2826-2830).
- [5] Zheng, X., **Ren, X.**, Chen, L., Xu, C., Wei, Y., Zhang, C., ... & Yu, B. (2023, June). High Quality Speech Restoration from Device Colorated and Reverberated Speech. In *National Conference on Sound and Music Technology* (pp. 73-83). Singapore: Springer Nature Singapore.
- [6] Chen, L., Xu, C., Zhang, X., **Ren, X.**, Zheng, X., Zhang, C., ... & Yu, B. (2022, May). Multi-stage and multi-loss training for fullband non-personalized and personalized speech enhancement. In *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 9296-9300). IEEE.
- [7] Zhang, X., Chen, L., Zheng, X., **Ren, X.**, Zhang, C., Guo, L., & Yu, B. (2022, May). A two-step backward compatible fullband speech enhancement system. In *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 7762-7766). IEEE.
- [8] Guizzo, E., Marinoni, C., Pennese, M., **Ren, X.**, Zheng, X., Zhang, C., ... & Communiello, D. (2022, May). L3DAS22 challenge: Learning 3D audio sources in a real office environment. In *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 9186-9190). IEEE.

授权专利 (共授权专利 13 篇)

- [1] 陈联武,张晨,郑羲光,任新蕾,张旭. 语音增强模型的训练方法和装置以及语音增强方法和装置: CN114121029B. 2024-12-10
- [2] 任新蕾,郑羲光,张晨. 模型训练方法、音频处理方法、装置、设备、介质及产品: CN114242110B. 2024-11-26
- [3] 郑羲光,李楠,任新蕾,张晨.语音处理模型训练方法及装置和语音处理方法及装置: CN112309426B. 2024-07-12
- [4] 陈联武,张晨,张旭,郑羲光,任新蕾. 语音增强模型的训练方法和设备及语音增强方法和设备: CN113593594B. 2024-03-08
- [5] 陈联武,张晨,张旭,郑羲光,任新蕾. 语音增强模型的训练方法及装置、语音增强方法及装置: CN113555031B. 2024-02-23
- [6] 任新蕾,郑羲光,陈联武,张晨. 音频处理方法、装置、电子设备及存储介质: CN115334349B. 2024-01-02
- [7] 任新蕾,郑羲光,李楠,张晨. 语音处理模型的训练方法和装置以及语音处理方法和装置: CN113035221B. 2023-12-19
- [8] 张旭,郑羲光,张晨,任新蕾. 音频处理模型的训练方法及装置、音频处理方法及装置: CN113314147B. 2023-07-25
- [9] 于利标,任新蕾,欧阳伟艳,张洪彬,李慧凤. 声音信号增强方法、装置、电子设备及存储介质: CN110609671B. 2023-07-14
- [10] 任新蕾,郑羲光,李楠,张晨. 语音处理模型的训练方法和装置以及语音处理方法和装置: CN112712816B. 2023-06-20
- [11] 崔凡,任新蕾,郑羲光,李楠,张晨. 声码器的训练方法和装置及合成音频信号的方法和声码器: CN113436603B. 2023-05-02
- [12] 任新蕾,于利标,李慧凤. 通话信号的处理方法、装置、电子设备及存储介质: CN110971769B. 2022-05-03
- [13] 任新蕾,张旭,郑羲光,陈联武,张晨. 语音增强模型的训练方法及装置、语音增强方法及装置: CN113241088B. 2021-10-22

学术活动

1. 受邀在 2021 年首届 SpeechHome 语音技术研讨会上作报告, 报告题目为《用于实时多通道语音增强的因果 U-Net 神经波束形成网络》