

钟立帆

E-mail: zhong_lifan@yahoo.co.jp

电话: +86-17302200223 / +81-090-5679-8104

LinkedIn: www.linkedin.com/in/lifanzhong/

微信: 17302200223

教育经历

❖ 东京大学 工学硕士

日本 东京都

工学系研究科 电子工程与信息系统

2020 年 4 月 - 2023 年 3 月 (预计)

指导老师: Prof. Nobuaki Minematsu

主要课程: 多媒体处理中的机器学习, 声学语音学, 语言科学, 认知多媒体论

❖ 天津大学 工学学士

天津市

精密仪器与光电子工程学院 电子科学与技术 (光电子)

2015 年 9 月 - 2019 年 7 月

实习经历

❖ 平安科技 算法工程师 实习生

深圳市

分析建模算法团队 机器学习前沿算法组

2021 年 7 月 - 2021 年 8 月

- ▶ 基于Transformer的中文普通话语音合成系统 (FastSpeech2)。

❖ 国立情报研究所 大学院学生研究员

日本 东京都

内容科学研究系 Yamagishi Lab

2022 年 9 月 - 2022 年 12 月 (预计)

- ▶ 说话人识别技术启发的乐器声音编码器在多声部音乐乐器识别中的应用。

研究经历

❖ 基于迁移学习的多声部音乐中的主要乐器识别 东京大学工学系研究科 峯松·齋藤研究室 硕士毕业研究

- 复现并改进先行研究中的主要乐器识别系统, 用于识别多声部音乐中主要乐器的种类和数量。
- 使用ImageNet预训练的图像识别模型ResNet50对音频的梅尔时频谱进行识别, 采用了多种方法使梅尔谱适配图像识别模型, 对比基线系统实现了7.9%的精度提升。
- 目前发表成果
 - ▶ L. Zhong, D. Saito and N. Minematsu, "Fine-tuning ImageNet Pre-Trained Models to be Transferred for Predominant Instrument Recognition in Polyphonic Music." 日本音响学会2022秋季研究发表会

项目经历

❖ 专业歌唱中的声乐技巧识别

- 通过MFCC对包含使用气声, Belt, 普通歌唱声, 和颤音唱法的歌声数据进行识别。

❖ 歌唱者分割聚类

- 制作并标注了一个迷你数据集, 测试用说话人语音数据集预训练的TDNN说话人信息提取模型是否适用于动态范围更大的歌声的识别/分割聚类, 并分析不同元音发声对分割聚类效果的影响。

工作技能

◆ 语言

- 编程语言: Python, Bash
- 自然语言: 英语 (流利), 日语 (日常对话)

◆ 框架: PyTorch, Librosa, Scikit-learn, Tensorflow, Matplotlib, Pandas 等

◆ 工具: FFmpeg, Spleeter, SpeechBrain, PyTorch-Lightning 等

◆ 其他: Linux, Latex