

声音自主构建-合成音库

声音构建介绍

基于科大讯飞最新语音合成深度学习技术，只需提供少量的干净录音数据，机器就可以快速学习并生成可使用的语音合成音库进行快速应用。

(注：该定制音库只为该定制客户提供使用。)

录音说明

需要准备一段5-60分钟的录音（建议不低于10分钟），内容及格式要求如下：

- 1、录音为单个人的高质量、无噪音、无背景音的人声语音数据，多人需分开独立训练；
- 2、录音需要为普通话，尽量不包含方言/英文/字母；
- 3、录音格式支持MP3、WAV格式的文件，并提供自动转码；
- 4、请在音频文件开头朗读声音ID注册文案，用于认证你的声音ID，若音频中缺失这段文案，则会训练失败；
- 5、录音尽量自然流畅，按照日常应用场景自然录制，且对录音效果满意；
- 6、发音尽量清晰准确，无错读、无口误、无杂音，给到训练的录音无过多异常；

声音ID注册文案：

本次录音是我公司参与科大讯飞人工智能语音音库训练模型的录音数据采集。

录音设备建议

- 1、有条件下建议进入录音棚进行录制，保证录音质量；
- 2、如采用手机或其他录音设备建议采用话筒录制，录制过程中话筒距离嘴部10cm左右；

Q&A

Q: 定制发音人的版权属于谁?

A: 定制发音人只有在申请商用通过后, 发音人的版权才属于音频提供方

Q: 录音如何确定环境音影响?

A: 尽量选择中等或者较小面积的会议室/房间/车内, 关闭空调以及电器进行录制; 如不确定, 可以先录制1min左右的音频跟我们确认环境是否合适;

Q: 手机需要注意什么?

A: 尽量选择苹果手机, 下载ShurePlus MOTIV, 选择16k/48khz进行录制收音效果更好。手机自带录音压缩较大影响质量;

Q: 录制内容文本应该怎么选?

A: 建议您按照使用场景的文本来录制, 尽量还原真人在使用场景下满意的音频效果。比如您用在儿童故事场景, 尽量选用儿童故事来作为文本来录制, 且对录制音频满意。

Q: 训练的等待时间?

A: 提交音频后, 预计耗时3-5小时。

Q: 申请商用的条件?

A: 选择该发音人对应的平台应用以及上传版权声明或联系我们处理;

感谢观看



讯飞智作