

语音识别ASR

产品名称	语音识别ASR
公司名称	昆山昱唯网络科技有限公司
价格	3.20/小时
规格参数	品牌:腾讯云
公司地址	花桥国际商务城曹新路70号
联系电话	17601404160

产品详情

腾讯云语音识别（Automatic Speech Recognition，ASR）为开发者提供语音转文字服务的最佳体验。语音识别服务经微信、腾讯视频、王者荣耀等大量内部业务验证，同时也在线上大量外部客户业务场景下成功落地，具备识别准确率高、接入便捷、性能稳定等特点。腾讯云语音识别服务开放实时语音识别、一句话识别和录音文件识别三种服务形式，满足不同类型开发者需求。

经公司内部微信、QQ、腾讯视频、王者荣耀等大体量业务充分验证，也在大量互联网、金融、教育等领域的外部客户业务场景成功落地，日服务亿级用户。具有海量数据支撑、算法业界领先、支持语种丰富、服务性能稳定、抗噪音能力强、识别准确率高等优势。

功能

语音识别开放实时语音识别、一句话识别和录音文件识别三种服务形式，满足不同类型开发者需求。

实时语音识别

对实时音频流进行识别，实时识别为文字。

适用于有一定实时性要求的场景，例如语音输入、语音机器人、会议现场记录、直播内容审核、视频实时添加字幕等场景。

语言和方言：目前支持中文普通话、英语、粤语和带有一定方言口音的中文普通话，后续将会支持维语。

音频属性：支持 wav、pcm、speex、silk 的音频格式，支持8k、16k采样率的单声道音频流，支持16bit的数据采样精度。

音频数据长度：建议每个数据包的音频分片最大不能超过200KB。

一句话识别

对60秒之内的短音频文件进行识别。

适用于语音消息转写场景，例如语音短信、语音搜索等。

音频属性：支持 wav、mp3 的音频格式，支持8k、16k采样率的单声道音频，支持16bit的数据采样精度。

音频数据长度：若采用直接上传音频数据方式，建议整个数据包不能大于600KB；若采用上传音频url方式，建议音频时长不能大于60s。

录音文件识别

对不超过一小时的录音文件进行识别。

适用于语音时间较长、对实时性要求低的场景，例如客服质检、视频字幕生成、音频节目字幕生成等。

语言和方言：目前支持中文普通话和带有一定方言口音的中文普通话，后续将会支持英语、粤语。

音频属性：支持 wav、pcm、mp3、silk、speex、amr、m4a 等主流的音频格式，支持8k、16k采样率的单声道或双声道音频识别，支持16bit的数据采样精度。

音频数据长度支持：若采用直接上传音频数据方式，建议音频数据不能大于5MB；若采用上传音频 url 方式，建议音频时长不能大于1小时。