

# 标贝科技语音识别服务

## Android SDK 使用说明文档（1.0）

Name	Date	Reason For Changes	Version
	2020.02.25	创建文档，编写使用说明。	1.0
	2020.08.05	增加文件识别	1.1

标贝（北京）科技有限公司 DataBaker(Beijing)technology co.,LTD

北京市海淀区西小口路 66 号中关村东升科技园 B-2 号楼 A203 室，010-58465943

## 目录

1. Android Studio 集成 aar（参考 demo） .....	3
2. SDK 关键类.....	3
3. 调用说明.....	4
4. 参数说明.....	5
4.1 基本参数说明.....	5
4.2 BakerRecognizerCallback 回调类方法说明.....	6
4.3 失败时返回的 code 对应表.....	6

# 1.Android Studio 集成 aar（参考 demo）

**1.1 集成 aar 包有多种方式，本文提供一种作为参考，采用其他正确集成方式都可行。将 aar 包拷贝至 libs 目录下，并执行下一步（1.2）。注意新增\*.aar 格式。**

**1.2 在主 module 的 build.gradle 文件里，添加以下代码。**

```
dependencies {  
    implementation fileTree(dir: 'libs', include: ['*.jar', '*.aar'])  
    implementation 'com.squareup.okhttp3:okhttp:4.2.2'  
    implementation 'com.google.code.gson:gson:2.8.6'  
}
```

**1.3 在主 Module 的 AndroidManifest.xml 文件中添加网络权限。**

```
<uses-permission android:name="android.permission.INTERNET" />  
<uses-permission android:name="android.permission.ACCESS_NETWORK_STATE" />  
<uses-permission android:name="android.permission.READ_PHONE_STATE" />  
<uses-permission android:name="android.permission.RECORD_AUDIO" />  
<uses-permission android:name="android.permission.READ_EXTERNAL_STORAGE" />  
<uses-permission android:name="android.permission.WRITE_EXTERNAL_STORAGE" />
```

**1.4 在主 Module 的 AndroidManifest.xml 文件中的 application 节点添加以下属性。**

```
android:usesCleartextTraffic="true"
```

Eclipse 环境也遵循相关集成 aar 包的方式即可。

## 2. SDK 关键类

1. BakerRecognizer: 语音识别关键业务处理类，全局只需一个实例即可。

2. BakerRecognizerCallback: 实时识别结果源数据回调类。您可以在这些个回调方法中获取到当前识别的各类结果和状态，或者错误信息。您可以在回调方法中处理识别结果。设置参数时请将此 callback 提交给 BakerRecognizer 实例。

## 3. 调用说明

### 1.1 初始化 BakerRecognizer 类，得到 BakerRecognizer 的实例。

如果是调用我们公有云公共识别接口，请采用以下示例代码初始化 BakerRecognizer 类，传入由我们授权邮件里面给到的账户相关的 `clientId` 和 `clientSecret`。

```
BakerRecognizer bakerRecognizer = BakerRecognizer.getInstance  
(VoiceSpeechActivity.this, clientId, clientSecret);
```

如果是私有化部署请采用以下示例代码初始化 BakerRecognizer 类。并注意需要在参数中设置私有化部署的识别服务器地址，仅私有化部署时需要设置该参数。

```
BakerRecognizer bakerRecognizer = BakerRecognizer.getInstance  
(VoiceSpeechActivity.this);  
bakerRecognizer.setUrl("ws://xxxx"); //设置私有化部署的识别服务器地址
```

### 1.2 设置 BakerRecognizer 识别所需参数，包括必填参数和非必填参数。请参考 demo 和注释信息。

1.3 调用 BakerRecognizer.startRecognize()方法开启语音识别，该方法返回 int 值，返回 0 代表启动成功。识别结果反馈在 callback 类的各个回调方法中，各回调方法的介绍请参考以下 BakerRecognizerCallback 回调类方法说明。正常识别一般无需调用 BakerRecognizer.stopRecognition()方法结束识别，SDK 中的 VAD 会自动检测，当符合条件时会自动结束识别，比如声音停顿时间超过设置的语音后端点静音检测时间。

文件识别功能：BakerRecognizer.startRecognize(String filePath)在此方法中的参数传入文件路径即可，支持原始 PCM 文件。

### 1.4 如果需要发起新的识别请求，可以重复第 1.2-1.3 步。

## 4. 参数说明

### 4.1 基本参数说明

参数	参数名称	是否必填	说明
setCallback	设置识别回调	是	设置识别回调类
setUrl	识别服务器地址	否	仅私有化部署，需要设置服务器 URL。
setVadSos	设置语音前端点	否	静音超时时间，即用户多长时间不说话则当做超时处理，参数请勿随意设置，因为此处设置的参数是按照帧计算的。
setVadEos	设置语音后端点	否	后端点静音检测时间，即用户停止说话多长时间即认为不再输入，自动停止录音，参数请勿随意设置，因为此处设置的参数是按照帧计算的。
setVadWait	设置语音最大识别时长	否	设置语音最大识别时长，最长 60s=1800，参数请勿随意设置，因为此处设置的参数是按照帧计算的。
setVadPause	设置语句间停顿间隔	否	设置语句间停顿间隔，默认 2s=45，参数请勿随意设置，因为此处设置的参数是按照帧计算的。
setSample	设置识别声音的采样率	否	设置采样率 目前支持 16000 默认 16000
addPct	是否添加标点	否	是否添加标点 true=返回标点，false=不返回标点，默认 false
enableItn	是否执行归一化处理	否	是否执行归一化处理,true=需要处理，false=不需要处理，默认=false。
setRecognizeType	识别类型	否	识别模式， 0= 一句话识别，sdk 端做 vad 1=长语音识别，服务端做 vad，默认为 0。

## 4.2 BakerRecognizerCallback 回调类方法说明

参数	参数名称	说明
onReadyOfSpeech	准备就绪	准备就绪回调
onBeginOfSpeech	识别开始	当识别到有声音时，会唤起此回调方法。即 mic 输出声音流时。
onResult	流式识别持续返回数据的接口回调	此回调方法会被多次唤起。持续返回实时的识别数据信息。List<String> nbest 参数是一个识别结果集合，最优识别结果理论上是 nbest(0)，List<String> uncertain 参数是一个预测的可能出现的语音识别内容的结果集，供选择使用。boolean isLast 当前结果是否是最终识别数据，false: 否，true: 是。
onEndOfSpeech	识别正常结束。	正常识别结束时，会唤起此回调方法。
onVolumeChanged	识别音量信息	float volume 是取值 0-1 之间的浮点数，代表声音的量能。byte[] data 是声音的二进制信息，也可以自己根据源数据流计算处理声音量能信息。
onError	识别出现错误	int code 是错误码，错误码具体请参考以下错误码对照表。String message 是具体的错误信息日志。识别过程中出现各种错误情况都会回调此方法。

## 4.3 失败时返回的 code 对应表

错误码	含义
90000	接口正常，正确返回的识别码
91001	录音被电话打断
91002	没有录音权限
91003	录音出现未知错误
91004	本地网络不可用
91005	webSocket failure
91006	解析服务器返回信息出错
91007	本地网络未知错误
91008	授权信息为空
91009	token 初始化失败
92010	vad 设置参数错误
92011	vad 发送数据错误
92012	vad 状态错误
92015	超过 vad 识别最长时限
92016	vad 未知错误
92017	vad 数据返回错误
92018	buffer_closed

92019	没有说话
92020	声音太短
30001	HTTP 请求参数错误
30002	服务内部错误
30003	识别结果解析出错
30004	应用包名未知
30005	语音质量问题
30006	输入语音过长，服务器检测到识别超过 60s
30007	连接识别引擎失败
30008	会话 id 不存在
30009	Rpc 调用非法
30010	redis rpop 操作返回空
30011	redis rpop 值不合法
30012	rpc 调用识别引擎失败
30013	Redis rpop 操作失败
30014	redis lpush 操作失败
30015	单个语音分片过长
30016	回调 url 失败
40001	json 解析失败
40002	json 字段不全
40003	版本错误
40004	json 字段值类型错误
40005	参数错误
40006	idx 超时
40007	idx 顺序错误
40008	token 校验失败
40009	token 处于未激活状态
40010	token 已过期
40011	使用量已超过购买量
40012	qps 错误
50001	处理超时
50002	内部 rpc 调用失败
50003	服务端繁忙
50004	其他内部错误