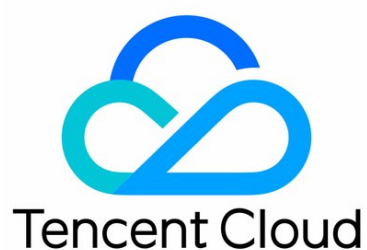


语音识别 快速入门



Copyright Notice

©2013–2025 Tencent Cloud. All rights reserved.

The complete copyright of this document, including all text, data, images, and other content, is solely and exclusively owned by Tencent Cloud Computing (Beijing) Co., Ltd. ("Tencent Cloud"); Without prior explicit written permission from Tencent Cloud, no entity shall reproduce, modify, use, plagiarize, or disseminate the entire or partial content of this document in any form. Such actions constitute an infringement of Tencent Cloud's copyright, and Tencent Cloud will take legal measures to pursue liability under the applicable laws.

Trademark Notice

 Tencent Cloud

This trademark and its related service trademarks are owned by Tencent Cloud Computing (Beijing) Co., Ltd. and its affiliated companies ("Tencent Cloud"). The trademarks of third parties mentioned in this document are the property of their respective owners under the applicable laws. Without the written permission of Tencent Cloud and the relevant trademark rights owners, no entity shall use, reproduce, modify, disseminate, or copy the trademarks as mentioned above in any way. Any such actions will constitute an infringement of Tencent Cloud's and the relevant owners' trademark rights, and Tencent Cloud will take legal measures to pursue liability under the applicable laws.

Service Notice

This document provides an overview of the as-is details of Tencent Cloud's products and services in their entirety or part. The descriptions of certain products and services may be subject to adjustments from time to time.

The commercial contract concluded by you and Tencent Cloud will provide the specific types of Tencent Cloud products and services you purchase and the service standards. Unless otherwise agreed upon by both parties, Tencent Cloud does not make any explicit or implied commitments or warranties regarding the content of this document.

Contact Us

We are committed to providing personalized pre-sales consultation and technical after-sale support. Don't hesitate to contact us at 4009100100 or 95716 for any inquiries or concerns.

Contents

快速入门

操作指引

一分钟接入服务端 API

一分钟跑通集成 SDK

可视化调用语音识别服务

新手常见问题

快速入门

操作指引

Last updated: 2023-12-22 18:00:32

登录控制台

注册并通过实名认证后，登录腾讯云控制台。如果没有账号，请参考 [账号注册教程](#)。

开通服务

单击进入 [语音识别控制台](#)，即可开通相应服务。

免费额度和购买

开通了语音识别某项服务，该项服务即可享受对应额度的免费调用额度，以资源包的形式发放到您的腾讯云账号中，并在计费结算时优先扣减。免费额度耗尽后，语音识别提供预付费和后付费两种计费模式，您可以查看语音识别的 [计费概述](#)。

使用语音识别服务

您可以根据以下四种方式的介绍选择合适的使用方式：

使用方式	适用对象	说明	相关文档
通过 语音识别控制台 进行功能体验	非业务开发者，没有代码编写基础	此方式只能用于体验，不能用于开发	功能体验操作步骤
通过 API 3.0 Explorer 在线调用功能	开发初学者，有代码编写基础	此方式能够实现在线调用、签名验证、SDK 代码生成和快速检索接口等能力	语音识别 API 概览
通过编写代码 调用 API 进行开发	开发工程师，熟悉代码编写	腾讯云已编写好的开发工具集（SDK），支持通过调用语音识别服务 API 开发功能。目前 SDK 已支持多种语言，包括 Python、Java、PHP、Go、Node.js、.Net 等，可在每个服务的文档中下载对应的 SDK	一分钟接入服务端 API
通过 客户端 SDK 进行集成	客户端开发工程师，熟悉代码编写	腾讯云支持通过已编写好的开发组件（SDK）在客户端集成语音识别的能力。目前客户端 SDK 主要支持的 Android、iOS 平台，您可以在 客户端 SDK 中心 获取对应的 SDK 包	客户端与服务端 SDK 概览

说明：
若您需调用 API 实现功能开发，请提前对 HTTP 请求和 API 调用进行了解。

查看调用量

登录腾讯云 [语音识别控制台](#)，查看语音识别各服务的使用情况。

一分钟接入服务端 API

Last updated: 2023-12-08 15:49:59

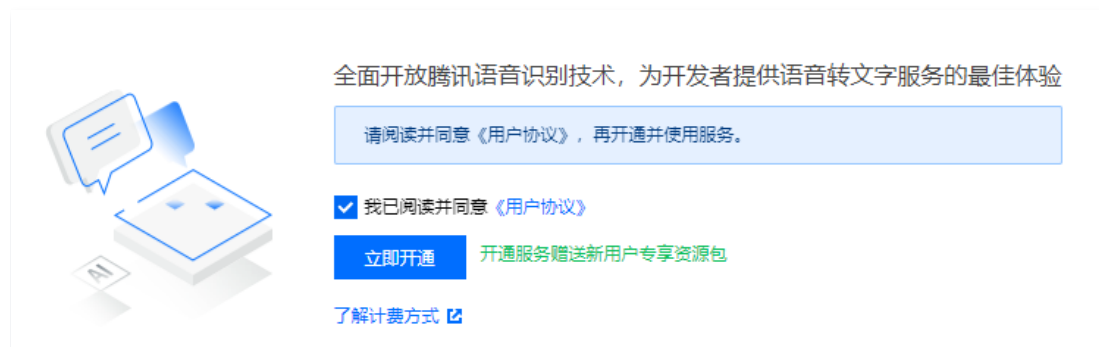
操作场景

本文将为您介绍如何使用 API 3.0 Explorer 在线调试语音识别相关接口，并快速将该接口对应的腾讯云开发者工具套件（SDK）集成到本地项目中。

操作步骤

开通语音识别服务

在调用语音识别相关接口前，您需要进入 [语音识别控制台](#)，进行实名认证和人脸认证，认证完成后，阅读《用户协议》后勾选“我已阅读并同意《用户协议》”，然后单击【立即开通】，即可一键开通[录音文件识别](#)、[实时语音识别](#)、[一句话识别](#)、[录音文件识别极速版](#)、[语音流异步识别服务接口](#)，如需开通营业执照核验或增值税发票核验功能，可前往官网页服务介绍页申请开通，审核通过后即可使用该服务。



服务开通成功后，您将获得各项服务对应的免费调用额度，可在 [资源包管理页](#) 查看。同时您也可以[在语音识别购买页](#)中购买对应语音识别服务的资源包，若免费额度以及资源包调用次数耗尽，接口计费将自动转为后付费方式按月/日进行结算，具体计费标准可以查看语音识别的 [计费概述](#)。

调试语音识别接口

语音识别服务开通成功后，进入语音识别 [API 3.0 Explorer](#) 在线接口调试页面，选择需要调用的接口，并填写输入参数。输入参数在 API 3.0 Explorer 界面的“参数说明”选项卡中可以查看对应接口输入参数的具体含义。

说明

平台将对登录用户提供临时 Access Key，以便进行调试。

语音流异步识别相关接口

语音流异步识别任务创建

语音流异步识别任务列表

语音流异步识别任务关闭

录音文件识别相关接口

一句话识别相关接口

自学习相关接口

热词相关接口

CreateAsyncRecogni... 查看API文档 2019-06-14

个人密钥

平台将对登录用户提供临时Access Key，以便进行测试

更多选项

输入参数

只看必填参数

Region (地域参数，建议您阅读文档了解地域以及计费情况)

本接口不需要传递该参数

EngineType

string

Url

string

CallbackUrl

string

SignToken (选填)

string

FilterDirty (选填)

integer

FilterModal (选填)

integer

FilterPunc (选填)

代码生成 在线调用 签名串生成 参数说明 问题反馈 查看文档

Engine Type

必填：是。

类型：String。

描述：

引擎模型类型。• 16k_zh: 16k 中文普通话通用；• 16k_zh_video: 16k 音视频领域；• 16k_en: 16k 英语；• 16k_ca: 16k 粤语；

Url

必填：是。

类型：String。

描述：

语音流地址，支持rtmp、hls、rtsp等流媒体协议，以及各类基于http协议的直播流

CallbackUrl

必填：是。

类型：String。

描述：

支持HTTP和HTTPS协议，用于接收识别结果，您需要自行搭建公网可调用的服务。回调格式&内容详见：[语音流异步识别回调说明](#)

填写输入参数后，选择“代码生成”选项卡，可以看到自动生成的不同编程语言代码（可支持 Java、Python、Node.js、PHP、GO、.NET、C++ 语言），生成代码中的部分字段信息和填写内容是关联的，如需调整传入参数，可在左侧修改参数值后重新生成代码。

CreateAsyncRecogni... 查看API文档 2019-06-14

代码生成

在线调用

签名串生成

参数说明

问题反馈

查看文档

个人密钥

⚠ 平台将对登录用户提供临时Access Key，以便进行测试

更多选项 ▾

输入参数 ☐ 只看必填参数

Region (地域参数，建议您阅读文档了解地域以及计费情况)
本接口不需要传递该参数 ▾

EngineType ?
string

Url ?
string

CallbackUrl ?
string

SignToken ? (选填)
string

FilterDirty ? (选填)
integer

FilterModal ? (选填)
integer

tool cli asr CreateAsyncRecognitionTask --cli-unfold-argument

☒ Java

☐ Python

☐ Node.js

☐ PHP

☐ GO

☐ .NET

☐ C++

JAVA SDK使用说明

```
import com.tencentcloudapi.common.Credential;
import com.tencentcloudapi.common.profile.ClientProfile;
import com.tencentcloudapi.common.profile.HttpProfile;
import com.tencentcloudapi.common.exception.TencentCloudSDKException;

import com.tencentcloudapi.asr.v20190614.AsrClient;
import com.tencentcloudapi.asr.v20190614.models.*;

public class CreateAsyncRecognitionTask
{
    public static void main(String [] args) {
        try{

            Credential cred = new Credential("SecretId", "SecretKey");

            HttpProfile httpProfile = new HttpProfile();
            httpProfile.setEndpoint("asr.tencentcloudapi.com");

            ClientProfile clientProfile = new ClientProfile();
            clientProfile.setHttpProfile(httpProfile);

            AsrClient client = new AsrClient(cred, "", clientProfile);

            CreateAsyncRecognitionTaskRequest req = new CreateAsyncRecognitionTaskRequest();

            CreateAsyncRecognitionTaskResponse resp = client.CreateAsyncRecognitionTask(req);

            System.out.println(CreateAsyncRecognitionTaskResponse.toJsonString(resp));
        } catch (TencentCloudSDKException e) {
            System.out.println(e.toString());
        }
    }
}
```

选择“在线调用”选项卡，单击【发送请求】可进行真实请求，供您调试、参考。

代码生成

在线调用

签名串生成

参数说明

问题反馈

查看文档

注意：通过API发送请求等同于真实操作，请小心进行

在线调用

点击下面的“发送请求”按钮，系统会以POST的请求方法发送您在左侧填写的参数到对应的接口，该操作等同于真实操作，建议您仔细阅读产品计费文档了解费用详情，同时系统会给您展示请求之后的结果、响应头等相关信息，供您调试、参考。

发送请求

集成语音识别 SDK

确认本地依赖环境满足以下条件：

编程环境	SDK 集成要求
Node.js	需要 NODEJS 10.0.0 版本及以上

Python	需要 Python 2.7、3.6至3.9版本
Java	需要 JDK 7 版本及以上
Go	需要 Go 1.9 版本及以上（如使用 go mod 需要 Go 1.14）
.Net	需要 .NET Framework 4.5+ 或者 .NET Core 2.1
PHP	需要 PHP 5.6.0 版本及以上
C++	需要 C++ 11 或更高版本的编译器 GCC 4.8 或以上版本。暂时仅支持 Linux 环境，不支持 Windows 环境
Ruby	需要 Ruby 2.3 及以上版本

安装与本地依赖环境对应的腾讯云语音识别 SDK，下面将以 Node.js 为例说明 SDK 安装、使用的方法，其他语言的 SDK 使用方法请参考 [腾讯云 SDK 使用手册](#)。

通过 npm 安装（推荐）

通过 npm 获取安装是使用 NODEJS SDK 的推荐方法，npm 是 NODEJS 的包管理工具。关于 npm 详细可参考 [npm 官网](#)。

1. 执行以下安装命令：

```
npm install tencentcloud-sdk-nodejs --save
```

2. 在您的代码中引用对应模块代码，可参考示例。
3. 如上引用方式会将腾讯云所有产品 sdk 下载到本地，可以将 `tencentcloud-sdk-nodejs` 换成 `tencentcloud-sdk-nodejs-cvm/cbs/vpc` 等，即可引用特定产品的 sdk，代码中可将 `require("tencentcloud-sdk-nodejs")` 改为 `require("tencentcloud-sdk-nodejs-cvm/cbs/vpc")`，其余不变，可参考示例，可大大节省存储空间。

通过源码包安装

- 前往 [GitHub 代码托管地址](#) 或者 [快速下载地址](#)，下载源码压缩包。
- 解压源码包到您项目合适的位置。
- 在您的代码中引用对应模块代码，可参考示例。

示例

SDK 安装完成后，可在您的项目代码中引用 API 3.0 Explorer 自动生成的代码，以 Node.js 为例，简易 demo 示例如下：

```
const tencentcloud = require("tencentcloud-sdk-nodejs")

// 导入对应产品模块的 client models。
const CvmClient = tencentcloud.cvm.v20170312.Client

const clientConfig = {
  // 腾讯云认证信息
  credential: {
    secretId: "secretId",
    secretKey: "secretKey",
  },
  // 产品地域
  region: "ap-shanghai",
  // 可选配置实例
  profile: {
    signMethod: "HmacSHA256", // 签名方法
    httpProfile: {
      reqMethod: "POST", // 请求方法
      reqTimeout: 30, // 请求超时时间，默认60s
    },
  },
}
```

```
}  
// 实例化要请求产品(以 cvm 为例)的 client 对象  
const client = new CvmClient(clientConfig)  
// 通过 client 对象调用想要访问的接口, 需要传入请求对象以及响应回调函数  
client.DescribeZones().then(  
  (data) => {  
    console.log(data)  
  },  
  (err) => {  
    console.error("error", err)  
  }  
)
```

在支持 typescript 项目中, 采用如下方式调用:

```
import * as tencentcloud from "tencentcloud-sdk-nodejs"  
  
// 导入对应产品模块的 client models。  
const CvmClient = tencentcloud.cvm.v20170312.Client  
  
const clientConfig = {  
  // 腾讯云认证信息  
  credential: {  
    secretId: "secretId",  
    secretKey: "secretKey",  
  },  
  // 产品地域  
  region: "ap-shanghai",  
  // 可选配置实例  
  profile: {  
    signMethod: "HmacSHA256", // 签名方法  
    httpProfile: {  
      reqMethod: "POST", // 请求方法  
      reqTimeout: 30, // 请求超时时间, 默认60s  
    },  
  },  
}  
// 实例化要请求产品(以 cvm 为例)的 client 对象  
const client = new CvmClient(clientConfig)  
// 通过 client 对象调用想要访问的接口, 需要传入请求对象以及响应回调函数  
client.DescribeZones().then(  
  (data) => {  
    console.log(data)  
  },  
  (err) => {  
    console.error("error", err)  
  }  
)
```

实例化 Client 的入参支持 clientConfig 数据结构和说明, 详情请参见 [ClientConfig](#)。

一分钟跑通集成 SDK

Last updated: 2024-12-23 11:36:44

- iOS
 - [实时语音识别](#)
 - [一句话识别](#)
 - [录音文件识别极速版](#)
- Android
 - [实时语音识别](#)
 - [一句话识别](#)
 - [录音文件识别极速版](#)
- FLUTTER
 - [实时语音识别](#)
 - [一句话识别](#)
 - [录音文件识别极速版](#)

可视化调用语音识别服务

Last updated: 2025-01-20 16:43:52

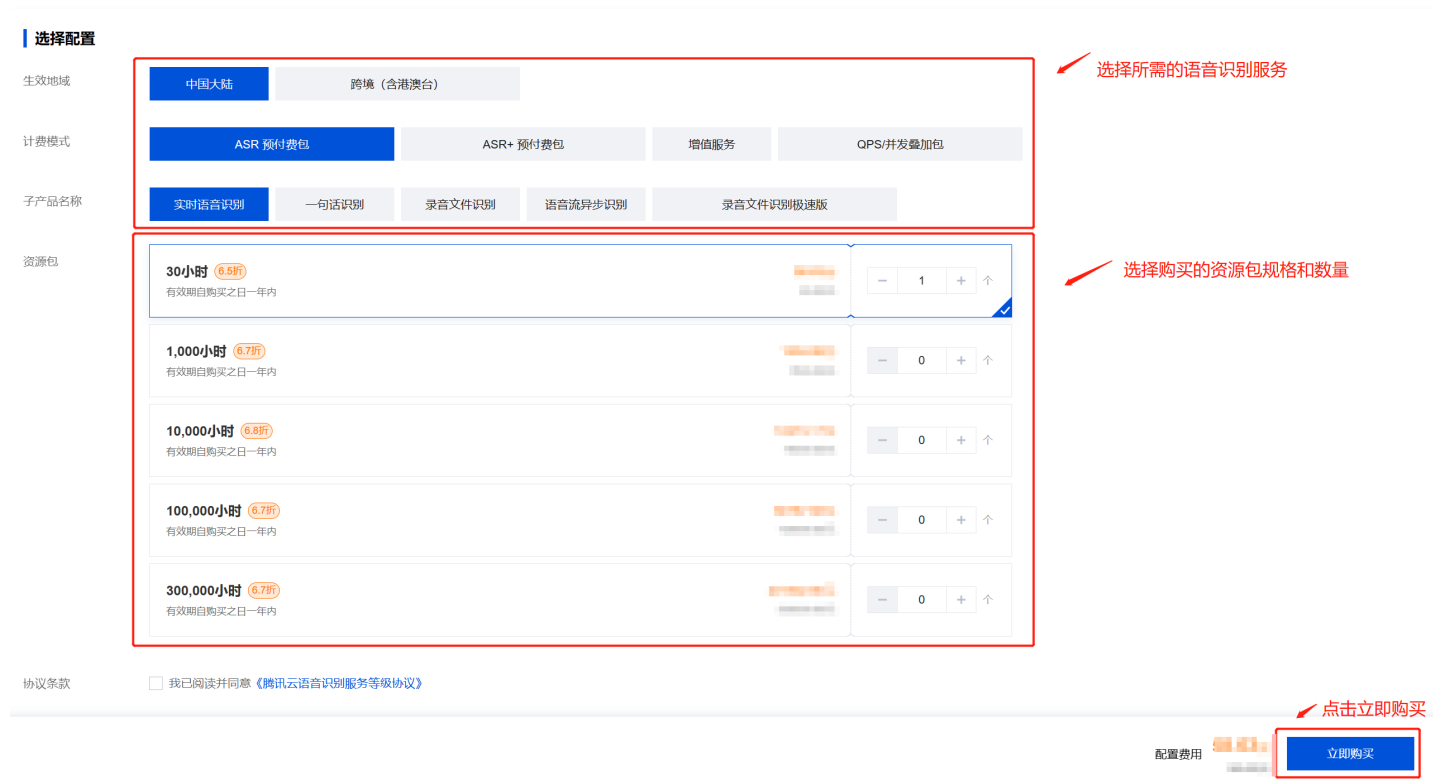
本文将为您介绍如何可视化调用语音识别服务，包括两种方式：[API 3.0 Explorer](#) 以及 Postman，建议您使用 API 3.0 Explorer 方式。具体步骤如下。

开通语音识别服务

在调用语音识别相关接口前，您需要进入 [语音识别控制台](#)，阅读《[语音识别服务条款](#)》后勾选同意并单击**立即开通**，即可一键开通语音识别、热词、自学习模型等服务接 。



服务开通成功后，您将获得各项服务对应的免费调用额度，可在 [资源包管理页](#) 查看。同时您也可以 在 [语音识别购买页](#) 中购买对应语音识别服务的资源包，若免费额度以及资源包调用次数耗尽，接口计费将自动转为后付费方式按月进行结算，具体计费标准可以查看语音识别的 [购买指南](#)。



说明

语音识别服务开通成功后，您可根据喜好选择如下两种方式来调用语音识别服务，建议选择 API 3.0 Explorer。

使用 API 3.0 Explorer 调用语音识别接口

输入参数

进入语音识别 [API 3.0 Explorer](#) 在线接口调试页面，选择需要调用的接口，并填写输入参数。

API Explorer 语音识别 (ASR)

搜索接口，支持中英文搜索

SentenceRecognition
asr 2019-06-14 [查看API文档](#) [点赞](#) [吐槽](#)

以确保是您本人操作；该操作等同于真实操作，建议您仔细阅读相关产品文档了解费用等详情，谨慎操作！

更多选项 ▾

输入参数

Region ①
本接口不需要传递该参数

参数输入方式
 [参数推荐](#)

EngServiceType [\[*\]](#) ①
string

SourceType [\[*\]](#) ①
integer

VoiceFormat [\[*\]](#) ①
string

Uri (选填) [\[*\]](#) ①
string

Data (选填) [\[*\]](#) ①
string

DataLen (选填) [\[*\]](#) ①
integer

WordInfo (选填) [\[*\]](#) ①
integer

FilterDirty (选填) [\[*\]](#) ①
integer

FilterModal (选填) [\[*\]](#) ①
integer

FilterPunc (选填) [\[*\]](#) ①
integer

ConvertNumMode (选填) [\[*\]](#) ①
integer

☐ 展示英文接口 [吐槽](#)

[发起调用](#) [调用历史](#) [展示所有参数 ▾](#)

在 API 3.0 Explorer 界面的文档说明选项卡中可以查看对应接口输入参数的具体含义。

CreateAsyncRecognitionTask
asr 2019-06-14 [查看API文档](#)

点赞 吐槽

在线调用 代码示例 CLI示例 签名示例 **文档说明** 数据模拟 问题反馈

ⓘ

- 在线调用模块中当您发起请求时，平台通过已登录用户信息获取当前账号临时Access Keys，对当前账号发起操作。
- 发起请求为敏感操作，在您进行敏感操作前，需要先完成身份验证以确保是您本人操作；该操作等同于真实操作，建议您仔细阅读相关产品文档了解费用等详情，谨慎操作！

更多选项 ▾

输入参数

Region ⓘ

本接口不需要传递该参数

参数输入方式

表单 JSON

参数推荐

EngineType [*](#) ⓘ

string

Url [*](#) ⓘ

string

CallbackUrl [*](#) ⓘ

string

SignToken (选填) [*](#) ⓘ

string

FilterDirty (选填) [*](#) ⓘ

integer

FilterModal (选填) [*](#) ⓘ

integer

FilterPunc (选填) [*](#) ⓘ

EngineType

必填：是。

类型：String。

描述：

引擎模型类型。• 16k_zh：中文普通话通用；• 16k_en：英语；• 16k_yue：粤语；• 16k_id：印度尼西亚语；• 16k_fil：菲律宾语；• 16k_th：泰语；• 16k_fr：法语；

Url

必填：是。

类型：String。

描述：

语音流地址，支持rtmp、rtsp等流媒体协议，以及各类基于http协议的直播流(不支持hls、m3u8)

CallbackUrl

必填：是。

类型：String。

描述：

支持HTTP和HTTPS协议，用于接收识别结果，您需要自行搭建公网可调用的服务。回调格式&内容详见：[语音流异步识别回调说明](#)

SignToken

必填：不。

填写好输入参数之后，选择代码生成选项卡，可以看到自动生成的不同编程语言代码（可支持 Java、Python、Node.js、PHP、GO、.NET、C++ 语言），生成代码中的部分字段信息和填写内容是关联的，如需调整传入参数，可在左侧修改参数值后重新生成代码。

CreateAsyncRecognitionTask
asr 2019-06-14 [查看API文档](#)

在线调用 **代码示例** CLI示例 签名示例 文档说明 数据模拟 问题反馈

选择代码示例

选择接入方式
开发语言

SDK SDK Common Client HTTP Request

Golang Python Java C++ Node.js PHP .Net

调试代码 下载工程

根据左侧输入参数自动生成代码，需要手动修改代码中的SecretId和SecretKey参数

```
package main

import (
    "fmt"

    "github.com/tencentcloud/tencentcloud-sdk-go/tencentcloud/common"
    "github.com/tencentcloud/tencentcloud-sdk-go/tencentcloud/common/errors"
    "github.com/tencentcloud/tencentcloud-sdk-go/tencentcloud/common/profile"
    asr "github.com/tencentcloud/tencentcloud-sdk-go/tencentcloud/asr/v20190614"
)

func main() {
    // 实例化一个认证对象，入参需要传入腾讯云账户 SecretId 和 SecretKey，此处还需注意密钥对的保密
    // 代码泄露可能会导致 SecretId 和 SecretKey 泄露，并威胁账号下所有资源的安全性。以下代码示例仅供参考，建议采用更安全的方式来
    // 密钥可前往官网控制台 https://console.cloud.tencent.com/cam/capi 进行获取
    credential := common.NewCredential(
        "SecretId",
        "SecretKey",
    )

    // 实例化一个client选项，可选的，没有特殊需求可以跳过
    cpf := profile.NewClientProfile()
    cpf.HttpProfile.Endpoint = "asr.tencentcloudapi.com"
    // 实例化要请求产品的client对象，clientProfile是可选的
    client, _ := asr.NewClient(credential, "", cpf)

    // 实例化一个请求对象，每个接口都会对应一个request对象
    request := asr.NewCreateAsyncRecognitionTaskRequest()

    // 返回的resp是一个CreateAsyncRecognitionTaskResponse的实例，与请求对象对应
    response, err := client.CreateAsyncRecognitionTask(request)
    if _, ok := err.(*errors.TencentCloudSDKError); ok {
        fmt.Printf("An API error has returned: %s", err)
        return
    }
    if err != nil {
        panic(err)
    }
    // 输出json格式的字符串回包
    fmt.Printf("%s", response.ToJsonString())
}
```

发起请求

填写好输入参数之后，选择在线调用选项卡，单击发送请求按钮可进行真实请求，供您调试、参考。

在线调用 代码示例 CLI示例 签名示例 文档说明 数据模拟 问题反馈

选择在线调用

注意：通过API发送请求等同于真实操作，请小心进行
点击下面的“发送请求”按钮，系统会以POST的请求方法发送您在左侧填写的参数到对应的接口，该操作等同于真实操作，建议您仔细阅读产品计费

点击发送请求按钮，进行真实请求

发送请求

使用 Postman 调用语音识别接口

Postman 下载安装

1. 下载地址：[Download Postman | Get Started for Free](#)，按照电脑类型选择安装包。
2. 下载完成后，双击打开安装包，按照提示安装，如无特殊需求选择默认即可。初次打开需要登录，若无账号，选择跳过。

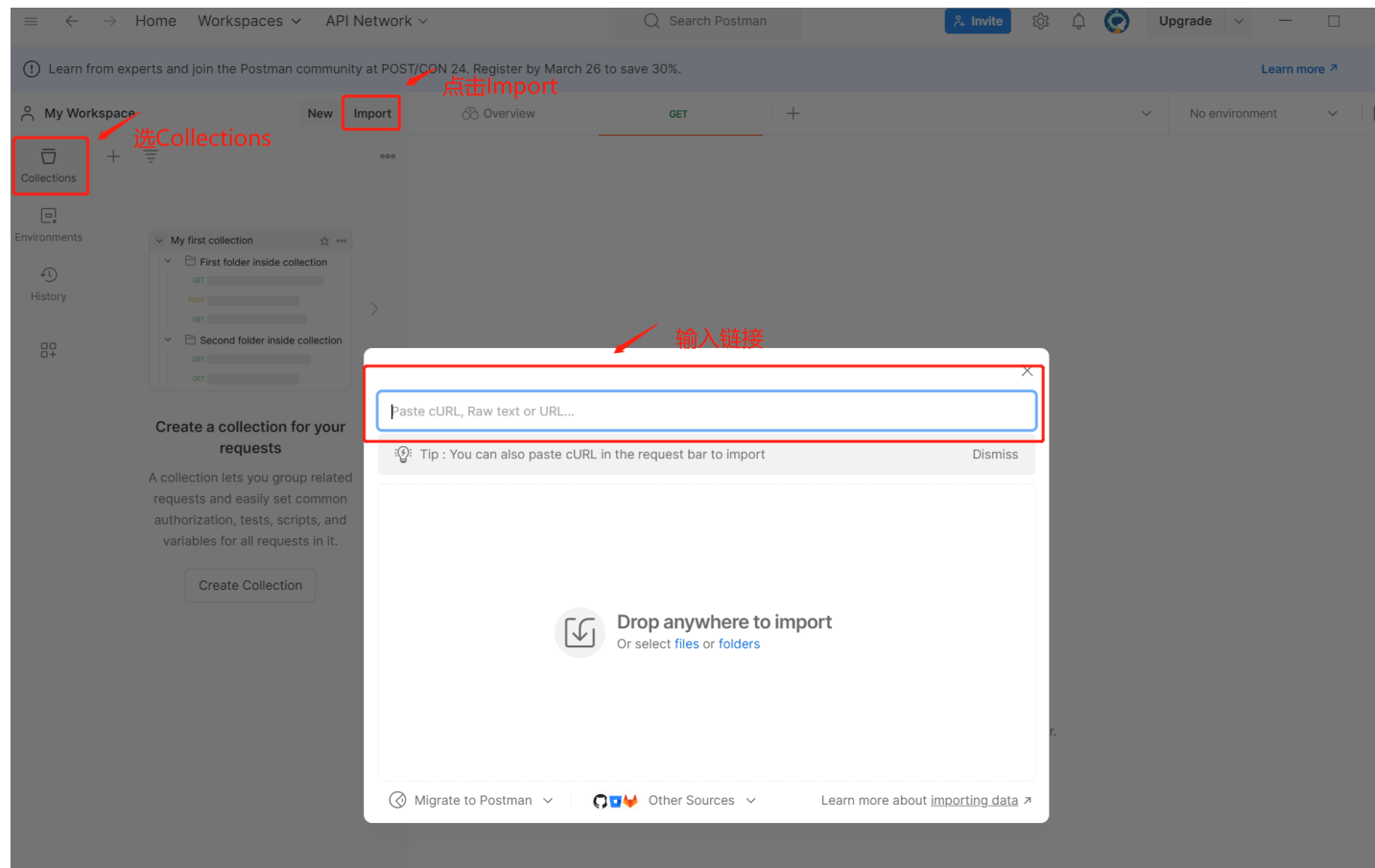
接口调用

导入 Postman 配置

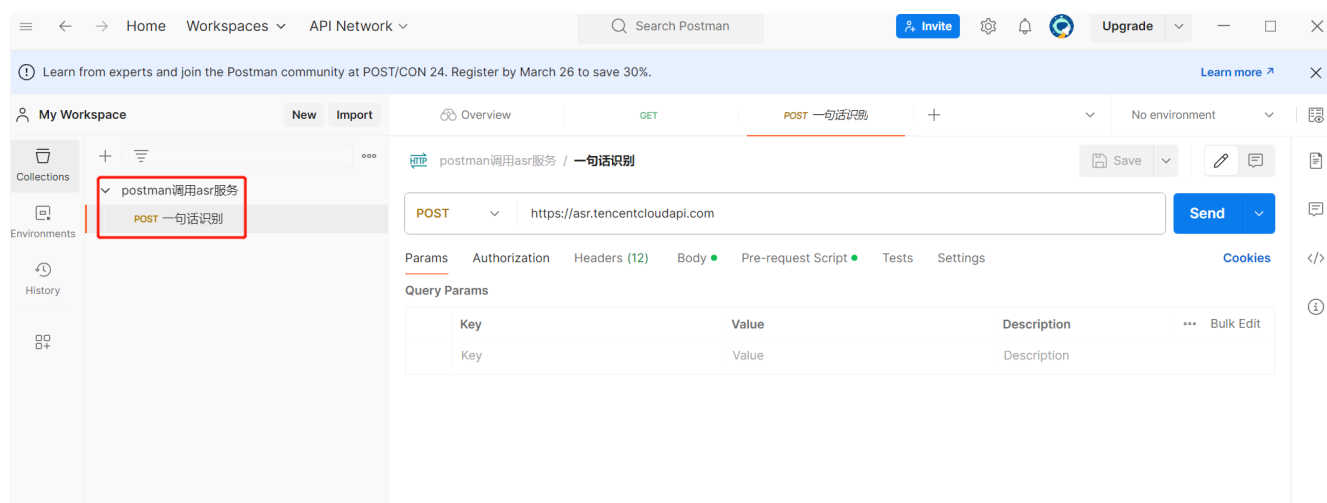
借助 Postman 的导入功能来初始化配置，配置链接：

https://postman-demo-1256237915.cos.ap-guangzhou.myqcloud.com/postman_collection.json

具体导入步骤如下：



初始配置包含一个接口示例，其所需配置信息都已包含在内。看到如下内容表示导入成功。



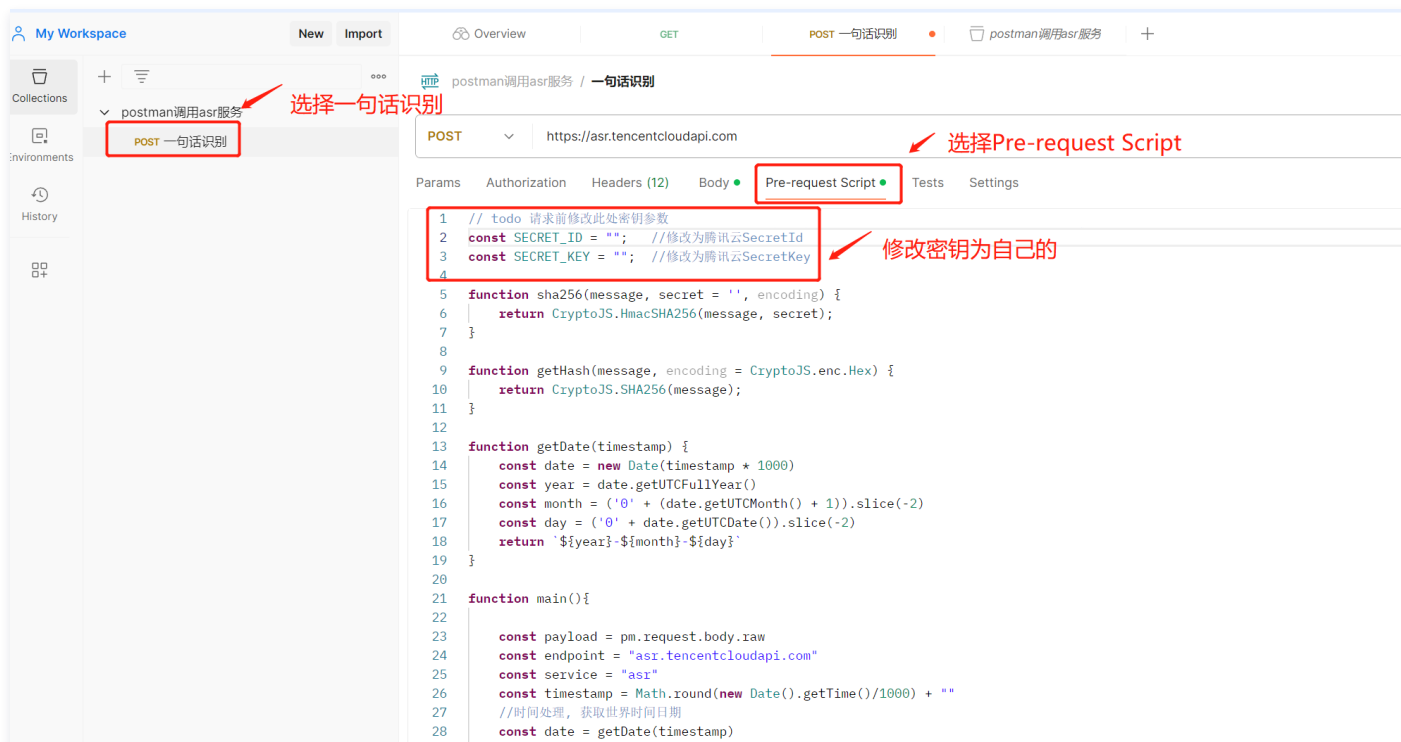
修改腾讯云密钥

发起请求前，需先将该请求对应 Pre-request-Script 中的密钥修改为自己的，这样脚本才能正确计算鉴权参数。

1. 登录访问管理控制台 [获取密钥](#)。



2. 例如我们要调用 [一句话识别接口](#)，修改其预处理脚本中的密钥。



发起请求

密钥修改完后，选择 Body 完善请求参数。单击 **Send** 发送请求，看到如下内容表示请求成功。

The screenshot shows the Postman interface for a POST request to `https://asr.tencentcloudapi.com`. The request body is in JSON format, containing fields like `EngServiceType`, `SourceType`, `VoiceFormat`, and `Url`. The response body is also in JSON format, containing fields like `RequestId`, `Result`, `AudioDuration`, `WordSize`, and `WordList`. Red arrows point to the request and response bodies with labels "请求Body" and "请求成功" respectively.

```
1 {
2   "EngServiceType": "16k_zh",
3   "SourceType": 0,
4   "VoiceFormat": "m4a",
5   "Url": "https://..."
6 }
```

```
1 {
2   "Response": {
3     "RequestId": "977bc31e-6122-450a-b9b4-0fe2a4c2b913",
4     "Result": "腾讯云语音欢迎您。",
5     "AudioDuration": 4649,
6     "WordSize": 0,
7     "WordList": null
8   }
9 }
```

如何修改请求参数和请求接口

参数介绍

ASR 服务（云 API）参数分为公共参数和非公共参数，以 [一句话识别接口](#) 为例：

产品简介

购买指南

快速入门

热词和自学习定制模型说明

操作指南

API 文档

更新历史

简介

API 概览

调用方式

语音流异步识别相关接口

录音文件识别相关接口

一句话识别相关接口

一句话识别

自学习相关接口

热词相关接口

说话人验证相关接口

实时语音识别相关接口

录音文件识别极速版相关接口

虚拟写真人判定相关接口

数据结构

错误码

SDK文档

最佳实践

常见问题

视频专区

相关协议

联系我们

词汇表

2. 输入参数

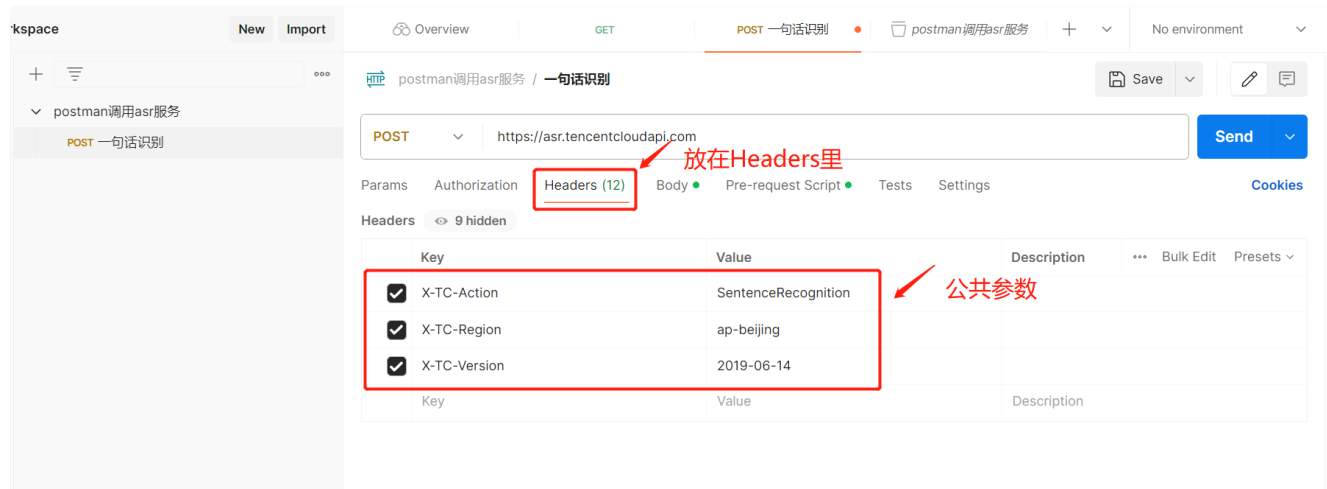
以下请求参数列表仅列出了接口请求参数和部分公共参数，完整公共参数列表见 [公共请求参数](#)。

参数名称	必选	类型	描述
Action	是	String	公共参数，本接口取值：SentenceRecognition。
Version	是	String	公共参数，本接口取值：2019-06-14。
Region	否	String	公共参数，本接口不需要传递此参数。
EngServiceType	是	String	引擎模型类型。 电话场景： • 8k_zh：中文电话通用； • 8k_en：英文电话通用； 非电话场景： • 16k_zh：中文通用； • 16k_zh-PY：中英粤； • 16k_zh_medical：中文医疗； • 16k_en：英语； • 16k_yue：粤语； • 16k_ja：日语； • 16k_ko：韩语； • 16k_vi：越南语； • 16k_ms：马来语； • 16k_id：印度尼西亚语； • 16k_fil：菲律宾语； • 16k_th：泰语； • 16k_pt：葡萄牙语； • 16k_tr：土耳其语； • 16k_ar：阿拉伯语； • 16k_es：西班牙语； • 16k_hi：印地语； • 16k_fr：法语； • 16k_zh_dialect：多方言，支持23种方言（上海话、四川话、武汉话、贵阳话、昆明话、西安话、郑州话、太原话、兰州话、银川话、西宁话、南京话、合肥话、南昌话、长沙话、苏州话、杭州话、济南话、天津话、石家庄话、黑龙江话、吉林话、辽宁话）； 示例值：16k_en
SourceType	是	Integer	语音数据来源。0：语音 URL；1：语音数据（post body）。 示例值：1

公共参数，放在Header中传递

非公共参数，放在Body中传递

对应到 Postman 当中：



注意：

公共参数传递时参数名需加上前缀 “X-TC-”，详情见 [API 中心](#)。



请求参数变更

如果想解析不同的录音或者设置不同的请求参数，步骤如下：

1. 查看接口入参定义，例如 [一句话识别](#)。

语音识别

产品动态

新手指引

产品简介

购买指南

快速入门

热词和自学习定制模型说明

操作指南

API 文档

更新历史

简介

API 概览

调用方式

语音流异步识别相关接口

录音文件识别相关接口

录音文件识别结果查询

录音文件识别请求

录音文件识别回调说明

一句话识别相关接口

一句话识别

自学习相关接口

热词相关接口

说话人验证相关接口

实时语音识别相关接口

录音文件识别极速版相关接口

虚拟号真人判定相关接口

数据结构

错误码

SDK文档

最佳实践

常见问题

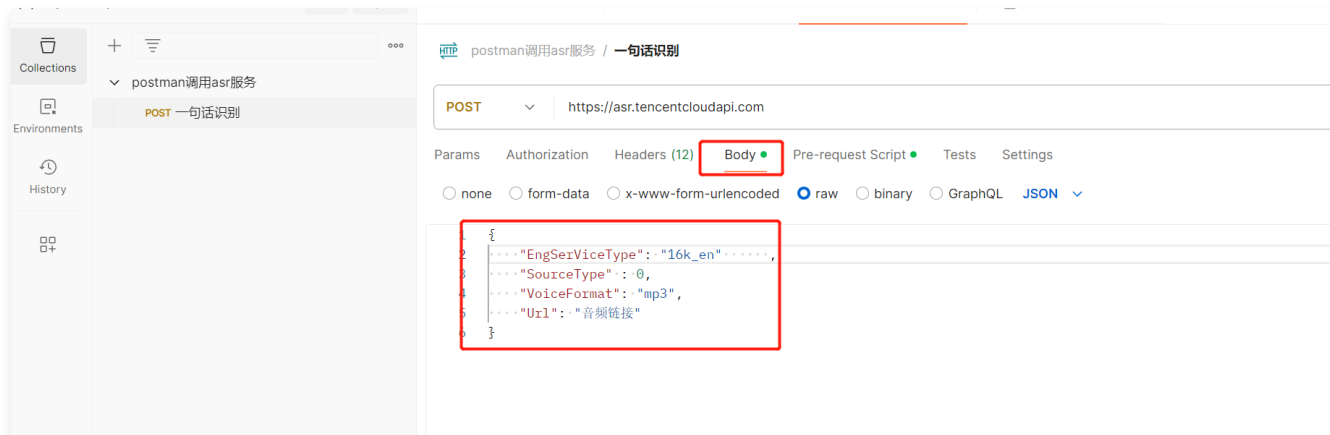
视频专区

2. 输入参数

以下请求参数列表仅列出了接口请求参数和部分公共参数，完整公共参数列表见 [公共请求参数](#)。

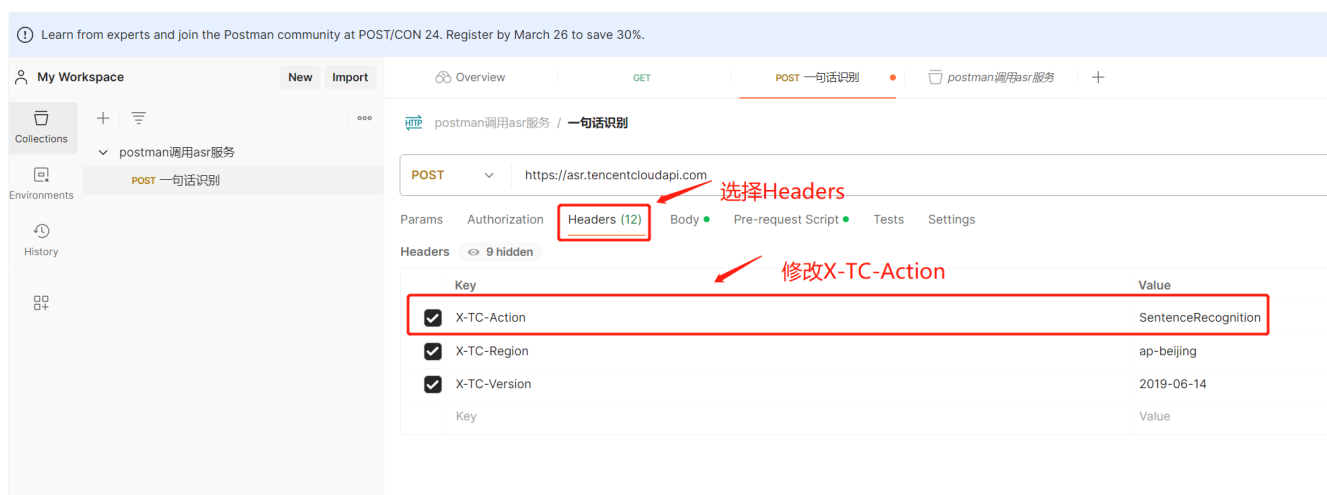
参数名称	必选	类型	描述
Action	是	String	公共参数，本接口取值：SentenceRecognition。
Version	是	String	公共参数，本接口取值：2019-06-14。
Region	否	String	公共参数，本接口不需要传递此参数。
EngSerViceType	是	String	引擎模型类型。 电话场景： • 8k_zh：中文电话通用； • 8k_en：英文电话通用； 非电话场景： • 16k_zh：中文通用； • 16k_zh-PY：中英粤； • 16k_zh_medical：中文医疗； • 16k_en：英语； • 16k_yue：粤语； • 16k_ja：日语； • 16k_ko：韩语； • 16k_vi：越南语； • 16k_ms：马来语； • 16k_id：印度尼西亚语； • 16k_fil：菲律宾语； • 16k_th：泰语； • 16k_pt：葡萄牙语； • 16k_tr：土耳其语； • 16k_ar：阿拉伯语； • 16k_es：西班牙语； • 16k_hi：印地语； • 16k_fr：法语； • 16k_zh_dialect：多方言，支持23种方言（上海话、四川话、武汉话、贵阳话、昆明话、西安话、郑州话、太原话、兰州话、银川话、西宁话、南京话、合肥话、南昌话、长沙话、苏州话、杭州话、济南话、天津话、石家庄话、黑龙江话、吉林话、辽宁话）； 示例值：16k_en
SourceType	是	Integer	语音数据来源。0：语音 URL；1：语音数据（post body）。 示例值：1
VoiceFormat	是	String	识别音频的音频格式，支持wav、pcm、ogg-opus、speex、silk、mp3、m4a、aac、amr。 示例值：wav

2. 在 Body 中变更参数，例如我们想识别的引擎模型为16k_en 英语而不是16k_zh 中文，则需传递 16k_en 参数，如一句话识别示例接口的传参。



请求接口变更

如需请求 ASR 服务的其他接口，需修改 Header 中的 X-TC-Action。



X-TC-Action 参数对应接口文档的 Action 入参，如录音文件识别接口的 Action 为 CreateRecTask。

API 文档

更新历史

简介

API 概览

调用方式

语音流异步识别相关接口

录音文件识别相关接口

录音文件识别结果查询

录音文件识别请求

录音文件识别回调说明

一句话识别相关接口

一句话识别

自学习相关接口

热词相关接口

说话人验证相关接口

实时语音识别相关接口

录音文件识别极速版相关接口

虚拟号真人判定相关接口

数据结构

错误码

SDK文档

最佳实践

常见问题

2. 输入参数

以下请求参数列表仅列出了接口请求参数和部分公共参数，完整公共参数列表见 [公共请求参数](#)。

参数名称	必选	类型	描述
Action	是	String	公共参数 ，本接口取值：CreateRecTask。
Version	是	String	公共参数 ，本接口取值：2019-06-14。
Region	否	String	公共参数 ，此参数为可选参数。
			引擎模型类型 识别引擎采用分级计费方案，标记为“大模型版”的引擎适用大模型计费方案， 点击这里 查看产品计费说明 电话通讯场景引擎： 注意：电话通讯场景，请务必使用以下8k引擎 • 8k_zh：中文电话通讯； • 8k_en：英文电话通讯； 如您有电话通讯场景识别需求，但发现需求语种仅支持16k，可将8k音频传入下方16k引擎，亦能获取识别结果。但16k引擎并非基于电话通讯数据训练，无法承诺此种调用方式的识别效果，需由您自行验证识别结果是否可用 通用场景引擎： 注意：除电话通讯场景以外的其它识别场景，请务必使用以下16k引擎 • 16k_zh：中文普通话通用引擎，支持中文普通话和少量英语，使用丰富的中文普通话语料训练，覆盖场景广泛，适用于除电话通讯外的所有中文普通话识别场景； • 16k_zh_large：中文普通话通用引擎【大模型版】。模型参数量极大，语言模型性能增强，针对噪声大、回音大、人声小、人声远、等低质量音频的识别准确率极大提升， 点击这里 对比中文普通话常规版本与大模型版本的识别效果； • 16k_zh_dialect：中文普通话+多方言混合引擎，除普通话外支持23种方言（上海话、四川话、武汉话、贵阳

修改 Action 之后，再按照接口文档修改请求 Body，发起请求即可。

新手常见问题

Last updated: 2024-12-23 14:57:33

语音识别如何接入？

语音识别目前支持 API 和 SDK 接入，推荐 SDK 接入，详情可参见 [一分钟接入服务端 API](#) 和 [一分钟跑通集成 SDK](#)。

语音识别怎么进行功能体验？

可通过微信搜索“腾讯云 AI 语音”小程序，选择语音识别进行体验。也可在 [语音识别控制台](#) 功能体验模块，通过上传文件或者 URL 进行体验。详情可参考 [体验功能](#)。

语音识别控制台功能体验怎样上传大于 5M 的文件？

可在 [语音识别控制台-功能体验](#) 中采用上传音频 URL 方式上传体验，建议音频时长不能大于五个小时。

不同使用场景对应的是语音识别哪种服务？

- 实时语音识别适用于有实时性要求的场景，例如语音输入、语音机器人、会议现场记录等场景。
- 一句话识别适用于对60秒之内的短音频文件进行识别的场景，例如语音短信、语音搜索等场景。
- 录音文件识别适用于语音时间较长、实时性要求低的场景，例如客服质检、视频字幕生成等场景。

支持远场和离线的语音识别吗？

支持离线语音识别，详情可参见 [离线SDK文档](#)。

语音识别支持中英文混合场景和地方方言吗？

- [实时语音识别](#)、[一句话识别](#)、[录音文件识别](#)、[录音文件识别极速版](#)和 [语音流异步识别](#) 支持中英文混合识别（使用中文引擎的时候，在少量英文的情况下可以支持中英文混合识别，大量英文可能出现识别率降低的情况），且支持带口音的中文普通话语音识别。
- [实时语音识别](#)、[一句话识别](#)、[录音文件识别](#)和 [录音文件识别极速版](#) 支持上海话、四川话、武汉话、贵阳话、昆明话、西安话、郑州话、太原话、兰州话、银川话等23种方言的识别。

语音识别支持的输入音频时长是多少？

- 一句话识别每次调用支持60秒之内的音频。
- 录音文件识别每次调用支持五小时之内的音频。
- 实时语音音频流中每个数据包的音频分片为200ms。

语音识别接口的 HTTP 请求返回鉴权失败？

请用户对照参数表检查自己的参数是否正确上传。如果想快速接入，推荐使用官网提供的 SDK。

语音识别接口会限制音频文件的采样率吗？

接口不会限制，但是采样率不符合标准，会影响到识别效果。