

# 接口说明

图像识别（OCR）可以将图片中的文字转换为计算机可编程的文字。该能力通过WebSocket API的方式提供给开发者，相较于SDK，该方式具有轻量、跨平台、跨开发语言的特点。

## 接口要求

项目	说明
请求地址	ws://api.baller-tech.com/v1/service/ws/v1/ocr
字符编码	UTF-8
WebSocket版本	13 ( <a href="#">RFC 6455</a> )
响应格式	统一采用JSON格式
图像格式	jpg; jpeg; bmp; png; gif; tif; tiff; pdf
图像大小	图像大小不超过4M

## 调用流程

1. 通过hmac-sha256计算签名，向服务器端发送WebSocket协议握手请求。
2. 握手成功之后，通过WebSocket连接上传和接收数据。
3. 请求方接收到服务器端推送的结果返回结束标记后断开WebSocket连接

## 握手和接口鉴权

在WebSocket的握手阶段，请求方需要对请求进行签名，服务端会根据签名检查请求的合法性。握手时请求方将签名相关的参数经过url编码后加到请求地址的后面，具体的参数和示例如下：

```
ws://api.baller-tech.com/v1/service/ws/v1/ocr?authorization=xxxx&host=xxxx&date=xxx
```

参数	类型	说明	示例
host	string	请求的主机	api.baller-tech.com
date	string	当前GMT格式的时间	Fri, 10 Jan 2020 07:31:50 GMT
authorization	string	鉴权信息Base64编码后的数据	-

## 握手和鉴权参数详细介绍

## date介绍

- 1. date必须是GMT+0时区的符合RFC1123格式的日期和时间，星期和月份只能使用英文表示
- 2. 服务端允许date的最大偏差为300秒，超出此偏差请求会被拒绝

## authorization介绍

authorization使用base64编码前的格式如下json格式

```
{
  "app_id": "1172448516240310275",
  "signature": "qaIpgE3Ecs78g6GRFXQBJKgdna28b7ronAcSDCSO+ZW="
}
```

## app\_id介绍

- 1. 由北京大牛儿科技发展有限公司统一分配。

## signature介绍

- 1. signautre 是使用hmac-sha256对参数进行签名后并base64编码的字符串。
- 2. signautre 使用hmac-sha256签名前的原始字段由三部分构成，分别为app\_id、date、host。每一部分使用换行符(\n)进行分割，“:”号前后无空格。

```
app_id:1172448516240310275
date:Fri, 10 Jan 2020 07:31:50 GMT
host:api.baller-tech.com
```

- 3. 使用hmac-sha256算法，结合app\_key（由北京大牛儿科技发展有限公司统一分配）对signautre的原始字段进行签名。
- 4. 对签名数据进行base64编码，生成signature的字段值。

## 握手和鉴权消息响应

- 1. 接口鉴权成功时，WebSocket握手回复报文的状态码为101。
- 2. 接口鉴权失败时，WebSocket握手回复报文的状态码为403，可以通过响应行的原因短语查看接口鉴权失败原因。
- 3. 接口鉴权失败时，响应报文的主体中会返回json格式的数据，包含了以下信息

参数	类型	说明
task_id	string	本次任务的标识，如果对请求有疑问，可以将task_id提供给我公司进行排查
message	string	接口鉴权失败的原因，与响应行中的原因短语相同

## 数据的发送和接收

握手成功之后，请求方和服务端会建立WebSocket的连接，请求方将数据通过WebSocket发送给服务器，服务器有识别结果的时候，会通过WebSocket连接推送识别结果到请求方。请求方和服务端通过json的格式交换数据。

### 请求方发送数据时使用的参数

参数名	类型	是否每帧必须	描述
business	obj	否	业务参数，仅在握手成功后首帧中上传
data	obj	是	数据流参数，握手成功后所有帧中都需要上传

#### 业务参数(business)

参数名	类型	是否必须	描述
image_mode	string	是	传入图片的文本模式
language	string	否	识别语种参见 <a href="#">支持的图像识别语种</a>
file_format	string	否	仅当识别PDF文件时需要填写，填写值为pdf。请参考 <a href="#">PDF识别注意事项</a>

#### image\_mode 介绍

image\_mode表明图像的模式，有以下几种可以设定的值

可设置值	描述
multi_row	多行

#### 数据流参数 (data)

参数名	类型	是否必须	描述
image	string	是	经过base64编码后的图片数据
input_mode	string	否	仅当识别PDF文件时需要填写。请参考 <a href="#">PDF识别注意事项</a> once continue end

```
{
  "data": {
    "image":
"AAAFAAoADWAXAB0AJgA0AEIATABPAE8AUQBRAEgAOWa0AC8AJwACABUAEQAJAAIAAgADAAAA+P="
  },
  "business": {
    "image_mode": "multi_row",
  }
}
```

## 服务器推送结果的参数

图像识别时，会将图像按一定的规则（目前为按行）分为不同的子项，每次返回的识别结果是一个或多个子项的识别结果。响应数据为json格式，具体字段的含义如下

参数名	类型	描述
task_id	string	本次任务的id，仅在第一帧中返回，如果对请求有疑问，可以将task_id提供给我公司进行排查
code	int	请求处理的结果码
message	string	错误提示
is_end	int	结果返回是否结束（0-未结束; 1-结束），当为1时，请求方需关闭WebSocket
data	数组	每一个子项的识别结果， <a href="#">详见</a>

### data字段介绍

参数	类型	说明
order	int	当前响应报文中子句的顺序（是一次推送帧中多个子句的顺序，不是整个识别事务的）
result	string	该子句的识别结果
page	int	仅识别PDF文件时有效，表示页数的索引（从0开始）

```
{
  "task_id": "2903dc7e3ab65879b4fc66055720ec09",
  "code": 0,
  "message": "success",
  "is_end": 1,
  "data": [
    {
      "order": 0,
      "result": "谢谢惠顾",
      "page": 0,
    },
    {
```

```
        "order": 1,
        "result": "期待下次再见",
        "page": 0,
    },
]
}
```

## PDF识别注意事项

本接口支持对10M以内的PDF文件进行识别，识别PDF时业务参数中的file\_format需设置为\*\*pdf\*\*。<br>如果pdf文件较小（4M以内），可以一次将整个PDF文件发送到服务器，此时数据参数中的input\_mode字段可以不设置，或设置为once；如果数据文件较大，不能一次将整个PDF文件发送到服务器，可以将PDF文件切分成多段，并分多次发送给服务器，在这种情况下如果不是最后一段需设置input\_mode为continue，如果是最后一段需设置input\_mode为end。

## 支持的语种

图像识别语种	对应的language字段
简体中文	zho
繁体中文	cht
藏文	tib
蒙文(传统)	mon_i
蒙文(西里尔)	mon_o
维文	uig
彝文	iii
壮文	zha
韩文	kor
哈萨克文(传统)	kaz_i