

文字识别服务

# 用户指南

文档版本 01  
发布日期 2024-05-22



版权所有 © 华为云计算技术有限公司 2024。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

## 商标声明



HUAWEI和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

## 注意

您购买的产品、服务或特性等应受华为云计算技术有限公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为云计算技术有限公司对本文档内容不做任何明示或暗示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

# 华为云计算技术有限公司

地址：贵州省贵安新区黔中大道交兴功路华为云数据中心 邮编：550029

网址：<https://www.huaweicloud.com/>

---

# 目录

---

<b>1 功能体验与试用</b>	<b>1</b>
<b>2 使用流程简介</b>	<b>2</b>
<b>3 开通文字识别服务</b>	<b>3</b>
<b>4 准备数据</b>	<b>6</b>
<b>5 自定义 OCR</b>	<b>7</b>
5.1 自定义 OCR 介绍	7
5.2 创建应用	8
5.3 通用单模板 workflow	10
5.3.1  workflow 介绍	10
5.3.2 上传模板图片	11
5.3.3 定义预处理	13
5.3.4 框选参照字段	14
5.3.5 框选识别区	15
5.3.6 评估	16
5.3.7 调用指南	17
5.3.8 编辑应用	18
5.3.9 删除应用	19
5.4 多模板分类 workflow	19
5.4.1  workflow 介绍	19
5.4.2 上传模板图片	21
5.4.3 定义预处理	22
5.4.4 框选参照字段	23
5.4.5 框选识别区	25
5.4.6 模板总览	25
5.4.7 评估	26
5.4.8 调用指南	28
5.4.9 编辑应用	29
5.4.10 删除应用	30
5.5 字段类型	30
5.5.1 默认字段类型	31
5.5.2 自定义字段类型	31
5.6 查看应用详情	33

5.7 常见问题.....	34
5.7.1 什么情况下可以使用自定义模板? .....	34
5.7.2 如何选取参照字段? .....	35
5.7.3 识别结果容易漏字或多识别出内容怎么办? .....	35
5.7.4 识别字符出现错误如何优化? .....	35
5.7.5 自定义模板 OCR 取消开通后, 是否会将之前创建好的模板删除? .....	36
<b>6 调用 API 或 SDK.....</b>	<b>37</b>
6.1 在线调试.....	37
6.2 本地调用.....	38
<b>7 获取识别结果.....</b>	<b>41</b>
7.1 解读识别结果.....	41
7.2 识别结果后处理.....	42
<b>8 统计 API 使用量.....</b>	<b>47</b>
<b>9 关闭服务.....</b>	<b>48</b>
<b>10 监控.....</b>	<b>49</b>
10.1 文字识别的监控指标.....	49
10.2 查看监控指标.....	50
<b>11 审计.....</b>	<b>51</b>
11.1 支持云审计的关键操作.....	51
11.2 如何查看审计日志.....	51
<b>12 权限管理.....</b>	<b>53</b>
12.1 权限管理.....	53
12.2 创建用户组与 IAM 用户.....	61
12.3 为 IAM 用户分配权限.....	62
12.4 IAM 用户登录并验证权限.....	64
12.5 OCR 自定义策略.....	64
<b>13 企业项目管理.....</b>	<b>66</b>
13.1 应用场景简介.....	66
13.2 创建企业项目.....	67
13.3 企业项目财务管理.....	70
<b>14 V1 版本 APIs 切换至 V2 版指导.....</b>	<b>71</b>

# 1 功能体验与试用

文字识别（Optical Character Recognition，OCR）是指对图像文件的打印字符进行检测识别，将图像中的文字转换成可编辑的文本格式。OCR通过API提供服务能力，用户需要通过编程来处理识别结果。

在开通OCR前，请先使用[OCR体验馆](#)体验服务功能。该方式无需编程，只需在网页端上传图片，即可体验识别效果。

开通OCR后，默认按API调用次数进行收费。API使用指导请参见[使用流程简介](#)。

## 说明

使用OCR体验馆前，请先注册华为账号并开通华为云，并实名认证。

图 1-1 OCR 体验馆

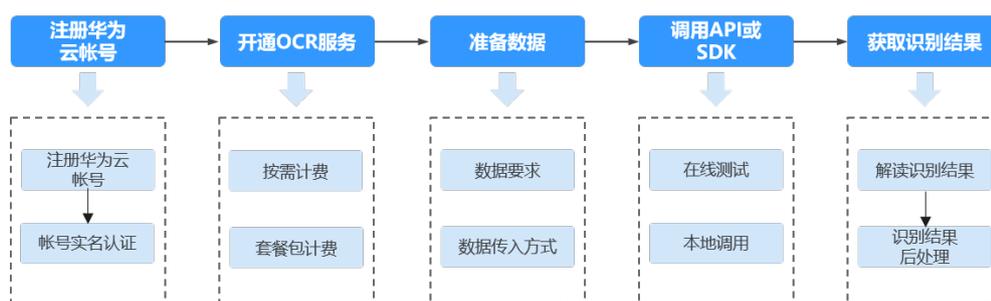


# 2 使用流程简介

文字识别是指对图像中的打印字符进行检测识别，通过调用API，识别图片中的文字，并返回JSON格式的识别结果。

## 使用流程

图 2-1 使用流程



## 使用前必读

用户需要具备编程能力，熟悉Java、Python、iOS、Android、Node.js编程语言。

OCR服务需要用户通过调用API接口，将图片或扫描件中的文字识别成可编辑的文本，然后返回JSON格式的识别结果，用户需要通过编码将识别结果对接到业务系统或保存为TXT、Excel等格式。

关于文字识别的相关声明请参见[文字识别服务声明](#)。

文字识别服务等级协议请参见[华为云服务等级协议](#)。

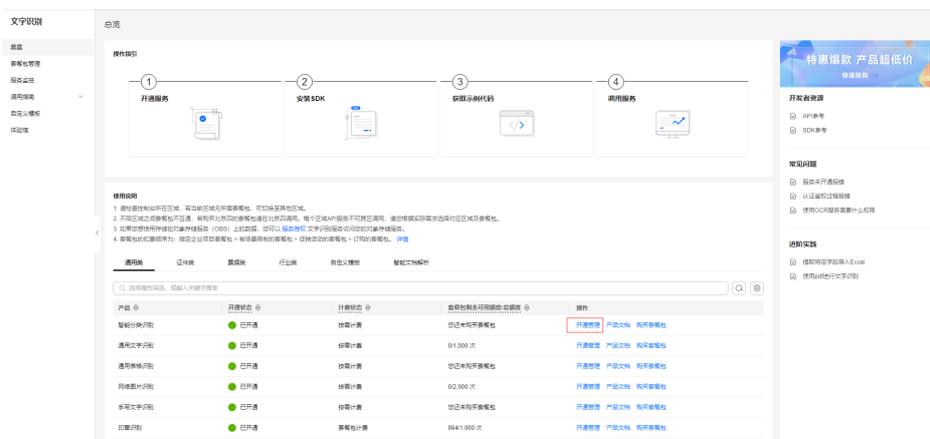
# 3 开通文字识别服务

OCR服务提供的开通方式有以下两种，用户可以任选其一进行开通服务。

## 按需计费开通服务

1. 进入[文字识别官网主页](#)，单击“立即使用”，进入文字识别控制台。
2. 在“总览”页面，选择需要使用的服务，执行开通操作，默认的计费方式采用按需计费。

图 3-1 服务开通



服务开通成功后，开通状态将显示为“已开通”。

3. 如果您需要调用对象存储服务（OBS）中的数据，请在开通服务页面进行服务授权。

图 3-2 OBS 服务授权



### 说明

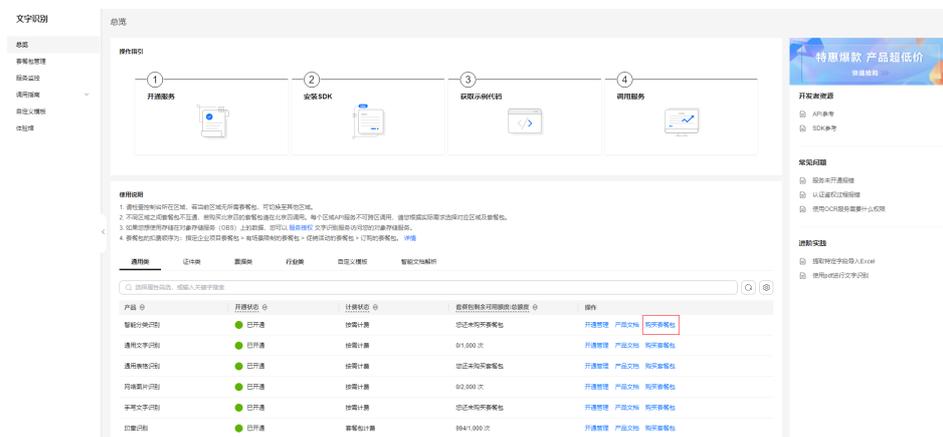
- OCR服务开通后，可在控制台进行关闭。开通服务时，计费规则默认为“按需计费”。只有调用成功才会计费，其中返回的2xx状态码表示调用成功，调用失败不计费，详细的计费价格及规则请参见[计费说明](#)。
- 如未开通服务，直接调用OCR API会提示ModelArts.4204报错。

4.服务开通后，您可以按照官方文档[数据要求](#)，[调用API或SDK](#)步骤继续完成实际服务需求。

## 购买套餐包开通服务

1. 进入[文字识别官网主页](#)，单击“立即使用”，进入文字识别控制台。
2. 在“总览”页面，选择需要使用的服务，在操作列单击“购买套餐包”，此时的计费方式采用套餐包计费，套餐包额度用完后自动转为按需计费。

图 3-3 购买套餐包



服务开通成功后，开通状态将显示为“已开通”。

3. 如果您需要调用对象存储服务（OBS）中的数据，请在开通服务页面进行服务授权。

图 3-4 OBS 服务授权



### 📖 说明

- 购买套餐包后只有调用成功才会计算调用次数，其中返回的2xx状态码表示调用成功，调用失败不计算次数。套餐包扣减规则包括计费模式、变更配置、续费、欠费类等问题请参见[计费说明](#)。
- 如未开通服务，直接调用OCR API会提示ModelArts.4204报错。

4.服务开通后，您可以按照官方文档[数据要求](#)，[调用API或SDK](#)步骤继续完成实际服务需求。

# 4 准备数据

## 数据要求

受技术与成本多种因素制约，文字识别服务存在一些约束限制。

以**通用文字识别API**为例，输入数据存在以下约束。其他API的使用约束请参见[约束与限制](#)。

- 只支持识别PNG、JPG、JPEG、BMP、TIFF格式的图片。
- 图像各边的像素大小在15px到8192px之间。
- 图像中识别区域有效占比超过80%，保证所有文字及其边缘包含在图像内。
- 支持图像任意角度的水平旋转。
- 目前不支持复杂背景（如户外自然场景、防伪水印等）和文字扭曲图像的文字识别。

## 数据传入方式

使用OCR API或SDK时，数据主要通过以下两种方法传入。

- image  
传入图片的base64编码。  
您可以通过在线转码工具，[浏览器](#)，使用Python中的base64.b64encode函数等方法，获取图片的base64编码。
- url  
传入图片的url路径。
  - 使用公网http/https url，例如https://support.huaweicloud.com/api-ocr/zh-cn\_image\_0288038400.png
  - 将图片上传至华为云对象存储服务（OBS）中，使用OBS提供的url。使用OBS数据需要进行授权。包括对服务授权、临时授权、匿名公开授权，详情参见[配置OBS访问权限](#)。

# 5 自定义 OCR

## 5.1 自定义 OCR 介绍

### 📖 说明

OCR自定义模板已正式停售，为避免影响您目前的业务，请您及时进行相关迁移工作，迁移指导详见[OCR自定义模板停售公告](#)。

自定义OCR基于丰富的文字识别算法和行业知识积累，帮助客户快速构建满足不同业务场景需求的文字识别服务，实现多种版式图像的文字信息结构化提取。

### workflows

自定义OCR当前提供了单模板工作流和多模板工作流，自主构建文字识别模板，识别模板图片中的文字，提供高精度的文字识别模型，保证结构化信息提取精度。

- 通用单模板工作流  
通过构建文字识别模板，识别单个板式图片中的文字，提供高精度的文字识别模型，保证结构化信息提取精度。
- 多模板分类工作流  
支持用户自定义多个文字识别模板，通过AI分类算法，自动识别图片所需使用的模板，从而支持从大量不同板式图像中提取结构化信息。

### 使用工作流开发应用流程

自定义OCR提供了单模板工作流和多模板工作流，您可以通过预置的工作流，自主构建文字识别模板，识别模板图片中的文字，提供高精度的文字识别模型，保证结构化信息提取精度。

图 5-1 使用预置 workflow 开发应用

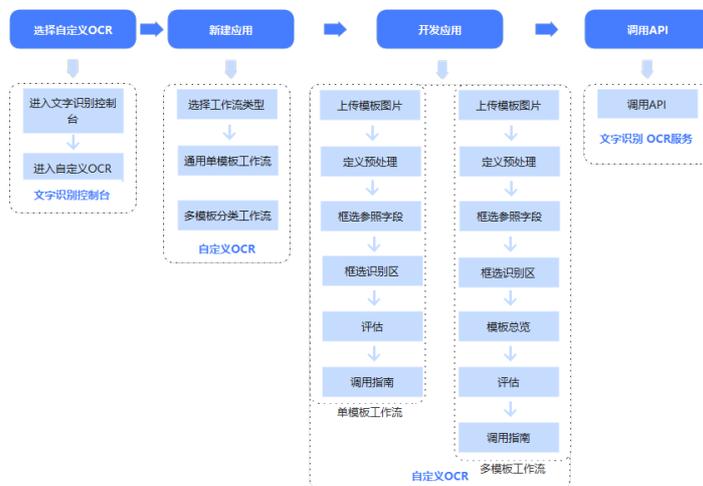


表 5-1 使用 workflow 开发应用流程

流程	说明	详细指导
自定义OCR	根据您的实际使用需求选择自定义OCR。您可以通过查看产品文档选择所需使用的工作流。	在“文字识别”控制台界面，单击左侧菜单“自定义OCR”，进入自定义OCR应用页面。
创建应用	基于已选择的工作流创建应用，填写应用基本信息、选择工作流。	<a href="#">创建应用</a>
开发应用	通用单模板工作流	根据工作流指引，创建单模板服务，通过模板训练文字识别模型，实现单模板图像的文字信息结构化提取。
	多模板分类工作流	根据工作流指引，创建多模板服务，通过多模板训练模板分类器和文字识别模型，实现多模板图像的文字信息结构化提取。
调用API	调用指南，支持通过API调用当前模板服务。	<a href="#">调用API</a>

## 5.2 创建应用

在使用自定义OCR开发应用之前，需要创建应用，填写应用基本信息、选择工作流。

## 前提条件

- 保证华为云账号不欠费。在使用自定义OCR产品时，需在“文字识别”控制台页面订阅自定义OCR服务。
- 如果您的账号是IAM账号，在创建应用前，请先使用管理员账号为IAM用户授予文字识别服务（Optical Character Recognition，简称OCR）操作权限。

## 操作步骤

1. 在文字识别控制台左侧导航栏选择“自定义OCR”。默认进入“我的应用”页签。

图 5-2 文字识别控制台



2. 在“我的应用”页签下，单击“创建应用”。进入“创建应用”页面。

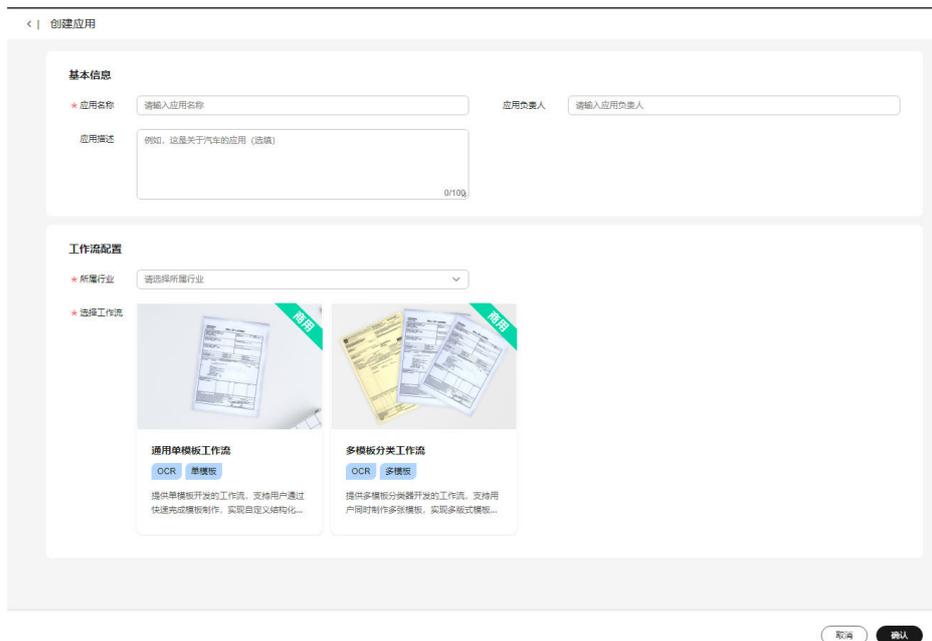
图 5-3 创建应用



您也可以单击“我的工作流”，切换至“我的工作流”页签，选择工作流并单击卡片中的“创建应用”。

3. 根据业务需求填写“应用名称”、“应用负责人”和“应用描述”，仅支持英文、中文、数字、下划线、中划线。选择“所属行业”及工作流。

图 5-4 创建应用



4. 确认信息后，单击“确认”。  
成功创建应用。

## 后续操作

根据所选的预置工作流开发应用。

[通用单模板工作流](#)

[多模板分类工作流](#)

## 5.3 通用单模板工作流

### 5.3.1 工作流介绍

#### 工作流简介

- 功能介绍  
支持构建文字识别模板，识别单个板式图片中的文字，提供高精度的文字识别模型，保证结构化信息提取精度。
- 适用场景
  - 用户认证识别  
识别证件中关键信息，节省人工录入，提升效率，降低用户实名认证成本，准确快速便捷。
  - 快递单自动填写  
识别图片中联系人信息并自动填写快递单，减少人工输入。
  - 合同录入与审核  
自动识别结构化信息与提取签名盖章区域，有助快速审核。

- 优势
  - 解决手工录入投入大、效率低、语种多等问题，提升业务效率。
  - 一键式部署，快速输出高精度结构化数据。

## 工作流流程

在“文字识别控制台>自定义OCR”页面选择“通用单模板工作流”创建应用，详细操作请见[6.2 新建应用](#)。您可以创建单模板服务，通过模板配置文字识别模型，实现单模板图像的文字信息结构化提取。

图 5-5 创建单模板流程



表 5-2 创建单模板流程说明

流程	说明	详细指导
上传模板图片	在使用单模板工作流开发应用之前，需要上传模板图片，明确以哪张图片作为模板训练文字识别模型。	<a href="#">上传模板图片</a>
定义预处理	上传模板图片后，需要对模板图片进行预处理，去掉冗余部分，将图片旋转至水平，保证模型识别的准确性。	<a href="#">定义预处理</a>
框选参照字段	在图片模板中框选参照字段，用于矫正图片的方向，进而在正确的方向上，识别图片中的结构化信息。	<a href="#">框选参照字段</a>
框选识别区	在图片模板中框选识别区，确定模板图片中需要识别的文字位置。	<a href="#">框选识别区</a>
评估	通过上传与模板图片同一板式的测试图片，评估图片识别结果是否正确。	<a href="#">评估</a>
调用指南	生成模板ID和自定义文字识别接口，用于识别自己所上传的模板样式的图片。提供标准API和常用编程语言的SDK，可快速开发属于自己的文字识别应用。	<a href="#">调用指南</a>

### 5.3.2 上传模板图片

在使用单模板工作流开发应用之前，必须确保需要识别的图片都属于同一类型的模板。同一类模板的图片，它们有相同的版面排布，且要识别的文字类型和位置均要求固定。如果您的图片具有多种模板，请考虑使用“多模板工作流”。

确认模板类型后，要准备一张文字清晰的图作为模板图片，模板图中要包含业务需要定制识别的字段。例如上传某一格式的发票图片作为模板，配置文字识别模型就能识别并提取同格式发票上的关键字段。

## 前提条件

- 已在“自定义OCR”控制台选择“通用单模板工作流”创建应用，详情请见[6.2新建应用](#)。
- 提前准备待识别的图片，图片要求请见[图片要求](#)。

## 图片要求

- 保证图片质量：不能有损坏的图片；目前支持的格式包括JPG、JPEG、PNG、BMP。
- 图片最大边不大于4096px，最小边不小于100px，且大小不超过5M。
- 图片中的文字清晰可认。

## 进入应用开发页面

1. 进入“文字识别控制台>自定义OCR”页面。  
默认进入“应用开发>工作台”页面。
1. 在“我的应用”页签下，选择应用并单击“操作”列的“查看”。  
进入“应用资产”页面。
2. 单击“创建模板”。  
进入“应用开发”页面，开始开发应用。

图 5-6 应用开发



## 上传模板图片

在“应用开发>上传模板图片”页面，配置如下信息：

图 5-7 上传模板图片



“模板名称”：输入模板名称。

“语种”：单击图片上方“语种”下拉选择框，选择模板对应的语种。不选择的情况下，默认为中英文。

“上传图片”：单击“上传图片”区域，或鼠标直接拖拽图片至“上传图片”区域，上传本地一张图片作为模板，用于业务场景的文字结构化识别。

确认信息后，单击“下一步”，进入[定义预处理](#)步骤，对上传的模板图片进行自动旋转预处理。

### 5.3.3 定义预处理

使用通用单模板 workflow 上传模板图片后，需要对模板图片进行预处理，通过旋转、裁剪、降噪等操作。图片预处理的目的是保留图片的关键内容，去掉冗余部分，保持图片内容清晰可见，保证模型识别的准确性。

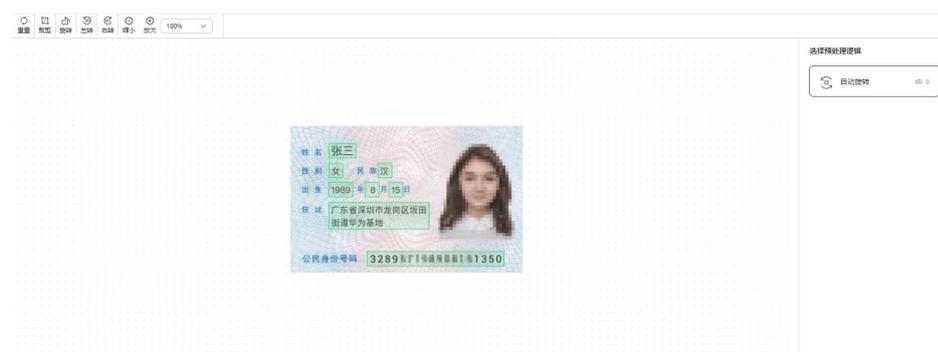
#### 前提条件

已在自定义 OCR 控制台选择“通用单模板 workflow”创建应用，并上传模板图片，详情请见[上传模板图片](#)。

#### 定义预处理

在“应用开发>定义预处理”页面，您可以选择“预处理”逻辑快速处理图片，也可以按左上角操作指标调整图片。

图 5-8 定义预处理



右侧“选择预处理逻辑”区域勾选对应操作，当前仅支持“自动旋转”操作，系统自动旋转文字方向不正确的图片，保持图片中的文字方向水平。

单击预处理区域左上方的操作图标，调整模板图片的大小、方向等。

：单击图标重置图片为初始状态，即未进行任何处理的状态。

：单击图标，在“图片裁剪”窗口调整图片裁剪范围，然后单击“裁剪”，调整图片的大小。

：单击图标，在“图片旋转”窗口，鼠标选中调整图片的方向，然后单击“确认”，调整图片的方向。

：单击图标，向左旋转图片90°。

：单击图标，向右旋转图片90°。

：单击图标，缩小图片大小。

：单击图标，放大图片大小。

：单击选择图片放大比例。

处理完后，单击“下一步”，进入[框选参照字段](#)步骤。

### 5.3.4 框选参照字段

在文字识别过程中，套件会检查所识别图片与模板图片是否为同一种模板，并将识别图片校正后再提取结构化信息，支持图片平移、旋转与拉伸变换。

为了检查并校正待识别的图片，这就需要在模板图片中指定参照字段。通过参照字段的文字内容来判断是否属于同一种模板，通过参照字段的位置来校正待识别图片。

#### 基本概念

参照字段为模板图片和待识别图片中的公共文字部分，所有需要识别的图片中都要包含参照字段，且位置必须固定。

套件提供了自动搜索参照字段和手动框选参照字段这两种模式。

- 自动搜索参照字段：未手动框选任何参照字段的情况下，默认激活自动搜索参照字段模式。
- 手动框选参照字段：若手动框选了任意参照字段，将激活手动框选模式。

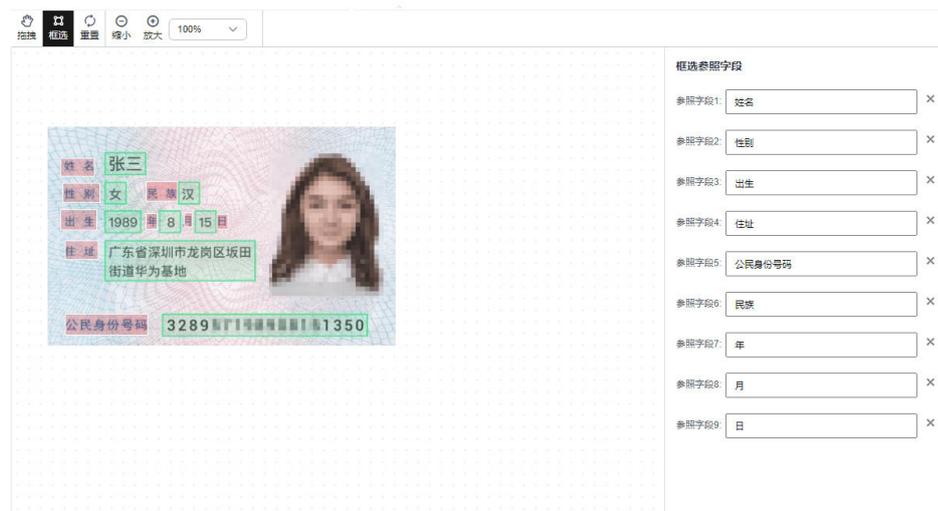
当识别图片的场景比较单一时，即只有一种模板，且参照字段不容易与其他文字混淆时，可以使用自动参照字段来简化模板制作过程，否则建议手动框选，详细步骤请参见[操作步骤](#)。

#### 前提条件

已在“自定义OCR控制台”选择“通用单模板工作流”创建应用，并对上传的模板图片进行预处理，详情请见[定义预处理](#)。

#### 操作步骤

图 5-9 框选参照字段



1. 在“应用开发>框选参照字段”页面，通过单击左上方操作图标，调整模板图片到合适大小。
2. 单击框选操作图标 ，单击参照字段左上角，移动鼠标，使得矩形框覆盖参照字段。

- 参照字段是文字内容、位置固定不变的文本框区域。
  - 参照字段为单行文本框，不可以框选竖版文字或跨行框选。
  - 框选参照字段个数须建议大于4个，越多越好，并尽量分散在图片的四周。
  - 参考字段尽量沿着文字边缘框选，精确框住对应文本行为佳。
3. 核对右侧“框选参照字段”中的参照字段是否与框选的参照字段一致。
  4. 框选并核对完所有参照字段后，单击“下一步”，进入[框选识别区](#)步骤。

### 5.3.5 框选识别区

在文字识别过程中，需要确定图片中识别的文字位置，这就需要在图片模板中框选识别区。

识别区指图片中待识别的文字位置。所有需要识别的图片中都会包含此识别区的字段，且位置固定不变，因此模型可以通过识别区找到需要识别内容的位置。

#### 前提条件

已在自定义OCR控制台选择“通用单模板 workflow”创建应用，并框选参照字段，详情请见[框选参照字段](#)。

#### 操作步骤

图 5-10 框选识别区



1. 在“应用开发>框选识别区”页面，通过单击左上方操作图标，调整模板图片到合适大小。
2. 单击框选操作图标 ，单击待识别文字的左上角，移动鼠标框选识别区，使得矩形框覆盖待识别的文字。  
框选识别区应尽量扩大识别区范围，使所框选识别区覆盖字段值可能出现的区域。
3. 在右侧“框选识别区”中填写“框选字段”，选择“字段类型”。  
“字段类型”指待识别文字的内容，您可以在默认字段类型中选择，当前可选择的默认字段类型包括“日期”、“小写金额”、“大写金额”、“数字”、“数字和英文”、“编号”、“性别”，每个识别区可选择多个字段类型。  
如果“默认字段类型”不能满足您的业务需求，您可以在字段类型下拉框单击“创建新字段类型”，[创建新字段类型](#)。

4. 框选并核对完所有识别区后，单击“下一步”，进入6.3.6-评估步骤。

## 5.3.6 评估

确定模板图片的参照字段和识别区后，需要对模板图片进行评估和考察。您可以通过上传测试图片，在线评估模板的识别情况，保证能正确识别同样模板下其他图片中的识别区文字。

### 前提条件

已在自定义OCR控制台选择“通用单模板 workflow”创建应用，并完成框选识别区步骤，详情请见[框选识别区](#)。

### 进入评估页面

登录“文字识别>自定义OCR”控制台，默认进入“应用开发>工作台”页面。选择应用并单击“操作”列的“查看”，进入“应用资产”页面。

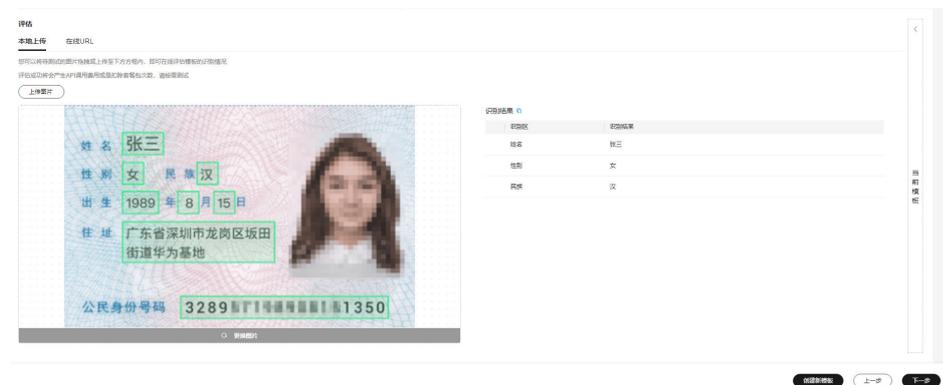
图 5-11 进入评估页面



- 在“应用资产”页面，单击已有模板操作列的“评估”，进入“应用开发>评估”页面。
- 在“应用资产”页面，单击已有模板操作列的“编辑”，进入“应用开发”页面，依次完成“上传模板图片”、“定义预处理”、“框选参照字段”、“框选识别区”步骤，单击“下一步”，进入“评估”页面。

### 本地上传图片

图 5-12 评估模板



1. 在“应用开发>评估”页面，默认进入“本地上传”页签。
2. 单击“上传图片”，或者拖拽测试图片至虚线框内上传图片区域，上传本地的图片作为测试图片。

3. 测试图片上传成功后，右侧会显示识别结果。  
您可以核对识别结果是否正确。  
如果不正确，可单击“创建新模板”，重新创建一个模板，或者单击“上一步”，对当前模板进行修改。  
如果识别结果正确，可单击“下一步”，部署模板。

## 上传在线图片

1. 在“应用开发>评估”页面，单击“在线URL”，切换至“在线URL”页签。

图 5-13 上传在线图片



2. 在“开始识别”左侧输出框中输入待测试的图片URL地址，或者拖拽测试图片至虚线框内上传图片区域。
3. 单击“开始识别”，右侧会显示识别结果。  
如果不正确，可单击“创建新模板”，重新创建一个模板，或者单击“上一步”，对当前模板进行修改。  
如果识别结果正确，可单击“下一步”，部署模板。

## 5.3.7 调用指南

评估模板应用后，就完成模板制作，开发属于自己的文字识别应用，此应用用于识别自己所上传的模板样式的图片。

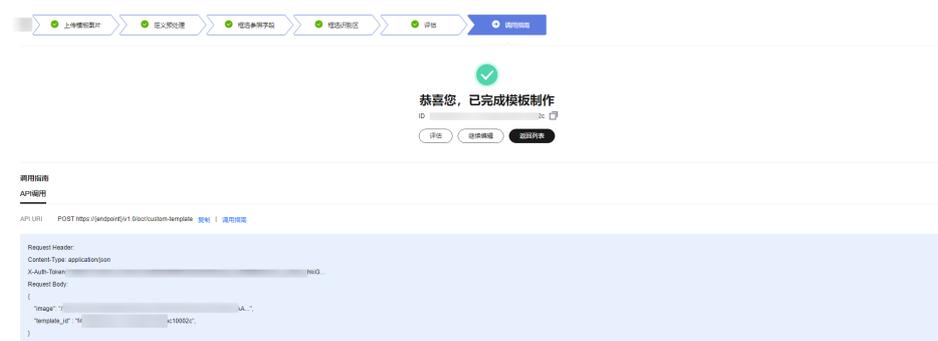
### 前提条件

已在自定义OCR控制台选择“通用单模板 workflows”创建应用，并完成评估模板步骤，详情请见[评估](#)。

### 操作步骤

在“应用开发>评估”页面完成模板评估后，单击“下一步”，进入“创建模板>调用指南”页面，页面显示“恭喜您，已完成模板制作”。

图 5-14 调用指南



完成模板制作后，在“应用开发>调用指南”页面，您可以进行如下操作。

“评估”：单击“评估”，继续上传测试图片评估模板。

“继续编辑”：单击“继续编辑”，返回“应用开发>框选参照字段”页面，继续编辑模板。

“返回列表”：单击“返回列表”，返回至应用详情页。

“调用API”：下方调用指南区域，可直接调用API或SDK轻松调用当前模板服务，详情请见调用API。

## 5.3.8 编辑应用

对于已经创建的模板应用，您可以修改模板的配置信息以匹配业务变化。

### 前提条件

已存在创建的模板应用。

### 编辑模板配置信息

1. 登录“文字识别>自定义OCR”控制台。  
默认进入“应用开发>工作台”页面。
2. 在“我的应用”页签下，选择应用并单击“操作”列的“查看”。  
进入“应用资产”页面。

图 5-15 进入应用资产



3. 在“模板列表”页签，选择模板名称，单击操作列的“编辑”。

图 5-16 模板列表



进入“应用开发”页面，您可以依次修改“上传模板图片”、“定义预处理”、“框选参照字段”、“框选识别区”、“评估”步骤的信息，重新编辑模板。操作指引如下：

- [上传模板图片](#)
- [定义预处理](#)
- [6.3.4-框选参照字段](#)

- [6.3.5-框选识别区](#)
- [6.3.6-评估](#)
- [6.3.7-调用指南](#)

## 5.3.9 删除应用

如果已创建的模板应用不再使用，您可以删除模板释放资源。

### 操作步骤

1. 登录“文字识别>自定义OCR”控制台。  
默认进入“应用开发>工作台”页面。
2. 在“我的应用”页签下，选择应用并单击“操作”列的“查看”。  
进入“应用资产”页面。

图 5-17 进入应用资产



3. 在“模板列表”页签，选择模板名称，单击操作列的“删除”。  
弹出“确认删除”对话框。

图 5-18 删除模板



4. 单击“确认”，删除模板。

## 5.4 多模板分类 workflow

### 5.4.1 workflow 介绍

#### workflow 简介

- 功能介绍  
支持用户自定义多个文字识别模板，通过模型分类，自动识别图片所需使用的模板，从而支持从大量不同板式图像中提取结构化信息。
- 适用场景
  - 用户认证识别  
识别证件中关键信息，节省人工录入，提升效率，降低用户实名认证成本，准确快速便捷。

- 快递单自动填写  
识别图片中联系人信息并自动填写快递单，减少人工输入。
- 合同录入与审核  
自动提取合同结构化信息，有助快速审核。
- 优势
  - 解决手工录入投入大、效率低、语种多等问题，提升业务效率。
  - 一键式部署，快速输出高精度结构化数据。
  - 解决单据复杂、单据板式多、语种多问题，支持自定义多个图像板式，快速适配新板式，快速接入业务。
  - 支持从多个不同板式图像中提取结构化信息。

## 工作流程

在“文字识别>自定义OCR”控制台选择“多模板分类 workflow”创建应用，详细操作请见[创建应用](#)。您可以创建多模板服务，通过多模板训练模板分类模型和文字识别模型，实现多模板图像的文字信息结构化提取。

图 5-19 创建多模板流程



表 5-3 创建多模板流程说明

流程	说明	详细指导
上传模板图片	在使用多模板 workflow 开发应用之前，需要上传模板图片，明确以哪些图片作为模板训练文字识别模型。	<a href="#">上传模板图片</a>
定义预处理	上传模板图片后，需要对模板图片进行预处理，去掉冗余部分，将图片旋转至水平，保证模型识别的准确性。	<a href="#">定义预处理</a>
框选参照字段	在图片模板中框选参照字段，用于矫正图片的方向，进而在正确的方向上，识别图片中的结构化信息。对于多模板应用而言，参照字段还将用于区分不同的模板。	<a href="#">框选参照字段</a>
框选识别区	在图片模板中框选识别区，确定模板图片中需要识别的文字位置。	<a href="#">框选识别区</a>
模板总览	检查所有模板是否已正确配置。	<a href="#">模板总览</a>
评估	通过上传测试图片，在线评估模板分类情况和模板的文字识别情况，保证能在多个模板情况下正确分类测试图片的模板，并且能正确识别测试图片中的识别区文字。	<a href="#">评估</a>

## 5.4.2 上传模板图片

在使用多模板分类 workflow 开发应用之前，必须要明确文字识别的模板类型，明确以哪几种板式图片作为模板训练文字识别模型，基于自己的业务需求制定针对性的文字识别模型。例如上传两种不同格式的发票图片作为模板，训练的文字识别模型就能识别并提取这两种格式发票上的关键字段。

### 前提条件

- 已在自定义OCR控制台选择“多模板分类 workflow”创建应用，详情请见[创建应用](#)。
- 提前准备模板图片以及训练分类器的数据集，其要求请见[数据要求](#)。

### 数据要求

- 保证图片质量：不能有损坏的图片；目前支持的格式包括JPG、JPEG、PNG、BMP。
- 图片最大边不大于4096px，最小边不小于100px，且大小不超过5M。
- 图片中的文字清晰可认。

### 进入应用开发页面

1. 登录“文字识别>自定义OCR”控制台。  
默认进入“应用开发>工作台”页面。
2. 在“我的应用”页签下，选择应用并单击“操作”列的“查看”。  
进入“应用资产”页面。

图 5-20 进入应用资产



3. 单击“创建分类器”。  
进入“应用开发”页面，开始开发应用。

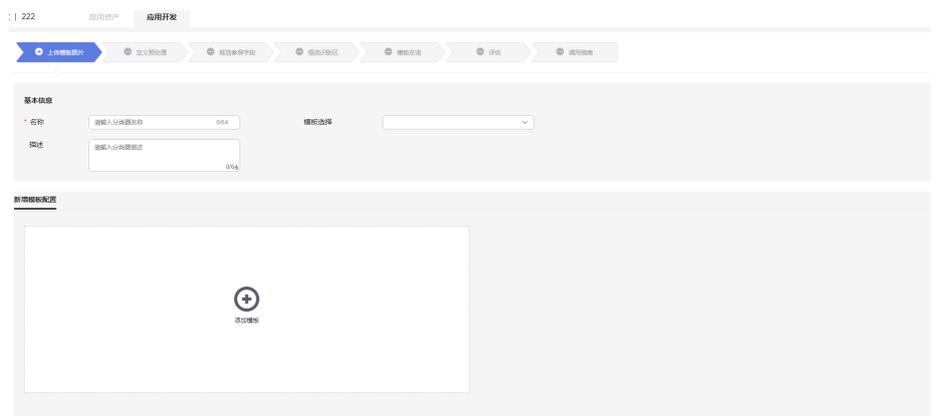
图 5-21 创建分类器



### 填写基本信息

在“应用开发>上传模板图片”页面，输入多模板分类器的“名称”和“描述”。

图 5-22 上传模板图片



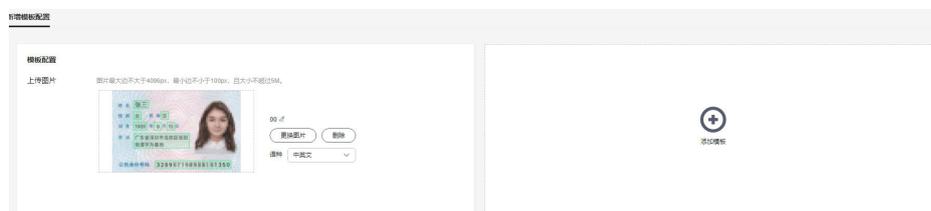
然后上传模板图片，可选择新增模板配置。

- **新增模板配置**

## 新增模板配置

默认进入“新增模板配置”页签。

图 5-23 新增模板配置



单击 $\oplus$ 添加多个模板，针对每个模板，选择模板类型，并且上传图片。

“上传图片”：单击“上传图片”区域，或鼠标直接拖拽图片至“上传图片”区域，上传本地一张图片作为模板，用于业务场景的文字结构化识别。

“修改模板名称”：单击图片右侧的 $\ell$ ，在弹出的输入框中输入新的模板名称，单击“确认”。

“删除模板”：单击图片右侧“删除”，在弹出的确认删除对话框中单击“确认”。

“语种”：单击图片右侧“语种”下拉选择框，选择模板对应的语种。不选择的情况下，默认为中英文。

确认信息后，单击“下一步”，进入**定义预处理**步骤，对上传的模板图片进行自动旋转、裁剪等预处理。

### 5.4.3 定义预处理

使用多模板分类 workflow 上传模板图片后，需要对模板图片进行预处理，通过旋转、裁剪等操作。图片预处理的目的是保留图片的关键内容，去掉冗余部分，保持图片内容清晰可见，保证模型识别的准确性。

## 前提条件

已在自定义OCR控制台选择“多模板分类 workflow”创建应用，并上传模板图片，详情请见[上传模板图片](#)。

## 定义预处理

在“应用开发>定义预处理”页面上方    或者“选择模板”下拉框中选择模板，对每个模板图片进行预处理。

图 5-24 多模板定义预处理



在右侧“选择预处理逻辑”区域勾选对应操作，当前仅支持“自动旋转”操作，系统自动旋转文字方向不正确的图片，保持图片中的文字方向正确。

单击预处理区域左上方的操作图标，调整模板图片的大小、方向等。

：单击图标重置图片为初始状态，即未进行任何处理的状态。

：单击图标，在“图片裁剪”窗口调整图片裁剪范围，然后单击“裁剪”，调整图片的大小。

：单击图标，在“图片旋转”窗口，鼠标选中  调整图片的方向，然后单击“确认”，调整图片的方向。

：单击图标，向左旋转图片90°。

：单击图标，向右旋转图片90°。

：单击图标，缩小图片大小。

：单击图标，放大图片大小。

：单击 ▾ 选择图片放大比例。

处理完后，单击“下一步”，进入[框选参照字段](#)步骤。

### 5.4.4 框选参照字段

在识别文字的过程中，套件会先对图片进行模板分类，再检查所识别图片与模板图片是否为同一种模板，并将识别图片校正后再提取结构化信息，支持图片平移、旋转与拉伸变换。

为了正确分类并校正待识别的图片，这就需要在模板图片中指定参照字段。通过参照字段的文字内容来判断识别图片属于哪一种模板，最后通过参照字段的位置来校正待识别图片。

## 基本概念

参照字段为模板图片和待识别图片中的公共文字部分，所有需要识别的图片中都要包含参照字段，且位置必须固定。

套件提供了自动搜索参照字段和手动框选参照字段这两种模式。

- 自动搜索参照字段：未手动框选任何参照字段的情况下，默认激活自动搜索参照字段模式。
- 手动框选参照字段：若手动框选了任意参照字段，将激活手动框选模式。

当识别图片的场景比较单一时，即只有一种模板，且参照字段不容易与其他文字混淆时，可以使用自动参照字段来简化模板制作过程，否则建议手动框选，详细步骤请参见[操作步骤](#)。

## 前提条件

已在自定义OCR控制台选择“多模板分类工作流”创建应用，并对上传的模板图片进行预处理，详情请见[定义预处理](#)。

## 操作步骤

图 5-25 多模板框选参照字段



1. 在“应用开发>框选参照字段”页面上方   或者“选择模板”下拉框中选择模板，对每个模板图片进行预处理。
2. 单击左上方操作图标，调整当前模板图片到合适大小和位置。
3. 单击框选操作图标 ，单击参照字段左上角，移动鼠标，使得矩形框覆盖参照字段。
  - 参照字段是文字内容、位置固定不变的文本框区域。
  - 参照字段为单行文本框，不可以框选竖版文字或跨行框选。
  - 框选参照字段个数须建议大于4个，越多越好，并尽量分散在图片的四周。
  - 参考字段尽量沿着文字边缘框选，精确框住对应文本行为佳。
4. 核对右侧“框选参照字段”中的参照字段是否与框选的参照字段一致。
5. 框选并核对完所有模板图片的参照字段后，单击“下一步”，进入[框选识别区](#)步骤。

## 5.4.5 框选识别区

在文字识别过程中，需要确定图片中识别的文字位置，这就需要在图片模板中框选识别区。

识别区指图片中待识别的文字位置。所有需要识别的图片中都会包含此识别区的字段，且位置固定不变，因此模型可以通过识别区找到需要识别内容的位置。

### 前提条件

已在自定义OCR控制台选择“多模板分类 workflow”创建应用，并框选参照字段，详情请见[框选参照字段](#)。

### 操作步骤

图 5-26 多模板框选识别区



1. 在“应用开发>框选参照字段”页面上方 或者“选择模板”下拉框中选择模板，对每个模板图片进行预处理。
2. 单击左上方操作图标，调整当前模板图片到合适大小。
3. 单击框选操作图标 ，单击待识别文字的左上角，移动鼠标框选识别区，使得矩形框覆盖待识别的文字。  
框选识别区应尽量扩大识别区范围，使所框选识别区覆盖字段值可能出现的区域。
4. 在右侧“框选识别区”中填写“框选字段”，选择“字段类型”。  
“字段类型”指待识别文字的内容，您可以在默认字段类型中选择，当前可选择的默认字段类型包括“日期”、“小写金额”、“大写金额”、“数字”、“数字和英文”、“编号”、“性别”，每个识别区可选择多个字段类型。  
如果“默认字段类型”不能满足您的业务需求，您可以在字段类型下拉框单击“创建新字段类型”，[自定义字段类型](#)。
5. 框选并核对完所有模板图片的识别区后，单击“下一步”，进入[模板总览](#)步骤。

## 5.4.6 模板总览

预览分类器下面所有模板的状态，如果有模板的识别区未设置，则无法进行后续的评价。

## 前提条件

已在自定义OCR控制台选择“多模板分类工作流”创建应用，并框选添加的多个模板的识别区，详情请见[框选识别区](#)。

## 删除模板

如果已创建的模板应用不再使用，您可以删除模板释放资源。

在“应用开发>模板总览”页面，选择“模板”，单击操作列的“删除”，弹出“确认删除”对话框，单击“确认”，删除模板。

## 5.4.7 评估

模板总览过后，需要对模板分类器和模板图片进行评估和考察。您可以通过上传测试图片，在线评估模板分类情况和模板的文字识别情况，保证能在多个模板情况下正确分类测试图片的模板，并且能正确识别测试图片中的识别区文字。

## 前提条件

已在自定义OCR控制台选择“多模板分类工作流”创建应用，并已框选识别，详情请见[模板总览](#)。

## 进入评估页面

登录“文字识别>自定义OCR”控制台，默认进入“应用开发>工作台”页面。在“我的应用”页签下，选择应用并单击操作列的“查看”，进入应用详情页，可从“应用资产”和“应用开发”两个页签进入评估模板页面。

- 在“应用资产”页面，默认进入“分类器列表”页签，单击已有模板操作列的“评估”。进入“评估”页面。

图 5-27 评估



- 在“应用开发”页面，依次完成“上传模板图片”、“定义预处理”、“框选参照字段”、“框选识别区”、“模板总览”步骤，单击“下一步”，进入“评估”页面。

图 5-28 评估

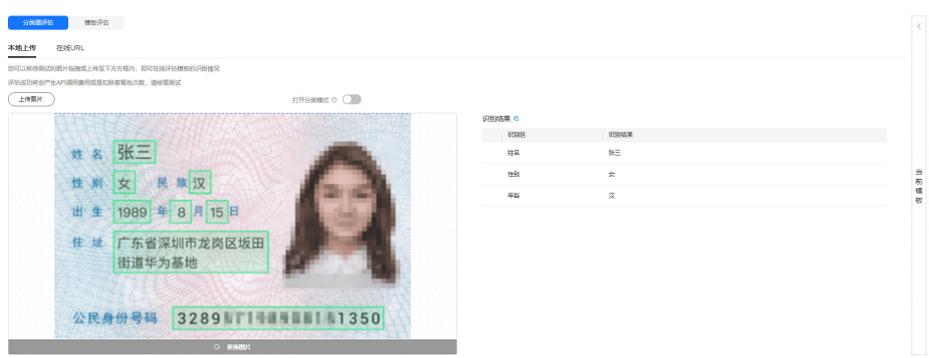


## 评估分类器

通过上传测试图片，在线评估模型训练的模板分类器是否能正确分类模板。

在“应用开发>评估”页面，您可以选择本地上传测试图片，也可以上传在线图片。

图 5-29 评估分类器



- 本地上传图片

默认进入“本地上传”页签，根据自身需要选择打开“分类模式”开关，单击“上传图片”，或者拖拽测试图片至虚线框内上传图片区域，上传本地的图片作为测试图片。

“分类模式”：打开“分类模式”开关时，单独对分类器的准确度进行评估。上传图片后，右侧会显示模板自动分类的结果，包括“模板ID”、“模板名”、“置信度”。关闭“分类模式”开关时，默认评估状态为端到端地对待识别图片自动分类并进行结构化识别。

- 上传在线图片

单击“在线URL”，切换至“在线URL”页签。在“开始识别”左侧输出框中输入待测试的图片URL地址，或者拖拽测试图片至虚线框内上传图片区域，上传在线图片作为测试图片。

上传图片后，右侧会显示模板识别结果，包括“模板ID”、“模板名”、“置信度”。

上传图片后，您可以核对识别结果是否正确。

如果不正确，单击“上一步”，对当前模板进行修改。

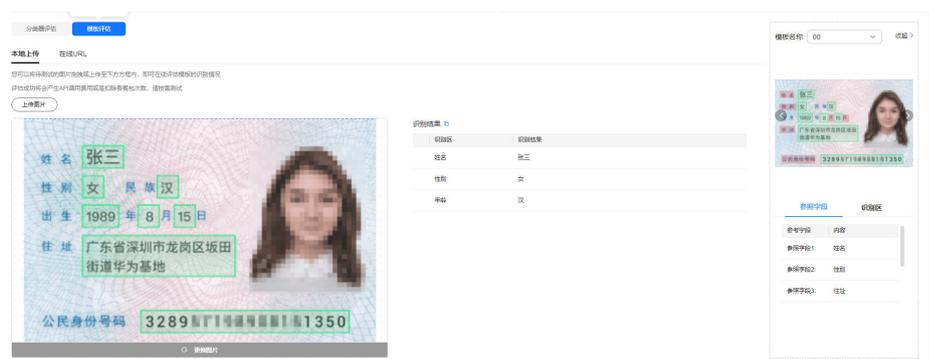
如果识别结果正确，可对模板作进一步评估，详情请见[评估模板](#)。

## 评估模板

通过上传测试图片，在线评估模板识别并提取结构化文字的能力。

首先在“应用开发>评估”页面右侧选择与待评估模板对应的原模板，然后选择本地上传测试图片或者上传在线图片，评估模板。

图 5-30 评估模板



- **本地上传图片**

默认进入“本地上传”页签，单击“上传图片”，或者拖拽测试图片至虚线框内上传图片区域，上传本地的图片作为测试图片。

上传图片后，右侧会显示文字识别结果，包括“识别区”和对应的“识别结果”。

- **上传在线图片**

单击“在线URL”，切换至“在线URL”页签。在“开始识别”左侧输出框中输入待测试的图片URL地址，或者拖拽测试图片至虚线框内上传图片区域，上传在线图片作为测试图片。

上传图片后，右侧会显示文字识别结果，包括“识别区”和对应的“识别结果”。

上传图片后，您可以核对识别结果是否正确。

如果不正确，单击“上一步”，对当前模板进行修改。

如果识别结果正确，可单击“下一步”，部署模板。

## 后续操作

评估分类器和模板之后，如果识别结果正确，可单击“下一步”，[调用指南](#)

### 5.4.8 调用指南

评估模板应用后，就完成模板制作，开发属于自己的文字识别应用，此应用用于识别自己所上传的图片属于哪种模板以及识别图片中的文字。

## 前提条件

已在自定义OCR控制台选择“多模板分类 workflow”创建应用，并完成评估模板步骤，详情请见[评估](#)。

## 操作步骤

在“应用开发>调用指南”页面完成模板评估后，单击“下一步”，进入“创建模板>调用指南”，页面显示“恭喜您，已完成模板制作”。

图 5-31 调用指南



在“应用开发>调用指南”页面，您可以进行如下操作。

“评估”：单击“评估”，继续上传测试图片评估模板。

“继续编辑”：单击“继续编辑”，返回“应用开发>框选参照字段”页面，重新编辑模板。

“返回列表”：单击“返回列表”，返回至应用详情页。

“调用API”：下方调用指南区域，提供了调用当前模板服务的请求样例，详情请见[调用API](#)。

## 5.4.9 编辑应用

对于已经创建的模板应用，您可以修改模板的配置信息以匹配业务变化。

### 前提条件

已存在创建的模板应用。

### 编辑模板配置信息

1. 登录“文字识别>自定义OCR”控制台。  
默认进入“应用开发>工作台”页面。
2. 在“我的应用”页签下，选择应用并单击“操作”列的“查看”。  
进入“应用资产”页面。

图 5-32 进入应用资产



3. 在“模板列表”页签，选择模板名称，单击操作列的“编辑”。

图 5-33 编辑模板



进入“应用开发”页面，您可以依次修改“上传模板图片”、“定义预处理”、“框选参照字段”、“框选识别区”、“模板总览”、“评估”、“调用指南”步骤的信息，重新部署模板。操作指引如下：

- [上传模板图片](#)
- [定义预处理](#)
- [框选参照字段](#)
- [框选识别区](#)
- [模板总览](#)

- [评估](#)
- [调用指南](#)

## 5.4.10 删除应用

如果已创建的模板应用不再使用，您可以删除模板释放资源。

### 操作步骤

1. 登录“文字识别>自定义OCR”控制台。  
默认进入“应用开发>工作台”页面。
2. 在“我的应用”页签下，选择应用并单击“操作”列的“查看”。  
进入“应用资产”页面。

图 5-34 进入应用资产



3. 在“分类器列表”页签，选择多模板名称，单击操作列的“删除”。  
弹出“确认删除”对话框。

图 5-35 删除分类器



4. 单击“确认”，删除分类器。

## 5.5 字段类型

在“框选识别区”页面，可以给每个字段指定字段类型来提高识别准确率，或是将文字识别结果处理成业务需要的格式。勾选字段类型后，后台会对文字识别结果进行相应的处理后再输出。每个识别区可选择多个字段类型，多个字段类型会按照勾选顺序执行。自定义OCR控制台提供了默认字段类型和自定义字段类型两种功能。

图 5-36 默认字段类型和自定义字段类型



### 5.5.1 默认字段类型

当前可选择的默认字段类型包括“日期”、“小写金额”、“大写金额”、“数字”、“数字和英文”、“编号”、“性别”。选择特定的字段类型后，将会按照指定的字段类型对字符串处理后再输出。比如，未激活任何字段类型时的输出为“发证日期：2021年09月12日”，为了精准提取日期信息，可以勾选“日期”类型，将得到输出“2021年09月12日”。每个字段类型所做的处理如下：

- “日期”：将字符串处理成日期格式后再输出，支持“yyyy年mm月dd日”，“yyyy-mm-dd”常见的日期格式。
- “小写金额”：将字符串处理成小写金额格式，比如“¥1.56”。
- “大写金额”：将字符串处理成大写金额格式，比如“壹圆伍角陆分”。
- “数字”：仅保留字符串中的数字。
- “数字和英文”：仅保留字符串中的数字或英文字母。
- “编号”：仅保留字符串中的数字、英文字母或特殊符号“.-/()\_<”。
- “性别”：仅保留性别的缩写，“男”、“女”、“M”、“F”。

### 5.5.2 自定义字段类型

如果“默认字段类型”不能满足您的业务需求，您可以创建新的字段类型。

1. 在“应用开发>框选识别区”页面单击框选操作图标，单击待识别文字的左上角，移动鼠标框选识别区。
2. 在右侧“框选识别区”中的“字段类型”下拉框中，单击“创建新字段类型”。

- 弹出“创建字段类型”对话框。
- 填写字段类型信息。

图 5-37 创建字段类型

“**字段类型名称**”：填写待创建的字段类型名称，如识别身份证照片，可新增字段类型“出生日期”。

“**字段属性**”：选择字段属性，可选择“自定义正则提取”和“自定义字典”，然后按表5-4填写相关信息。

表 5-4 字段属性参数说明

字段属性	参数	参数说明	举例
自定义字典	取值范围	在识别当前字段类型的文字时，选择字典的取值范围中最相似的取值为最后识别结果。	例如： “ <b>字段类型名称</b> ”： “城市” “ <b>字段属性</b> ”：“自定义字典” “ <b>取值范围</b> ”： “Shenzhen”、 “Beijing”、 “Xi'an” 在识别“ <b>字段类型</b> ”为“城市”的文字时，从“Shenzhen”、“Beijing”、“Xi'an”中找出与待识别文字最相似的取值，作为识别结果。

字段属性	参数	参数说明	举例
自定义正则提取	预过滤	<p>对初始的待识别文字进行预处理。</p> <p>左边输入框填写待识别文字中被替换字符的正则表达式。</p> <p>右边输入框填写所替换的新字符。</p> <p>不填写时，默认不做预处理。</p> <p>如果需要多次预处理，可单击 ，填写新增的预处理规则。</p>	<p>例如：</p> <p>“字段类型名称”： “出生日期”</p> <p>“字段属性”：“自定义正则提取”</p> <p>“预过滤”：“” / “”</p> <p>“提取”：“\d+年\d+月\d+日”</p> <p>“后处理”：“年” / “.”、“月” / “.”、 “日” / “”</p> <p>在识别字段类型为出生日期的文字“2020年1月1日出生”时，首先不做预过滤，然后提取关键字“2020年1月1日”，最后做后处理，最终提取内容为“2020.1.1”。</p>
	提取	<p>对经过“预处理”的文字进行关键字提取。</p> <p>在输入框中填写查找关键字的正则表达式。</p> <p>不填写时，默认提取全部字段。</p> <p>如果需要多个提取规则，单击  新增提取规则。提取时按从上到下优先级规则提取，选择第一个非空的提取内容作为提取后的内容。</p>	
	后处理	<p>根据实际情况，对经过“提取”后的文字进行“后处理”。</p>	

## 5.6 查看应用详情

创建应用后，您可以在“文字识别->自定义OCR”控制台查看应用详情，包括应用开发的配置信息和应用资产。

### 前提条件

已在“自定义OCR控制台”选择预置 workflow 创建应用，详情请见[创建应用](#)。

### 进入应用详情页

1. 登录“文字识别”管理控制台，单击左侧菜单的“自定义OCR”。  
进入自定义OCR控制台。默认进入“我的应用”页签。
2. 在“我的应用”页签下，选择已创建的应用，单击操作列的“查看”。进入应用详情页，默认进入“应用资产”页面，查看应用资产。  
您也可以单击“应用开发”，切换至“应用开发”页面，查看应用开发配置。

图 5-38 进入应用资产



## 查看应用资产

在“应用资产”页面下，可以查看当前应用的资产信息。

使用预置 workflow 开发应用，“应用资产”可查看“模板列表”和“自定义字段类型”。

图 5-39 应用资产



在“应用资产>模板列表”页面下可以查看已创建的模板列表，操作列可执行如下操作：

- “**评估**”：单击“评估”，进入“应用开发>评估”页面，对当前模板进行评估，详情请见[评估](#)。
- “**编辑**”：单击“编辑”，进入应用开发页面，对当前模板开发的配置信息进行编辑，通用单模板 workflow 操作指引请见[工作流介绍](#)，多模板 workflow 操作指引请见[工作流介绍](#)。
- “**删除**”：单击“删除”，弹出“确认删除”对话框，单击“确认”，删除当前模板。

在“应用资产>自定义字段类型”页面下，可以查看自己定义的字段类型，也可单击“创建字段类型”，在此处创建新的字段类型，详情请见[自定义字段类型](#)。

## 5.7 常见问题

### 5.7.1 什么情况下可以使用自定义模板？

自定义模板是一款用于提取卡证、票据和文档图片中自定义区域信息的文字识别产品。当您识别的证件或票据种类不在[华为云 OCR API](#)支持之中，并且识别图片的版式固定，则可以使用本产品提取自定义版式、自定义区域的文字信息。比如，目前华为云尚未推出毕业证书识别 API，则可以利用自定义模板打造出毕业证书识别接口，提取毕业证上的文字信息。

自定义模板支持识别单模板应用或是多模板应用，单模板仅支持一种版式，多模板支持识别多种版式。

### 单模板的前提:

- 识别图片需与模板图片版式相同，具体而言，图片中要存在文字内容和位置均固定不变的文字（参照字段）。
- 识别文字内容不能偏离设定的识别区域，包括打印偏移、由上下文长度变化引起的偏移等。

### 多模板的前提:

- 分类器中的任一模板，均需满足上述单模板的要求。
- 模板种类固定，识别不在多模板中的图片可能会返回错误的结果或无法识别。
- 模板间有足够的区分度，不同模板要有不同的参照字段，且参照字段差异尽量明显，相似的参照字段容易导致分类错误。
- 模板数量可穷举，建议不超过20种模板，过多的模板会降低分类效果（以实际测试为准，模板差异较大时允许支持更多模板）。

## 5.7.2 如何选取参照字段?

参照字段是在所有图片中，文字位置和内容均不发生变化的文字。

参照字段有两个作用:

- 在单模板应用中，用于矫正识别图片，从而找准识别字段；
- 在多模板应用中，参照字段的内容和位置将作为相应模板的分类特征。

在框选参照字段时，首先要确保所框选的文字位置和内容都固定不变，如果不满足此要求，可能会提取到错误的识别结果，或是返回AIS.0119（输入图片与模板匹配失败）、AIS.0120（输入图片分类失败）。

为了获得更好的识别效果:

- 尽可能多框选参照字段，建议不少于4个参照字段，并尽量分散在四周。
- 在多模板应用中，为了获得较好的分类效果，还应该框选各个模板中有独有的参照字段。

如果不确定哪些文字是参照字段，或是想快速评估模板效果的，可以跳过“框选参照字段”流程，当后台检测到模板没有配置任何参照字段时，会自动寻找合适的参照字段进行识别，此功能可以提高项目开发与验证的效率。

## 5.7.3 识别结果容易漏字或多识别出内容怎么办?

识别结果是根据识别区来进行提取的，识别区的位置和大小均会影响识别结果。

- 若结果漏字，可能是由于识别区太小导致的，需在相应模板的“框选识别区”页面，把漏字的识别区调大一些。
- 若多识别出文字，可能是识别区太大，将周边无关的文字也框进来了，需将识别区改小一点。
- 若上述办法均无法解决，请检查识别区的文字是否发生了偏移，或者识别图片是否跟模板属于相同版式。

## 5.7.4 识别字符出现错误如何优化?

不同的错误情形需要具体分析:

情形1：大部分文字识别正确，部分形近符号的识别错误。

解决办法：产品提供了预置字段类型可以对结果进行处理，同时也提供了自定义（正则）类型、字典类型，用于纠正识别结果中的错误，适用范围详见[字段类型](#)。此外，您也可以在调用程序中使用字符串替换、正则抽取等规则来修正识别结果。

情形2：单模板工作流识别出错，且大部分字段为空或错误。

解决办法：请检查参照字段是否框选正确，或尝试框选更多参照字段（框选办法参照[框选参照字段](#)），如果新增参照字段后仍有问题，请检查识别图片与模板图片是否为同一个版式，如果为新版式，请考虑创建新模板或使用多模板分类工作流。

情形3：多模板识别工作流出错，大部分字段为空或错误。

解决办法：先检查分类结果是否正确（返回结果中的template\_id表示分类结果），若分类错误，则说明分类器中存在相似的模板导致分类出错，请在参照字段中添加模板特有的参照字段，从而提高分类的鉴别能力。若分类结果正确，但识别结果错误，请参照情形2的解决办法。

## 5.7.5 自定义模板 OCR 取消开通后，是否会将之前创建好的模板删除？

自定义模板OCR取消开通后，模板不会被删除，后期如果需要再次使用可以开通后继续使用即可。

开通服务可选择按需购买（开通管理），也可购买套餐包，详情请参考[开通文字识别服务](#)。

图 5-40 开通自定义模板 OCR



# 6 调用 API 或 SDK

## 6.1 在线调试

**API Explorer** 在线调试工具提供 API 的检索、调试、代码示例生成功能。同时，集成开发环境 CloudIDE，可完成代码的构建、调试、运行。

本章节以**通用文字识别**为例，介绍如何使用 API Explorer 调试 API。

### 前提条件

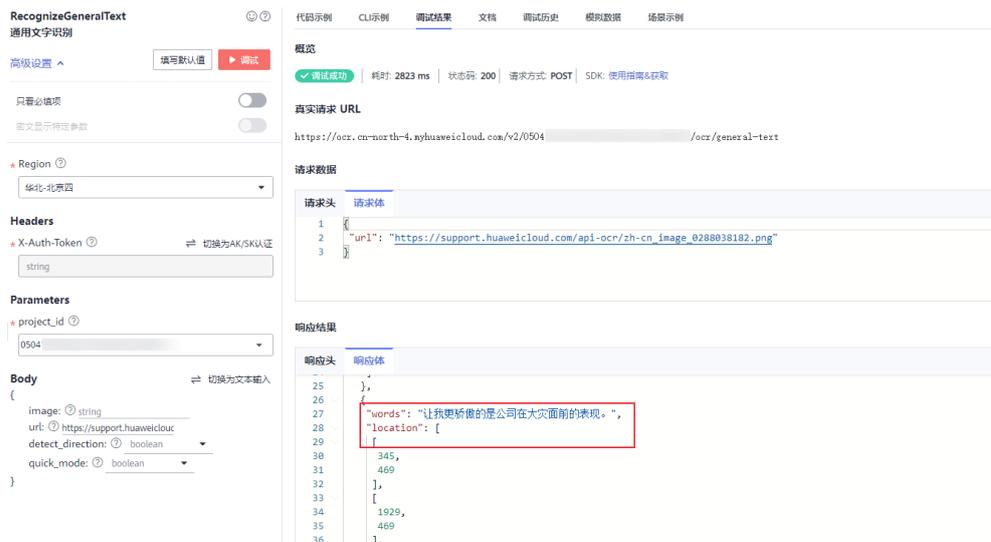
- 注册华为账号并开通华为云，并完成实名认证，账号不能处于欠费、冻结、被注销等异常状态。
- 已**开通通用文字识别**。

### 操作步骤

1. 登录 **API Explorer**。  
登录后，“X-Auth-Token”和“project\_id”参数会自动填充，无需填写。
2. 填写待检测图片数据。  
输入数据可选 image 或 url，本示例使用 url，例如 `https://support.huaweicloud.com/api-ocr/zh-cn_image_0288038182.png`。

```
image: ? string  
url: ? https://support.huaweicloud
```

3. 单击“调试”按钮，获取识别结果。



The screenshot displays the 'RecognizeGeneralText' interface. On the left, the 'Body' section is configured with the following parameters:

```
{
  "image": "string",
  "url": "https://support.huaweicloud.com/api-ocr/zh-cn_image_028803182.png",
  "detect_direction": "boolean",
  "quick_mode": "boolean"
}
```

The right-hand side shows the 'Request Data' and 'Response Results' sections. The request is a POST to the URL: `https://ocr.cn-north-4.myhuaweicloud.com/v2/0504/ocr/general-text`. The response body is a JSON object:

```
{
  "words": "让我更骄傲的是公司在灾面前的表现。",
  "location": [
    [
      345,
      469
    ],
    [
      1929,
      469
    ]
  ]
}
```

The 'words' field is highlighted with a red box in the original image.

## 6.2 本地调用

本章节以**通用表格识别**为例，介绍如何使用OCR Python SDK在本地进行开发。

该接口可以识别表格图片中的文字内容，并将识别结果以JSON格式返回给用户。返回结果包含两类：纯文本区（text）和表格区（table），并返回表格结构（row, column）和文本信息。

### 前提条件

- 注册华为账号并开通华为云，并完成实名认证，账号不能处于欠费、冻结、被注销等异常状态。
- 已**开通通用表格识别**。

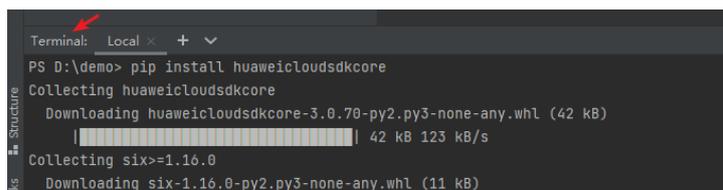
### 操作步骤

1. 获取并安装OCR Python SDK。

以PyCharm为例，介绍如何安装SDK。

在PyCharm中新建一个项目，并单击左下方“Terminal”按钮。分别执行以下命令安装SDK（该SDK支持Python3及以上版本）。

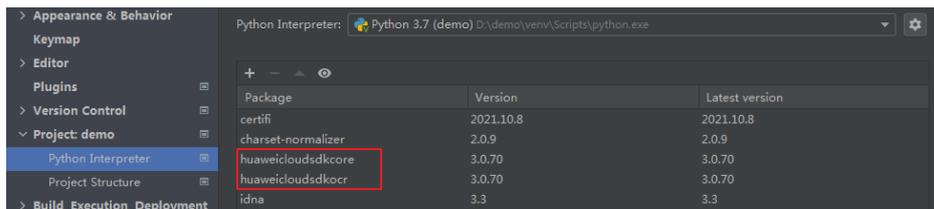
```
# 安装核心库
pip install huaweicloudsdkcore
# 安装OCR服务库
pip install huaweicloudsdkocr
```



```
Terminal: Local x + v
PS D:\demo> pip install huaweicloudsdkcore
Collecting huaweicloudsdkcore
  Downloading huaweicloudsdkcore-3.0.70-py2.py3-none-any.whl (42 kB)
    |#####| 42 kB 123 kB/s
Collecting six>=1.16.0
  Downloading six-1.16.0-py2.py3-none-any.whl (11 kB)
```

2. 检查SDK是否已正确安装。

在PyCharm中选择“File > Settings > Project > Python Interpreter”，检查SDK是否已正确安装。

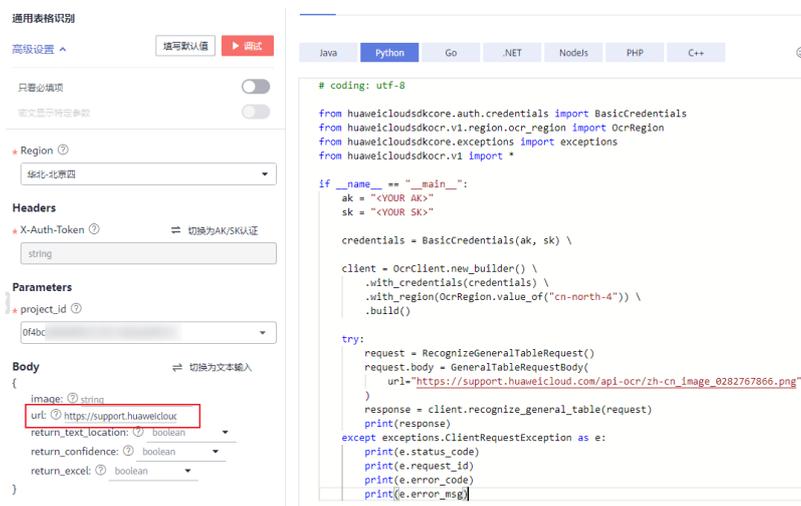


3. 获取通用表格识别SDK示例代码。

a. 登录API Explorer，在“代码示例”中选择“Python”。



b. 填写请求Body参数。输入数据可选image或url，本示例使用url，例如 [https://support.huaweicloud.com/api-ocr/zh-cn\\_image\\_0282767866.png](https://support.huaweicloud.com/api-ocr/zh-cn_image_0282767866.png)。



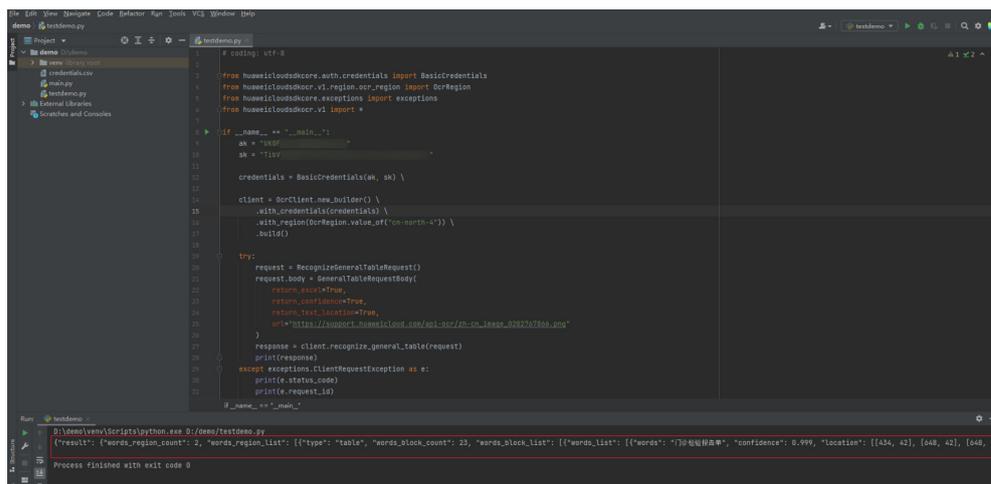
c. 复制代码示例至PyCharm中。

4. 获取AK/SK，替换代码示例中的“<YOUR AK>”、“<YOUR SK>”参数。

登录访问密钥页面，新增访问密钥，或使用已有的访问密钥。访问密钥为credentials.csv文件，包含AK/SK信息。

A	B	C	D	E
User Name	Access Key	Secret Access Key		
testuser	LSKM	rIZaQ		
	AK	SK		

5. 运行代码示例，获取识别结果。



```
demo testdemo.py
1 # coding: utf-8
2
3 from huaweicloudsdkcore.auth.credentials import BasicCredentials
4 from huaweicloudsdkocr.v1.region_ocr_region import OcrRegion
5 from huaweicloudsdkcore.exceptions import exceptions
6 from huaweicloudsdkocr.v1 import *
7
8 if __name__ == '__main__':
9     ak = 'AK'
10    sk = 'TSDF'
11
12    credentials = BasicCredentials(ak, sk)
13
14    client = OcrClient.new_builder() \
15        .with_credentials(credentials) \
16        .with_region(OcrRegion.value_of('cn-north-4')) \
17        .build()
18
19    try:
20        request = RecognizeGeneralTableRequest()
21        request_body = GeneralTableRequestBody(
22            source=True,
23            return_confidence=True,
24            return_text_location=True,
25            url="https://www.huaweicloud.com/en-us/qa/faq_428726366.pdf"
26        )
27        response = client.recognize_general_table(request)
28        print(response)
29    except exceptions.ClientRequestException as e:
30        print(e.status_code)
31        print(e.request_id)
32
33 if __name__ == '__main__':
34     sys.exit(0)
```

Run testdemo.py

```
D:\demo\venv\scripts\python.exe D:\demo/testdemo.py
{"result": {"words_region_count": 2, "words_region_list": [{"type": "table", "words_block_count": 23, "words_block_list": [{"words_list": [{"word": "\u7b80\u4f1d\u79ef\u52a0\u673a", "confidence": 0.999, "location": [{"x1": 436, "y1": 42, "x2": 648, "y2": 77}]}]}]}
```

Process finished with exit code 0

# 7 获取识别结果

## 7.1 解读识别结果

本章节通过[网络图片识别API](#)介绍如何解读调API返回的JSON格式识别结果。请参照[API参考](#)“响应参数”章节比对查看。

以下图识别结果为例，讲解图片内容如何与API的返回字段对应。



调用网络图片API成功后，在“JSON返回结果”中，可见result字段，该字段包含了图片中的文字块数目、文字块排列顺序、具体文本内容、所在位置、置信度等信息。

- result字段仅在API调用成功后返回。
- words\_block\_count表示文字块识别结果，本示例中，识别出2个文字块，分别代表图片中的2行文字。
- words\_block\_list表示文字块列表，按照图片文字从上到下、从左到右排列。
- words为识别出来的文本内容，confidence表示该字段的置信度。
- location表示文字块的区域位置信息，采用图像坐标系，第一个文字块左上角顶点为[140,45]。
- extracted\_data表示图片中包含的联系人信息，仅在设置了入参extract\_type，且图片中有联系人信息时支持提取。

```
{  
  "result": {  
    "words_block_count": 2,  
    "words_block_list": [  

```

```
{
  "words": "真材实料/舒柔养足",
  "confidence": 0.9491,
  "location": [
    [
      140,
      45
    ],
    [
      380,
      45
    ],
    [
      380,
      75
    ],
    [
      140,
      75
    ]
  ]
},
{
  "words": "选自优质头层牛皮，皮质柔软富弹性，耐折抗拉伸，穿着不变形",
  "confidence": 0.9897,
  "location": [
    [
      58,
      84
    ],
    [
      462,
      84
    ],
    [
      462,
      102
    ],
    [
      58,
      102
    ]
  ]
}
],
"extracted_data": {}
}
```

## 7.2 识别结果后处理

### 提取特定字段导入 Excel

本示例调用身份证识别API，并从获取到的JSON结果中，提取所需的字段，填入至Excel。

- 前提条件
  - [开通身份证识别](#)。
  - 参考[本地调用](#)，安装OCR Python SDK。并执行`pip install xlsxwriter`命令安装依赖包。
  - 登录[访问密钥](#)页面，获取AK、SK。可以新增访问密钥，或使用已有的访问密钥。访问密钥为credentials.csv文件，包含AK/SK信息。

A	B	C	D	E
User Name	Access Key	Secret Access Key		
testuser	LSKM	CrIZaQ		
	AK	SK		

- 代码示例

```
# -*- coding: utf-8 -*-
import base64
import xlswriter

from huaweicloudsdkcore.auth.credentials import BasicCredentials
from huaweicloudsdkocr.v1.region.ocr_region import OcrRegion
from huaweicloudsdkcore.exceptions import exceptions
from huaweicloudsdkocr.v1 import *
from huaweicloudsdkcore.http.http_config import HttpConfig

def recognize_id_card_request():
    try:
        request = RecognizeIdCardRequest()
        request.body = IdCardRequestBody(
            image=image_base64
        )
        response = client.recognize_id_card(request)
        return response
    except exceptions.ClientRequestException as e:
        print(e.status_code)
        print(e.request_id)
        print(e.error_code)
        print(e.error_msg)

def get_credential():
    return BasicCredentials(ak, sk)

def get_client():
    config = HttpConfig.get_default_config()
    config.ignore_ssl_verification = True
    return OcrClient.new_builder(OcrClient) \
        .with_credentials(credentials) \
        .with_region(OcrRegion.CN_NORTH_4) \
        .with_http_config(config) \
        .build()

def image_to_base64(imagepath):
    """
    将本地图片转化为base64编码
    """
    with open(imagepath, "rb") as bin_data:
        image_data = bin_data.read()
        base64_data = base64.b64encode(image_data).decode("utf-8")
    return base64_data

def response_to_excel(save_file, data):
    """
    :param save_file: 文件名
    :param data: result data
    """
    # 处理调用API返回的结果数据
    keys_list = list(data["result"].keys())
    values_list = list(data["result"].values())

    options = {'in_memory': True}
    with xlswriter.Workbook(save_file, options) as workbook:
        worksheet = workbook.add_worksheet()
```

```
worksheet.set_column('A1:A20', 23)
worksheet.set_column('B1:B20', 100)
worksheet.write_column('A1', keys_list)
worksheet.write_column('B1', values_list)
workbook.close()

if __name__ == '__main__':
    # 填写访问密钥AK、SK
    ak = "填写AK"
    sk = "填写SK"

    # Init Auth Info
    credentials = get_credential()

    # Create OcrClient
    client = get_client()

    image_base64 = image_to_base64(r"图片的路径, 例如D:\local\test.png")

    # request id card service
    response = recognize_id_card_request().to_dict()

    # 数据保存在excel上
    response_to_excel(r"excel路径, 例如D:\local\test.xlsx", response)
```

## 使用 PDF 进行文字识别

本示例将PDF文件转换为图片，并调用网络图片识别API，获取识别结果。

- 前提条件
  - [开通网络图片识别](#)。
  - 参考[本地调用](#)，安装OCR Python SDK。并执行**pip install fitz**命令和**pip install PyMuPDF==1.18.0**命令安装依赖包。
  - 登录[访问密钥](#)页面，获取AK、SK。可以新增访问密钥，或使用已有的访问密钥。访问密钥为credentials.csv文件，包含AK/SK信息。

A	B	C	D	E
User Name	Access Key	Secret	Access	Key
testuser	LSKM	CrIZaQ		
	AK	SK		

- 代码示例

```
# -*- coding: utf-8 -*-
import os
import base64
import fitz
import io
from PIL import Image
from glob import glob

from huaweicloudsdkcore.auth.credentials import BasicCredentials
from huaweicloudsdkocr.v1.region.ocr_region import OcrRegion
from huaweicloudsdkcore.exceptions import exceptions
from huaweicloudsdkocr.v1 import *
from huaweicloudsdkcore.http.http_config import HttpConfig

class CovertPdfToJpg:
    def __init__(self, file_path, save_root):
        self.file_path = file_path
        self.save_root = save_root

    @staticmethod
```

```
def open_pdf(file):
    return fitz.open(file)

@staticmethod
def get_trans(doc, page, min_side=0, max_side=0, rotate=0.0):
    """ 创建scale对象 """
    region = doc[page].rect
    scale = 1
    if max_side > min_side > 0:
        scale = min_side / min(region.width, region.height)
        if max(region.width, region.height) * scale > max_side:
            scale = max_side / max(region.width, region.height)
    trans = fitz.Matrix(scale, scale).preRotate(rotate)
    return trans

def page2pix(self, doc, page, trans):
    """ 根据给定的参数，解析当前页作为图片数据 """
    return doc[page].getPixmap(matrix=trans, alpha=False)

def pdf_to_jpg(self, width=1024, height=1400):
    """ 将pdf格式图片，转化为jpg格式 """
    doc = self.open_pdf(self.file_path)
    save_dir = os.path.join(self.save_root)
    if not os.path.exists(save_dir):
        os.makedirs(save_dir)
    print("document", len(doc), doc.pageCount)
    for i in range(len(doc)):
        trans = self.get_trans(doc, i, width, height, rotate=0)
        try:
            pdf = self.page2pix(doc, i, trans)
        except:
            continue
        image = pdf.getPNGData()
        image = Image.open(io.BytesIO(image))
        print(os.path.join(
            save_dir, os.path.basename(self.file_path).replace('.pdf', '') + '_' + str(i + 1) + '.jpg'))
        image.save(
            os.path.join(save_dir, os.path.basename(self.file_path).replace('.pdf', '') + '_' + str(i + 1) +
'.jpg'))
    return

def recognize_general_text_request():
    try:
        request = RecognizeGeneralTextRequest()
        request.body = GeneralTextRequestBody(
            image=image_base64
        )
        response = client.recognize_general_text(request)
        print(response)
    except exceptions.ClientRequestException as e:
        print(e.status_code)
        print(e.request_id)
        print(e.error_code)
        print(e.error_msg)

def get_credential():
    return BasicCredentials(ak, sk)

def get_client():
    config = HttpConfig.get_default_config()
    config.ignore_ssl_verification = True
    return OcrClient.new_builder(OcrClient) \
        .with_credentials(credentials) \
        .with_region(OcrRegion.CN_NORTH_4) \
        .with_http_config(config) \
        .build()
```

```
def url_to_base64(imagepath):
    with open(imagepath, "rb") as bin_data:
        image_data = bin_data.read()
        base64_data = base64.b64encode(image_data).decode("utf-8")
    return base64_data

if __name__ == '__main__':
    # 填写访问密钥AK、SK
    ak = "填写AK"
    sk = "填写SK"

    # Init Auth Info
    credentials = get_credential()

    # Create OcrClient
    client = get_client()

    df_path = r"本地pdf文件路径，例如D:\local\test.pdf"
    save_path = r"转换后的图片路径，例如D:\local"

    covert_pdf_to_jpg = CovertPdfToJpg(df_path, save_path)
    covert_pdf_to_jpg.pdf_to_jpg()

    jpgs = glob(os.path.join(save_path, "*.jpg"))
    for jpg in jpgs:
        image_base64 = url_to_base64(jpg)
        recognize_general_text_request()
```

#### 说明

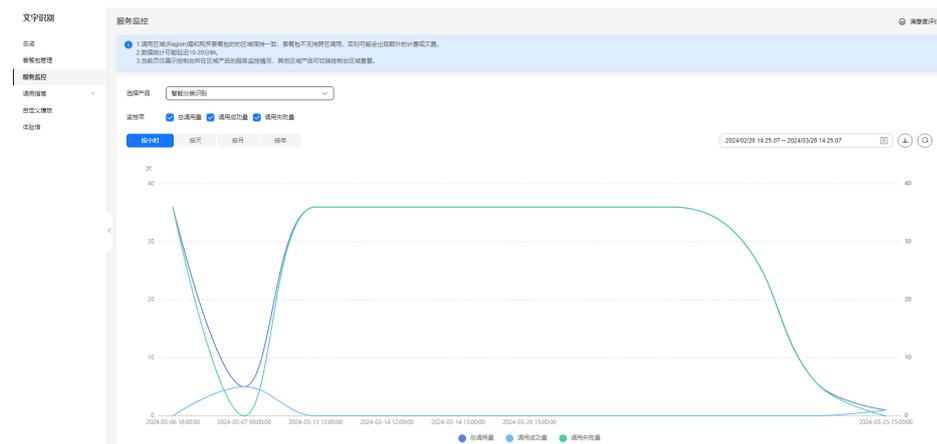
如果运行过程出现“AttributeError: 'Document' object has no attribute 'pageCount'”报错，可以卸载原有的PyMuPDF依赖包，并安装1.18.0版本解决。

# 8 统计 API 使用量

文字识别[控制台](#) > [服务监控](#)支持查看API的总调用量、调用成功量、调用失败量。

如果您购买了套餐包，可在“套餐包管理”页面，查看已有的套餐包及使用详情。套餐包费用扣减规则请参见“产品介绍 > 计费说明”。

图 8-1 服务监控



# 9 关闭服务

OCR服务开通后，如不使用，可登录[控制台](#)进行关闭。

如果您是主账户，也可以控制子用户的服务开通、关闭状态。

图 9-1 关闭服务



图 9-2 控制子用户的开通、关闭



# 10 监控

## 10.1 文字识别的监控指标

### 功能说明

本节定义了文字识别服务上报云监控服务的监控指标的命名空间，监控指标列表和维度定义，用户可以通过云监控服务提供管理控制台或[API接口](#)来检索文字识别服务产生的监控指标和告警信息。

### 命名空间

SYS.OCR

### 监控指标详情

表 10-1 OCR 支持的监控指标

指标ID	指标名称	指标含义	取值范围	测量对象	监控周期（原始指标）
successful_calls_of_service	调用服务成功次数	统计周期内调用服务成功次数。 单位：次/分钟	$\geq 0$ times/min	接口	1分钟
failed_call_times_of_service	调用服务失败次数	统计周期内调用服务失败次数。 单位：次/分钟	$\geq 0$ times/min	接口	1分钟

## 维度

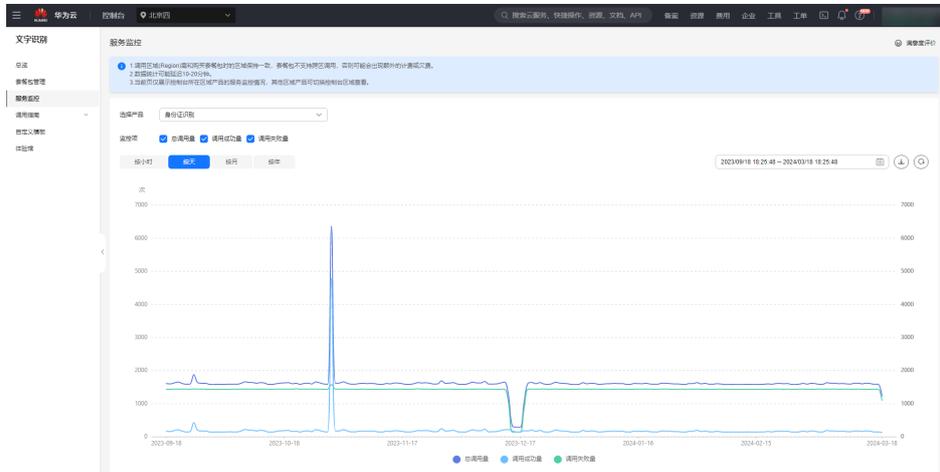
Key	Value
call_of_interface	接口

## 10.2 查看监控指标

操作步骤如下：

1. 登录[管理控制台](#)。
2. 在左侧导航栏选择“服务监控”。

图 10-1 服务监控



# 11 审计

## 11.1 支持云审计的关键操作

### 操作场景

平台提供了云审计服务。通过云审计服务，您可以记录与OCR相关的操作事件，便于日后的查询、审计和回溯。

### 前提条件

已开通云审计服务。

### 支持审计的关键操作列表

表 11-1 云审计服务支持的 OCR 操作列表

操作名称	资源类型	事件名称
开通服务	ai_service	subscribeAISubService

## 11.2 如何查看审计日志

### 操作场景

在您开启了云审计服务后，系统开始记录云服务器的相关操作。云审计服务管理控制台保存最近7天的操作记录。

本节介绍如何在云审计服务管理控制台查看最近7天的操作记录。

### 操作步骤

1. 登录云审计服务[控制台](#)。
2. 单击左侧导航树的“事件列表”，进入事件列表信息页面。

3. 事件列表支持通过筛选来查询对应的操作事件。当前事件列表支持四个维度的组合查询，详细信息如下：
  - 事件来源、资源类型和筛选类型。  
在下拉框中选择查询条件。  
其中筛选类型选择事件名称时，还需选择某个具体的事件名称。  
选择资源ID时，还需选择或者手动输入某个具体的资源ID。  
选择资源名称时，还需选择或手动输入某个具体的资源名称。
  - 操作用户：在下拉框中选择某一具体的操作用户，此操作用户指用户级别，而非租户级别。
  - 事件级别：可选项为“所有事件级别”、“normal”、“warning”、“incident”，只可选择其中一项。
  - 时间范围：可选择查询最近七天内任意时间段的操作事件。
4. 展开需要查看的事件，查看详细信息。

图 11-1 展开记录



事件名称	云服务	资源类型	资源名称	资源ID	操作用户	事件级别	事件时间
ModelKit		AI_Service				normal	2024/03/26
IAM		user				normal	2024/03/26

5. 在需要查看的记录右侧，单击“查看事件”，弹窗中显示了该操作事件结构的详细信息。

# 12 权限管理

## 12.1 权限管理

如果您需要对华为云上购买的OCR资源，为企业中的员工设置不同的访问权限，以达到不同员工之间的权限隔离，您可以使用统一身份认证服务（Identity and Access Management，简称IAM）进行精细的权限管理。该服务提供用户身份认证、权限分配、访问控制等功能，可以帮助您安全的控制华为云资源的访问。

如果华为云账号已经能满足您的要求，不需要创建独立的IAM用户进行权限管理，您可以跳过本章节，不影响您使用OCR服务的其它功能。

通过IAM，您可以在华为云账号中给员工创建IAM用户，并授权控制他们对华为云资源的访问范围。例如，您的员工中有负责软件开发的人员，您希望他们拥有OCR的使用权限，但是不希望他们拥有删除OCR等高危操作的权限，那么您可以使用IAM为开发人员创建用户，通过授予仅能使用OCR，但是不允许删除的权限，控制对资源的使用范围。

IAM是华为云提供权限管理的基础服务，无需付费即可使用，您只需要为您账号中的资源进行付费。关于IAM的详细介绍，请参见[IAM产品介绍](#)。

### OCR 权限

默认情况下，管理员创建的IAM用户没有任何权限，需要将其加入用户组，并给用户组授予策略或角色，才能使得用户组中的用户获得对应的权限，这一过程称为授权。授权后，用户就可以基于被授予的权限对云服务进行操作。

OCR部署时通过物理区域划分，为项目级服务。授权时，“作用范围”需要选择“区域级项目”，然后在指定区域（如华北-北京四）对应的项目（cn-north-4）中设置相关权限，并且该权限仅对此项目生效；如果在“所有项目”中设置权限，则该权限在所有区域项目中都生效。访问OCR时，需要先切换至授权区域。

如表12-1所示，包括了OCR的所有系统角色。

表 12-1 OCR 系统策略

策略名称	描述	策略类型	依赖关系
OCR FullAccess	所有权限	系统策略	无

策略名称	描述	策略类型	依赖关系
OCR ReadOnlyAccess	只读访问权限	系统策略	无

**表12-2**列出了OCR常用操作与系统策略的授权关系，您可以参照该表选择合适的系统策略。

**表 12-2** OCR 操作与系统策略关系

Action	Action说明	OCR FullAccess	OCR ReadOnlyAccess
ocr:financialStatement:subscribe	开通财报识别服务	√	×
ocr:financialStatement:unsubscribe	取消开通财报识别服务	√	×
ocr:financialStatement:getSubscribeUserList	查询开通财报识别的用户列表	√	√
ocr:financialStatement:subscribeAllUsers	为其余子用户开通财报识别服务	√	×
ocr:financialStatement:unsubscribeAllUsers	为其余子用户取消开通财报识别服务	√	×
ocr:autoClassification:subscribe	开通智能分类识别服务	√	×
ocr:autoClassification:unsubscribe	取消开通智能分类识别服务	√	×
ocr:autoClassification:getSubscribeUserList	查询开通智能分类识别的用户列表	√	√
ocr:autoClassification:subscribeAllUsers	为其余子用户开通智能分类识别服务	√	×
ocr:autoClassification:unsubscribeAllUsers	为其余子用户取消开通智能分类识别服务	√	×
ocr:idCard:subscribe	开通身份证识别服务	√	×
ocr:idCard:unsubscribe	取消开通身份证识别服务	√	×
ocr:idCard:getSubscribeUserList	查询开通身份证识别的用户列表	√	√
ocr:idCard:subscribeAllUsers	为其余子用户开通身份证识别服务	√	×

Action	Action说明	OCR FullAccess	OCR ReadOnlyAccess
ocr:idCard:unsubscribeAllUsers	为其余子用户取消开通身份证识别服务	√	×
ocr:driverLicense:subscribe	开通驾驶证识别服务	√	×
ocr:driverLicense:unsubscribe	取消开通驾驶证识别服务	√	×
ocr:driverLicense:getSubscribeUserList	查询开通驾驶证识别的用户列表	√	√
ocr:driverLicense:subscribeAllUsers	为其余子用户开通驾驶证识别服务	√	×
ocr:driverLicense:unsubscribeAllUsers	为其余子用户取消开通驾驶证识别服务	√	×
ocr:vehicleLicense:subscribe	开通行驶证识别服务	√	×
ocr:vehicleLicense:unsubscribe	取消开通行驶证识别服务	√	×
ocr:vehicleLicense:getSubscribeUserList	查询开通行驶证识别的用户列表	√	√
ocr:vehicleLicense:subscribeAllUsers	为其余子用户开通行驶证识别服务	√	×
ocr:vehicleLicense:unsubscribeAllUsers	为其余子用户取消开通行驶证识别服务	√	×
ocr:vatInvoice:subscribe	开通增值税发票识别服务	√	×
ocr:vatInvoice:unsubscribe	取消开通增值税发票识别服务	√	×
ocr:vatInvoice:getSubscribeUserList	查询开通增值税发票识别的用户列表	√	√
ocr:vatInvoice:subscribeAllUsers	为其余子用户开通增值税发票识别服务	√	×
ocr:vatInvoice:unsubscribeAllUsers	为其余子用户取消开通增值税发票识别服务	√	×
ocr:invoiceVerification:subscribe	开通发票验真服务	√	×
ocr:invoiceVerification:unsubscribe	取消开通发票验真服务	√	×

Action	Action说明	OCR FullAccess	OCR ReadOnlyAccess
ocr:invoiceVerification:getSubscribeUserList	查询开通发票验真的用户列表	√	√
ocr:invoiceVerification:subscribeAllUsers	为其余子用户开通发票验真服务	√	×
ocr:invoiceVerification:unsubscribeAllUsers	为其余子用户取消开通发票验真服务	√	×
ocr:mvsInvoice:subscribe	开通机动车销售发票识别服务	√	×
ocr:mvsInvoice:unsubscribe	取消开通机动车销售发票识别服务	√	×
ocr:mvsInvoice:getSubscribeUserList	查询开通机动车销售发票识别的用户列表	√	√
ocr:mvsInvoice:subscribeAllUsers	为其余子用户开通机动车销售发票识别服务	√	×
ocr:mvsInvoice:unsubscribeAllUsers	为其余子用户取消开通机动车销售发票识别服务	√	×
ocr:generalText:subscribe	开通通用文字识别服务	√	×
ocr:generalText:unsubscribe	取消开通通用文字识别服务	√	×
ocr:generalText:getSubscribeUserList	查询开通通用文字识别的用户列表	√	√
ocr:generalText:subscribeAllUsers	为其余子用户开通通用文字识别服务	√	×
ocr:generalText:unsubscribeAllUsers	为其余子用户取消开通通用文字识别服务	√	×
ocr:generalTable:subscribe	开通通用表格识别服务	√	×
ocr:generalTable:unsubscribe	取消开通通用表格识别服务	√	×
ocr:generalTable:getSubscribeUserList	查询开通通用表格识别的用户列表	√	√
ocr:generalTable:subscribeAllUsers	为其余子用户开通通用表格识别服务	√	×
ocr:generalTable:unsubscribeAllUsers	为其余子用户取消开通通用表格识别服务	√	×

Action	Action说明	OCR FullAccess	OCR ReadOnlyAccess
ocr:tollInvoice:subscribe	开通车辆通行费发票识别服务	√	×
ocr:tollInvoice:unsubscribe	取消开通车辆通行费发票识别服务	√	×
ocr:tollInvoice:getSubscribeUserList	查询开通车辆通行费发票识别的用户列表	√	√
ocr:tollInvoice:subscribeAllUsers	为其余子用户开通车辆通行费发票识别服务	√	×
ocr:tollInvoice:unsubscribeAllUsers	为其余子用户取消开通车辆通行费发票识别服务	√	×
ocr:quotaInvoice:subscribe	开通定额发票识别服务	√	×
ocr:quotaInvoice:unsubscribe	取消开通定额发票识别服务	√	×
ocr:quotaInvoice:getSubscribeUserList	查询开通定额发票识别的用户列表	√	√
ocr:quotaInvoice:subscribeAllUsers	为其余子用户开通定额发票识别服务	√	×
ocr:quotaInvoice:unsubscribeAllUsers	为其余子用户取消开通定额发票识别服务	√	×
ocr:flightItinerary:subscribe	开通飞机行程单识别服务	√	×
ocr:flightItinerary:unsubscribe	取消开通飞机行程单识别服务	√	×
ocr:flightItinerary:getSubscribeUserList	查询开通飞机行程单识别的用户列表	√	√
ocr:flightItinerary:subscribeAllUsers	为其余子用户开通飞机行程单识别服务	√	×
ocr:flightItinerary:unsubscribeAllUsers	为其余子用户取消开通飞机行程单识别服务	√	×
ocr:taxiInvoice:subscribe	开通出租车发票识别服务	√	×
ocr:taxiInvoice:unsubscribe	取消开通出租车发票识别服务	√	×
ocr:taxiInvoice:getSubscribeUserList	查询开通出租车发票识别的用户列表	√	√

Action	Action说明	OCR FullAccess	OCR ReadOnlyAccess
ocr:taxiInvoice:subscribeAllUsers	为其余子用户开通出租车发票识别服务	√	×
ocr:taxiInvoice:unsubscribeAllUsers	为其余子用户取消开通出租车发票识别服务	√	×
ocr:trainTicket:subscribe	开通火车票识别服务	√	×
ocr:trainTicket:unsubscribe	取消开通火车票识别服务	√	×
ocr:trainTicket:getSubscribeUserList	查询开通火车票识别的用户列表	√	√
ocr:trainTicket:subscribeAllUsers	为其余子用户开通火车票识别服务	√	×
ocr:trainTicket:unsubscribeAllUsers	为其余子用户取消开通火车票识别服务	√	×
ocr:bankCard:subscribe	开通银行卡识别服务	√	×
ocr:bankCard:unsubscribe	取消开通银行卡识别服务	√	×
ocr:bankCard:getSubscribeUserList	查询开通银行卡识别的用户列表	√	√
ocr:bankCard:subscribeAllUsers	为其余子用户开通银行卡识别服务	√	×
ocr:bankCard:unsubscribeAllUsers	为其余子用户取消开通银行卡识别服务	√	×
ocr:passport:subscribe	开通护照识别服务	√	×
ocr:passport:unsubscribe	取消开通护照识别服务	√	×
ocr:passport:getSubscribeUserList	查询开通护照识别的用户列表	√	√
ocr:passport:subscribeAllUsers	为其余子用户开通护照识别服务	√	×
ocr:passport:unsubscribeAllUsers	为其余子用户取消开通护照识别服务	√	×
ocr:transportationLicense:subscribe	开通道路运输证识别服务	√	×
ocr:transportationLicense:unsubscribe	取消开通道路运输证识别服务	√	×
ocr:transportationLicense:getSubscribeUserList	查询开通道路运输证识别的用户列表	√	√

Action	Action说明	OCR FullAccess	OCR ReadOnlyAccess
ocr:transportationLicense:subscribeAllUsers	为其余子用户开通过路运输证识别服务	√	×
ocr:transportationLicense:unsubscribeAllUsers	为其余子用户取消开通过路运输证识别服务	√	×
ocr:businessLicense:subscribe	开通营业执照识别服务	√	×
ocr:businessLicense:unsubscribe	取消开通营业执照识别服务	√	×
ocr:businessLicense:getSubscribeUserList	查询开通营业执照识别的用户列表	√	√
ocr:businessLicense:subscribeAllUsers	为其余子用户开通营业执照识别服务	√	×
ocr:businessLicense:unsubscribeAllUsers	为其余子用户取消开通营业执照识别服务	√	×
ocr:licensePlate:subscribe	开通车牌识别服务	√	×
ocr:licensePlate:unsubscribe	取消开通车牌识别服务	√	×
ocr:licensePlate:getSubscribeUserList	查询开通车牌识别的用户列表	√	√
ocr:licensePlate:subscribeAllUsers	为其余子用户开通车牌识别服务	√	×
ocr:licensePlate:unsubscribeAllUsers	为其余子用户取消开通车牌识别服务	√	×
ocr:webImage:subscribe	开通网络图片识别服务	√	×
ocr:webImage:unsubscribe	取消开通网络图片识别服务	√	×
ocr:webImage:getSubscribeUserList	查询开通网络图片识别的用户列表	√	√
ocr:webImage:subscribeAllUsers	为其余子用户开通网络图片识别服务	√	×
ocr:webImage:unsubscribeAllUsers	为其余子用户取消开通网络图片识别服务	√	×
ocr:handwriting:subscribe	开通手写文字识别服务	√	×
ocr:handwriting:unsubscribe	取消开通手写文字识别服务	√	×

Action	Action说明	OCR FullAccess	OCR ReadOnlyAccess
ocr:handwriting:getSubscribeUserList	查询开通手写文字识别的用户列表	√	√
ocr:handwriting:subscribeAllUsers	为其余子用户开通手写文字识别服务	√	×
ocr:handwriting:unsubscribeAllUsers	为其余子用户取消开通手写文字识别服务	√	×
ocr:vin:subscribe	开通vin码识别服务	√	×
ocr:vin:unsubscribe	取消开通vin码识别服务	√	×
ocr:vin:getSubscribeUserList	查询开通vin码识别的用户列表	√	√
ocr:vin:subscribeAllUsers	为其余子用户开通vin码识别服务	√	×
ocr:vin:unsubscribeAllUsers	为其余子用户取消开通vin码识别服务	√	×
ocr:businessCard:subscribe	开通名片识别服务	√	×
ocr:businessCard:unsubscribe	取消开通名片识别服务	√	×
ocr:businessCard:getSubscribeUserList	查询开通名片识别的用户列表	√	√
ocr:businessCard:subscribeAllUsers	为其余子用户开通名片识别服务	√	×
ocr:businessCard:unsubscribeAllUsers	为其余子用户取消开通名片识别服务	√	×
ocr:insurancePolicy:subscribe	开通保险单识别服务	√	×
ocr:insurancePolicy:unsubscribe	取消开通保险单识别服务	√	×
ocr:insurancePolicy:getSubscribeUserList	查询开通保险单识别的用户列表	√	√
ocr:insurancePolicy:subscribeAllUsers	为其余子用户开通保险单识别服务	√	×
ocr:insurancePolicy:unsubscribeAllUsers	为其余子用户取消开通保险单识别服务	√	×
ocr:transportationQualificationCertificate:subscribe	开通道路运输从业资格证服务	√	×

Action	Action说明	OCR FullAccess	OCR ReadOnlyAccess
ocr:transportationQualificationCertificate:unsubscribe	取消开通过路运输从业资格证服务	√	×
ocr:transportationQualificationCertificate:getSubscribeUserList	查询开通过路运输从业资格证的用户列表	√	√
ocr:transportationQualificationCertificate:subscribeAllUsers	为其余子用户开通过路运输从业资格证服务	√	×
ocr:transportationQualificationCertificate:unsubscribeAllUsers	为其余子用户取消开通过路运输从业资格证服务	√	×

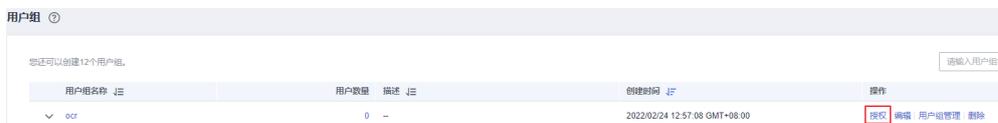
## 12.2 创建用户组与 IAM 用户

### 创建用户组

1. 使用主账号登录IAM服务控制台。
2. 左侧导航窗格中，选择“用户组”页签，单击右上方的“+创建用户组”。



3. 在“创建用户组”界面，输入“用户组名称”，创建用户组。
4. 返回用户组列表，单击列表中的“授权”。
5. 勾选需要授予用户组的权限，单击“确定”，完成用户组授权。



6. 选择权限的作用服务。此处选择“文字识别（OCR）”，并为该用户组设置所需的权限。

OCR FullAccess: 所有权限，可以使IAM用户具备开通/关闭/使用服务能力。

OCR ReadOnlyAccess: 只读访问权限，使IAM用户仅具备使用服务能力，开通/关闭服务需要主账户操作。



- 单击“下一步”，设置用户组的最小授权范围。单击“确定”，完成用户组授权。
  - 所有资源：不设置最小授权范围，授权后，用户根据权限使用账号中所有资源。
  - 指定区域项目资源：授权后，用户根据权限使用已选区域项目中的资源。
  - 选择指定企业项目：IAM用户可以根据权限使用该企业项目中的资源。仅开通企业项目后可选。



## 创建 IAM 用户，并加入用户组

- 在IAM控制台中创建用户。单击右上方的“+创建用户”，配置用户基本信息。



- 单击“下一步”，将用户加入至用户组。

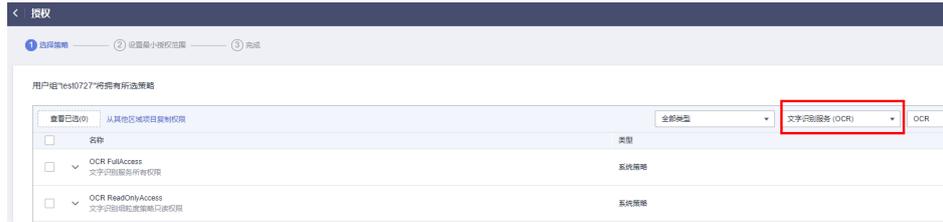


- 单击“创建用户”，完成IAM用户创建。

## 12.3 为 IAM 用户分配权限

### IAM 用户自行开通 OCR 服务

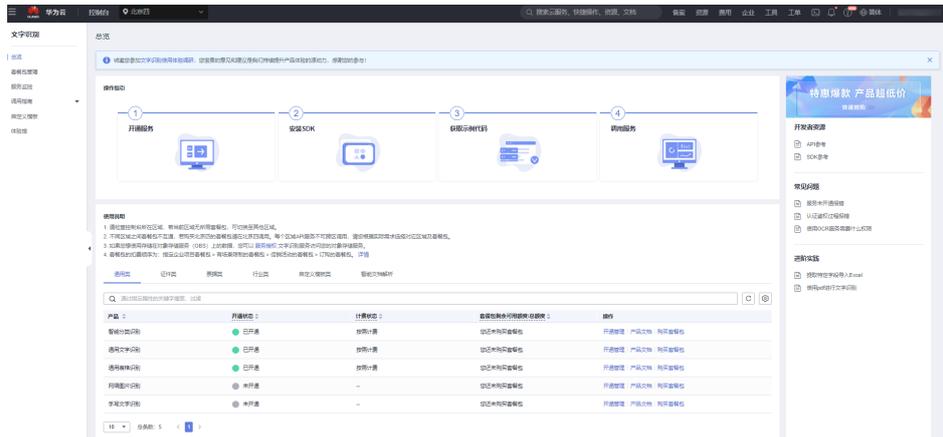
- 使用主账号登录IAM服务控制台
- 打开用户组列表，找到子用户所在的用户组，单击该用户组的“授权”操作。
- 勾选需要授予用户组的权限，单击“确定”，完成用户组授权。OCR全部系统权限请参见[权限管理](#)。



当现有策略不满足时，可使用**OCR自定义策略**



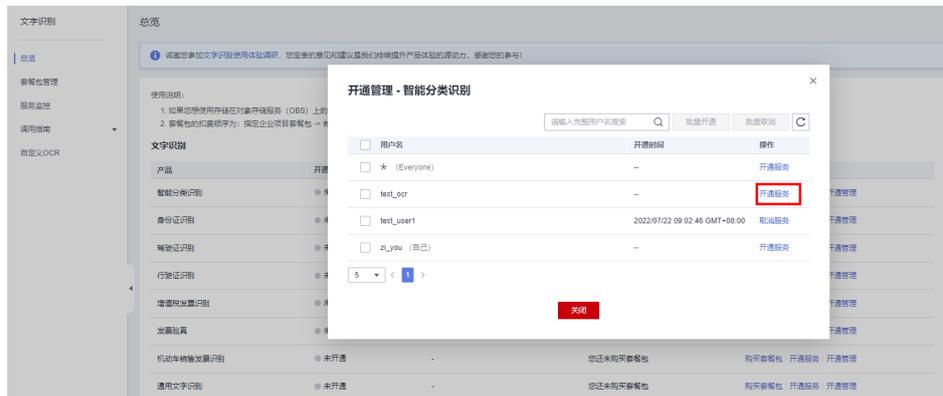
4. IAM用户被赋予相应权限后，可使用IAM用户登录OCR控制台，开通所需的服务。



## 主账号帮 IAM 用户开通 OCR 服务

1. 使用主账号登录OCR服务控制台。
2. 单击相应子服务后的“开通管理”，可以为某个IAM用户或全部IAM用户开通服务。

图 12-1 主账号帮 IAM 用户开通服务



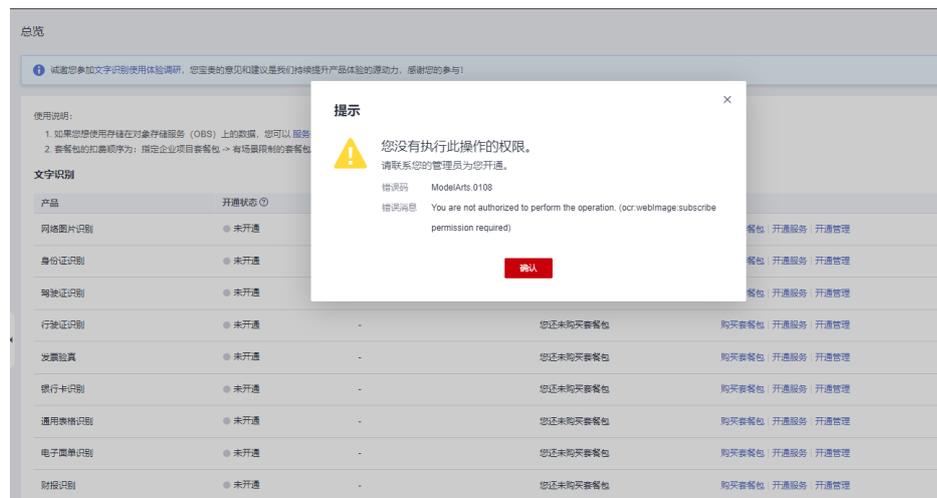
## 说明

当\*(Everyone)特殊用户开通后，其余IAM子用户均可使用当前开通的子服务，不需要单独开通，IAM子用户开通服务后，在调用文字识别服务期间，调用节点需与开通服务节点保持一致。

## IAM 用户缺少某个特定权限现象

当创建的IAM用户缺少某个特定权限时，执行开通操作时，会报错误原因和缺少的权限。例如，下图报ocr:weblImage:subscribe permission required错误（开通网络图片识别的权限）。

图 12-2 IAM 用户缺少网络图片识别权限



出现此类报错，IAM用户可联系主账户开通相应的服务。或者由主账户使用OCR自定义策略，帮助IAM用户开通权限后，IAM用户自行开通服务。

## 12.4 IAM 用户登录并验证权限

使用新创建的IAM账户登录控制台，切换至授权区域，进行验证权限。

- 假设IAM用户只具备“OCR ReadOnlyAccess”权限，即只读访问权限，使IAM用户仅具备使用服务能力，开通/关闭服务需要主账户操作。

进入OCR控制台界面，单击“开通服务”，如果无法开通，提示权限不足，则表示“OCR ReadOnlyAccess”已生效。

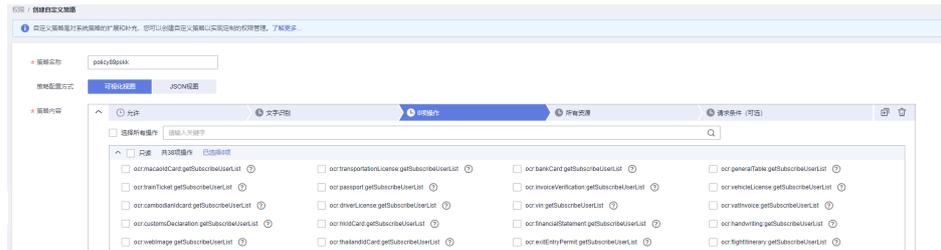
- IAM用户不管是否具备“OCR ReadOnlyAccess”权限，只要具备“OCR FullAccess”权限，便可以在控制台进行开通、关闭服务。

## 12.5 OCR 自定义策略

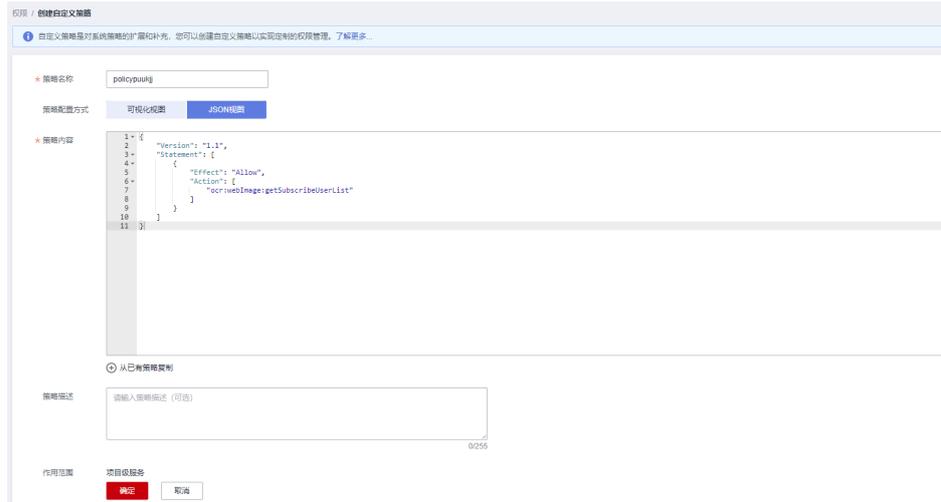
如果系统预置的权限，不满足您的授权要求，可以创建自定义策略。自定义策略中可以添加的授权项（Action）请参见[权限管理](#)。

目前华为云支持以下两种方式创建自定义策略：

- 可视化视图创建自定义策略：无需了解策略语法，按可视化视图导航栏选择云服务、操作、资源、条件等策略内容，可自动生成策略。



- **JSON视图创建自定义策略：**可以在选择策略模板后，根据具体需求编辑策略内容；也可以直接在编辑框内编写JSON格式的策略内容。



具体创建步骤请参见：[创建自定义策略](#)。

# 13 企业项目管理

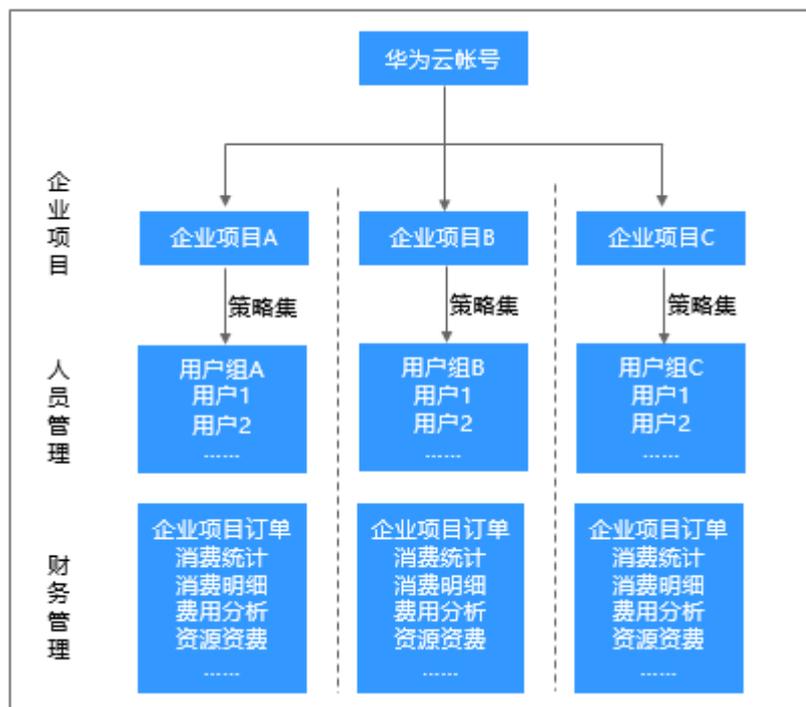
## 13.1 应用场景简介

OCR支持通过企业项目管理（EPS）对不同用户组和用户的资源使用，进行分账。企业可以根据组织架构规划不同的企业项目，并为每个企业项目设置拥有不同权限的用户组 and 用户，多个企业项目之间相互独立，资源分开结算。

### 📖 说明

- **企业账号**可申请开通企业项目。账号需要进行[企业实名认证](#)。如果企业账号注册成为华为云合作伙伴，将无法进入企业项目管理页面。
- 开通企业项目后，企业项目内的用户在调用OCR API时，只有在请求Header参数中传入Enterprise-Project-Id（企业项目ID）后，才支持按不同企业项目进行财务统计。传参方式详见API文档。
- 该功能仅支持开通在华北-北京一、华北-北京四、华东-上海一区域的API，不同API的部署情况请参见[终端节点](#)。

图 13-1 企业项目管理示意图



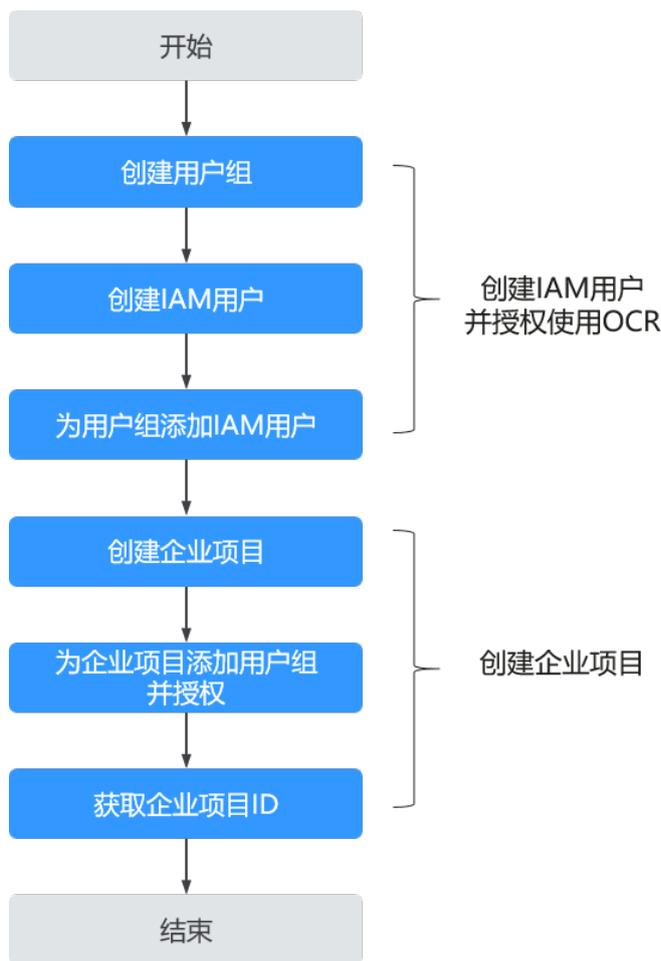
## 13.2 创建企业项目

创建企业项目的用户必须是管理员，或在IAM侧已被授予EPS FullAccess权限的用户。

### 📖 说明

**企业账号**可申请开通企业项目。账号需要进行**企业实名认证**。如果企业账号注册成为华为云合作伙伴，将无法进入企业项目管理页面。

图 13-2 操作流程



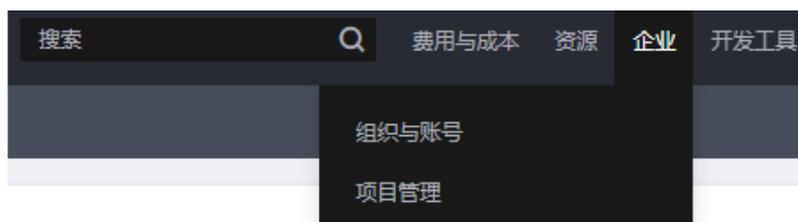
## 创建 IAM 用户并授权使用 OCR

创建IAM用户步骤请参见[创建用户组与IAM用户](#)，为IAM用户授权使用OCR请参见[为IAM用户分配权限](#)。

## 创建企业项目

1. 登录[控制台](#)，单击界面右上方的“企业 > 项目管理”。  
屏幕分辨率低的情况下单击页面右上方的“更多 > 企业 > 项目管理”。

图 13-3 企业项目入口



2. 在企业项目管理页面，单击页面右上方“创建企业项目”。依据界面指引完成企业项目的创建。

图 13-4 创建企业项目

✕

### 创建企业项目

**\* 项目类型**     商用生产项目     测试类项目

商用生产项目：供生产环境部署资源的企业项目使用，资源按大客户商务优惠进行批价。  
测试类项目：供测试企业项目使用，测试项目通常由云服务商提供测试代金券，测试的资源按官网价格进行批价，不使用大客户商务优惠进行批价；当测试代金券全消耗完，后续产生的费用会直接从您的账户扣除。

**\* 名称**

描述

0/512

确定
取消

## 为企业项目关联用户组并授权

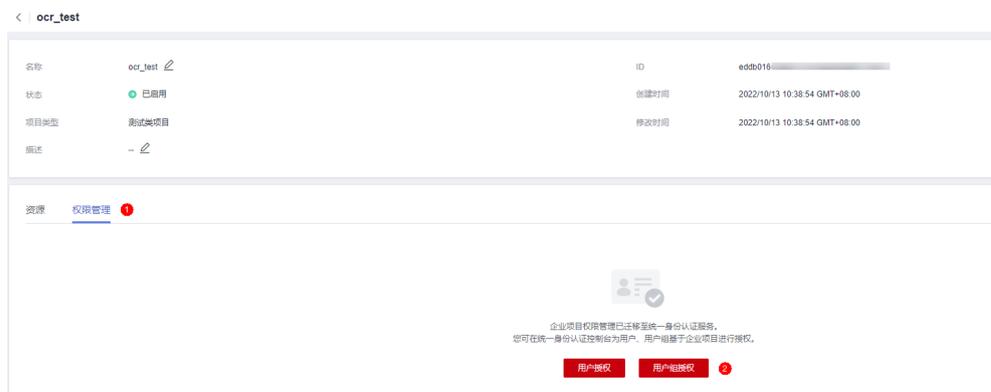
1. 在企业项目管理页面，单击待操作的企业项目的名称。

图 13-5 选择企业项目



2. 系统进入企业项目详情页面，在“权限管理”页签中单击“用户组授权”。系统将跳转至IAM的用户组页面，在“用户组”页签中为企业项目关联用户组并授权。

图 13-6 用户组授权



## 获取企业项目 ID

### 说明

开通企业项目后，企业项目内的用户在调用OCR API时，只有在请求Header参数中传入 Enterprise-Project-Id（企业项目ID）后，才支持按不同企业项目进行财务统计。传参方式详见 API文档。

进入项目管理的“[企业项目管理](#)”页面，单击企业名称，在企业项目详情页获取 Enterprise-Project-Id（企业项目ID）。

图 13-7 获取企业项目 ID



## 13.3 企业项目财务管理

开通了企业项目的客户，可以按企业项目查看各个企业项目的月度消费汇总、消费明细、产品费用走势和费用分布。

进入项目管理的“[企业项目管理](#)”页面，单击企业项目操作列的“查看消费”，查看各企业项目的费用信息。

图 13-8 查看费用信息





图 14-2 获取项目 ID

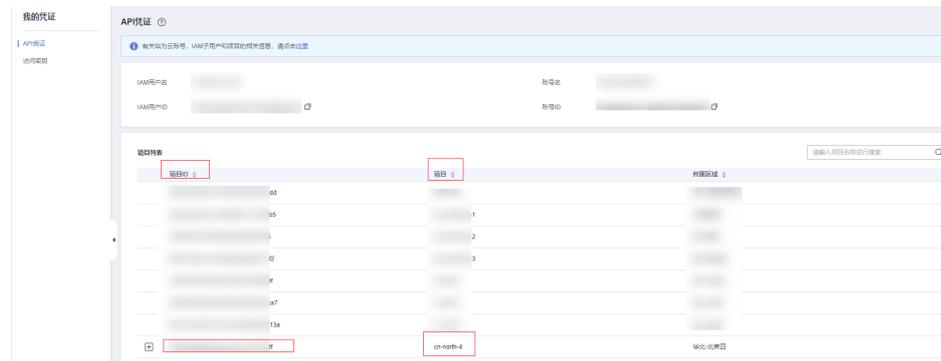


表 14-1 APIs 清单

名称	V1版本API	V2版本API
通用表格识别	/v1.0/ocr/general-table	/v2/{project_id}/ocr/general-table
通用文字识别	/v1.0/ocr/general-text	/v2/{project_id}/ocr/general-text
网络图片识别	/v1.0/ocr/web-image	/v2/{project_id}/ocr/web-image
智能分类识别	/v1.0/ocr/auto-classification	/v2/{project_id}/ocr/auto-classification
手写文字识别	/v1.0/ocr/handwriting	/v2/{project_id}/ocr/handwriting
身份证识别	/v1.0/ocr/id-card	/v2/{project_id}/ocr/id-card
行驶证识别	/v1.0/ocr/vehicle-license	/v2/{project_id}/ocr/vehicle-license
驾驶证识别	/v1.0/ocr/driver-license	/v2/{project_id}/ocr/driver-license
护照识别	/v1.0/ocr/passport	/v2/{project_id}/ocr/passport
银行卡识别	/v1.0/ocr/bankcard	/v2/{project_id}/ocr/bankcard
营业执照识别	/v1.0/ocr/business-license	/v2/{project_id}/ocr/business-license
车牌识别	/v1.0/ocr/license-plate	/v2/{project_id}/ocr/license-plate
增值税发票识别	/v1.0/ocr/vat-invoice	/v2/{project_id}/ocr/vat-invoice
出租车发票识别	/v1.0/ocr/taxi-invoice	/v2/{project_id}/ocr/taxi-invoice

名称	V1版本API	V2版本API
火车票识别	/v1.0/ocr/train-ticket	/v2/{project_id}/ocr/train-ticket
定额发票识别	/v1.0/ocr/quota-invoice	/v2/{project_id}/ocr/quota-invoice
车辆通信费发票识别	/v1.0/ocr/toll-invoice	/v2/{project_id}/ocr/toll-invoice
飞机行程单识别	/v1.0/ocr/flight-itinerary	/v2/{project_id}/ocr/flight-itinerary
道路运输证识别	/v1.0/ocr/transportation-license	/v2/{project_id}/ocr/transportation-license
机动车销售发票识别	/v1.0/ocr/mvs-invoice	/v2/{project_id}/ocr/mvs-invoice
名片识别	/v1.0/ocr/business-card	/v2/{project_id}/ocr/business-card
VIN码识别	/v1.0/ocr/vin	/v2/{project_id}/ocr/vin
保险单识别	/v1.0/ocr/insurance-policy	/v2/{project_id}/ocr/insurance-policy
电子面单识别	/v1.0/ocr/waybill-electronic	/v2/{project_id}/ocr/waybill-electronic