

内容审核

# 快速入门

文档版本 01  
发布日期 2024-04-23



版权所有 © 华为技术有限公司 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。

# 华为技术有限公司

地址： 深圳市龙岗区坂田华为总部办公楼 邮编： 518129

网址： <https://www.huawei.com>

客户服务邮箱： [support@huawei.com](mailto:support@huawei.com)

客户服务电话： 4008302118

# 安全声明

## 漏洞处理流程

华为公司对产品漏洞管理的规定以“漏洞处理流程”为准，该流程的详细内容请参见如下网址：

<https://www.huawei.com/cn/psirt/vul-response-process>

如企业客户须获取漏洞信息，请参见如下网址：

<https://securitybulletin.huawei.com/enterprise/cn/security-advisory>

---

## 目录

---

<b>1 内容审核服务使用简介</b> .....	<b>1</b>
<b>2 开通服务步骤说明</b> .....	<b>2</b>
<b>3 调用 API 实现内容审核功能</b> .....	<b>3</b>
3.1 内容审核-文本.....	3
3.2 内容审核-图像.....	8
<b>4 调用 SDK 实现内容审核功能</b> .....	<b>14</b>
4.1 内容审核-文本.....	14
4.2 内容审核-图像.....	18

# 1 内容审核服务使用简介

内容审核（Content Moderation），是基于图像、文本、音视频的检测技术，可自动检测涉黄图文违规等内容，对用户上传的图片、文字、音视频进行内容审核，以满足上传要求，帮助客户降低业务违规风险。

内容审核以开放API（Application Programming Interface，应用程序编程接口）的方式提供给用户，用户通过调用API获取推理结果，帮助用户打造智能化业务系统，提升业务效率。目前内容审核包括内容审核-图像、内容审核-文本。

您可以根据以下介绍选择合适的使用方式：

- 通过可视化工具（如curl、Postman）发送请求调用内容审核服务API。  
如果您是开发工程师，熟悉代码编写，熟悉HTTP请求与API调用，您可以通过postman调用、调试API。使用方法请参见[如何使用Postman调用华为云Moderation服务](#)。
- 通过软件开发工具包（SDK）调用内容审核服务API。  
如果您是开发工程师，熟悉代码编写，内容审核服务为您提供Java、Python、.NET、GO等版本的SDK，方便您快速集成。使用方法请参见[如何使用SDK调用内容审核服务](#)。

# 2 开通服务步骤说明

## 说明

本服务暂时仅面向企业用户开放。

内容审核服务申请开通您可以按照如下步骤操作：

1. 已注册华为账号，并完成实名认证。
2. 登录内容审核管理控制台，控制台左上角默认显示服务部署区域，请您根据业务需要选择对应区域，服务部署的区域具体请参见[终端节点](#)。
3. 单击页面右下角的“联系客服”按钮，联系客服帮助您开通本服务。
4. 商用服务申请成功后，在“服务管理”页面，“我的服务”中显示已经申请开通成功的服务，此时，您可以通过调用API的方式使用内容审核服务。
5. 开通服务后，单击右上角的“预付套餐包”按钮，进入到本服务套餐包购买页面，按需选择想要购买的功能类型和规格，选择完成后单击“立即购买”，确认购买信息无误后完成付款即可开始使用本服务。

图 2-1 预付套餐包



## 说明

目前内容审核服务提供两种计费模式供您选择：按需计费和预付套餐包计费。具体介绍请参见[计费说明](#)。如果您想使用按需计费的方式，详细费用价格请参见[内容审核价格详情](#)。

# 3 调用 API 实现内容审核功能

## 3.1 内容审核-文本

本章节提供了通过Postman调用“内容审核-文本”的样例，帮助您快速体验并熟悉使用本服务，具体步骤如下：

**步骤一：开通服务**，用户在内容审核控制台，申请开通内容审核-文本服务。

**步骤二：配置自定义词库**，用户可配置自定义白名单词库和自定义黑名单词库。

**步骤三：配置环境**，把准备的配置文件导入到开发环境中。

**步骤四：Token认证鉴权**，用户调用API接口时，需要使用Token进行鉴权。

**步骤五：调用服务**，调用API接口使用服务，使用过程中可以随时查看状态码与错误码。

**步骤六：查看调用次数（可选）**，在控制台查看调用详情和调用次数统计。

### 步骤一：开通服务

内容审核-文本服务开通步骤请参考[开通服务步骤说明](#)。

### 步骤二：配置自定义词库

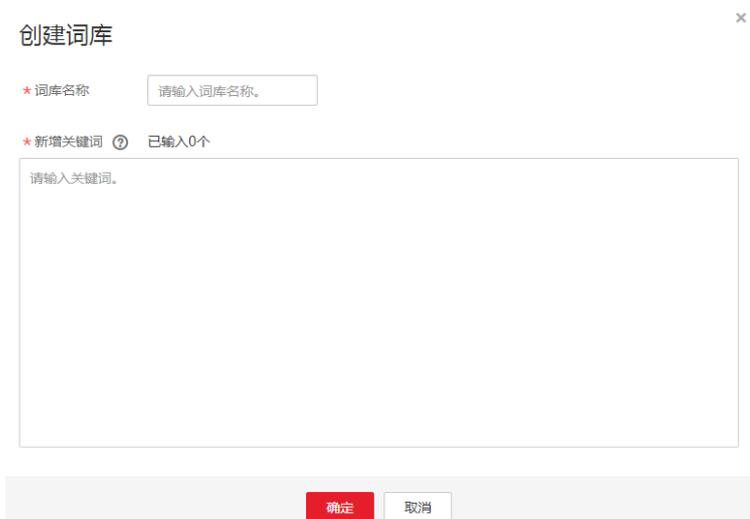
文本内容审核服务可支持用户配置自定义白名单词库和自定义黑名单词库进行文本审核。

- **创建自定义白名单词库**：在使用文本内容审核服务时，您在白名单词库配置的内容可以不被检测出。

具体操作步骤如下：

- a. 登录内容审核服务管理控制台，在左侧菜单栏中选择“设置>自定义词库”。
- b. 在自定义词库界面，单击“白词库”页签中“创建词库”按钮，进入“创建词库”页面。

图 3-1 创建词库



- c. 填写词库名称，输入新增的白名单关键词。
  - 名称：长度小于40个字符，只能包含字母、数字、中划线和下划线。
  - 新增关键词：长度不大于40个字符，最多支持5000个关键词，每一行一个关键词。
- d. 单击“确定”，完成词库的创建。
- **创建自定义黑名单词库**：在使用文本内容审核服务时，您在黑名单词库配置的内容会被检测出。  
具体操作步骤如下：
  - a. 登录内容审核服务管理控制台，在左侧菜单栏中选择“设置>自定义词库”。
  - b. 在自定义词库界面，单击“黑词库”页签中“创建词库”按钮，进入“创建词库”页面。
  - c. 填写词库名称，输入新增的黑名单关键词。
    - 名称：长度小于40个字符，只能包含字母、数字、中划线和下划线。
    - 新增关键词：长度不大于40个字符，最多支持5000个关键词，每一行一个关键词。
  - d. 单击“确定”，完成词库的创建。

#### 📖 说明

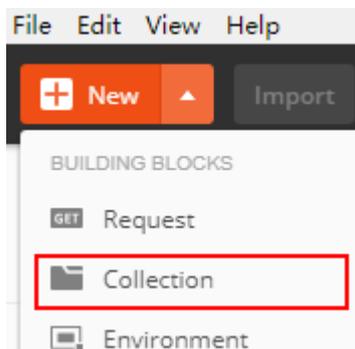
- 词库名称不支持修改。
- V2版本的文本审核，可以选择是否勾选“不指定检测场景”，默认不勾选。
  - 勾选：表示使用默认场景和用户自定义场景进行审核。
  - 不勾选：表示使用默认场景进行审核。

### 步骤三：配置环境

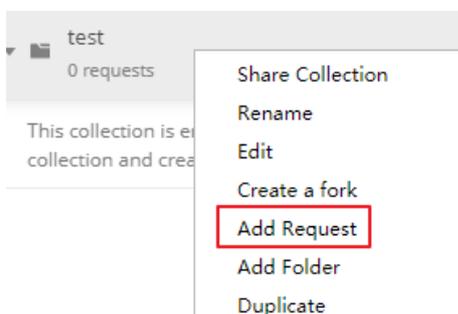
下载并安装Postman。Postman建议使用7.24.0版本。

## 步骤四：Token 认证鉴权

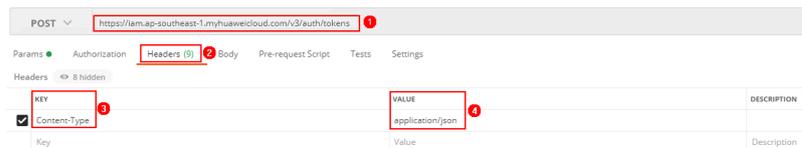
1. 在Postman界面，选择“New > Collection”，设置相应的名称并单击“Create”完成创建。



2. 选择创建的Collection，单击鼠标右键，选择“Add Request”，设置Request name并单击“Save”。

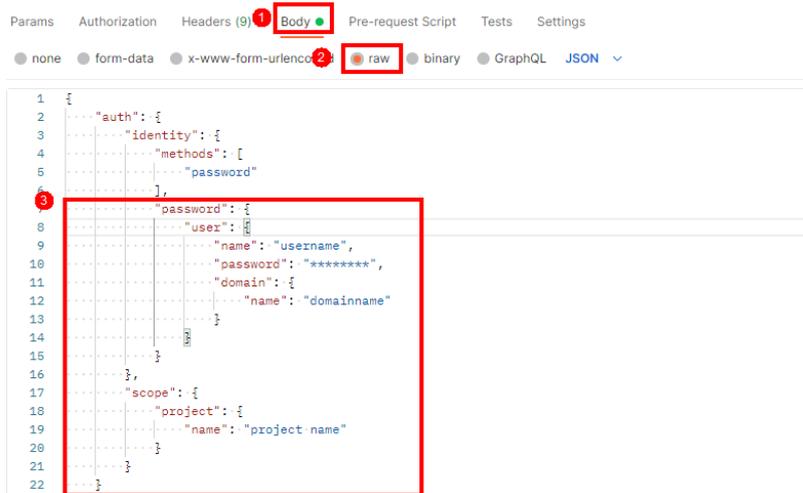


3. 请求方式修改为"POST"，输入URL。  
例如，以ap-southeast-1为例，URL为“https://iam.ap-southeast-1.myhuaweicloud.com/v3/auth/tokens”。
4. 在“Headers”列表中添加“KEY”为“Content-Type”，“VALUE”为“application/json”。

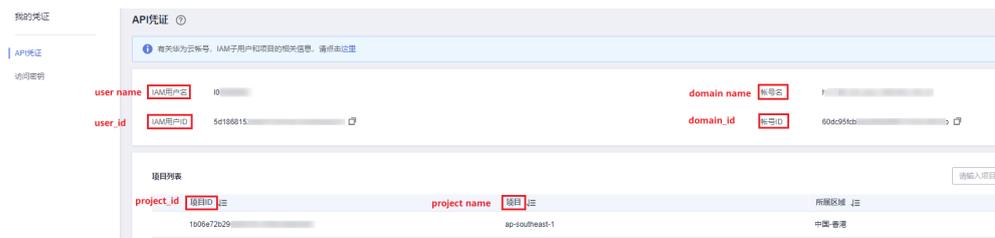


5. 选中“Body”的配置项，选中“raw”，在空白处添加以下代码。

图 3-2 Token 认证鉴权



图中代码红框内加粗的蓝色字段需要根据实际值填写，其中username为用户名，domainname为用户所属的账号名称，\*\*\*\*\*为用户登录密码，project name为服务的部署区域，获取方法请登录[我的凭证](#)获取。



内容审核服务部署的区域必须与调用的服务所在区域一致，本示例中为ap-southeast-1。

```

{
  "auth": {
    "identity": {
      "methods": [
        "password"
      ],
      "password": {
        "user": {
          "name": "username",
          "password": "*****",
          "domain": {
            "name": "domainname"
          }
        }
      }
    },
    "scope": {
      "project": {
        "name": "ap-southeast-1"
      }
    }
  }
}

```

- 单击右上角“Send”按钮发送请求。获取返回结果中的Token值（Token有效期为24小时）。

Header	Value
Content-Length	16647
Connection	keep-alive
X-IAM-Trace-Id	token_cn-north-4_null_02720ac73da1e27272edf6fab756e911
Cache-Control	no-cache, no-store, must-revalidate
Pragma	no-cache
Expires	Thu, 01 Jan 1970 00:00:00 GMT
X-Subject-Token	MIIZNgYjKoZlhvcNAQcCollZjpCCGSMCAQExDTALBgIghkg8ZQMEAgEwghdlBglqkhiG9
X-Request-Id	7d1dcfbac0e463dda61e7ba926279c7e
Server	api-gateway
Strict-Transport-Security	max-age=31536000; includeSubdomains;

## 步骤五：调用服务

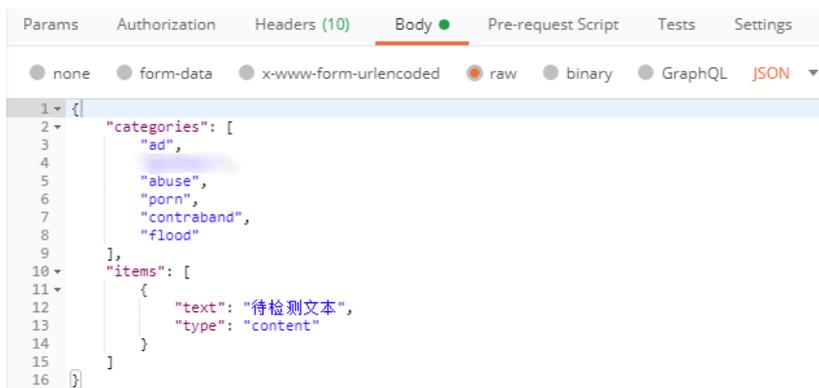
1. 新建请求，请求方式为修改“POST”，按要求填写URL。  
例如服务部署在“中国-香港”区域，请求URL为“https://moderation.ap-southeast-1.myhuaweicloud.com/v2/{project\_id}/moderation/text”。  
选中“Headers”配置项，复制Token值到“X-Auth-Token”。  
登录[我的凭证](#)，查询隶属于“中国-香港”的项目ID，替换URI中的{project\_id}

项目列表

项目ID	项目	所属区域
07f69	ap-southeast-1	中国-香港

2. 单击“Body”，将待检测的文本填入到“text”参数中。关于body体中其他参数说明，请参考[文本内容审核API](#)。

图 3-3 修改参数



3. 单击右上角“Send”按钮发送请求，查看识别结果。

## 步骤六：查看调用次数（可选）

**步骤1** 登录内容审核服务管理控制台。

**步骤2** 在左侧导航栏中选择“服务列表>文本审核”，可以查看识别统计详情。您可以设置时间范围，策略（事件类型）来观察这段时间内的调用次数变化情况。

- 识别结果统计：显示一段时间范围，内容审核的调用总数，拒绝数，疑似数和通过数，帮助您更好了解服务的调用情况和审核情况。
  - 总数：指的是审核调用总次数。
  - 拒绝数：指的是block总数，即文本中包含敏感信息，审核不通过的次数。
  - 疑似数：指的是review总数，即人工复查审核的次数。
  - 通过数：指的是pass总数，即通过审核的次数。

- 数据趋势：显示您设置的这段时间范围内，总数，拒绝数，疑似数和通过数的变化趋势。
- 拒绝数据原因分布：显示您设置的这段时间范围内，审核不通过的检测场景占比数。
- 疑似数据原因分布：显示您设置的这段时间范围内，需要人工复查的检测场景占比数。

**步骤3** 查看监控数据。您可以单击“查看监控指标”在云监控控制台查看服务调用成功和失败的次数等历史数据，如图3-4所示。

图 3-4 监控数据



----结束

## 3.2 内容审核-图像

本章节提供了通过Postman调用“内容审核-图像”的样例，帮助您快速体验并熟悉使用本服务，具体步骤如下：

**步骤一：开通服务**，用户在“服务列表”或“服务管理”页面选择内容审核-图像服务申请开通。

**步骤二：对象存储服务OBS授权（可选）**，如果需要对OBS中的数据进行内容审核请开通对象存储服务OBS授权。

**步骤三：配置审核规则（可选）**，可对涉黄内容进行不同场景的检测。

**步骤四：配置环境**，把准备的配置文件导入到开发环境中。

**步骤五：Token认证鉴权**，用户调用API接口时，需要使用Token进行鉴权。

**步骤六：调用服务**，调用API接口使用服务，使用过程中可以随时查看状态码与错误码。

**步骤七：查看调用次数（可选）**，在控制台查看调用详情和调用次数统计。

### 步骤一：开通服务

内容审核-图像服务开通步骤请参考[开通服务步骤说明](#)。

### 步骤二：对象存储服务 OBS 授权（可选）

如果您需要对OBS中的数据进行内容审核，请开通对象存储服务OBS授权。具体操作步骤如下：

1. 登录华为云内容审核管理控制台，单击进入“服务管理”页面。

2. 打开对象存储服务OBS授权的按钮 ，完成授权操作。

### 📖 说明

- Moderation不支持跨区域OBS，OBS的区域需要和服务保持一致。
- OBS授权需要使用IAM主账号进行授权，不支持使用子账号进行授权。
- 授权完成后，不支持取消授权操作。

## 步骤三：配置审核规则（可选）

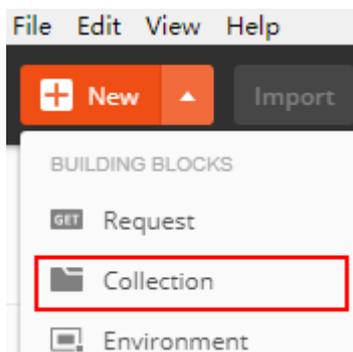
如果您需要对涉黄内容进行不同场景的检测，可进行审核规则的配置。具体操作步骤请参考[配置审核规则](#)。

## 步骤四：配置环境

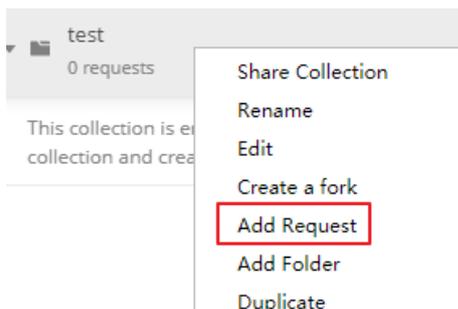
下载并安装Postman。Postman建议使用7.24.0版本。

## 步骤五：Token 认证鉴权

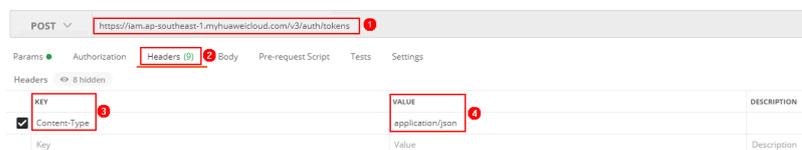
1. 在Postman界面，选择“New > Collection”，设置相应的名称并单击“Create”完成创建。



2. 选择创建的Collection，单击鼠标右键，选择“Add Request”，设置Request name并单击“Save”。

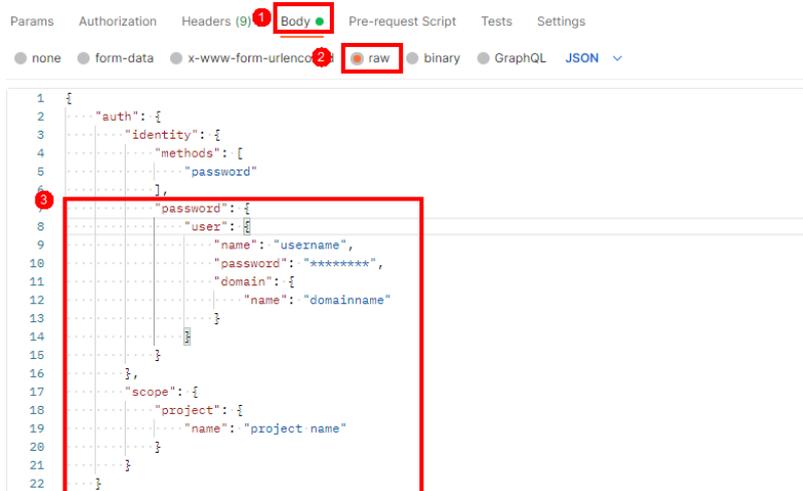


3. 请求方式修改为"POST"，输入URL。  
例如，以ap-southeast-1为例，URL为“https://iam.ap-southeast-1.myhuaweicloud.com/v3/auth/tokens”。
4. 在“Headers”列表中添加“KEY”为“Content-Type”，“VALUE”为“application/json”。

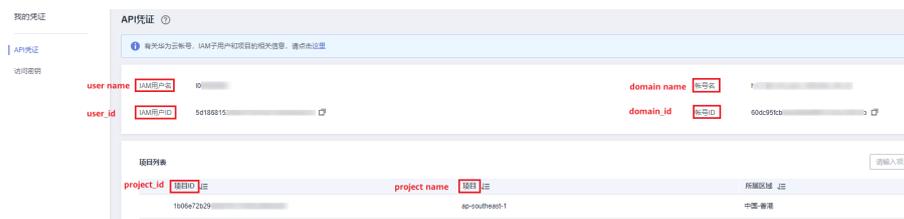


- 选中“Body”的配置项，选中“raw”，在空白处添加以下代码。

图 3-5 Token 认证鉴权



图中代码红框内加粗的蓝色字段需要根据实际值填写，其中username为用户名，domainname为用户所属的账号名称，\*\*\*\*\*为用户登录密码，project name为服务的部署区域，获取方法请登录[我的凭证](#)获取。



内容审核服务部署的区域必须与调用的服务所在区域一致，本示例中为ap-southeast-1。

```

{
  "auth": {
    "identity": {
      "methods": [
        "password"
      ],
      "password": {
        "user": {
          "name": "username",
          "password": "*****",
          "domain": {
            "name": "domainname"
          }
        }
      }
    },
    "scope": {
      "project": {
        "name": "ap-southeast-1"
      }
    }
  }
}

```

- 单击右上角“Send”按钮发送请求。获取返回结果中的Token值（Token有效期为24小时）。

Header	Value
Content-Length	16647
Connection	keep-alive
X-IAM-Trace-Id	token_cn-north-4_null_02720ac73da1e27272edf6fab756e911
Cache-Control	no-cache, no-store, must-revalidate
Pragma	no-cache
Expires	Thu, 01 Jan 1970 00:00:00 GMT
X-Subject-Token	MlIb+QYJKoZlhcNAQcCollZjCCGSMCAQExDALTALBgIghkgB2QMEAgEwghdlBglqhkiG9...
X-Request-Id	7d1dcfbac0e463dda61e7ba926279c7e
Server	api-gateway
Strict-Transport-Security	max-age=31536000; includeSubdomains;

## 步骤六：调用服务

- 新建请求，请求方式为修改"POST"，按要求填写URL。  
例如护照识别服务部署在“中国-香港”区域，请求URL为“https://ocr.ap-southeast-1.myhuaweicloud.com/v2/{project\_id}/ocr/passport”。  
选中“Headers”配置项，复制Token值到“X-Auth-Token”。  
登录[我的凭证](#)，查询隶属于“中国-香港”的项目ID，替换URI中的{project\_id}

KEY	VALUE	DESCRIPTION
<input checked="" type="checkbox"/> Content-Type	application/json	
<input checked="" type="checkbox"/> X-Auth-Token	MlIb+QYJKoZlhcNAQcCollZjCCGSMCAQExDALTALBgIghkgB2QMEAgEwghdlBglqhkiG9...	
Key	Value	Description

项目ID	项目	所属区域
07f59...	ap-southeast-1	中国-香港

- 单击“Body”，将待检测的文本填入到“text”参数中。关于body体中其他参数说明，请参考[图像内容审核API](#)。

图 3-6 修改参数

```

1 {
2   "categories": [
3     "ad",
4     "abuse",
5     "porn",
6     "contraband",
7     "flood"
8   ],
9   "items": [
10    {
11      "text": "待检测文本",
12      "type": "content"
13    }
14  ]
15 }
16

```

- 单击右上角“Send”按钮发送请求，查看识别结果。

## 步骤七：查看调用次数（可选）

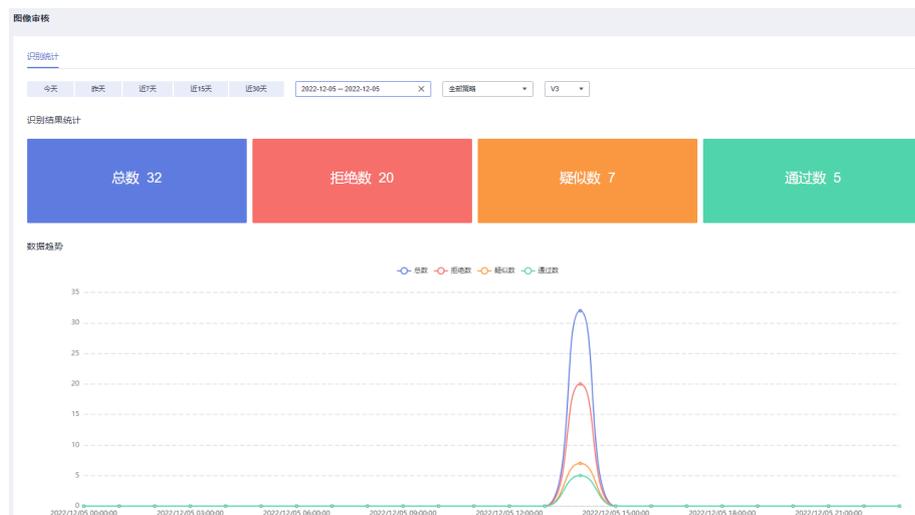
内容审核-图像API接口，分析并识别用户上传的图像内容是否有敏感内容（如色情等），并将识别结果返回给用户。

调用内容审核-图像API接口，您可以按照如下步骤操作：

**步骤1** 进入内容审核管理控制台。

**步骤2** 在左侧导航栏中选择“服务列表>图像审核”，可以查看识别统计详情，如图3-7所示。您可以设置时间范围，策略（事件类型）来观察这段时间内的调用次数变化情况。

图 3-7 识别统计



- 识别结果统计：显示一段时间范围，内容审核的调用总数，拒绝数，疑似数和通过数，帮助您更好了解服务的调用情况和审核情况。
  - 总数：指的是审核调用总次数。
  - 拒绝数：指的是block总数，即文本中包含敏感信息，审核不通过的次数。
  - 疑似数：指的是review总数，即人工复查审核的次数。
  - 通过数：指的是pass总数，即通过审核的次数。
- 数据趋势：显示您设置的这段时间范围内，总数，拒绝数，疑似数和通过数的变化趋势。

图 3-8 原因分布

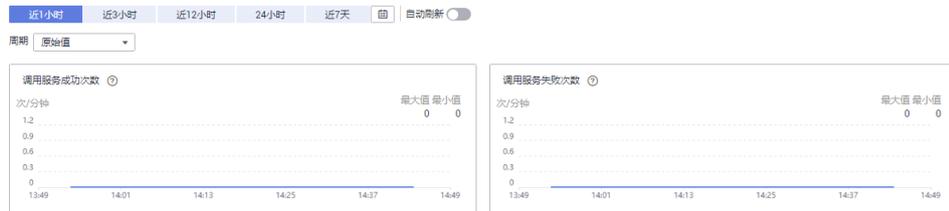


- 拒绝数据原因分布：显示您设置的这段时间范围内，审核不通过的检测场景占比数。

- 疑似数据原因分布：显示您设置的这段时间范围内，需要人工复查的检测场景占比数。

**步骤3** 查看监控数据。您可以单击“查看监控指标”在云监控控制台查看服务调用成功和失败的次数等历史数据，如图3-9所示。

图 3-9 监控数据



----结束

# 4 调用 SDK 实现内容审核功能

## 4.1 内容审核-文本

本章节提供了通过Java SDK调用“内容审核-文本”服务的样例，用户直接调用接口函数即可使用SDK功能。具体流程如下：

**步骤一：开通服务**，用户在“服务列表”或“服务管理”页面选择内容审核-文本服务申请开通。

**步骤二：配置自定义词库**，用户可配置自定义白名单词库和自定义黑名单词库。

**步骤三：配置环境**，获取SDK和样例工程，导入到开发环境中。

**步骤四：认证鉴权**，使用AK/SK方式进行认证。

**步骤五：调用服务**，调用API接口使用服务，使用过程中可以随时查看状态码与错误码。

### 步骤一：开通服务

内容审核-文本服务开通步骤请参考[开通服务步骤说明](#)。

### 步骤二：配置自定义词库

文本内容审核服务可支持用户配置自定义白名单词库和自定义黑名单词库进行文本审核。

1. 创建自定义白名单词库：在使用文本内容审核服务时，您在白名单词库配置的内容可以不被检测出。

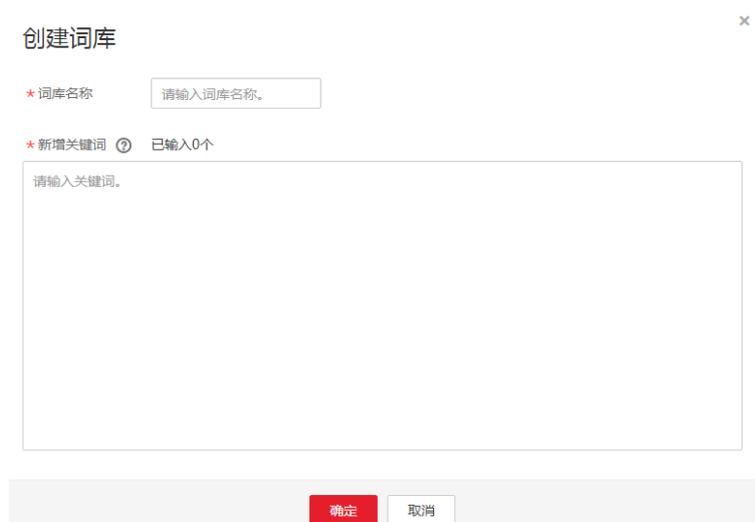
具体操作步骤如下：

**步骤1** 登录华为云，单击右上角的“控制台”，进入“管理控制台”界面。

**步骤2** 单击“服务列表”，选择“人工智能 > 内容审核”，进入“内容审核”管理页面。

**步骤3** 在左侧菜单栏中选择“服务列表>文本审核”服务。单击“自定义白名单词库”后的“配置”按钮，进入“创建词库”页面。

图 4-1 创建词库



**步骤4** 填写词库名称，输入新增的白名单关键词。单击“确定”，完成词库的创建

- 名称：长度小于40个字符，只能包含字母、数字、中划线和下划线。
- 新增关键词：长度不大于40个字符，最多支持5000个关键词，每一行一个关键词。

#### ----结束

2. 创建自定义黑名单词库：在使用文本内容审核服务时，您在黑名单词库配置的内容会被检测出。

具体操作步骤如下：

**步骤1** 登录华为云，单击右上角的“控制台”，进入“管理控制台”界面。

**步骤2** 单击“服务列表”，选择“人工智能 > 内容审核”，进入“内容审核”管理页面。

**步骤3** 在左侧菜单栏中选择“服务列表>文本审核”服务。单击“自定义黑名单词库”后的“配置”按钮，进入“自定义词库”页面。

**步骤4** 单击页面左上方“创建词库”，进入“创建词库”页面。填写词库名称，输入新增的黑名单关键词。

- 名称：长度小于40个字符，只能包含字母、数字、中划线和下划线。
- 新增关键词：长度不大于40个字符，最多支持5000个关键词，每一行一个关键词。

**步骤5** 单击“确定”，完成词库的创建。

#### 📖 说明

- 词库名称不支持修改。

#### ----结束

## 步骤三：配置环境

您可以基于内容审核SDK通过编写代码的方式调用内容审核-文本API。在使用SDK和调用API时您需要进行环境配置。具体操作步骤如下：

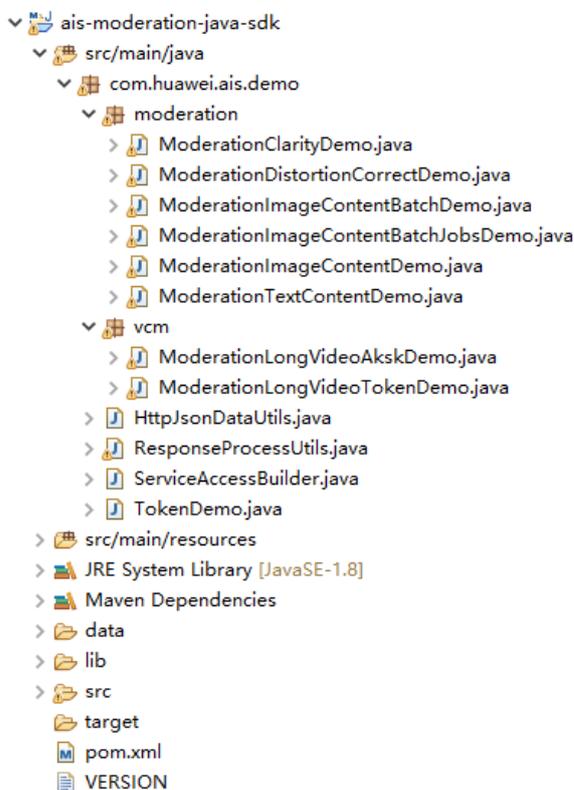
**步骤1** 最新版本内容审核SDK软件包和文档，请在[SDK中心](#)获取。

**步骤2** 环境配置请参见[Java开发环境配置](#)。

**步骤3** 以eclipse工具为例，导入SDK工程：

1. 打开Eclipse，确保“Windows > Preferences > Java > Installed JREs”配置正确的JRE路径。
2. 在左边导航上选择“General > Workspace”，在“Text file encoding”区域选中“Other”，并设置参数值为“UTF-8”，单击“Apply”后，单击“OK”。
3. 在菜单栏依次选择“File > Import”，弹出导入工程窗口，选择“Maven > Existing Maven Projects”，单击“Next”，单击“Browse”，选择“ais-moderation-java-sdk”所在的本地位置。
4. 单击Finish，导入SDK，导入后打开工程，工程目录如[图4-2](#)所示。

**图 4-2** 工程目录



----结束

## 步骤四：认证鉴权

内容审核服务认证方式有Token和AK/SK两种方式，本示例中使用AK/SK方式进行认证。

**步骤1** 获取AK/SK。

AK/SK即访问密钥，请登录[我的凭证](#)页面，选择“访问密钥 > 新增访问密钥”获取。

图 4-3 新增访问密钥



**步骤2** 配置Java SDK中的AK/SK，进行AK/SK认证鉴权。

根据获取的AK/SK，修改工程目录中“ModerationTextContentDemo.java”文件中函数的“AK”和“SK”的值。示例代码如下：

```
// 1. 配置好访问文本内容审核服务的基本信息,生成对应的一个客户端连接对象
AisAccess service = ServiceAccessBuilder.builder()
    // 认证用的ak和sk硬编码到代码中或者明文存储都有很大的安全风险,建议在配置文件或者环境变量中密文
    // 存放,使用时解密,确保安全
    // 本示例以ak和sk保存在环境变量中来实现身份验证为例,运行本示例前请先在本地环境中设置环境变量
    HUAWEICLOUD_SDK_AK和HUAWEICLOUD_SDK_SK
    .ak("HUAWEICLOUD_SDK_AK")
    .sk("HUAWEICLOUD_SDK_SK")
    .region("ap-southeast-1") // 内容审核服务支持中国-香港(ap-southeast-1)、亚太-新加坡(ap-
    southeast-3)的配置
    .connectionTimeout(5000) // 连接目标url超时限制
    .connectionRequestTimeout(1000) // 连接池获取可用连接超时限制
    .socketTimeout(20000) // 获取服务器响应数据超时限制
    .build();
```

----结束

## 步骤五：调用服务

**步骤1** 输入需要检测的文本，参考如下示例代码修改“ModerationTextContentDemo.java”文件中输入文本的位置（“6666666666”）。

```
//
// 2.构建访问文本内容审核服务需要的参数
//
String uri = "/v1.0/moderation/text";

JSONObject json = new JSONObject();
json.put("categories", new String[] { "porn" }); //检测内容

JSONObject text = new JSONObject();
text.put("text", "6666666666");
text.put("type", "content");

JSONArray items = new JSONArray();
items.add(text);
json.put("items", items);

StringEntity stringEntity = new StringEntity(json.toJSONString(), "utf-8");

// 3.传入文本内容审核服务对应的uri参数,传入文本内容审核服务需要的参数,
// 该参数主要通过JSON对象的方式传入,使用POST方法调用服务
HttpResponse response = service.post(uri, stringEntity);

// 4.验证服务调用返回的状态是否成功,如果为200,为成功,否则失败。
ResponseProcessUtils.processResponseStatus(response);
```

**步骤2** 执行“ModerationTextContentDemo.java”文件，控制台输出200即表示程序执行成功，**文本内容审核结果**输出到控制台，如图4-4所示。

图 4-4 运行结果

```
200
{
  "result": {
    "detail": {
      "flood": "666666666666"
    },
    "suggestion": "block"
  }
}
```

**步骤3** 查看调用次数。您可以在“服务列表”，“文本审核”页查看调用详情和调用次数统计。

- 识别结果统计：显示一段时间范围，内容审核的调用总数，拒绝数，疑似数和通过数，帮助您更好了解服务的调用情况和审核情况。
  - 总数：指的是审核调用总次数。
  - 拒绝数：指的是block总数，即文本中包含敏感信息，审核不通过的次数。
  - 疑似数：指的是review总数，即人工复查审核的次数。
  - 通过数：指的是pass总数，即通过审核的次数。
- 数据趋势：显示您设置的这段时间范围内，总数，拒绝数，疑似数和通过数的变化趋势。
- 拒绝数据原因分布：显示您设置的这段时间范围内，审核不通过的检测场景占比数。
- 疑似数据原因分布：显示您设置的这段时间范围内，需要人工复查的检测场景占比数。

**步骤4** 查看监控数据。您可以单击“查看监控指标”在云监控控制台查看服务调用成功和失败的次数等历史数据，如图4-5所示。

图 4-5 监控数据



---结束

## 4.2 内容审核-图像

本章节提供了通过Java SDK调用“内容审核-图像”服务的样例，用户直接调用接口函数即可使用SDK功能。具体流程如下：

**步骤一：开通服务**，用户在“服务列表”或“服务管理”页面选择内容审核-图像服务申请开通。

**步骤二：对象存储服务OBS授权（可选）**，如果需要对OBS中的数据内容进行内容审核请开通对象存储服务OBS授权。

**步骤三：配置审核规则（可选）**，可对涉黄内容进行不同场景的检测。

**步骤四：配置环境**，获取SDK和样例工程，导入到开发环境中。

**步骤五：认证鉴权**，使用AK/SK方式进行认证。

**步骤六：调用服务**，调用API接口使用服务，使用过程中可以随时查看状态码与错误码。

## 步骤一：开通服务

内容审核-图像服务开通步骤请参考[开通服务步骤说明](#)。

## 步骤二：对象存储服务 OBS 授权（可选）

如果您需要对OBS中的数据进行内容审核，请开通对象存储服务OBS授权。具体操作步骤如下：

1. 登录华为云内容审核管理控制台，单击进入“服务管理”页面。
2. 打开对象存储服务OBS授权的按钮 ，完成授权操作。

### 说明

- Moderation不支持跨区域OBS，OBS的区域需要和服务保持一致。
- OBS授权需要使用IAM主账号进行授权，不支持使用子账号进行授权。
- 授权完成后，不支持取消授权操作。

## 步骤三：配置审核规则（可选）

如果您需要对涉黄内容进行不同场景的检测，可进行审核规则的配置。具体操作步骤请参考[配置审核规则](#)。

## 步骤四：配置环境

您可以基于内容审核SDK通过编写代码的方式调用内容审核-文本API。在使用SDK和调用API时您需要进行环境配置。具体操作步骤如下：

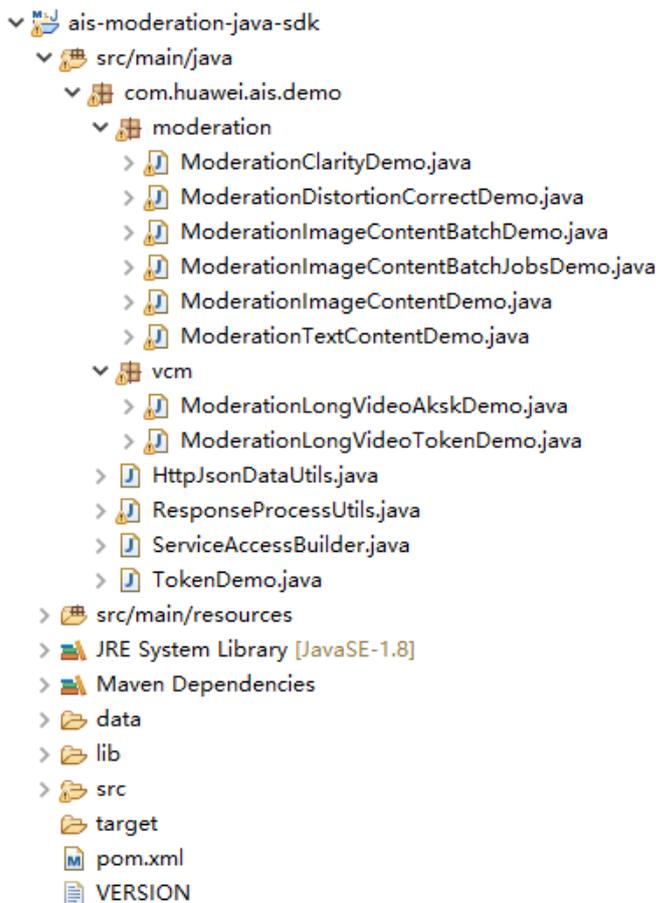
**步骤1** 最新版本内容审核SDK软件包和文档，请单击[SDK中心](#)。

**步骤2** 环境配置请参见[Java开发环境配置](#)。

**步骤3** 以eclipse工具为例，导入SDK工程：

1. 打开Eclipse，确保“Windows > Preferences > Java > Installed JREs”配置正确的JRE路径。
2. 在左边导航上选择“General > Workspace”，在“Text file encoding”区域选中“Other”，并设置参数值为“UTF-8”，单击“Apply”后，单击“OK”。
3. 在菜单栏依次选择“File > Import”，弹出导入工程窗口，选择“Maven > Existing Maven Projects”，单击“Next”，单击“Browse”，选择“ais-moderation-java-sdk”所在的本地位置。
4. 单击Finish，导入SDK，导入后打开工程，工程目录如[图4-6](#)所示。

图 4-6 工程目录



----结束

## 步骤五：认证鉴权

内容审核服务认证方式有Token和AK/SK两种方式，本示例中使用AK/SK方式进行认证。

### 步骤1 获取AK/SK。

AK/SK即访问密钥，请登录[我的凭证](#)页面，选择“访问密钥 > 新增访问密钥”获取。

图 4-7 新增访问密钥



### 步骤2 配置Java SDK中的AK/SK，进行AK/SK认证鉴权。

根据获取的AK/SK，修改工程目录中“ModerationImageContentDemo.java”文件中函数的“AK”和“SK”的值。示例代码如下：

```
// 1. 配置好访问图像内容审核服务的基本信息,生成对应的一个客户端连接对象
AisAccess service = ServiceAccessBuilder.builder()
    // 认证用的ak和sk硬编码到代码中或者明文存储都有很大的安全风险,建议在配置文件或者环境变量中密文
    // 存放,使用时解密,确保安全
    // 本示例以ak和sk保存在环境变量中来实现身份验证为例,运行本示例前请先在本地环境中设置环境变量
    HUAWEICLOUD_SDK_AK和HUAWEICLOUD_SDK_SK
    .ak("HUAWEICLOUD_SDK_AK")
    .sk("HUAWEICLOUD_SDK_SK")
    .region("ap-southeast-1") // 内容审核服务支持中国-香港(ap-southeast-1)、亚太-新加坡(ap-
    southeast-3)的配置
    .connectionTimeout(5000) // 连接目标url超时限制
    .connectionRequestTimeout(1000) // 连接池获取可用连接超时限制
    .socketTimeout(20000) // 获取服务器响应数据超时限制
    .build();
```

----结束

## 步骤六：调用服务

**步骤1** 选择本地图片或者使用Demo默认图片，参考如下示例代码修改

“ModerationImageContentDemo.java”文件中图片文件路径（“data/moderation-demo-1.jpg”）。

```
public static void main(String[] args) throws IOException {
    ModerationImageContentDemo tool = new ModerationImageContentDemo();
    tool.imageContentCheck("https://sdk-obs-source-save.obs.cn-north-4.myhuaweicloud.com/XXX.jpg");
    byte[] imageBytes = FileUtils.readFileToByteArray(new File("data/moderation-demo-1.jpg"));
    tool.imageContentCheck(imageBytes);
}
```

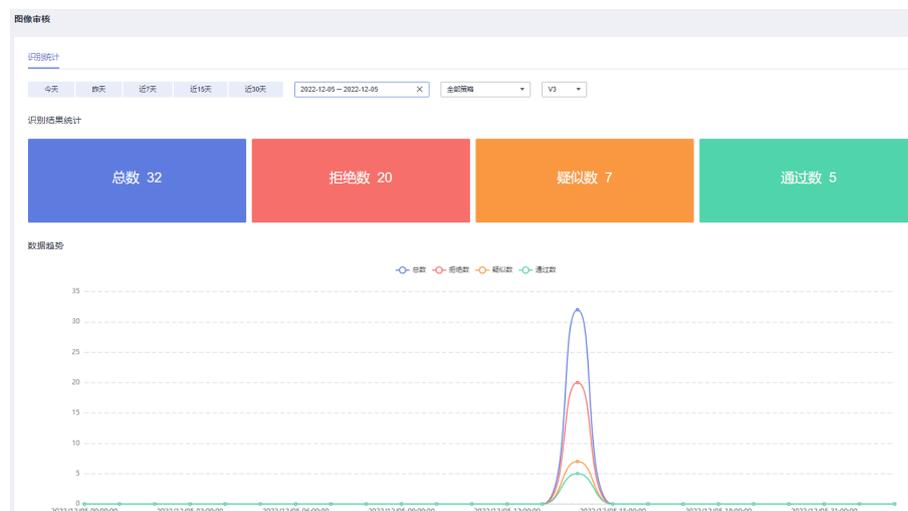
**步骤2** 执行“ModerationImageContentDemo.java”文件，控制台输出200即表示程序执行成功，**图像内容审核结果**输出到控制台，如图4-8所示。

图 4-8 运行结果

```
200
{"result":{"detail":{"terrorism":[{"confidence":0.0,"label":"bloody"}, {"confidence":0.0,"label":"fire"}, {"confidence":0.0,"label":"gun"}, {"confidence":0.0,"label":"knife"}, {"confidence":0.0,"label":"flag"}, {"confidence":0.0,"label":"symbol"}, {"confidence":0.0,"label":"dress"}, {"confidence":0.0,"label":"war"}, {"confidence":0.0,"label":"tiananmen"}], "suggestion":"pass"}}
```

**步骤3** 查看调用次数。您可以在“服务列表”，“图像审核”页查看调用详情和调用次数统计。如图4-9所示。

图 4-9 识别统计



- 识别结果统计：显示一段时间范围，内容审核的调用总数，拒绝数，疑似数和通过数，帮助您更好了解服务的调用情况和审核情况。
  - 总数：指的是审核调用总次数。
  - 拒绝数：指的是block总数，即文本中包含敏感信息，审核不通过的次数。
  - 疑似数：指的是review总数，即人工复查审核的次数。
  - 通过数：指的是pass总数，即通过审核的次数。
- 数据趋势：显示您设置的这段时间范围内，总数，拒绝数，疑似数和通过数的变化趋势。

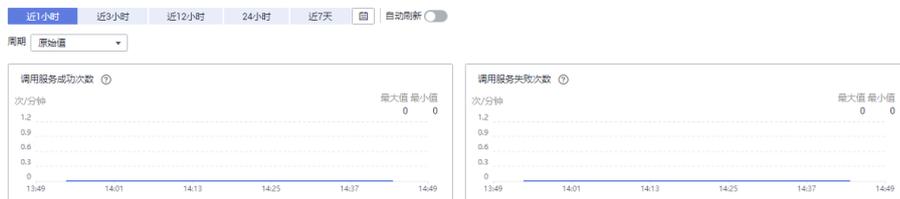
图 4-10 原因分布



- 拒绝数据原因分布：显示您设置的这段时间范围内，审核不通过的检测场景占比数。
- 疑似数据原因分布：显示您设置的这段时间范围内，需要人工复查的检测场景占比数。

**步骤4** 查看监控数据。您可以单击“查看监控指标”在云监控控制台查看服务调用成功和失败的次数等历史数据，如图4-11所示。

图 4-11 监控数据



---结束