

文字识别服务

产品介绍

文档版本 01
发布日期 2024-12-27



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

1 图解文字识别	1
2 什么是文字识别	3
3 功能介绍	5
3.1 通用类	5
3.2 证件类	5
4 约束与限制	7
5 与其他云服务的关系	11
6 安全	12
6.1 责任共担	12
6.2 身份认证与访问控制	13
6.3 数据保护技术	14
6.4 审计	14
6.5 服务韧性	14
6.6 监控安全风险	14

1 图解文字识别

2 什么是文字识别

文字识别（Optical Character Recognition，简称OCR）是指将图片中的打印字符进行检测识别成可编辑的文本格式，以JSON格式返回识别结果。

OCR以开放API（Application Programming Interface，应用程序编程接口）的方式提供给用户，用户使用Python、Java等编程语言调用OCR服务API将图片识别成文字，帮助用户自动采集关键数据，打造智能化业务系统，提升业务效率，API获取详情请参考《[API参考](#)》。

OCR还提供多种编程语言的SDK供您使用，SDK使用方法请参考《[SDK参考](#)》。

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。

OCR 能力

- **通用类 OCR**
通用类OCR（General OCR），支持网络图片等任意格式图片上文字信息的自动化识别，快速实现各种文档电子化。
- **证件类 OCR**
证件类OCR（Card OCR），支持身份证、驾驶证、护照等证件图片上有效信息的自动识别和关键字段结构化提取。

首次使用 OCR

如果您是首次使用OCR的用户，建议您学习并了解如下信息：

- 功能介绍
通过功能介绍章节的内容，了解OCR不同类别的文字识别功能，包括[通用类文字识别](#)、[证件类文字识别](#)。
- 入门使用
OCR以开放API的方式提供给用户，您可以参考《[快速入门](#)》学习并使用OCR服务。

- 使用方式
如果您是一个开发工程师，熟悉代码编写，想要直接调用OCR服务，您可以参考《[API参考](#)》或《[SDK参考](#)》获取详情。
- 由浅入深学习
您可以参考[成长地图](#)，由浅入深学习使用OCR。

3 功能介绍

3.1 通用类

功能介绍

- 通用表格识别
提取表格内的文字和所在行列位置信息，适应不同格式的表格。同时也识别表格外部的文字区域。用于各种单据和报表的电子化，恢复结构化信息。
- 通用文字识别
提取图片内的文字及其对应位置信息，并能够根据文字在图片中的位置进行结构化整理工作。
- 网络图片识别
自动识别网络图片内的所有文字及其对应位置信息，如果图片包含联系人信息，也支持联系人信息的提取，同时可供进一步的数据挖掘后处理操作。

应用场景

- 纸质文件电子化
自动识别结构化信息与提取签名盖章区域，有助快速审核。
- 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。

3.2 证件类

功能介绍

- 通用证件识别
识别身份证件图像，并将识别的结构化结果返回给用户。支持多个国家/地区的身份证、驾驶证和护照，具体国家/地区和证件列表详见[API接口文档](#)。

- **护照识别**
针对中国护照，可识别护照上的全部信息；针对其他国家护照，可根据两行国际标准化的机读码识别出6-7个关键字段信息，包括姓名、性别、出生日期、护照号码、签发国国家码、护照有效期等。
- **泰文身份证识别**
自动识别身份证上的全部信息，一次扫描即可识别身份证号码、姓名、地址等全部信息。
- **柬文身份证识别**
识别柬文身份证图片中的文字内容，一次扫描即可识别姓名、出生日期、性别、签发时间等全部信息。
- **缅文身份证识别**
自动识别身份证上的全部信息，支持身份证正反面识别，一次扫描即可识别身份证号码、姓名、地址等全部信息。
- **泰国车牌识别**
识别泰国车牌图片中的车牌信息，返回识别结果和车牌的区域位置信息。
- **缅文驾驶证识别**
自动识别缅文驾驶证的全部信息，一次扫描即可识别身份证号、名字、nrc号码、出生日期、血型、有效期等全部信息。
- **智利身份证识别**
自动识别智利身份证图片中的文字内容，一次扫描即可识别身份证号码、姓名、国籍、性别、出生日期、有效期等全部信息。
- **越南身份证识别**
自动识别越南身份证图片中的文字内容，一次扫描即可识别卡证编号、姓名、国籍、性别、出生日期、有效期等全部信息。
- **秘鲁身份证识别**
自动识别秘鲁身份证图片中的文字内容，一次扫描即可识别身份证号、第一姓氏、第二姓氏、名、性别等全部信息。

应用场景

- **认证场景**
核验用户是否为真实证件的持有者本人。
- **证件信息录入**
识别证件中关键信息，节省人工录入，提升效率。
- **人证核身**
核验用户是否为真实证件的持有者本人。

4 约束与限制

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。其中系统级约束限制，是所有子服务的约束。除系统级约束限制外，各子服务还有独立的约束条件。

智能文档解析

- 支持中英文及部分中文繁体字。
- 只支持识别PNG、JPG、JPEG、BMP、GIF、TIFF、WEBP、PCX、ICO、PSD格式的图片以及PDF文档，PDF只支持单页识别，可通过pdf_page_number参数指定页码。
- 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。
- 单页字符数不大于1800，以获得较优的识别效果。
- 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档解析。
- 文字识别服务属于公有云服务，线上用户资源共享，如果需要多并发请求，请提前[联系我们](#)。

通用证件识别

- 只支持JPEG、JPG、PNG、BMP、TIFF格式的图片。
- 图片最小边不小于100px，最长边不超过8192px。

通用表格识别

- 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%，保证整张表格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）和表格线扭曲图像的文字识别。
- 支持中英文以及部分繁体字。

通用文字识别

- 只支持识别PNG、JPG、JPEG、BMP、GIF、TIFF、WEBP、PCX、ICO、PSD、PDF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。
- 支持自动过滤浅色文字水印。
- 目前不支持复杂背景（如户外自然场景等）和文字扭曲图像的文字识别。
- 支持中英文以及部分繁体字、马来语、乌克兰语、印地语、俄语、越南语、印尼语、泰语、阿拉伯语、德语、拉丁语、法语、意大利语、西班牙语、葡萄牙语、罗马尼亚语、波兰语、阿姆哈拉语、日语、韩语、土耳其语、挪威语、丹麦语、瑞典语、柬埔寨语、希伯来语识别。

网络图片识别

- 支持中英文及部分中文繁体字。
- 只支持识别JPG、JPEG、PNG、BMP、TIFF、TGA、WEBP、ICO、PCX、GIF格式图片。
- 支持常见网络图片如：手机截图、电脑截图、电商产品图及广告设计图等互联网图片。
- 图像各边的像素大小在15到30000px之间。
- 图像中有效文字图片占比超过60%，避免有效文字图片占比过小。
- 支持图像中有效文字图片的任意角度的水平旋转（需开启方向检测）。

护照识别

- 支持中国大陆护照的全字段识别。
- 支持含有完整机读码的中国-港澳台地区及外国护照识别。
- 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15到8192px之间。
- 图像中护照首页区域有效占比超过25%，保证护照首页内容及其边缘包含在图像内。
- 支持图像中护照任意角度的水平旋转。
- 支持少量扭曲，扭曲后图像中的护照长宽比与实际护照相差不超过10%。
- 能处理反光、暗光等干扰的图片但影响识别精度。

泰文身份证识别

- 只支持泰国身份证的识别。
- 只支持识别PNG、JPG、JPEG、BMP、TIFF格式的图片。
- 泰国身份证识别图像各边的像素大小在15到8192px之间。
- 图像中身份证区域有效占比超过25%，保证整张身份证内容及其边缘包含在图像内。
- 支持图像中身份证任意角度的水平旋转。

- 支持少量扭曲，扭曲后图像中的身份证长宽比与实际身份证相差不超过10%。
- 能处理反光、暗光等干扰的图片但影响识别精度。
- 目前只支持识别单张身份证的正面或者反面。

柬文身份证识别

- 目前只支持单张柬文身份证正面识别。
- 只支持识别PNG、JPG、BMP、TIFF格式的图片。
- 图像各边的像素大小在15到8192px之间。
- 支持图像中柬文身份证任意角度的水平旋转。
- 能处理反光、暗光等干扰的图片但影响识别精度。

缅文身份证识别

- 只支持缅甸身份证的识别。
- 只支持识别PNG、JPG、JPEG、BMP、TIFF格式的图片。
- 缅甸身份证识别图像各边的像素大小在15到8192px之间。
- 图像中身份证区域有效占比超过25%，保证整张身份证内容及其边缘包含在图像内。
- 支持图像中身份证任意角度的水平旋转。
- 支持少量扭曲，扭曲后图像中的身份证长宽比与实际身份证相差不超过10%。
- 能处理反光、暗光等干扰的图片但影响识别精度。
- 目前只支持识别单张身份证的正面或者反面。

缅文驾驶证识别

- 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15到4096px之间。
- 目前只支持单张缅文驾驶证正面识别。
- 支持图像中缅文驾驶证任意角度的水平旋转。
- 支持少量扭曲，扭曲后图像中的驾驶证长宽比与实际驾驶证相差不超过10%。
- 能处理反光、暗光等干扰的图片但影响识别精度。

智利身份证识别

- 只支持识别PNG、JPG、BMP、TIFF格式的图片。
- 图像各边的像素大小在15到8192px之间。
- 目前只支持单张智利身份证正面识别。
- 支持图像中智利身份证任意角度的水平旋转。
- 能处理反光、暗光等干扰的图片但影响识别精度。

越南身份证识别

- 只支持越南身份证的识别。
- 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15到8192px之间。
- 支持图像中越南身份证任意角度的水平旋转。
- 能处理反光、暗光、防伪标识等干扰的图片但影响识别精度。

秘鲁身份证识别

- 只支持秘鲁身份证的识别。
- 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15到8192px之间。
- 图像中秘鲁身份证区域有效占比超过25%，保证整张秘鲁身份证内容及其边缘包含在图像内。
- 支持图像中秘鲁身份证任意角度的水平旋转。
- 支持少量扭曲，扭曲后图像中的秘鲁身份证长宽比与实际秘鲁身份证相差不超过10%。
- 能处理反光、暗光等干扰的图片但影响识别精度。

泰国车牌识别

- 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15到8192px之间。
- 支持图像中车牌任意角度的水平旋转。
- 能处理反光、暗光等干扰的图片但影响识别精度。

5 与其他云服务的关系

统一身份认证服务

统一身份认证（Identity and Access Management，简称IAM）服务，IAM为文字识别提供了用户认证和鉴权功能。

云监控

云监控（Cloud Eye）可以监控文字识别的相关指标，具体如表5-1所示。用户可以通过指标及时了解文字识别的使用情况。Cloud Eye的更多信息请参见《云监控用户指南》。

表 5-1 OCR 的监控指标

指标名称	指标含义	取值范围	测量对象
调用服务成功次数	该指标用于统计用户请求服务成功次数，以“次每分钟”为单位。	≥0次每分钟	文字识别
调用服务失败次数	该指标用于统计用户请求服务失败次数，以“次每分钟”为单位。	≥0次每分钟	文字识别

说明

每个子服务都有调用成功次数和失败次数两个指标。

对象存储服务

对象存储服务（Object Storage Service，简称OBS）是稳定、安全、高效、易用的云存储服务。文字识别接口都涉及到对用户的数据处理，用户的大量数据采用OBS批量方式处理，可以提升云上处理的总体效率。

文字识别接口支持从OBS上采用临时授权或者匿名公开授权的方式获取数据并进行处理。

6 安全

6.1 责任共担

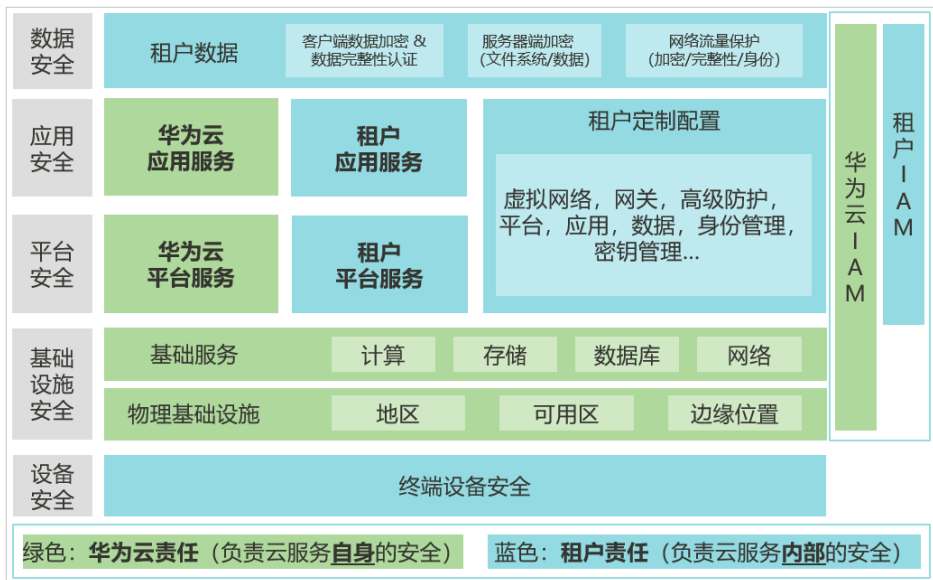
华为云秉承“将对网络和业务安全性保障的责任置于公司的商业利益之上”。针对层出不穷的云安全挑战和无孔不入的云安全威胁与攻击，华为云在遵从法律法规业界标准的基础上，以安全生态圈为护城河，依托华为独有的软硬件优势，构建面向不同区域和行业的完善云服务安全保障体系。

安全性是华为云与您的共同责任，如图6-1所示。

- **华为云**：负责云服务自身的安全，提供安全的云。华为云的安全责任在于保障其所提供的IaaS、PaaS和SaaS类云服务自身的安全，涵盖华为云数据中心的物理环境设施和运行其上的基础服务、平台服务、应用服务等。这不仅包括华为云基础设施和各项云服务技术的安全功能和性能本身，也包括运维运营安全，以及更广义的安全合规遵从。
- **租户**：负责云服务内部的安全，安全地使用云。华为云租户的安全责任在于对使用的IaaS、PaaS和SaaS类云服务内部的安全以及对租户定制配置进行安全有效的管理，包括但不限于虚拟网络、虚拟主机和访客虚拟机的操作系统，虚拟防火墙、API网关和高级安全服务，各项云服务，租户数据，以及身份账号和密钥管理等方面的安全配置。

《[华为云安全白皮书](#)》详细介绍华为云安全性的构建思路与措施，包括云安全战略、责任共担模型、合规与隐私、安全组织与人员、基础设施安全、租户服务与租户安全、工程安全、运维运营安全、生态安全。

图 6-1 华为云安全责任共担模型



6.2 身份认证与访问控制

身份认证

用户访问OCR的方式有多种，包括OCR控制台、API、SDK，无论访问方式封装成何种形式，其本质都是通过OCR提供的REST风格的API接口进行请求。

OCR的接口有如下两种认证方式，您可以选择其中一种进行认证鉴权。

- Token认证：通过Token认证调用请求。
- AK/SK认证：通过AK (Access Key ID) /SK (Secret Access Key)加密调用请求。经过认证的请求总是需要包含一个签名值，该签名值以请求者的访问密钥 (AK/SK) 作为加密因子，结合请求体携带的特定信息计算而成。通过访问密钥 (AK/SK) 认证方式进行认证鉴权，即使用Access Key ID (AK) /Secret Access Key (SK) 加密的方法来验证某个请求发送者身份。关于访问密钥的详细介绍及获取方式，请参见[认证鉴权](#)。

访问控制

OCR支持通过权限控制 (IAM权限) 进行访问控制。

表 6-1 表 1 OCR 访问控制

访问控制方式		简要说明
权限控制	IAM权限	IAM权限是作用于云资源的，IAM权限定义了允许和拒绝的访问操作，以此实现云资源权限访问控制。管理员创建IAM用户后，需要将用户加入到一个用户组中，IAM可以对这个组授予OCR所需的权限，组内用户自动继承用户组的所有权限。

访问控制方式		简要说明
	委托授权	如果用户想使用存储在对象存储服务（OBS）上的数据，他可以委托授权OCR服务访问对象存储服务。

6.3 数据保护技术

OCR服务处理以下类型的数据：

- OCR输入数据，包括图像（PNG、JPG、JPEG和BMP等）。
- OCR结果，包括从客户图像中提取的文本及其位置以及置信度等。

OCR通过多种数据保护手段和特性，保障客户在使用OCR服务时数据安全。

表 6-2 OCR 的数据保护手段和特性

数据保护手段	说明
传输加密（HTTPS）	OCR支持HTTPS传输协议，采用华为云官网标准网关，数据传输过程有全链路安全保障。API网关支持TLS 1.1及TLS 1.2版本，推荐您使用更加安全的TLS 1.2版本。
数据存储	华为云OCR承诺数据不落盘，用户的原始图片和识别结果数据均不作保留，识别返回后立即释放。

6.4 审计

云审计服务（Cloud Trace Service，CTS），是华为云安全解决方案中专业的日志审计服务，提供对各种云资源操作记录的收集、存储和查询功能，可用于支撑安全分析、合规审计、资源跟踪和问题定位等常见应用场景。

用户开通云审计服务并创建和配置追踪器后，CTS可记录OCR的管理事件和数据事件用于审计。

CTS的详细介绍和开通配置方法，请参见[CTS快速入门](#)。

CTS支持追踪的OCR管理事件列表，请参见[审计](#)。

6.5 服务韧性

OCR提供三级可靠性架构，通过AZ之间容灾、AZ内实例冗余、实例健康检测等技术方案，保障服务的可靠性。

6.6 监控安全风险

OCR提供基于云监控服务CES的监控能力，帮助用户监控账号下的OCR服务API，执行自动实时监控、告警和通知操作。用户可以实时掌握API所产生的调用成功次数、调用失败次数等信息。

关于OCR支持的监控指标，以及如何创建监控告警规则等内容，请参见[监控](#)。

同时，OCR控制台支持监控账号下的OCR服务每个API的用量统计，详情请参见[统计API使用量](#)。