Optical Character Recognition

Service Overview

Issue 01

Date 2024-12-27

HUAWEI CLOUD COMPUTING TECHNOLOGIES CO., LTD.

Trademarks and Permissions


and other Huawei trademarks are the property of Huawei Technologies Co., Ltd.
All other trademarks and trade names mentioned in this document are the property of their respective holders.

Notice

The purchased products, services and features are stipulated by the contract made between Huawei Cloud and the customer. All or part of the products, services and features described in this document may not be within the purchase scope or the usage scope. Unless otherwise specified in the contract, all statements, information, and recommendations in this document are provided "AS IS" without warranties, guarantees or representations of any kind, either express or implied.

The information in this document is subject to change without notice. Every effort has been made in the preparation of this document to ensure accuracy of the contents, but all statements, information, and recommendations in this document do not constitute a warranty of any kind, express or implied.

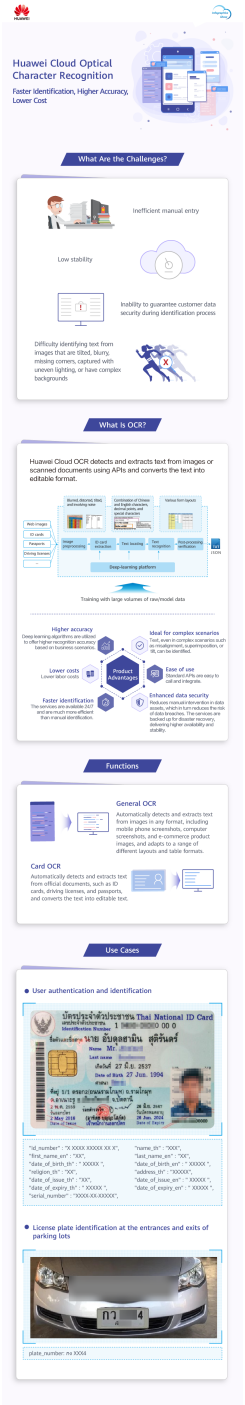
Huawei Cloud Computing Technologies Co., Ltd.

Address: Huawei Cloud Data Center Jiaoxinggong Road
Qianzhong Avenue
Gui'an New District
Gui Zhou 550029
People's Republic of China

Website: https://www.huaweicloud.com/intl/en-us/

Contents

1 OCR Infographics



2 What Is OCR?

Optical Character Recognition (OCR) detects and extracts text from images and converts the recognition results into an editable JSON format.

OCR provides open APIs, so you can use programming languages such as Python and Java to call OCR APIs to extract text from images. OCR allows you to automate the collection of key data. It helps you build an intelligent service system to improve efficiency. For details about how to obtain APIs, see **Optical Character Recognition API Reference**.

OCR also provides software development kits (SDKs) for multiple programming languages. For details about how to use SDKs, see the **Optical Character Recognition SDK Reference**.

Before You Start

You will need some basic programming skills. Familiarity with Java, Python, iOS, Android, and Node.js is recommended.

You need to call APIs to use OCR and transmit the results to the service system, or to convert the results from JSON to TXT or Excel form.

OCR Capabilities

- **General OCR**

  Text in images (including web images and more) can be automatically identified.
- **Card OCR**

  OCR automatically identifies information in images of certificates such as passports, ID cards, driving licenses, and converts the information into editable text.

Using OCR for the First Time

If you are a first-time user, the following sections are a good place to start:

- Function Description

  Learn about the different OCR functions, including **General OCR** and **Card OCR**.

- Getting Started

  Learn how to use OCR by referring to **Optical Character Recognition Getting Started**.

- Using OCR

  Learn how to call OCR services as a developer who feel more comfortable writing code, see **Optical Character Recognition API Reference** or **Optical Character Recognition SDK Reference**.

- Progressive Knowledge

  Learn how to get started using OCR.

3 Function Description

3.1 General OCR

Function Description

- General Table OCR

  Detects and extracts text and their row and column locations from images of tables in various formats, as well as the text areas outside tables. It is used to store information on documents and reports as structured data.

- General Text OCR

  Detects and extracts text and their locations from images and converts them into structured data.

- Web Image OCR

  Detects and extracts all text, their locations, and contact information (if any) from web images for data mining and post-processing.

Application Scenarios

- Electronic documentation archive

  Detects and extracts text, signatures, and seals from document images and converts them into structured data for faster review.

- Express waybill filling

  Detects and extracts contact information from images and fills in express waybills, eliminating the need for manual input.

- Contract upload and review

  Detects and extracts text, signatures, and seals from document images and converts them into structured data for faster review.

3.2 Card OCR

Function Description

- ID Document OCR

  Automatically detects and extracts text from images of identity documents and converts the text into a structured format. These documents include ID cards, driving licenses, and passports from multiple countries and regions. For details about the mapping between supported countries/regions and document types, see **ID Document**.

- Passport OCR

  Automatically detects and extracts all information from images of Chinese passports and six to seven key fields from images of passports issued by other countries based on the machine-readable code, including the name, gender, date of birth, passport number, country code, and date of expiry.

- Thailand ID Card OCR

  Automatically detects and extracts all information from ID card images, including the ID number, name, and address.

- Cambodian ID Card OCR

  Detects and extracts text from images of Cambodia-issued ID cards, including the name, date of birth, gender, and issuance date.

- Myanmar ID Card OCR

  Automatically detects and extracts all information from images of both sides of ID cards, including the ID number, name, and address.

- Thailand Plate Number OCR

  Automatically detects and extracts license plate information from images of Thailand license plates and returns the license plate number and location.

- Myanmar Driving License OCR

  Automatically detects and extracts text from Myanmar-issued driving licenses and returns information such as the ID card number, name, National Registration Card (NRC) number, date of birth, blood type, and validity period.

- Chile ID Card OCR

  Automatically detects and extracts text from Chile ID card images, including the ID card number, name, nationality, gender, date of birth, and validity period.

- Vietnam ID Card OCR

  Automatically detects and extracts text from images of Vietnam-issued ID cards, including the card number, name, nationality, gender, date of birth, and validity period.

- Peru ID Card OCR

  Automatically detects and extracts text from images of Peru-issued ID cards, including the ID number, first surname, second surname, other names, and gender.

Application Scenarios

- Authentication

  Verifies that the user is the certificate holder.

- Certificate information entry

  Automatically detects and extracts key information from certificate images, eliminating the need for manual entry.

- Identity verification

  Verifies that the user is the certificate holder.

4 Notes and Constraints

Due to various technological and cost issues, there are some constraints on how OCR can be used. System-wide constraints affect all sub-services. Additionally, sub-services are subject to their respective constraints.

Smart Document Analysis

- English and Chinese are supported but support for traditional Chinese characters is limited.
- Only images in PNG, JPG, JPEG, BMP, GIF, TIFF, WebP, PCX, ICO or PSD format and PDF files can be recognized. PDF files can only be recognized one page at a time, but you can use the **pdf_page_number** parameter to specify which page you want to recognize.
- No side of the image can be smaller than 15 or larger than 8,192 pixels.
- The area to be recognized must occupy more than 80% of the image. When scanning a table, ensure that all text and its surrounding area are included in the image.
- An image can be rotated to any angle.
- For more accurate recognition results, the number of characters on a single page must be limited to 1,800 or less.
- Text in images with complex backgrounds (such as outdoor scenery or anti-counterfeit watermarks) or distorted text cannot be analyzed.
- OCR is a public cloud service whose resources are sharable to all online users. If you need to call multiple APIs concurrently, **contact us**.

ID Document OCR

- Only images in JPEG, JPG, PNG, BMP, or TIFF format can be recognized.
- No side of the image can be smaller than 100 or larger than 8,192 pixels.

General Table OCR

- Only images in PNG, JPG, JPEG, BMP, or TIFF format can be recognized.
- No side of the image can be smaller than 15 or larger than 8,192 pixels.
- The area to be recognized must occupy more than 80% of the image. When scanning a table, ensure that the entire table and its surrounding area are included in the image.

- An image can be rotated to any angle.
- Text in images with complex backgrounds (such as outdoor scenery or anti-counterfeit watermarks) or distorted table lines cannot be recognized.
- English and Chinese are supported but support for traditional Chinese characters is limited.

General Text OCR

- Only images in PNG, JPG, JPEG, BMP, GIF, TIFF, WebP, PCX, ICO, PDF, or PSD format can be recognized.
- No side of the image can be smaller than 15 or larger than 8,192 pixels.
- The area to be recognized must occupy more than 80% of the image. When scanning a table, ensure that all text and its surrounding area are included in the image.
- An image can be rotated to any angle.
- Light-colored text watermarks can be automatically filtered out.
- Text in images with complex backgrounds (such as outdoor scenery) or distorted text cannot be recognized.
- Supported languages: Chinese, English, some traditional Chinese, Malay, Ukrainian, Hindi, Russian, Vietnamese, Indonesian, Thai, Arabic, German, Latin, French, Italian, Spanish, Portuguese, Romanian, Polish Amharic, Japanese, Korean, Turkish, Norwegian, Danish, Swedish, Khmer, and Hebrew.

Web Image OCR

- English and Chinese are supported but support for traditional Chinese characters is limited.
- Only images in JPG, JPEG, PNG, BMP, TIFF, TGA, WEBP, ICO, PCX, or GIF format can be recognized.
- Common image types are supported, such as mobile phone or desktop screenshots, e-commerce product images, and advertisement design drawings.
- No side of the image can be smaller than 15 or larger than 30,000 pixels.
- The characters to be recognized must occupy more than 60% of the image.
- The web image to be recognized can be rotated to any angle (direction detection must be enabled).

Passport OCR

- All fields on Chinese mainland passports can be recognized.
- Passports that are issued by China, Hong Kong (China), Macao (China), Taiwan (China), and other countries and regions and that are with complete machine-readable codes can be recognized.
- Only images in PNG, JPG, JPEG, BMP, or TIFF format can be recognized.
- No side of the image can be smaller than 15 or larger than 8,192 pixels.
- The information page of the passport to be recognized must occupy more than 25% of the image. When scanning a passport, ensure that the entire page is displayed in the image.
- A passport can be rotated to any angle.

- The passport in the image can be moderately distorted, but the aspect ratio cannot be distorted by more than 10%.
- Illuminated or dark images can be recognized, but the accuracy may be compromised.

Thailand ID Card OCR

- Only ID cards issued by Thailand can be recognized.
- Only images in PNG, JPG, JPEG, BMP, or TIFF format can be recognized.
- No side of the image can be smaller than 15 or larger than 8,192 pixels.
- An ID card to be recognized must occupy more than 25% of the image. When scanning an ID card, ensure that the entire ID card is displayed in the image.
- An ID card can be rotated to any angle.
- The ID card in the image can be moderately distorted, but the aspect ratio cannot be distorted by more than 10%.
- Illuminated or dark images can be recognized, but the accuracy may be compromised.
- Only the front or back of a single ID card can be identified each time.

Cambodian ID Card OCR

- Currently, only the front of an ID card can be recognized each time.
- Only images in PNG, JPG, BMP, or TIFF format can be recognized.
- No side of the image can be smaller than 15 or larger than 8,192 pixels.
- An ID card can be rotated to any angle.
- Illuminated or dark images can be recognized, but the accuracy may be compromised.

Myanmar ID Card OCR

- Only the national registration cards issued by Myanmar can be recognized.
- Only images in PNG, JPG, JPEG, BMP, or TIFF format can be recognized.
- No side of the image can be smaller than 15 or larger than 8,192 pixels.
- An ID card to be recognized must occupy more than 25% of the image. When scanning an ID card, ensure that the entire ID card is displayed in the image.
- An ID card can be rotated to any angle.
- The ID card in the image can be moderately distorted, but the aspect ratio cannot be distorted by more than 10%.
- Illuminated or dark images can be recognized, but the accuracy may be compromised.
- Only the front or back of a single ID card can be identified each time.

Myanmar Driving License OCR

- Only images in PNG, JPG, JPEG, BMP, or TIFF format can be recognized.
- No side of the image can be smaller than 15 or larger than 4,096 pixels.
- Currently, only the front of a driving license can be recognized each time.

- A driving license can be rotated to any angle.
- The driving license in the image can be moderately distorted, but the aspect ratio cannot be distorted by more than 10%.
- Illuminated or dark images can be recognized, but the accuracy may be compromised.

Chile ID Card OCR

- Only images in PNG, JPG, BMP, or TIFF format can be recognized.
- No side of the image can be smaller than 15 or larger than 8,192 pixels.
- Currently, only the front of an ID card can be recognized each time.
- An ID card can be rotated to any angle.
- Illuminated or dark images can be recognized, but the accuracy may be compromised.

Vietnam ID Card OCR

- Only ID cards issued by Vietnam can be recognized.
- Only images in PNG, JPG, JPEG, BMP, or TIFF format can be recognized.
- No side of the image can be smaller than 15 or larger than 8,192 pixels.
- An ID card can be rotated to any angle.
- Illuminated or dark images, or images with anti-counterfeit watermarks can be recognized, but the accuracy may be compromised.

Peru ID Card OCR

- Only ID cards issued by Peru can be recognized.
- Only images in PNG, JPG, JPEG, BMP, or TIFF format can be recognized.
- No side of the image can be smaller than 15 or larger than 8,192 pixels.
- An ID card to be recognized must occupy more than 25% of the image. When scanning an ID card, ensure that the entire ID card is displayed in the image.
- An ID card can be rotated to any angle.
- The ID card in the image can be moderately distorted, but the aspect ratio cannot be distorted by more than 10%.
- Illuminated or dark images can be recognized, but the accuracy may be compromised.

Thailand Plate Number OCR

- Only images in PNG, JPG, JPEG, BMP, or TIFF format can be recognized.
- No side of the image can be smaller than 15 or larger than 8,192 pixels.
- A license plate can be rotated to any angle.
- Illuminated or dark images can be recognized, but the accuracy may be compromised.

5 Related Services

IAM

Identity and Access Management (IAM) lets you control user authentication and access to OCR.

Cloud Eye

Cloud Eye monitors the metrics of OCR listed in **Table 5-1**. You use these metrics to monitor OCR usage. For more information about Cloud Eye, see *Cloud Eye User Guide*.

Table 5-1 OCR monitoring metrics

Metric	Description	Value Range	Monitored Object
Successful Calls	Counts the number of successful API calls. The unit is API calls/minute.	≥ 0 API calls/minute	OCR
Failed Calls	Counts the number of failed API calls. The unit is API calls/minute.	≥ 0 API calls/minute	OCR

NOTE

Each sub-service includes both the preceding metrics.

OBS

Object Storage Service (OBS) is a stable, secure, efficient, and easy-to-use cloud storage service. OCR APIs involve a lot of data processing. You can use OBS to improve processing efficiency by batch processing data on the cloud.

OCR APIs can be temporarily authenticated or anonymously and publicly authorized to obtain data from OBS for processing.

6 Security

6.1 Shared Responsibilities

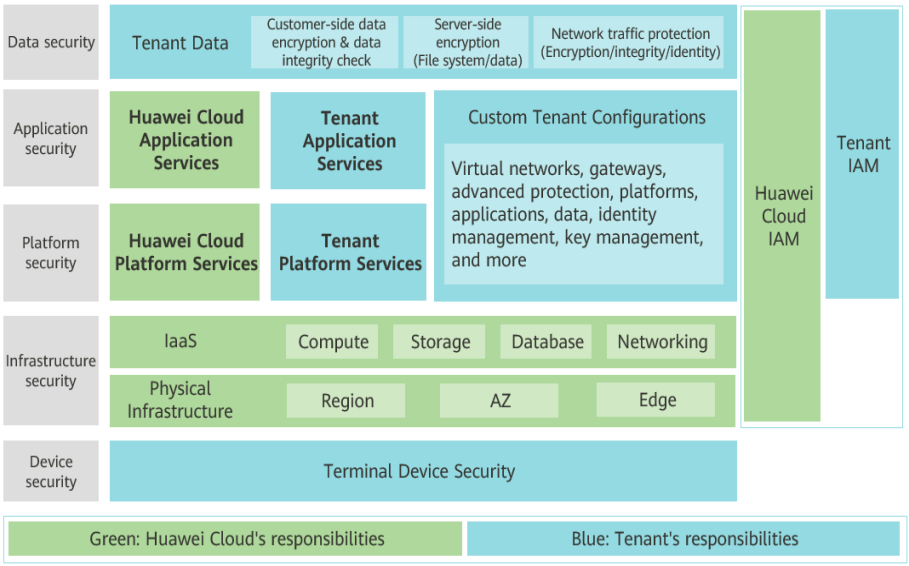
Huawei guarantees that its commitment to cyber security will never be outweighed by the consideration of commercial interests. To cope with emerging cloud security challenges and pervasive cloud security threats and attacks, Huawei Cloud builds a comprehensive cloud service security assurance system for different regions and industries based on Huawei's unique software and hardware advantages, laws, regulations, industry standards, and security ecosystem.

Figure 6-1 illustrates the responsibilities shared by Huawei Cloud and users.

- **Huawei Cloud**: Ensure the security of cloud services and provide secure clouds. Huawei Cloud's security responsibilities include ensuring the security of our IaaS, PaaS, and SaaS services, as well as the physical environments of the Huawei Cloud data centers where our IaaS, PaaS, and SaaS services operate. Huawei Cloud is responsible for not only the security functions and performance of our infrastructure, cloud services, and technologies, but also for the overall cloud O&M security and, in the broader sense, the security and compliance of our infrastructure and services.
- **Tenant**: Use the cloud securely. Tenants of Huawei Cloud are responsible for the secure and effective management of the tenant-customized configurations of cloud services including IaaS, PaaS, and SaaS. This includes but is not limited to virtual networks, the OS of virtual machine hosts and guests, virtual firewalls, API Gateway, advanced security services, all types of cloud services, tenant data, identity accounts, and key management.

Huawei Cloud Security White Paper elaborates on the ideas and measures for building Huawei Cloud security, including cloud security strategies, the shared responsibility model, compliance and privacy, security organizations and personnel, infrastructure security, tenant service and security, engineering security, O&M security, and ecosystem security.

Figure 6-1 Huawei Cloud shared security responsibility model



6.2 Identity Authentication and Access Control

Identity Authentication

You can use OCR services through the console, APIs, or SDKs. Essentially, access requests are sent through OCR RESTful APIs.

You can use either of the following authentication methods to call OCR APIs:

- Token authentication: Requests are authenticated using a token.
- AK/SK-based authentication. Requests are encrypted using access key ID (AK)/secret access key (SK). An authenticated request must contain a signature value. The signature value is calculated based on the requestor's access key (AK/SK) as the encryption factor and the specific information carried in the request body. OCR supports authentication using an access key (AK/SK pair). It uses AK/SK-based encryption to authenticate requests. For details about access keys and how to obtain them, see **Authentication**.

Access Control

OCR supports access control through permissions (IAM permissions).

Table 6-1 Table 1 OCR access control

Method		Description
Permission Control	IAM permissions	IAM permissions define which actions on your cloud resources are allowed and which actions are denied, to control access to your resources. After an IAM user is created, the administrator adds it to a user group. The administrator can grant the user group required OCR access permissions and all users in this group then inherit the granted permissions.
	Agency authorization	To use data stored on OBS, you can authorize OCR to access OBS.

6.3 Data Protection Technologies

OCR processes the following types of data:

- OCR input data, including images (PNG, JPG, JPEG, and BMP).
- OCR results, including the text extracted from customer images, its locations, and confidences.

OCR utilizes multiple data protection measures to ensure customer data security.

Table 6-2 Data protection measures

Measure	Description
Transmission encryption (HTTPS)	OCR uses HTTPS and Huawei Cloud API Gateway (APIG) to ensure full-link security during data transmission. APIG supports both TLS 1.1 and 1.2. TLS 1.2 is recommended because it is more secure.
Data storage	Huawei Cloud OCR promises that your data is not written to disks, and your data is deleted immediately after the recognition results are returned to you.

6.4 Auditing

Cloud Trace Service (CTS) records operations on the cloud resources in your account. You can use the logs generated by CTS to perform security analysis, track resource changes, audit compliance, and locate faults.

After you enable CTS and configure a tracker, CTS can record management and data traces of OCR for auditing.

For details about how to enable and configure CTS, see **Enabling CTS**.

For details about the OCR management and data traces that can be tracked by CTS, see **Audit**.

6.5 Service Resilience

OCR provides a three-level reliability architecture and uses technical solutions, such as inter-AZ disaster recovery (DR), intra-AZ instance redundancy, and instance health check, to ensure service reliability.

6.6 Security Risk Monitoring

OCR uses Cloud Eye to help you monitor your OCR APIs and receive alarms and notifications in real time. So you can obtain the number of successful and failed API calls in real time.

For details about OCR monitoring metrics and how to create alarm rules, see **OCR Monitoring Metrics**.

In addition, the OCR console can monitor the usage statistics of each OCR API under your account. For details, see **Collecting API Usage Statistics**.

