Microsoft Cognitive Services: Computer Vision API

UNDERSTANDING THE CORE ASPECTS OF THE COMPUTER VISION API

Eduardo Freitas
DATA CAPTURE SPECIALIST

https://edfreitas.me

Overview

What Are Microsoft Cognitive Services?

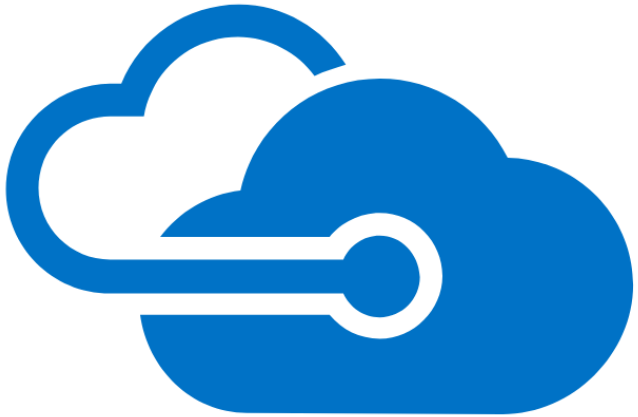
Fundamental Aspects of the Computer Vision API

What Is Image Classification & Identification?

What Is Text Recognition & Extraction?

What Is Thumbnail Generation?

What Are Microsoft Cognitive Services?



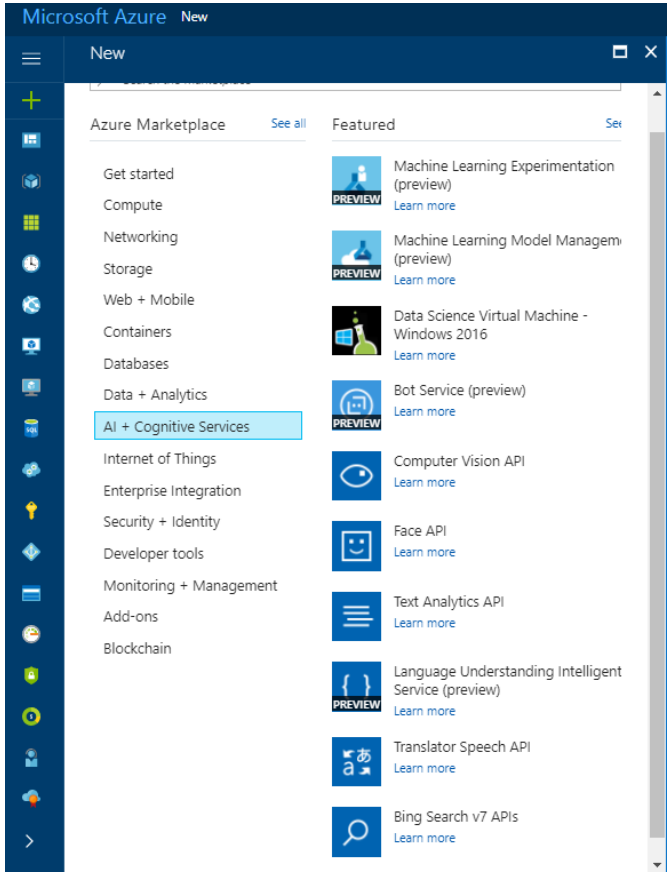
Set of APIs that perform specific AI features

Hosted on Microsoft Azure

Enable AI application development

Core AI Features of Cognitive Services

Getting Started With Azure



Azure subscription

API endpoint and subscription key

Send and receive JSON

SDKs for various programming languages

Fundamental Aspects of the Computer Vision API

Image Classification & Identification

Thumbnail Generation

Text Recognition

API Requirements

URL or raw binary image

Formats: JPEG, PNG, GIF, BMP

Less than 4MB

Dimensions > 50 x 50 pixels

Image Classification & Identification

- Content detection
- Content categorization
- 2000+ objects
- 86 taxonomy concepts
- Image type
- Image quality

Thumbnail Generation

Smart cropping

ROIs - recognizes regions of interest

Change aspect ratio

Text Recognition

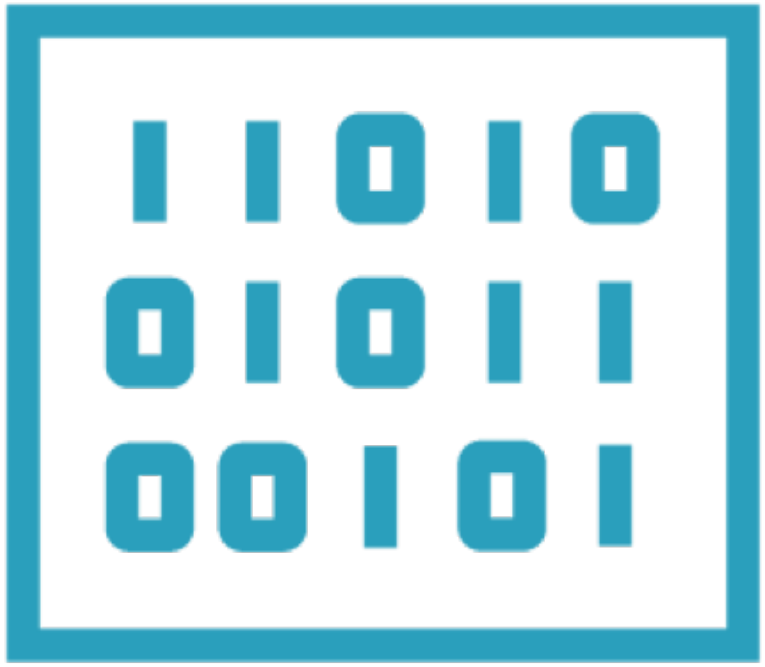
OCR - optical character recognition

25 languages

Rotation correction

Recognize handwritten text

What Is Image Classification & Identification?

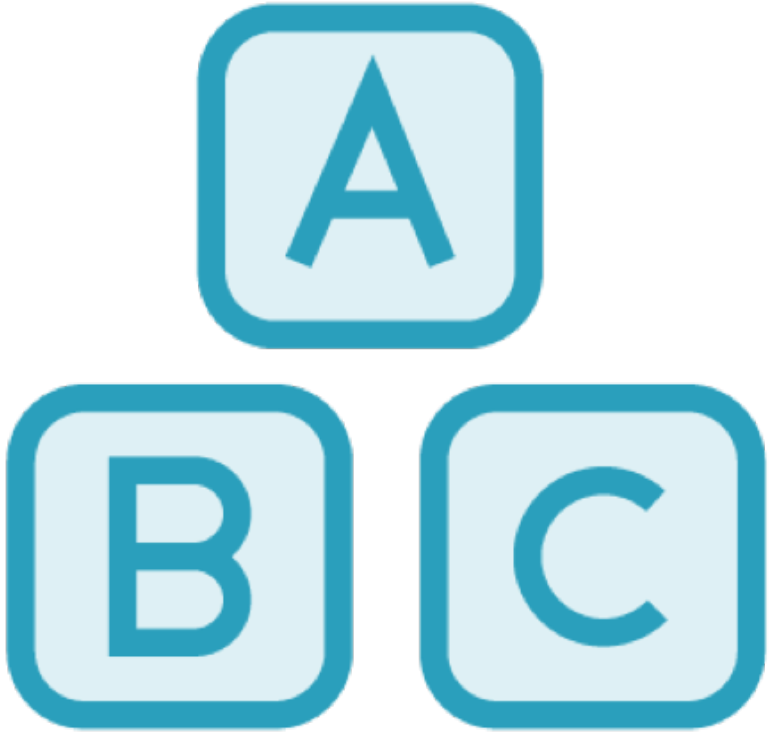


- Tagging & descriptions
- Flagging content
- Taxonomy based categories
- Image types and quality
- Face detection
- Perceiving color schemes

JSON Response Example

```
Returned Json
{
   'tags':[
      {
         "name":"grass",
         "confidence":0.999999761581421
      },
      {
         "name":"outdoor",
         "confidence":0.999970674514771
      },
      {
         "name":"sky",
         "confidence":999289751052856
      },
      {
         "name":"building",
         "confidence":0.996463239192963
      },
      {
         "name":"house",
         "confidence":0.992798030376434
      },
```

What Is Text Recognition & Extraction?



Optical Character Recognition (OCR)

- Medical records
- Insurance and banking
- Invoices

Recognize Handwritten Text

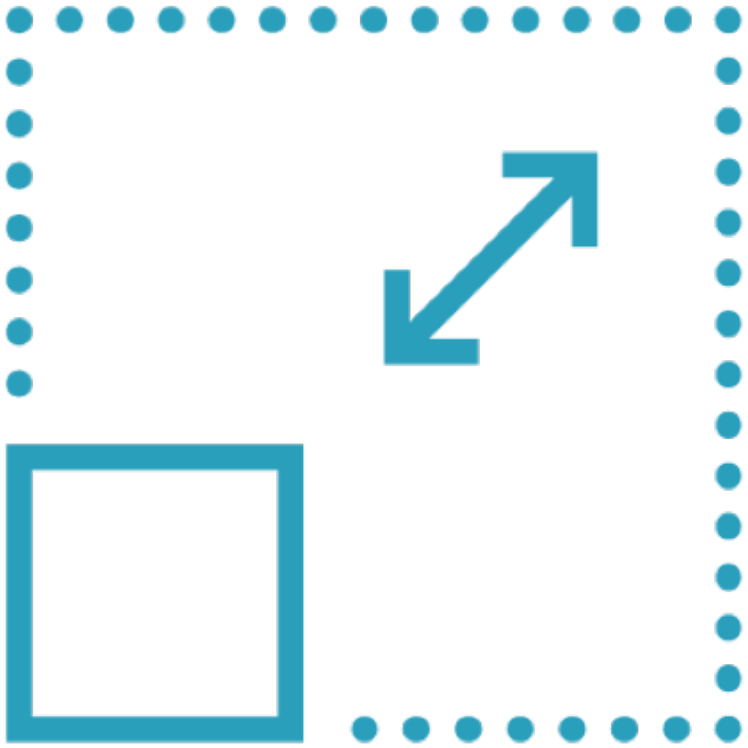
- Notes
- Essays
- Letters

What Is Text Recognition & Extraction?

Extract text from images

Save time

What Is Thumbnail Generation?



Identifies region of interest (ROI)

Crops image

Changes aspect ratio

Example of Thumbnail Generation

Summary

Microsoft Cognitive Services

- Subscribe to Azure
- Access the Computer Vision API

Three Fundamentals

- Image Identification & Classification
- Text Recognition & Extraction
- Thumbnail Generation