

Read text in images

1. Introduction

<https://learn.microsoft.com/en-us/training/modules/read-text-images-documents-with-computer-vision-service/1-introduction>

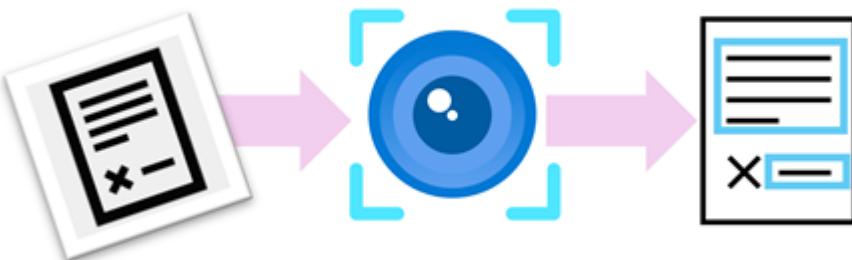
Introduction

Completed

- 2 minutes

We live in a digital world, in which data is increasingly captured as images. Often, those images contain text, which you need to be able to extract from their pixelated format in the image for processing, indexing, and other tasks. Everyday examples include:

- Meeting a new business associate and taking a photograph of their business card to store their contact details digitally.
- Scanning a document or ID card to include in an application for a government or commercial service.
- Taking a photo of a menu or recipe to store it in a digital notebook.
- Photographing street signs or store fronts so you can submit the text they contain to a translation app.
- Digitizing handwritten notes using a cellphone camera.



In this module, we'll explore the *optical character recognition* (OCR) capabilities of the Azure Vision *Image Analysis* API, which makes these scenarios, and more, possible.

2. Explore Azure AI options for reading text

<https://learn.microsoft.com/en-us/training/modules/read-text-images-documents-with-computer-vision-service/2-options-read-text>

Explore Azure AI options for reading text

Completed

- 3 minutes

There are multiple Azure services that read text from documents and images, each optimized for results depending on the input and the specific requirements of your application.

Azure Vision



Azure Vision includes an *image analysis* capability that supports *optical character recognition* (OCR). Consider using Azure Vision in the following scenarios:

- **Text location and extraction from scanned documents:** Azure Vision is a great solution for general, unstructured documents that have been scanned as images. For example, reading text in labels, menus, or business cards.
- **Finding and reading text in photographs:** Examples include photo's that include street signs and store names.
- **Digital asset management (DAM):** Azure Vision includes functionality for analyzing images beyond extracting text; including object detection, describing or categorizing an image, generating smart-cropped thumbnails and more. These capabilities make it a useful service when you need to catalog, index, or analyze large volumes of digital image-based content.

Azure Document Intelligence



Azure Document Intelligence is a service that you can use to extract information from complex digital documents. Azure Document Intelligence is designed for extracting text, key-value pairs, tables, and structures from documents automatically and accurately. Key considerations for choosing Azure Document Intelligence include:

- **Form processing:** Azure Document Intelligence is specifically designed to extract data from forms, invoices, receipts, and other structured documents.
- **Prebuilt models:** Azure Document Intelligence provides prebuilt models for common document types to reduce complexity and integrate into workflows or applications.
- **Custom models:** Creating custom models tailored to your specific documents, makes Azure Document Intelligence a flexible solution that can be used in many business scenarios.

Azure Content Understanding



Azure Content Understanding is a service that you can use to analyze and extract information from multiple kinds of content; including documents, images, audio streams, and video. It is suitable for:

- **Multimodal content extraction:** Extracting content and structured fields from documents, forms, audio, video, and images.
- **Custom content analysis scenarios:** Support for customizable analyzers enables you to extract specific content or fields tailored to business needs.

Note

In the rest of this module, we'll focus on the OCR image analysis feature in **Azure Vision**. To learn more about Azure Document Intelligence and Azure AI Content understanding, consider completing the following training modules:

- [Plan an Azure Document Intelligence solution](#)
- [Analyze content with Azure Content Understanding](#)

3. Read text with Azure Vision Image Analysis

<https://learn.microsoft.com/en-us/training/modules/read-text-images-documents-with-computer-vision-service/4-use-read-api>

Read text with Azure Vision Image Analysis

Completed

- 6 minutes

To use Azure Vision for image analysis, including optical character recognition, you must provision an Azure Vision resource in an Azure subscription. The resource can be:

- A **Foundry Tools** resource (either deployed as part of a Microsoft Foundry hub and project, or as a standalone resource).
- An **Azure Vision** resource.

To use your deployed resource in an application, you must connect to its *endpoint* using either key-based authentication or Microsoft Entra ID authentication. You can find the endpoint for your resource in the Azure portal, or if you're working in a Microsoft Foundry project, in the Microsoft Foundry portal. The endpoint is in the form of a URL, and typically looks something like this:

`https://<resource_name>.cognitiveservices.azure.com/`

After establishing a connection, you can use the OCR feature by calling the **ImageAnalysis** function (via the REST API or with an equivalent SDK method), passing the image URL or binary data, and optionally specifying the language the text is written in (with a default value of **en** for English).

`https://<endpoint>/computervision/imageanalysis:analyze?features=read&...`

To use the Azure Vision Python SDK to extract text from an image, install the **azure-ai-vision-imageanalysis** package. Then, in your code, use either key-based authentication or Microsoft Entra ID authentication to connect an **ImageAnalysisClient** object to an Azure Vision resource. To find and read text in an image, call the **analyze** (or **analyze_from_url**) method, specifying the **VisualFeatures.READ** enumeration.

```

from azure.ai.vision.imageanalysis import ImageAnalysisClient
from azure.ai.vision.imageanalysis.models import VisualFeatures
from azure.core.credentials import AzureKeyCredential

client = ImageAnalysisClient(
    endpoint=<YOUR_RESOURCE_ENDPOINT>,
    credential=AzureKeyCredential("<YOUR_AUTHORIZATION_KEY>")
)

result = client.analyze(
    image_data=<IMAGE_DATA_BYTES>, # Binary data from your image file
    visual_features=[VisualFeatures.READ],
    language="en",
)

```

To use the Azure Vision .NET SDK to extract text from an image, install the **Azure.AI.Vision.ImageAnalysis** package. Then, in your code, use either key-based authentication or Microsoft Entra ID authentication to connect an **ImageAnalysisClient** object to an Azure Vision resource. To find and read text in an image, call the **Analyze** method, specifying the **VisualFeatures.Read** enumeration.

```

using Azure.AI.Vision.ImageAnalysis;

ImageAnalysisClient client = new ImageAnalysisClient(
    "<YOUR_RESOURCE_ENDPOINT>",
    new AzureKeyCredential("<YOUR_AUTHORIZATION_KEY>"));

ImageAnalysisResult result = client.Analyze(
    <IMAGE_DATA_BYTES>, // Binary data from your image file
    VisualFeatures.Read,
    new ImageAnalysisOptions { Language = t"en" });

```

The results of the Read OCR function are returned synchronously, either as JSON or the language-specific object of a similar structure. These results are broken down in *blocks* (with the current service only using one block), then *lines*, and then *words*. Additionally, the text values are included at both the *line* and *word* levels, making it easier to read entire lines of text if you don't need to extract text at the individual *word* level.

```
{
  "metadata": {
    {
      "width": 500,
      "height": 430
    },
    "readResult": {
      "blocks": [
        {
          "text": "The quick brown fox jumps over the lazy dog."
        }
      ]
    }
  }
}
```

```

[
  {
    "lines": [
      {
        "text": "Hello World!",
        "boundingPolygon": [
          {"x":251,"y":265},
          {"x":673,"y":260},
          {"x":674,"y":308},
          {"x":252,"y":318}
        ],
        "words": [
          {
            "text": "Hello",
            "boundingPolygon": [
              {"x":252,"y":267},
              {"x":307,"y":265},
              {"x":307,"y":318},
              {"x":253,"y":318}
            ],
            "confidence": 0.996
          },
          {
            "text": "World!",
            "boundingPolygon": [
              {"x":318,"y":264},
              {"x":386,"y":263},
              {"x":387,"y":316},
              {"x":319,"y":318}
            ],
            "confidence": 0.99
          }
        ]
      },
      ]
    }
  }
]
}

```

4. Exercise - Read text in images

Exercise - Read text in images

Completed

- 30 minutes

Now it's your turn to try using the OCR capabilities of Azure Vision.

In this exercise, you use the Azure Vision Image Analysis SDK to develop a client application that extracts text from images.

Note

To complete this lab, you need an [Azure subscription](#) in which you have administrative access.

Launch the exercise and follow the instructions.

[Launch Exercise](#)

Tip

After completing the exercise, if you've finished exploring Foundry Tools, delete the Azure resources that you created during the exercise.

5. Module assessment

Module assessment

Completed

- 3 minutes

6. Summary

<https://learn.microsoft.com/en-us/training/modules/read-text-images-documents-with-computer-vision-service/7-summary>

Summary

Completed

- 1 minute

In this module, you learned how to provision an Azure Vision resource and use it from a client application to extract text from images.

To learn more about using Azure Vision for OCR, see the [OCR - Optical Character Recognition](#) in the Azure Vision documentation.