

Get started with Vision

Recent resources I've worked on



You don't have any recently used resources yet. Start with one of the scenarios below as a starting point, then use one of your resources to create a custom solution tailored to your own needs. The list of recent resources you've worked on will then appear here.

[View all resources](#)

Coming soon: Video analytics features from Azure AI Vision, using models like GPT-4 Turbo and Phi-3-V. Sign up for more information at aka.ms/video-analytics

Featured

Optical character recognition

Spatial analysis

Face

Image analysis



Video Retrieval and Summary Preview

Generate a brief summary of the main points shown in video. Locate specific keywords and jump to the relevant section.

[Try it out](#)



Recognize products on shelves Preview

Identify products on shelves, gaps in product availability, and compliance for planograms.

[Try it out](#)



Customize models with images Preview

Create custom image classification and object detection models with images using Vision Studio and Azure ML.

[Start a project](#)

Directories

✓ LODSPRODMCA.onmicrosoft.com

[Change Theme](#)

[Sign Out](#)

Analyze images in Vision Studio

53 Minutes Remaining

Instructions Resources Help 100%

that I have read and understood all the terms below: **Selected.**

- ✓ 5. Select **Review + create** then **Create** and wait for deployment to complete.

Connect your Azure AI service resource to Vision Studio

Next, connect the Azure AI service resource you provisioned above to Vision Studio.

- ✓ 1. In another browser tab, navigate to **Vision Studio** <https://portal.vision.cognitive.azure.com?azure-portal=true>.
- ✓ 2. Sign in with your account and making sure you are using the same directory as the one where you have created your Azure AI services resource.
- ☐ 3. On the Vision Studio home page, select **View all resources** under the **Getting started with Vision** heading.

Get started with Vision

Recent resources I've worked on

You don't have any recently used resources yet. Start with one of the scenarios below as a starting point, then use one of your resources to create a custom solution tailored to your own needs. The list of recent resources you've worked on will then appear here.

[View all resources](#)

- ☐ 4. On the **Select a resource to work with** page, hover your mouse cursor over the resource you created above in the list and then check the box to the left of the resource name, then select **Select as default resource**.

Note : If your resource is not listed, you may need to **Refresh** the page.

Select a resource to work with

Learn more about working with Azure resources

Current resource: [Azure AI Vision \(US\)](#)

Current subscription: [Azure AI Vision](#)

All resources

[Refresh](#)

21% Tasks Complete



Search



ENG
US



7:58 PM
1/19/2025

End >

Select a resource to work with

[Learn more about creating an Azure resource](#)

Current resource: nehalmrinstance-02 (East US, S0)


Current subscription: MOC Subscription-lod50075926

All resources

Search

[+ Create a new resource](#)

 Refresh

Resource name ↑	Azure subscription	Region	Pricing tier	Type
 nehalmrinstance-02 	MOC Subscription-Iod50075926	East US	S0	CognitiveServ

☐ Select as default resource

Analyze images in Vision Studio

48 Minutes Remaining

Instructions Resources Help 100%

created your Azure AI services resource.

3. On the Vision Studio home page, select **View all resources** under the **Getting started with Vision** heading.

Get started with Vision

Recent resources I've worked on

 You don't have any recently used resources yet. Start with one of the scenarios below as a starting point, then use one of your resources to create a custom solution tailored to your own needs. The list of recent resources you've worked on will then appear here.

Table 1

4. On the **Select a resource to work with** page, hover your mouse cursor over the resource you created above in the list and then check the box to the left of the resource name, then select **Select as default resource**.

Note : If your resource is not listed, you may need to **Refresh** the page.

Select a resource to work with

my-new-iam-role (X-ray-GS-RG)

Current resource: my-new-iam-role (X-ray-GS-RG)

Current subscription: Dedicated (Production)

API resources

Search

Resources name Basic subscription Region Pricing tier Type

my-new-iam-role	Basic	us-east-1	Basic	IAM role
-----------------	-------	-----------	-------	----------

5. Close the settings page by selecting the "x" at the top right of the screen.

Generate captions for an image

Now you are ready to use Vision Studio to analyze images taken by a camera in the *Northwind Traders* store.

Let's look at the image captioning functionality of Azure AI Vision. Image captions are available through the **Caption** and **Dense Captions** features.

- 1. In a web browser, navigate to **Vision Studio** T
<https://portal.vision.cognitive.azure.com?azure-portal=true>.

31% Tasks Complete

Microsoft.CognitiveServicesAlln X Vision Studio X Vision Studio

https://portal.vision.cognitive.azure.com/gallery/imageanalysis

Azure AI | Vision Studio

Vision Studio

Featured Optical character recognition Spatial analysis Face Image analysis

Recognize products on shelves
Preview

Identify products on shelves, gaps in product availability, and compliance for planograms.

Try it out

Customize models with images
Preview

Create custom image classification and object detection models with images using Vision Studio and Azure ML.

Start a project

Search photos with image retrieval

Retrieve specific moments within your photo album. For example, you can search for: a wedding you attended last summer, your pet, or your favorite city.

Try it out

Add dense captions to images

Generate human-readable captions for all important objects detected in your image.

Try it out

Remove backgrounds from images
Preview

Easily remove the background and preserve foreground elements in your image.

Try it out

Add captions to images

Generate a human-readable sentence that describes the content of an image.

Try it out

https://portal.vision.cognitive.azure.com/demo/image-captioning

Analyze images in Vision Studio

46 Minutes Remaining

Instructions Resources Help 100%

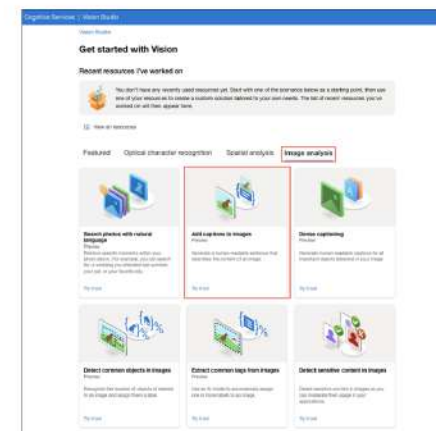
- 5. Close the settings page by selecting the "x" at the top right of the screen.

Generate captions for an image

Now you are ready to use Vision Studio to analyze images taken by a camera in the *Northwind Traders* store.

Let's look at the image captioning functionality of Azure AI Vision. Image captions are available through the **Caption** and **Dense Captions** features.

- 1. In a web browser, navigate to **Vision Studio** <https://portal.vision.cognitive.azure.com?azure-portal=true>.
- 2. On the **Getting started with Vision** landing page, select the **Image analysis** tab and then select the **Add captions to images** tile.



- 3. Under the **Try It Out** subheading, acknowledge the resource usage policy by reading and checking the box.
- 4. Select

37% Tasks Complete

End >

Microsoft.CognitiveServicesAlln

Vision Studio

Vision Studio

https://portal.vision.cognitive.azure.com/demo/image-captioning

A

☆

📖

☆

🔍

👤

...

🔗

Azure AI | Vision Studio

Vision Studio

>

Add captions to images

+

Add captions to images

×

Generate a human-readable sentence that describes the content of an image.

View documentation

View SDK reference

Use the REST API

View samples on Github

Try it out

☒

I acknowledge that this demo will incur usage to resource nehalmrinstance-02 in my Azure account.

Choose a different resource.

Use one of your own files or choose from a sample below.

📁

Drag and drop a file here or

Browse for a file or

Take a photo

Detected attributes

JSON

Windows Taskbar

Search

🌸

📁

🗨

📅

🔍

ENG US

8:05 PM 1/19/2025

Analyze images in Vision Studio

46 Minutes Remaining

Instructions

Resources

Help

100%

2. On the **Getting started with Vision** landing page, select the **Image analysis** tab and then select the **Add captions to images** tile.

3. Under the **Try It Out** subheading, acknowledge the resource usage policy by reading and checking the box.

4. Select

https://aka.ms/mslearn-images-for-analysis

to download **image-analysis.zip**. Open the folder on your computer and locate the file named **store-camera-1.jpg**; which contains the following image:

40% Tasks Complete

End >



Detected attributes JSON

A woman and a girl in a grocery store

Analyze images in Vision Studio

43 Minutes Remaining

Instructions Resources Help 100%

4. Select

<https://aka.ms/mslearn-images-for-analysis> to download **image-analysis.zip**. Open the folder on your computer and locate the file named **store-camera-1.jpg**; which contains the following image:



5. Upload the **store-camera-1.jpg** image by dragging it to the **Drag and drop files here** box, or by browsing to it on your file system.

6. Observe the generated caption text, visible in the **Detected attributes** panel to the right of the image.

The **Caption** functionality provides a single, human-readable English sentence describing the image's content.

7. Next, use the same image to perform **Dense captioning**. Return to the **Vision Studio** home page, and as you did before, select the **Image analysis** tab, then select the **Dense captioning** tile.

The **Dense Captions** feature differs from the **Caption** capability in that it provides multiple human-readable captions for an image, one describing the image's content and others, each covering the essential objects detected in the picture. Each detected object includes a bounding box, which defines the pixel coordinates within the image associated with the object.

50% Tasks Complete



Search



ENG US



8:07 PM 1/19/2025

End >



Recognize products on shelves

Preview

Identify products on shelves, gaps in product availability, and compliance for planograms.

[Try it out](#)



Customize models with images

Preview

Create custom image classification and object detection models with images using Vision Studio and Azure ML.

[Start a project](#)



Search photos with image retrieval

Retrieve specific moments within your photo album. For example, you can search for: a wedding you attended last summer, your pet, or your favorite city.

[Try it out](#)



Add dense captions to images

Generate human-readable captions for all important objects detected in your image.

[Try it out](#)



Remove backgrounds from images

Preview

Easily remove the background and preserve foreground elements in your image.

[Try it out](#)



Add captions to images

Generate a human-readable sentence that describes the content of an image.

[Try it out](#)

4. Select

<https://aka.ms/mslearn-images-for-analysis>

to download **image-analysis.zip**. Open the folder on your computer and locate the file named **store-camera-1.jpg**; which contains the following image:



5. Upload the **store-camera-1.jpg** image by dragging it to the **Drag and drop files here** box, or by browsing to it on your file system.

6. Observe the generated caption text, visible in the **Detected attributes** panel to the right of the image.

The **Caption** functionality provides a single, human-readable English sentence describing the image's content.

7. Next, use the same image to perform **Dense captioning**. Return to the **Vision Studio** home page, and as you did before, select the **Image analysis** tab, then select the **Dense captioning** tile.

The **Dense Captions** feature differs from the **Caption** capability in that it provides multiple human-readable captions for an image, one describing the image's content and others, each covering the essential objects detected in the picture. Each detected object includes a bounding box, which defines the pixel coordinates within the image associated with the object.

Vision Studio > Add dense captions to images

☒ I acknowledge that this demo will incur usage to resource nenaimrinstance-uz in my Azure account. [Choose a different resource.](#)

Use one of your own files or choose from a sample below.



Detected attributes JSON

- A woman and a girl in a grocery store
- A woman smiling while holding a phone
- A person holding a phone case
- A woman wearing a purple hat
- A close-up of a person wearing a purple hat
- A person holding a piece of ginger
- A close-up of a person wearing a purple hat
- A woman smiling at a cell phone
- A row of jars on a shelf
- A blurry image of a person standing in a room

Analyze images in Vision Studio

41 Minutes Remaining

Instructions Resources Help 100%

6. Observe the generated caption text, visible in the **Detected attributes** panel to the right of the image.

The **Caption** functionality provides a single, human-readable English sentence describing the image's content.

7. Next, use the same image to perform **Dense captioning**. Return to the **Vision Studio** home page, and as you did before, select the **Image analysis** tab, then select the **Dense captioning** tile.

The **Dense Captions** feature differs from the **Caption** capability in that it provides multiple human-readable captions for an image, one describing the image's content and others, each covering the essential objects detected in the picture. Each detected object includes a bounding box, which defines the pixel coordinates within the image associated with the object.

8. Hover over one of the captions in the **Detected attributes** list and observe what happens within the image.



Move your mouse cursor over the other captions in the list, and notice how the bounding box shifts in the image to highlight the portion of the image used to generate the caption.

Tagging images

The next feature you will try is the **Extract Tags** functionality. Extract tags is based on thousands of

56% Tasks Complete

End >

Microsoft.CognitiveServicesAlln X Vision Studio X Vision Studio

https://portal.vision.cognitive.azure.com/gallery/imageanalysis

Azure AI | Vision Studio

Vision Studio

Add dense captions to images

Generate human-readable captions for all important objects detected in your image.

[Try it out](#)

Remove backgrounds from images

Preview


Easily remove the background and preserve foreground elements in your image.

[Try it out](#)

Add captions to images

Generate a human-readable sentence that describes the content of an image.


[Try it out](#)



Detect common objects in images

Recognize the location of objects of interest in an image and assign them a label.


[Try it out](#)



Extract common tags from images

Use an AI model to automatically assign one or more labels to an image.

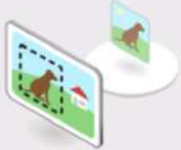
[Try it out](#)



Detect sensitive content in images

Detect sensitive content in images so you can moderate their usage in your applications.

[Try it out](#)



Create smart-cropped images

https://portal.vision.cognitive.azure.com/demo/generic-image-tagging on

Search

ENG US 8:11 PM 1/19/2025

Analyze images in Vision Studio


40 Minutes Remaining

Instructions Resources Help 100%

Move your mouse cursor over the other captions in the list, and notice how the bounding box shifts in the image to highlight the portion of the image used to generate the caption.

Tagging images

The next feature you will try is the **Extract Tags** functionality. Extract tags is based on thousands of recognizable objects, including living beings, scenery, and actions.

- Return to the home page of Vision Studio, then select the **Extract common tags from images** tile under the **Image analysis** tab.
- In the **Choose the model you want to try out**, leave **Prebuilt product vs. gap model** selected. In the **Choose your language**, select **English** or a language of your preference.
- Open the folder containing the images you downloaded and locate the file named **store-image-2.jpg**, which looks like this:

- Upload the **store-camera-2.jpg** file.
- Review the list of tags extracted from the image and the confidence score for each in the detected

59% Tasks Complete

End >

[Vision Studio](#) > [Extract common tags from images](#)

[illegible]

Prebuilt product vs. gap model

- Train your own custom model

Choose your language

English

Drag and drop a file here
or
Browse for a file
or
Take a photo



Detected attributes JSON



- clothing (99.04%)
- person (98.26%)
- convenience store (98.00%)
- retail (97.88%)
- supermarket (95.77%)
- grocery store (94.18%)
- shopping (91.96%)
- customer (91.77%)
- text (90.62%)
- trade (90.15%)
- human face (90.05%)
- indoor (89.03%)
- woman (88.37%)
- selling (88.15%)

39 Minutes Remaining

Instructions Resources Help 100%

a language of your preference.

3. Open the folder containing the images you downloaded and locate the file named **store-image-2.jpg**, which looks like this:



- ☒ 4. Upload the **store-camera-2.jpg** file.
- ☐ 5. Review the list of tags extracted from the image and the confidence score for each in the detected attributes panel. Here the confidence score is the likelihood that the text for the detected attribute describes what is actually in the image. Notice in the list of tags that it includes not only objects, but actions, such as *shopping*, *selling*, and *standing*.

Detected attributes JSON



- clothing (99.04%)
- person (98.26%)
- convenience store (98.00%)
- retail (97.88%)
- supermarket (95.77%)
- grocery store (94.18%)
- shopping (91.96%)
- customer (91.77%)
- text (90.62%)
- trade (90.13%)
- human face (90.09%)
- indoor (89.83%)
- woman (88.37%)
- selling (88.15%)
- shop (88.11%)

Object detection

In this task, you use the **Object detection** feature of

68% Tasks Complete

Vision Studio > Extract common tags from images

Drag and drop a file here
or
Browse for a file
or
Take a photo



Detected attributes JSON



- clothing (99.04%)
- person (98.26%)
- convenience store (98.00%)
- retail (97.88%)
- supermarket (95.77%)
- grocery store (94.18%)
- shopping (91.96%)
- customer (91.77%)
- text (90.62%)
- trade (90.15%)
- human face (90.05%)
- indoor (89.03%)
- woman (88.37%)
- selling (88.15%)
- shop (88.11%)
- market (86.82%)
- store (81.26%)
- shelf (69.05%)
- standing (65.64%)
- marketplace (51.18%)

Analyze images in Vision Studio

38 Minutes Remaining

Instructions Resources Help 100%

a language of your preference.

3. Open the folder containing the images you downloaded and locate the file named **store-image-2.jpg**, which looks like this:



4. Upload the **store-camera-2.jpg** file.
5. Review the list of tags extracted from the image and the confidence score for each in the detected attributes panel. Here the confidence score is the likelihood that the text for the detected attribute describes what is actually in the image. Notice in the list of tags that it includes not only objects, but actions, such as *shopping*, *selling*, and *standing*.

Detected attributes JSON



- clothing (99.04%)
- person (98.26%)
- convenience store (98.00%)
- retail (97.88%)
- supermarket (95.77%)
- grocery store (94.18%)
- shopping (91.96%)
- customer (91.77%)
- test (90.62%)
- trade (90.15%)
- human face (90.09%)
- indoor (89.83%)
- woman (88.37%)
- selling (88.15%)
- shop (88.11%)

Object detection

In this task, you use the **Object detection** feature of

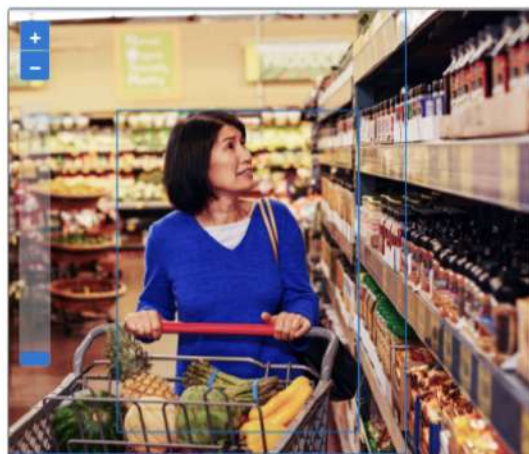
71% Tasks Complete

Vision Studio > Detect common objects in images

Choose the model you want to try out

Prebuilt product vs. gap model Train your own custom model

Drag and drop a file here
or
Browse for a file
or
Take a photo



Detected attributes JSON

Threshold value

15

person (84.80%)

supermarket (51.70%)

Analyze images in Vision Studio

37 Minutes Remaining

Instructions Resources Help 100%

Object detection

In this task, you use the **Object detection** feature of Image Analysis. Object detection detects and extracts bounding boxes based on thousands of recognizable objects and living beings.

- Return to the home page of Vision Studio, then select the **Detect common objects in images** tile under the **Image analysis** tab.
- In the **Choose the model you want to try out**, leave **Prebuilt product vs. gap model** selected.
- Open the folder containing the images you downloaded and locate the file named **store-camera-3.jpg**, which looks like this:



- Upload the **store-camera-3.jpg** file.
- In the **Detected attributes** box, observe the list of detected objects and their confidence scores.
- Hover your mouse cursor over the objects in the **Detected attributes** list to highlight the object's bounding box in the image.
- Move the **Threshold value** slider until a value of 70 is displayed to the right of the slider. Observe what happens to the objects in the list. The

87% Tasks Complete

End >

Vision Studio > Detect common objects in images

Choose the model you want to try out

Prebuilt product vs. gap model Train your own custom model

Drag and drop a file here
or
Browse for a file
or
Take a photo



Detected attributes JSON

Threshold value

70

person (84.80%)

Analyze images in Vision Studio

37 Minutes Remaining

Instructions Resources Help 100%

In this task, you use the **Object detection** feature of Image Analysis. Object detection detects and extracts bounding boxes based on thousands of recognizable objects and living beings.

1. Return to the home page of Vision Studio, then select the **Detect common objects in images** tile under the **Image analysis** tab.
2. In the **Choose the model you want to try out**, leave **Prebuilt product vs. gap model** selected.
3. Open the folder containing the images you downloaded and locate the file named **store-camera-3.jpg**, which looks like this:



4. Upload the **store-camera-3.jpg** file.
5. In the **Detected attributes** box, observe the list of detected objects and their confidence scores.
6. Hover your mouse cursor over the objects in the **Detected attributes** list to highlight the object's bounding box in the image.
7. Move the **Threshold value** slider until a value of 70 is displayed to the right of the slider. Observe what happens to the objects in the list. The threshold slider specifies that only objects identified with a confidence score or probability greater than the threshold should be displayed.

90% Tasks Complete

End >

Analyze images in Vision Studio


36 Minutes Remaining

Instructions

Resources

Help

100%



- ✓

 4. Upload the **store-camera-3.jpg** file.
- ✓

 5. In the **Detected attributes** box, observe the list of detected objects and their confidence scores.
- ✓

 6. Hover your mouse cursor over the objects in the **Detected attributes** list to highlight the object's bounding box in the image.
- ✓

 7. Move the **Threshold value** slider until a value of 70 is displayed to the right of the slider. Observe what happens to the objects in the list. The threshold slider specifies that only objects identified with a confidence score or probability greater than the threshold should be displayed.

Clean up

If you don't intend to do more exercises, delete any resources that you no longer need. This avoids accruing any unnecessary costs.

- ✓

 1. Open the **Azure portal**] <https://portal.azure.com> and select the resource group that contains the resource you created.
- ✓

 2. Select the resource and select **Delete** and then **Yes** to confirm. The resource is then deleted.

Learn more

To learn more about what you can do with this service, see the [Azure AI Vision page](#).

100% Tasks Complete

End >