

# HOS01A-4 Introduction to Rekognition

AI620 Fall 2024 by Verónica Elze


8. Create Project

image\_classified\_project\_velze

Info

▼ How it works

Creating your dataset




1. Create dataset

A dataset is a collection of images, and image labels, that you use to train or test a model.

Create dataset


2. Label images



Labels identify objects, scenes, or concepts on an entire image, or they identify object locations on an image.

Add labels

Training your model




3. Train model

Depending on the training dataset, the training model finds image-level scenes and concepts, or it finds object locations.

Train model

Evaluating your model



4. Check performance metrics

Performance metrics tell you if your model needs additional training before you can use it.

Check metrics

Project details

<div>Project name</div> <div>image_classified_project_velze</div>	<div>Created</div> <div>October 11, 2024 at 20:48:54 (UTC-06:00)</div>	<div>Dataset</div> <div>0 training labels, 0 training images, 0 test labels, 0 test images</div>	<div>Models</div> <div>0</div>
---	--	--	--------------------------------

## 9. Create Dataset

[Custom Labels](#) > [image\\_classified\\_project\\_velze](#) > **Create dataset**

# Create dataset [Info](#)

## Starting configuration

### Configuration options



#### Start with a single dataset

When you train your model, the dataset is split to create the training dataset (80%) and test dataset (20%) for your project.



#### Start with a training dataset and a test dataset

Recommended for most users. Start with the highest control over training, testing, and performance tuning.



### What are training datasets and test datasets?

- A training dataset teaches your model to identify scenes or objects in images.
- A test dataset evaluates the performance of your trained model.

## 10. Single Dataset from S3 bucket

### Create dataset [Info](#)

#### Starting configuration

##### Configuration options

☒ Start with a single dataset

When you train your model, the dataset is split to create the training dataset (80%) and test dataset (20%) for your project.

☐ Start with a training dataset and a test dataset

Recommended for most users. Start with the highest control over training, testing, and performance tuning.



#### What are training datasets and test datasets?

- A training dataset teaches your model to identify scenes or objects in images.
- A test dataset evaluates the performance of your trained model.

#### Training dataset details

##### Import training images [Info](#)

Import images from one of the sources below.

☒ Import images from S3 bucket

Use images from an existing S3 bucket by entering the S3 bucket URI. You can automatically add labels based on your S3 bucket folder names.



☐ Upload images from your computer

Add images by uploading files from your local computer. You're limited to uploading 30 images at one time.



☐ Copy an existing Amazon Rekognition Custom Labels dataset

Use an existing dataset as a starting point for your new dataset. Your original dataset will remain unchanged.



☐ Import images labeled by SageMaker Ground Truth

Provide the location of your manifest file. If you have a labeled datasets in a different format, convert them to a manifest format.



## 11. Auto-label S3 URI images

### S3 URI

s3://custom-labels-console-us-east-2-1323f944c6/flowers/flowers/

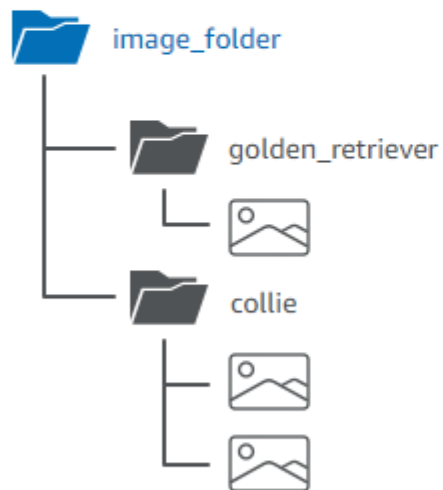
Supported image formats: JPG, PNG. Maximum images per dataset: 250,000. Maximum image size: 15 MB, Minimum size (px): 64 x 64. Maximum size (px): 4096 x 4096. Images must have the same dimensions.

For best results, we recommend uploading images from folders within the **S3 bucket** created for you during first-time setup.

### Automatic labeling

If you've organized the images in your S3 bucket by folder name (/Golden-Retriever/01.jpeg), Custom Labels can automatically label these images.

☒ Automatically assign image-level labels to images based on the folder name



Cancel

Create Dataset

## 12. Train Model

### Train model



#### Train model



To train your model, Amazon Rekognition Custom Labels uses your project's training dataset and test dataset. You can add tags to help track your models. You can also encrypt your images with your own AWS Key Management Service key.

#### Training details [Info](#)

##### Choose project

Amazon Rekognition Custom Labels trains a new version of the model within the project you choose.



#### Tags [Info](#)

A tag is a label that you can assign to your model. Each tag consists of a key and an optional value.

No tags associated with the resource.

Add new tag

You can add up to 50 more tags.

#### Image Data Encryption

Your data is encrypted by default with a key that AWS owns and manages for you. To choose a different key, customize your encryption settings. [Learn More](#)

☐ Customize encryption settings (advanced)

Cancel

Train Model

## Do you want to train your model?



Your dataset will be split into a training dataset (80%) and test dataset (20%).

Typically, training takes from 30 minutes to 24 hours to complete. For more information, see [Training hours](#).

You are charged for the amount of time it takes to successfully train your model and for the amount of time your model runs. You aren't charged if model training fails.

Cancel

Train model

## image\_classified\_project\_velze<sup>Info</sup>

### ▼ How it works

#### Creating your dataset



##### 1. Create dataset

A dataset is a collection of images, and image labels, that you use to train or test a model.

✓ Created



##### 2. Label images

Labels identify objects, scenes, or concepts on an entire image, or they identify object locations on an image.

Add labels

#### Training your model



##### 3. Train model

Depending on the training dataset, the training model finds image-level scenes and concepts, or it finds object locations.

Train model

#### Evaluating your model



##### 4. Check performance metrics

Performance metrics tell you if your model needs additional training before you can use it.

Check metrics

### Project details

Project name

image\_classified\_project\_velze

Created

October 11, 2024 at 20:48:54 (UTC-06:00)

Dataset

5 training labels, 400 training images, 5 test labels, 100 test images

Models

1

Models (1)

Delete model

Download validation results ▼

Q

Find resources

<

1

...

>

<input type="checkbox"/>	Name ▼	Date created ▼	Training dataset ▼	Test dataset ▼	Model performance (F1 score) ▼	Model status ▼	Status message ▼
<input type="checkbox"/>	image_classified_project_velze.2024-10-11T20.55.52	October 11, 2024			N/A	TRAINING_IN_PROGRESS	The model is being trained.

Models (1)

Delete model

Download validation results ▼

Q

Find resources

<

1

...

>

<input type="checkbox"/>	Name ▼	Date created ▼	Training dataset ▼	Test dataset ▼	Model performance (F1 score) ▼	Model status ▼	Status message ▼
<input type="checkbox"/>	<a href="#">image_classified_project_velze.2024-10-11T20.55.52</a>	October 11, 2024			0.980	TRAINING_COMPLETED	The model is ready to run.

13. Check Metrics

image\_classified\_project\_velze.2024-10-11T20.55.52

Info

Delete model

Evaluation

Model details

Use model

Tags

?

Evaluation

The Evaluation tab shows the testing results for your trained model. This helps you understand the overall performance of your model. To view the results for an image, choose the View test results button.

Evaluation results

View test results

F1 scoreInfo

0.980

Date completed

October 12, 2024

Trained in 0.950 hours

Average precisionInfo

0.972

Training dataset

5 labels, 400 images

Overall recallInfo

0.990

Testing dataset

5 labels, 100 images

Per label performance (5)

Q Find labels

< 1 >

Label name	▲	F1 score	▼	Test images	▼	Precision	▼	Recall	▼	Assumed threshold	▼
daisy		1.000		20		1.000		1.000		0.622	
dandelion		1.000		20		1.000		1.000		0.479	
rose		0.976		20		0.952		1.000		0.332	
sunflower		0.974		20		1.000		0.950		0.693	
tulip		0.952		20		0.909		1.000		0.164	