

Amazon Rekognition Custom Labels Lab

Objectives

In this workshop, we'll learn the fundamentals of [Amazon Rekognition Custom Labels](#) by building an AWS logo detector. This serves as an example of how to build models to enable brand monitoring and analysis through rich media channels like social media.

Learn how to:

1. Use the Amazon Rekognition Custom Labels console.
2. Prepare classification or object detection training and test datasets using Rekognition.
3. Train a custom object detection model with the help of Rekognition Custom Labels (AutoML) through few shot learning.
4. Deploy a Rekognition Custom Labels Demo app.
5. Manage the continuous improvement mechanics.

Setup your Lab Environment

1. Pre-requisites:
 - [Access](#) to AWS with permissions to the lab resources.
 - [AWS CLI](#) installed on your workstation.
2. Launch the Amazon Rekognition Custom Labels demo app from [this](#) GitHub repository. Use the launch template buttons as displayed below from your preferred region.

Region	Launch Template
US East (N. Virginia) (us-east-1)	Launch on AWS
US East (Ohio) (us-east-2)	Launch on AWS
US West (Oregon) (us-west-2)	Launch on AWS
EU (Ireland) (eu-west-1)	Launch on AWS

You don't have to wait for the CloudFormation template to complete. Later you will log into the web app that it creates. You can access login URL from the **Output** sub tab.

3. Create an S3 bucket. You can use an existing S3 bucket, but you're expected to have ownership and full access.
4. Copy the aws-logos dataset over to your S3 bucket using the AWS CLI:

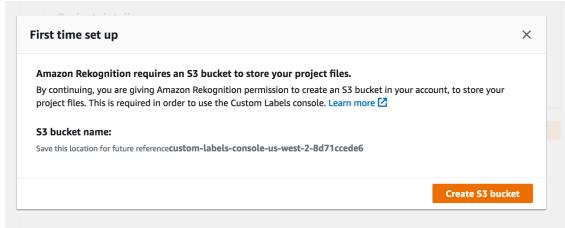
```
aws s3 sync s3://dtong-public-fileshare/machine-learning/datasets/aws-logos/ <>REPLACE WITH YOUR S3 URI>
eg) <>REPLACE WITH YOUR S3 URI>: s3://dtong-ml-datasets/aws-logos
```

Amazon Rekognition Custom Labels Console

5. Log into the AWS console and navigate to the Amazon Rekognition console. You can use the search bar.
6. Navigate to the Custom Labels console from the navigation panel on the left.

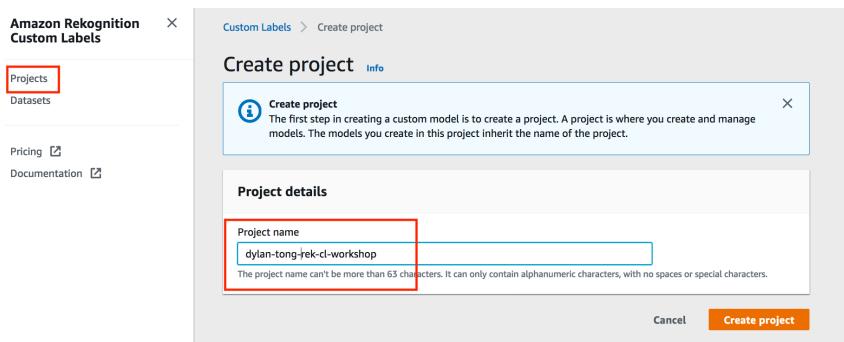
7. Ensure you're accessing Amazon Rekognition in the same region as where you've provisioned the CloudFormation template and S3 bucket (Steps 2-3).

8. If this is your first-time accessing Amazon Rekognition Custom Labels, you'll be prompted to provision an S3 bucket, which the service will be using to store meta-data. Click on the **Create S3 bucket** button to proceed.



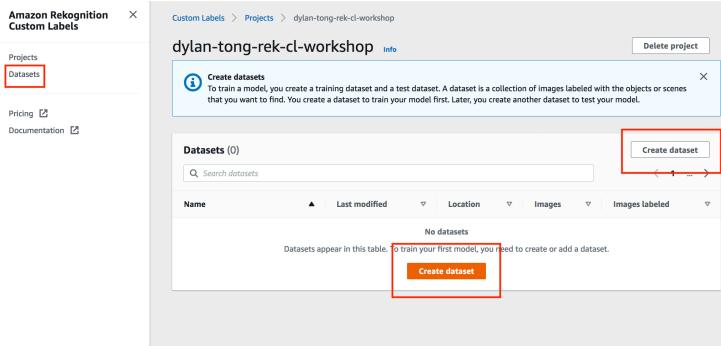
Create a Project

9. Select projects from the navigation pane and click on the **Create project** button.
10. Enter a unique name for your project.

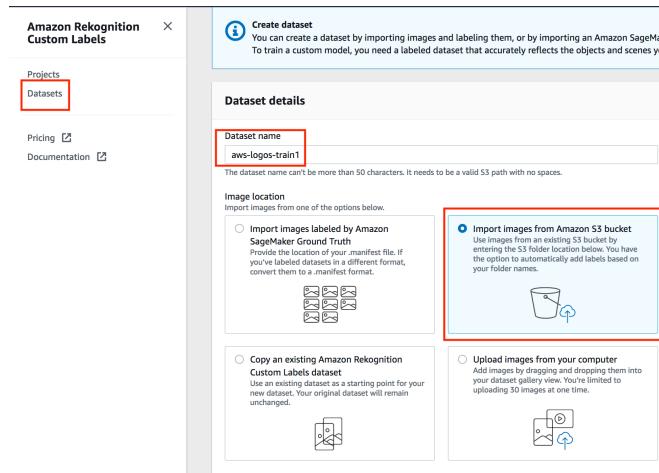


Prepare Datasets

11. Navigate to the Datasets page and click on **Create dataset**.



12. We'll start by creating our training dataset. Provide a name to identify this dataset and select **Import images from Amazon S3 bucket**.



13. Provide the S3 URI location of the training dataset. This location is the S3 URI that you provided as the destination of your sync command followed by the /train1/ prefix:
`s3://<Your S3 URI>/train1/.`

S3 folder location
S3 bucket location of your images.

Supported image formats: JPEG, PNG. Maximum images per dataset: 250,000. Maximum image size: 15 MB, Minimum size (px): 64 x 64. Maximum size (px): 4096 x 4096.
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), Amazon Rekognition Custom Labels can automatically label these images.
 Automatically attach a label to my images based on the folder they're stored in.

14. Amazon Rekognition requires access to your S3 bucket, so that we can do things such as storing image annotations and reading images for training and testing. Follow the provided instructions and copy the provided bucket policy as shown below:

Make sure that your S3 bucket is correctly configured

You've specified an external S3 bucket: dtong-ml-datasets.

If the images in this bucket, copy the policy below (to copy, choose the preceding link text). Paste the policy into the "Bucket Policy" section of dtong-ml-datasets.

⚠️ If you don't apply this policy, you won't be able to take a snapshot on this dataset.

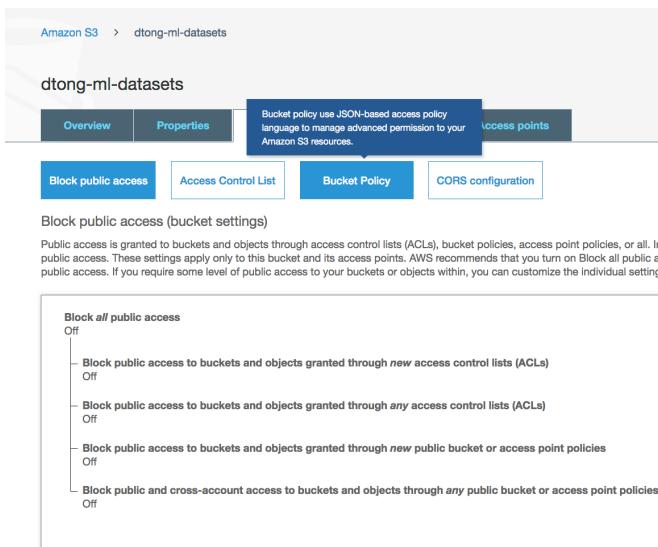
```

1  {
2    "Version": "2012-10-17",
3    "Statement": [
4      {
5        "Sid": "AMRekognitionS3McBucketRead0B93B11",
6        "Effect": "Allow",
7        "Principal": "*",
8        "Service": "rekognition.amazonaws.com"
9        "Action": [
10          "s3:GetBucketAcl",
11          "s3:GetBucketLocation",
12          "s3:GetBucketMetrics"
13        ],
14        "Resource": "arn:aws:s3:::dtong-ml-datasets"
15      },
16      {
17        "Sid": "AMRekognitionS3GetBucket20B93B11",
18        "Effect": "Allow",
19        "Principal": "*",
20        "Service": "rekognition.amazonaws.com"
21        "Action": [
22          "s3:GetBucketAcl",
23          "s3:GetBucketMetrics",
24          "s3:GetObject",
25          "s3:GetObjectVersion",
26          "s3:GetObjectTagging"
27        ],
28        "Resource": "arn:aws:s3:::dtong-ml-datasets/*"
29      },
30      {
31        "Sid": "AMRekognitionS3MCBucketWrite0B93B11",
32        "Effect": "Allow",
33        "Principal": "*",
34        "Service": "rekognition.amazonaws.com"
35        "Action": [
36          "s3:PutObject"
37        ],
38        "Resource": "arn:aws:s3:::dtong-ml-datasets"
39      },
40      {
41        "Sid": "AMRekognitionS3PutObject0B93B11",
42        "Effect": "Allow",
43        "Principal": "*",
44        "Service": "rekognition.amazonaws.com"
45        "Action": [
46          "s3:PutObject"
47        ],
48        "Resource": "arn:aws:s3:::dtong-ml-datasets/*",
49        "Condition": {
50          "StringLike": {
51            "s3:x-amz-acl": "bucket-owner-full-control"
52          }
53        }
54      }
    ]
  }

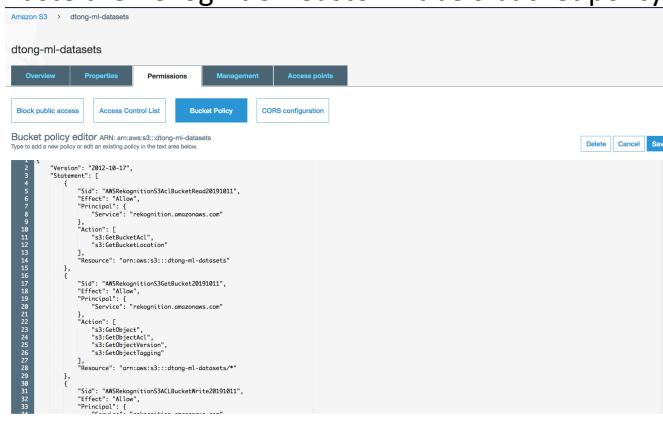
```

C VonAWS2 (1)

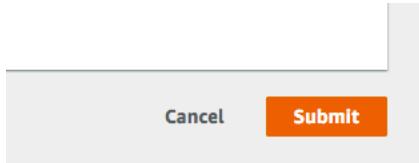
15. In a new browser tab, navigate to the S3 console and select your S3 bucket. Click on the **Bucket Policy** button under the **Permissions** tab.



16. Paste the Rekognition Custom Labels bucket policy and click on **Save**.



17. Return to the Rekognition Custom Labels console. Rekognition should now have access to your S3 bucket, so when you click on **Submit**, the dataset should be successfully created.

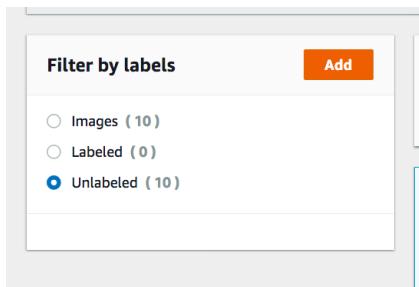


Annotate Training Data using Amazon Rekognition Custom Labels

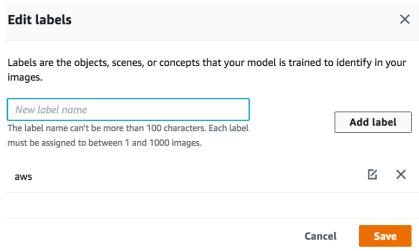
Amazon Rekognition requires annotations in a specific [format](#). The easiest way to label your dataset is to use the integrated labeling tool, or Amazon GroundTruth. [Amazon GroundTruth](#) is the recommend option for production projects where you need to scale by manage labeling workforces and large training sets.

In this lab, we're engaging in rapid prototyping and operating on a very small dataset. The integrated labeling tool is more practical for our purpose. Further down the developing process, you might choose to migrate over to Amazon GroundTruth as you transition into production.

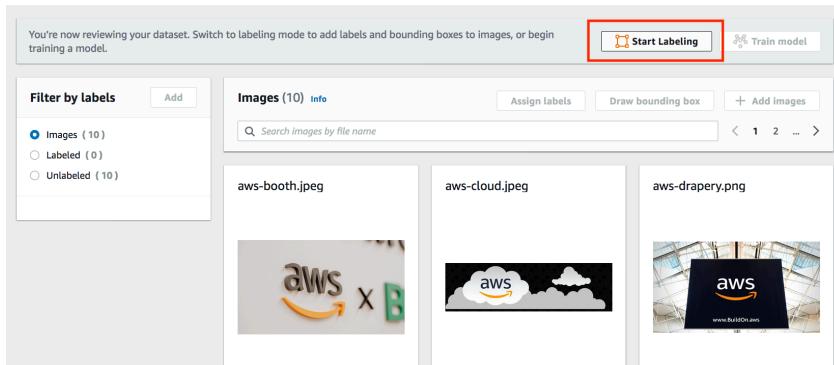
18. After you create your training dataset, you should be re-directed to the details pane of the dataset. Click on the **Add** button the left-hand side panel to create a label.



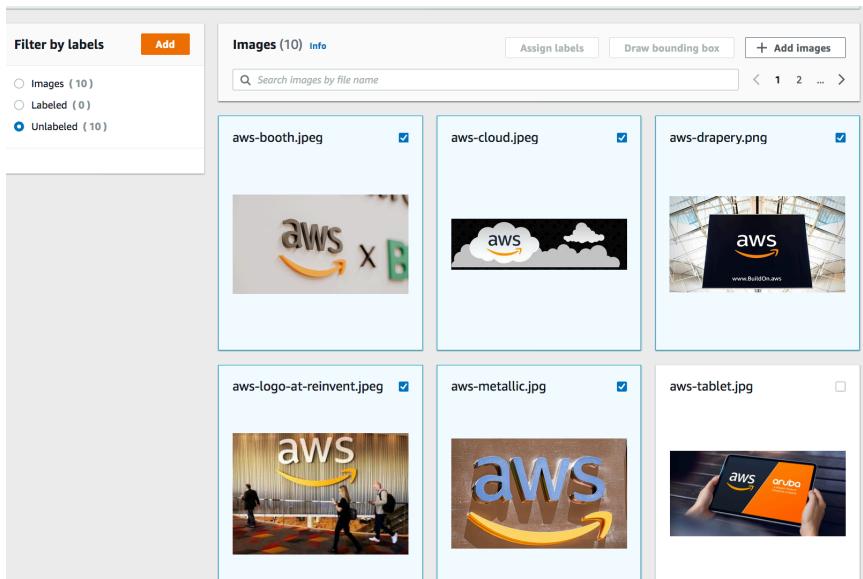
19. Type in “aws” and click on **Add label** to create a single label. Click on the **Save** button.



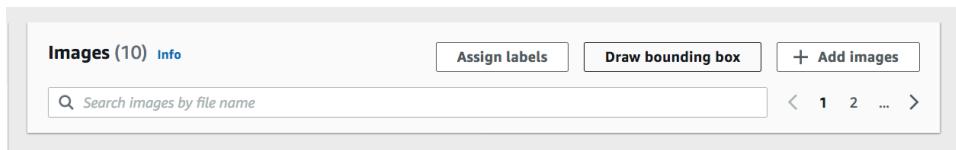
20. Click on the **Start Labeling** button to enter the embedded labeling tool for Rekognition Custom Labels.



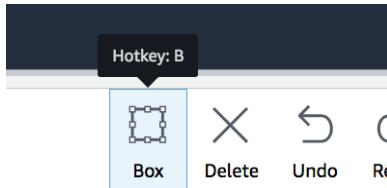
21. Select each of the images displayed on the page. Unfortunately, you can't select images on the next page, so select the first 9 images.



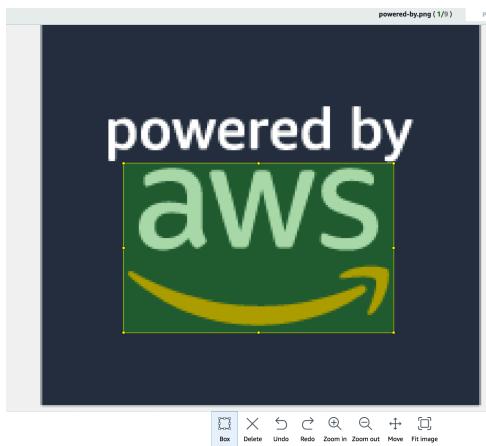
22. Click on the **Draw bounding box** button.



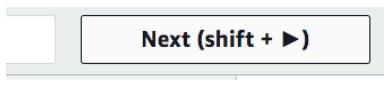
23. You're now in the annotation tool context. The bounding box tool at the bottom of the editor should be selected by default.



24. Click, drag and release to create a bounding box around the AWS logos.



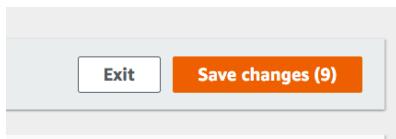
25. Click on the **Next** button after you've completed annotation.



26. After you've annotated the last selected image, you can press the **Done** button.



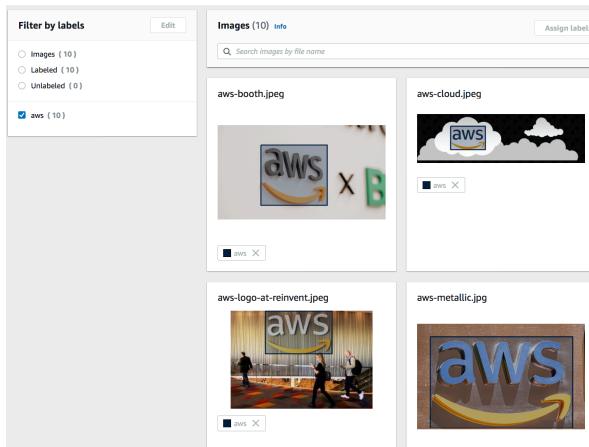
27. **IMPORTANT:** after pressing the done button, you'll return to the datasets pane. Click the **Save changes** button. Any other action could result in losing your work.



28. We have 10 images to label. You can select the **Unlabeled** radio button to display the remaining image. Select the image and repeat the process to label the last image.



29. You should now have 10 labeled images, which you can preview from the dataset details pane as shown below:

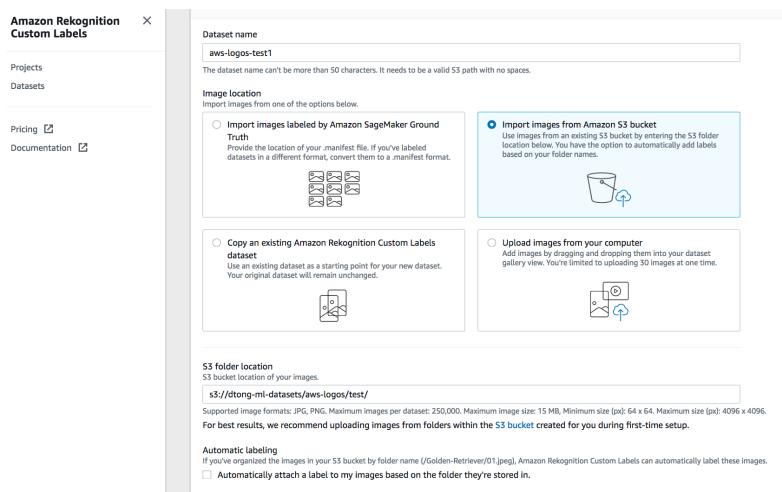


Import your Test Dataset

Rekognition Custom Labels can perform a split and generate a validation dataset for you automatically during training, or you can provide a separate test dataset. We're going to do the latter, so that we have a standard dataset that we can use as a test across multiple models.

30. Select **Datasets** from the navigation panel and click on the **Create dataset** button. As before, select **Import images from Amazon S3 bucket**.

The test data is under the /test/ prefix: `s3://<Your S3 URI>/test/`.



Note that you don't need to copy the bucket policy again. The default bucket policy that you copied over previously provides permissions to all the contents in the S3 bucket.

Train Your Model

31. Navigate to the **Projects** panel and select the project that you created. Your project should look similar to the image below. Click on the **Train new model** button.

The screenshot shows the 'Custom Labels' interface under the 'Projects' tab. A project named 'dylan-tong-rek-cl-workshop' is selected. The 'Models' section is currently empty. The 'Datasets' section lists two entries:

Name	Last modified	Location	Images	Images labeled
aws-logos-train1	August 31, 2020	S3 bucket	10	10
aws-logos-test	August 31, 2020	S3 bucket	25	0

A prominent orange 'Train new model' button is located at the bottom of the page.

32. Under **Choose training dataset**, select the name of the training dataset that you created.

Select **Choose an existing test dataset**. Select the name of the test dataset that you created.

Lastly, Click on the **Train** button.

The screenshot shows the 'Train model' configuration page. In the 'Training details' section, a project is selected: 'arn:awsrekognition:us-east-1:803235869972:project/dylan-tong-rek-cl-workshop/1598913093885'. In the 'Create test set' section, the 'Choose an existing test dataset' radio button is selected, with 'aws-logos-test' listed in the dropdown. There are also options for 'Create a new test dataset' (using a file icon) and 'Split training dataset' (using a grid icon).

33. From your Projects detailed pane. You should see your training job running:

The job will take 1-3 hours to complete.

Models (1)							
<input type="text"/> Find resources		Date created	Training dataset	Testing dataset	Model performance	Status	Status message
<input type="checkbox"/>	Name	August 31, 2020	aws-logos-train1	aws-logos-test	N/A	TRAINING_IN_PROGRESS	The model is being trained.
<input type="checkbox"/>	dylan-tong-rek-cl-workshop.2020-08-31T15.55.51						

Evaluate your Model

34. Once your model is trained, you can click on the model name from the console.



35. The models detailed pane provides a summary of evaluation results. You can click on **View test results** to analyze the predictions on the test dataset.

Evaluation results			View test results
F1 score Info	Average precision Info	Overall recall Info	
0.802	0.768	0.840	
Date completed	Training dataset	Testing dataset	
August 25, 2020	1 labels, 15 images	1 labels, 27 images	
Trained in 0.828 hours			

Flip through the images to perform error analysis.

Filter by label

Choose labels
Choose labels to filter images

[Find label](#)

- True positive
- False positive
- False negative

Images (27) [Info](#)

[Search images by file name](#)

amazon-branding-versions.jpg

Box # Labels Confidence

0	aws True positive	94.2%
---	----------------------	-------

aws-australia.jpg

Better organisation and scaling of operations
Reliability
Security of data
Optimization of processes and data analysis

Box # Labels Confidence

0	aws True positive	96.1%
1	aws False positive	28.1%

Deploy your Model

36. From the model details pane, you find API code to deploy your model using the AWS CLI. Alternatively, you can use the [SDK](#).

Use your model

Amazon Resource Name (ARN)
arn:aws:rekognition:us-east-1:407247006381:project/aws-brand-detector/version/aws-brand-detector.2020-08-24T19.06.38/1598321198705

▼ API Code

Use your model aws-brand-detector.2020-08-24T19.06.38 by calling the following AWS CLI commands or Python scripts. You can start and stop the model, and analyze custom labels in new images.

AWS CLI command

Python
coming soon

Start model
Command used to start the aws-brand-detector.2020-08-24T19.06.38 model.

```
1 aws rekognition start-project-version \
2   --project-version-arn "arn:aws:rekognition:us-east-1:407247006381:project/aws-brand-detector/version/aws-brand-
3   --min-inference-units 1 \
4   --region us-east-1"
```

Analyze image
Command used to analyze an image with the aws-brand-detector.2020-08-24T19.06.38 model. Replace MY_BUCKET and PATH_TO_MY_IMAGE with your S3 bucket name and image path.

```
1 aws rekognition detect-custom-labels \
2   --project-version-arn "arn:aws:rekognition:us-east-1:407247006381:project/aws-brand-detector/version/aws-brand-
3   --image '{"S3Object": {"Bucket": "MY_BUCKET", "Name": "PATH_TO_MY_IMAGE"}}" \
4   --region us-east-1"
```

37. We're going to use the Demo app that you deployed previously to start and test the model. Check your email to acquire your temporary password.

[EXTERNAL] Your temporary password for Amazon Rekognition Custom Labels Demo

no-reply@verificationemail.com <no-reply@verificationemail.com>

Tong, Dylan

Monday, August 31, 2020 at 3:55 PM

Show Details

CAUTION: This email originated from outside of the organization. Do not click links or open attachments unless you can confirm.

Your Amazon Rekognition Custom Labels Demo username is dylatong@amazon.com and the temporary password is v5Cwu&

38. Return to the CloudFormation console. Select your template and click on the URL under the **Output** sub panel.

CloudFormation > Stacks > CustomLabelsDemo

Stacks (7)

CustomLabelsDemo

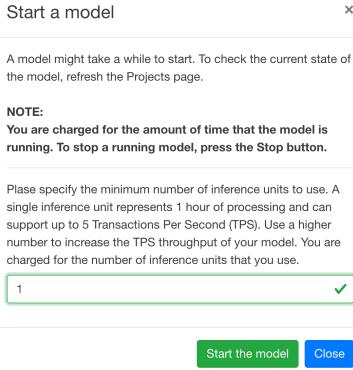
CREATE_COMPLETE

Outputs (1)

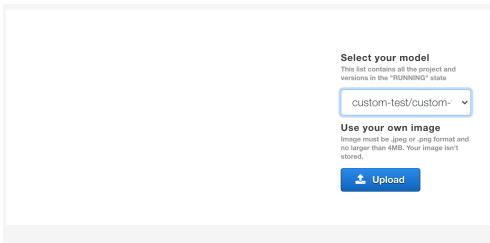
Key	Value	Description	Export name
url	https://c1km3h0ewyeo.cloudfront.net	Amazon Rekognition Custom Labels Demo Url	

39. Log into the demo app with your temporary credentials and change your password.

40. Click on the **Start Model** button from the main page. Enter **1** as the number of inference units and click on the **Start Model** button. This will take *5-15 minutes* to complete.

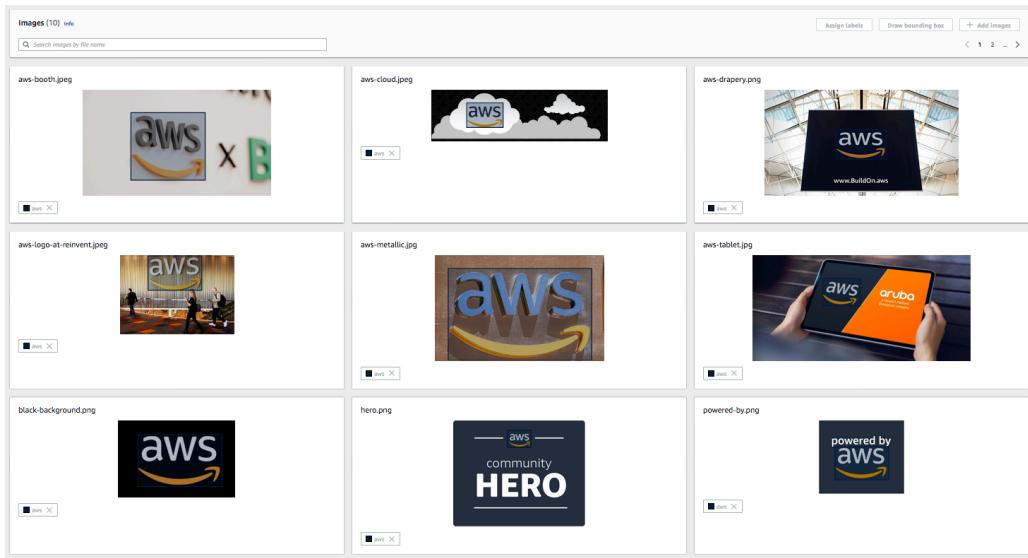


- Once your model is running, click on the name. You can upload an image to test your model.



Improve your Model

If you inspect the training set, you'll notice that the images are basic. Many of the images are high quality images of the logos. These images don't represent how you'll use the model in your production application.



Rekognition Custom Labels performs various image augmentation and processing techniques to create a more robust model from your dataset. However, training images more like what you expect to see in production will improve your model.

42. Create a new training dataset by repeating the same process as before.

Note that there are multiple ways to import new datasets. You could create this new dataset using the first training dataset then add and label images ad-hoc.

At production scale, this process would be better managed through Amazon GroundTruth. GroundTruth datasets can be imported into Rekognition Custom Labels.

The overlap between the first and second training dataset is minimal. Since there're only 14 images, it's quicker to create and label a new dataset from scratch through the Rekognition Custom Labels tooling.

The second training dataset is located at: `s3://<Your S3 URL>/train2/`. Label all the images by following the same process as before.

Dataset details

Dataset name
`aws-logos-train2`
The dataset name can't be more than 50 characters. It needs to be a valid S3 path with no spaces.

Image location
Import Images from one of the options below.

Import images labeled by Amazon SageMaker Ground Truth
Provide the location of your .manifest file. If you've labeled datasets in a different format, convert them to a .manifest format.

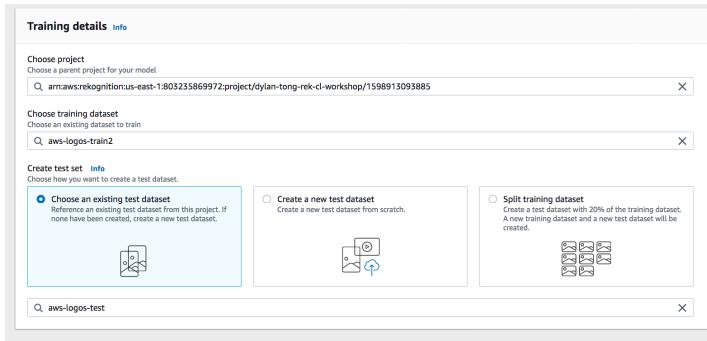

Import images from Amazon S3 bucket
Use images from an existing S3 bucket by entering the S3 folder location below. You have the option to automatically add labels based on your folder names.


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.

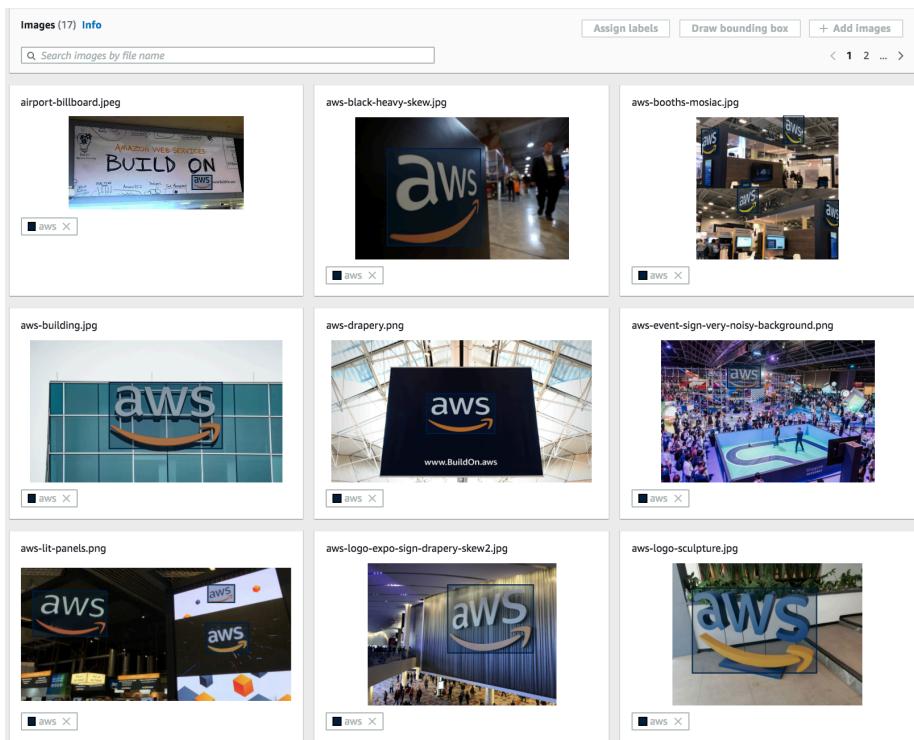

Upload images from your computer
Add images by dragging and dropping them into your dataset gallery view. You're limited to uploading 30 images at one time.


S3 folder location
S3 bucket location of your images.
`s3://dtong-ml-datasets/aws-logos/train2/`
Supported image formats: JPEG, PNG. Maximum images per dataset: 250,000. Maximum image size: 15 MB, Minimum size (px): 64 x 64. Maximum size (px): 4096 x 4096
For best results, we recommend uploading images from folders within the [S3 bucket](#) created for you during first-time setup.

43. Click on **Train Model** again. This will train and track a second version of your model under the same project. Use the same test dataset and select the new training dataset that you just created. *This will take another 1-3 hours to complete.*



44. Take a second look at this second training dataset and the first. Notice that it's more robust than the previously.



45. As you improve your dataset through error analysis and human feedback in production, you can re-train, track and evaluate model improvements.

	Name	Versions	Date created	F1 score	Status	Status message
<input type="radio"/>	aws-brand-detector	2	2020-08-25			
<input type="radio"/>	aws-brand-detector.2020-08-24T19.06.38		2020-08-25	0.854	STOPPED	The model has stopped running.
<input type="radio"/>	aws-brand-detector.2020-08-24T17.58.31		2020-08-25	0.802	STOPPED	The model has stopped running.

Next: Productionizing Your Models

Congratulations, you've completed this lab!

As you take your Rekognition Custom Label models into production consider the following:

1. In production, you should roll out your training and deployment process in a repeatable and automated process. A workflow engine like [AWS Step Functions](#) could serve this need.
2. You should consider more granular and automated tested. Rekognition provides summary evaluation metrics and a way to manually inspect the predictions on your test set. In production, it's desirable to be able to perform granular error analysis—for instance, be able to compare the training and test performance for each class in an automated fashion to detect issues between model versions. This is necessary as you scale your datasets. You can automate this step using AWS Step Functions and the [Rekognition SDK](#).
3. Manage your labels with [Amazon GroundTruth](#) as your datasets grow.
4. Having a human feedback loop is important for identify issues with your model in production and capture errors so that this data can be used for re-training and improving your model. You can use [Amazon Augmented AI](#) (A2I) for this purpose. Here's a Jupyter Notebook that [demonstrates](#) how to use A2I with Rekognition.

Additional Learning Resources

[Documentation](#)

[Human Feedback for Rekognition Lab](#)

[Media Insights Solution using Amazon Rekognition](#)