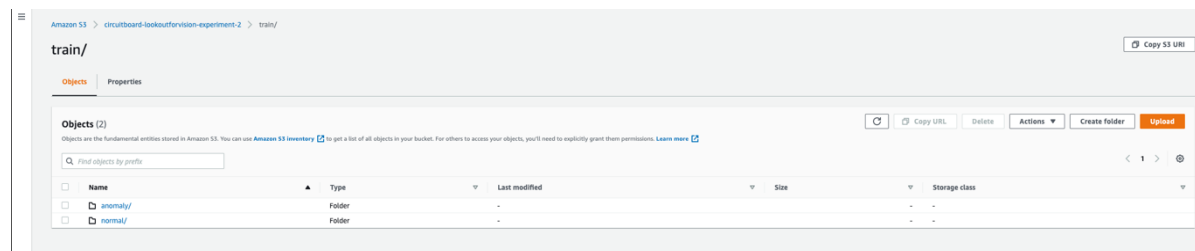
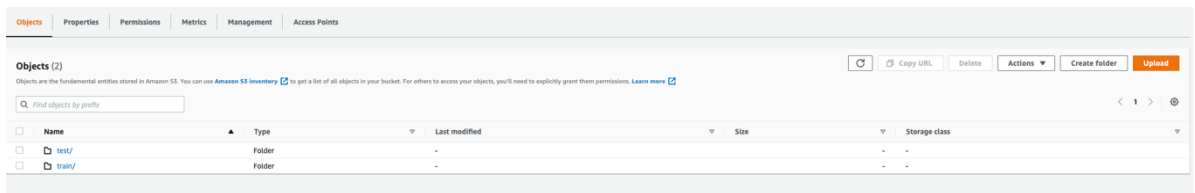
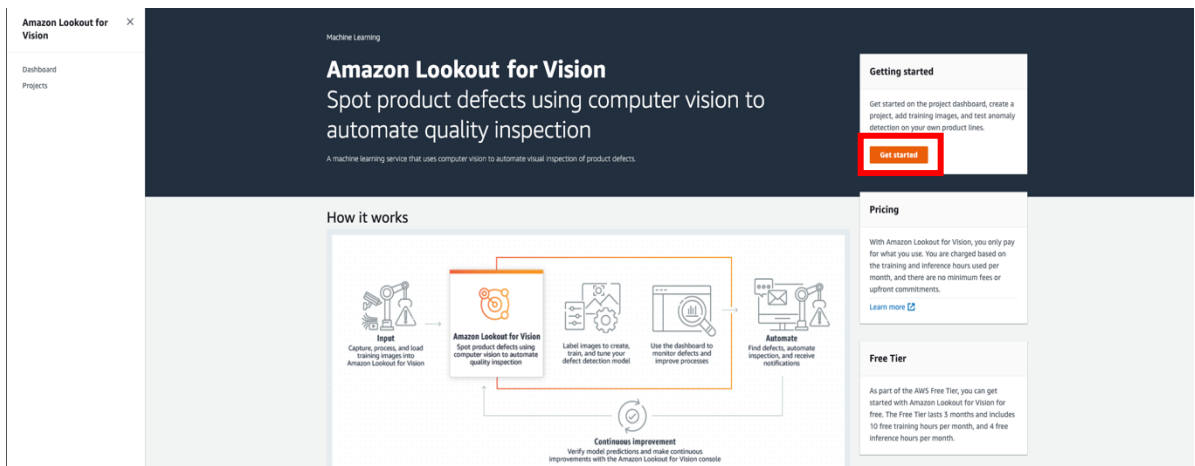


Amazon Lookout for Vision Model Training

1. **Prepare the dataset:** We start by training the model in Amazon Lookout for Vision to learn the differences between normal and anomalous circuit board images.
 - a. Open the Amazon Simple Storage Service (Amazon S3), and create a new bucket in the same region as the Amazon Lookout for Vision service. Upload the two folders, **test** and **train**, from the downloaded dataset into the S3 bucket. You should see 60 images in **train** and 20 images in **the test** folder.



- b. Open the [Amazon Lookout for Vision](#) service on the AWS management console in one of the supported regions. Choose **Get Started** shown in the figure below. The service will ask you to create an S3 bucket if you are using Amazon Lookout for Vision for the first time



2. **Create Project and Create dataset:** In the Amazon Lookout for Vision console page,


- a. Choose **Create project** to create a new project with a unique name.
- b. Choose **Create dataset**, and select the **Create a training dataset and a test dataset** radio button under **Dataset Configuration**. The remaining selections should match the below figure. Enter the S3 URL pointing to the location of the train folder & select the checkbox for **Automatic Labeling**

Dataset configuration

Configuration option

☐ **Create a single dataset**
Simplify model training by using a single dataset. Recommended for most use cases. Later, you can add a test dataset for finer control over training images, test images, and performance tuning.

☒ **Create a training dataset and a test dataset**
Use separate training and test datasets to get advanced control over training, testing, and performance tuning. Later, you can revert to a single dataset project by deleting the test dataset.


 **What are training datasets and test datasets?**


- A training dataset teaches your model to find anomalies in images.
- A test dataset evaluates the performance of your trained model.


Training dataset details

Import training images [Info](#)

Import the images for your training dataset from one of the following locations.

☒ **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.


☐ **Import images labeled by 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.


S3 URI


s3://circuitboard-lookoutforvision-experiment-2/train/

Supported image formats: JPG, PNG. Maximum images per dataset: 20,000. Maximum image size: 8 MB, Minimum size (px): 64 x 64. Maximum size (px): 4096 x 4096. Images must have the same dimensions.

Automatic labeling

To automatically label your images, create the following folder structure. Place anomalous images in the anomaly folder. Place normal images in the normal folder. Images in other folders are added as unlabeled images.

☒ **Automatically attach labels to images based on the folder name**


 **image_folder**

- c. Repeat the same process for the test dataset details, but make sure the S3 URL is pointing to the test folder in the S3 bucket. Select **Create dataset**


Test dataset details

Import test images [Info](#)
Import the images for your test dataset from one of the following locations.


☒ **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.



☐ **Import images labeled by 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.




S3 URI






s3://circuitboard-lookoutforvision-experiment-2/test/

Supported image formats: JPG, PNG. Maximum images per dataset: 20,000. Maximum image size: 8 MB, Minimum size (px): 64 x 64. Maximum size (px): 4096 x 4096. Images must have the same dimensions.

Automatic labeling
To automatically label your images, create the following folder structure. Place anomalous images in the anomaly folder. Place normal images in the normal folder. Images in other folders are added as unlabeled images.

☒ Automatically attach labels to images based on the folder name

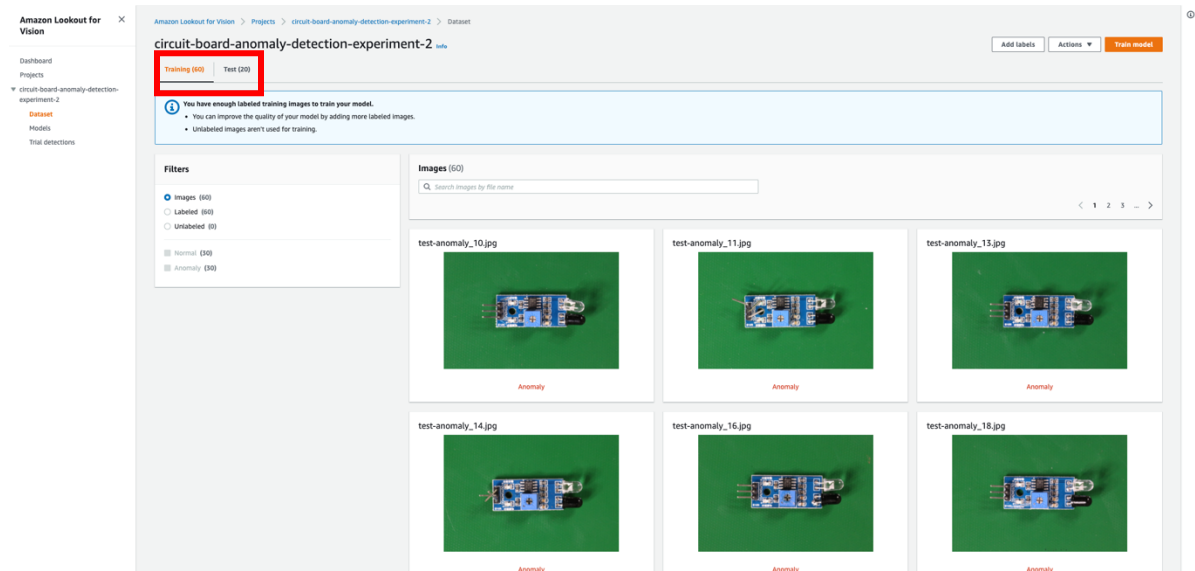
 **image_folder**

-  **anomaly**
 - 
-  **normal**
 - 
 - 

Cancel

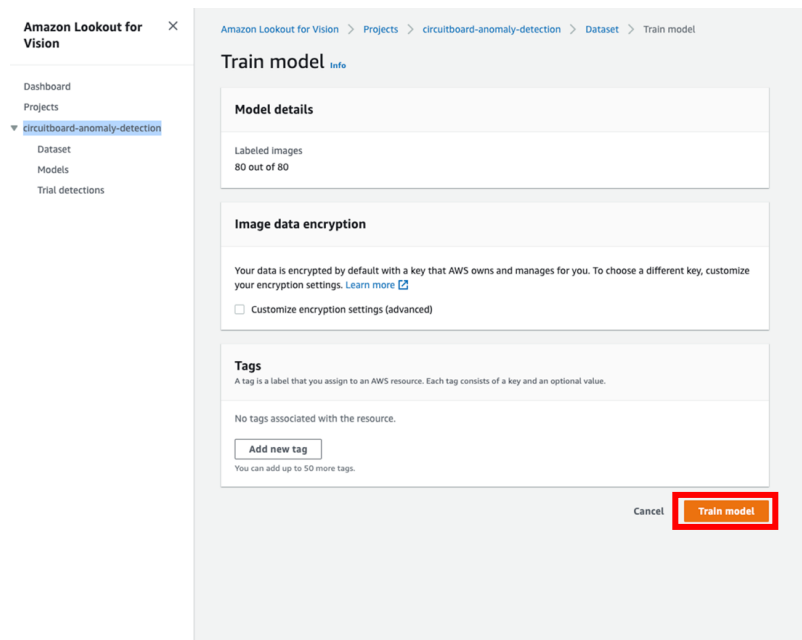
Create dataset

- d. Check the Dataset in Amazon Lookout for Vision to confirm that 60 labeled images in train and 20 in the test are created with labels as **Anomaly** and **Normal** as highlighted by the figure below.



3. Train the Amazon Lookout for Vision Model:

- a. In the Dataset section of the project, select the **Train Model** button as shown in the figure below. Select **Train Model** again on the prompt. Return to the **Models** section of the project and select **Model 1**. The model status should show **Training in progress**



Once the training is complete, Amazon Lookout will show the performance metrics of the test images against the trained model.