AWS Academy Machine Learning Foundations

Module 5: Introducing Computer Vision

Module overview

Sections

1. Introducing computer vision
2. Analyzing images and videos
3. Preparing custom datasets for computer vision
4. Module wrap-up

Demonstrations

1. Introducing Amazon Rekognition
2. Labeling Images with Amazon SageMaker Ground Truth

Lab

Guided Lab: Facial Recognition

Knowledge check

Module objectives

At the end of this module, you should be able to:

- Describe the use cases for computer vision
- Describe the Amazon managed machine learning (ML) services for image and video analysis
- List the steps required to prepare a custom dataset for object detection
- Describe how Amazon SageMaker Ground Truth can be used to prepare a custom dataset
- Use Amazon Rekognition to perform facial detection

Module 5: Introducing Computer Vision

Section 1: Introducing computer vision

aws academy

Computer vision overview

Computer vision is the automated extraction of information from digital images.

Computer vision applications

Public safety
and home security

Authentication and enhanced
computer-human interaction

Content management
and analysis

Autonomous driving

Medical imaging

Manufacturing process control

Computer vision problems

Content recognition

Image analysis

- Object classification

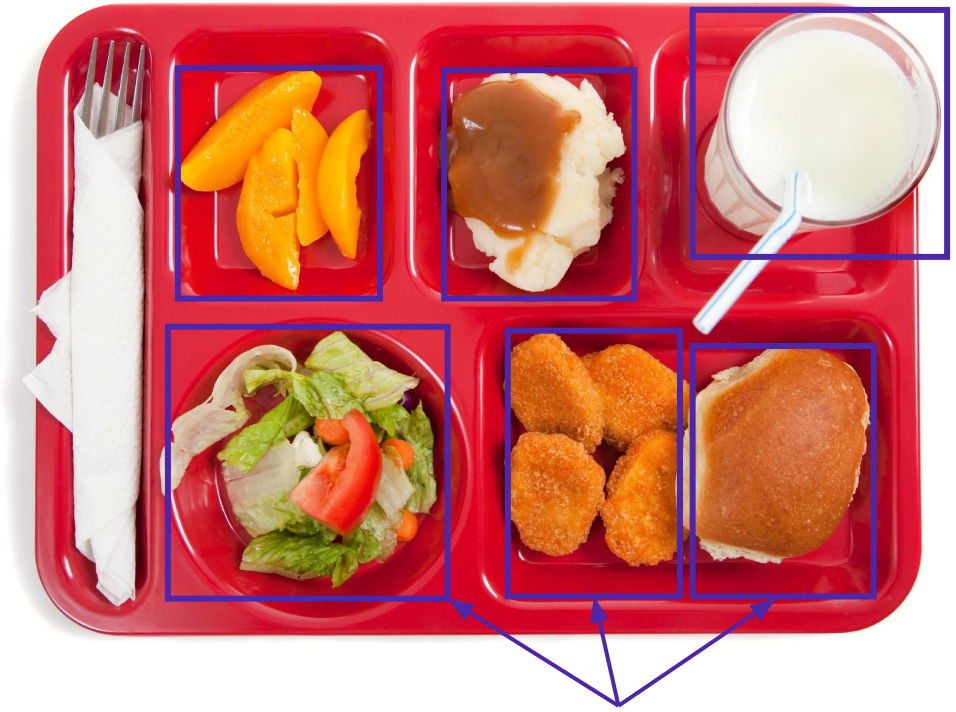
Food?
Breakfast?
Lunch?
Dinner?

Computer vision problems

Content recognition

Image analysis

- Object classification
- Object detection



	Confidence
Milk	97.1
Peaches	92.3
Ice Cream	97.1
Salad	69.5
Nuggets	77.5
Bread Roll	94.5

Bounding boxes

(top,left,width,height)

Computer vision problems

Content recognition

Image analysis

- Object classification
- Object detection
- Object segmentation

Milk

Peaches

Ice Cream

Salad

Nuggets

Bread Roll

Computer vision use cases

Content recognition

Video analysis

- Instance tracking

Pathing – You can capture the path of people in the scene. For example, you can use the movement of athletes during a game to identify plays for post-game analysis.

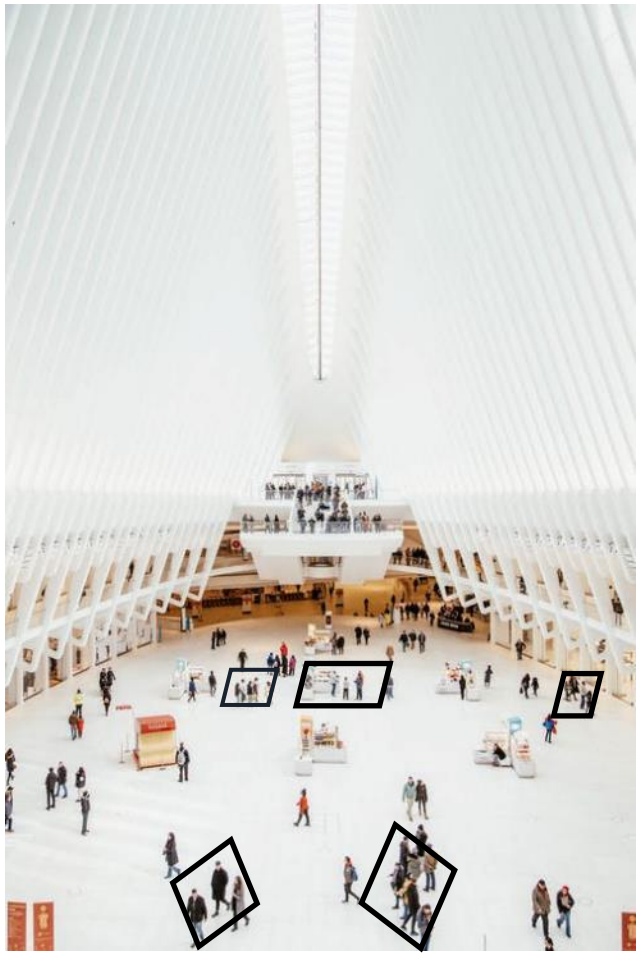
Computer vision use cases

Content recognition

Video analysis

- Instance tracking
- Action recognition

Analyze shopper behavior and density in your retail store by studying the path that each person follows



Computer vision use cases

Content recognition

Video analysis

- Instance tracking
- Action recognition
- Motion estimation

Section 1 key takeaways

- Computer vision is the automated extraction of information from images
- Image analysis includes object classification, detection, and segmentation
- Video analysis includes instance tracking, action recognition, and motion estimation

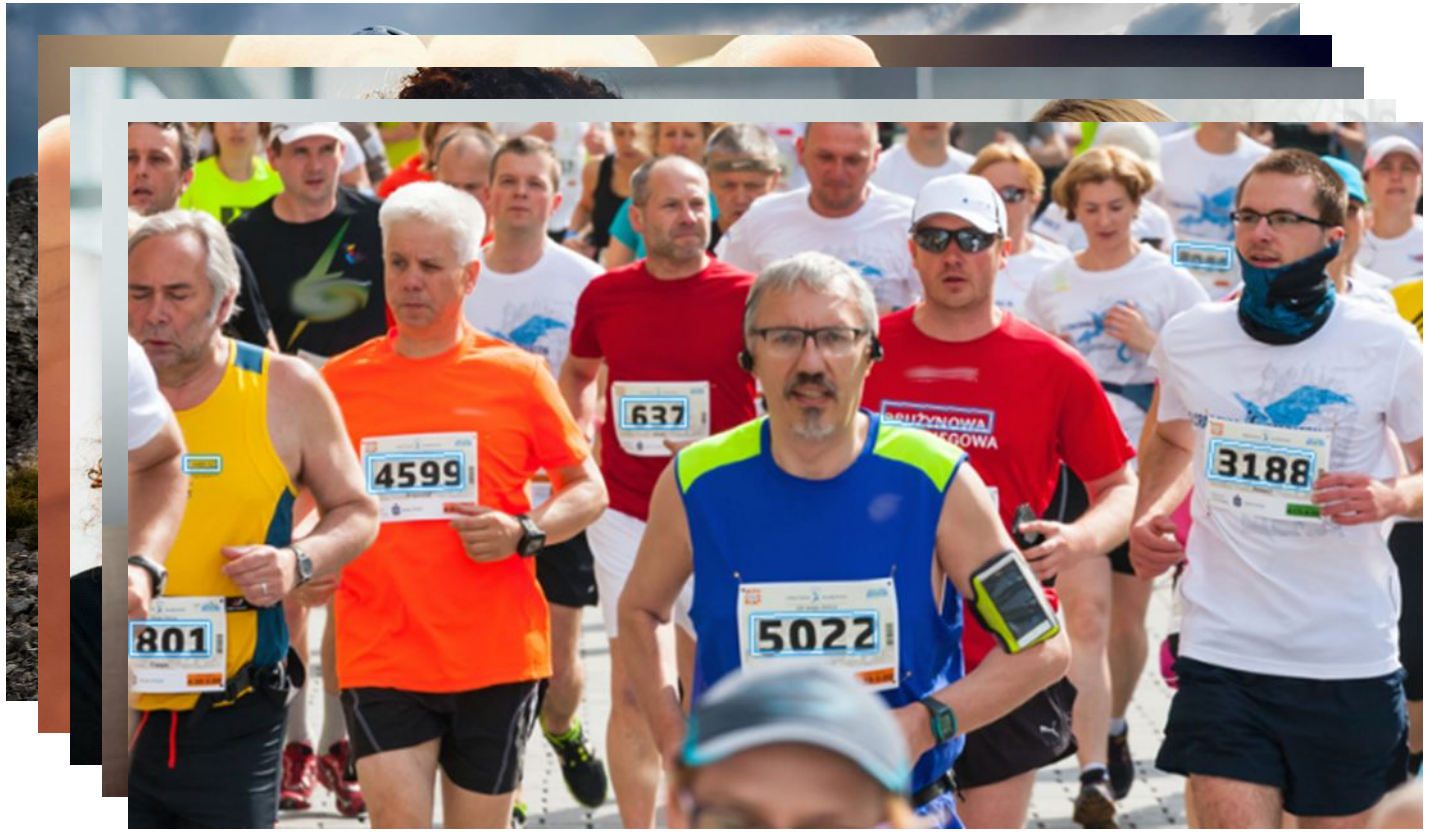
Module 5: Introducing Computer Vision

Section 2: Analyzing images and videos

aws academy

Amazon Rekognition

- Managed service for image and video analysis
- Types of analysis
  - Searchable image and video libraries
  - Face-based user verification
  - Sentiment and demographic analysis
  - Unsafe content detection
  - Text detection
- Security and compliance



Amazon Rekognition

Can add powerful visual analysis to your application

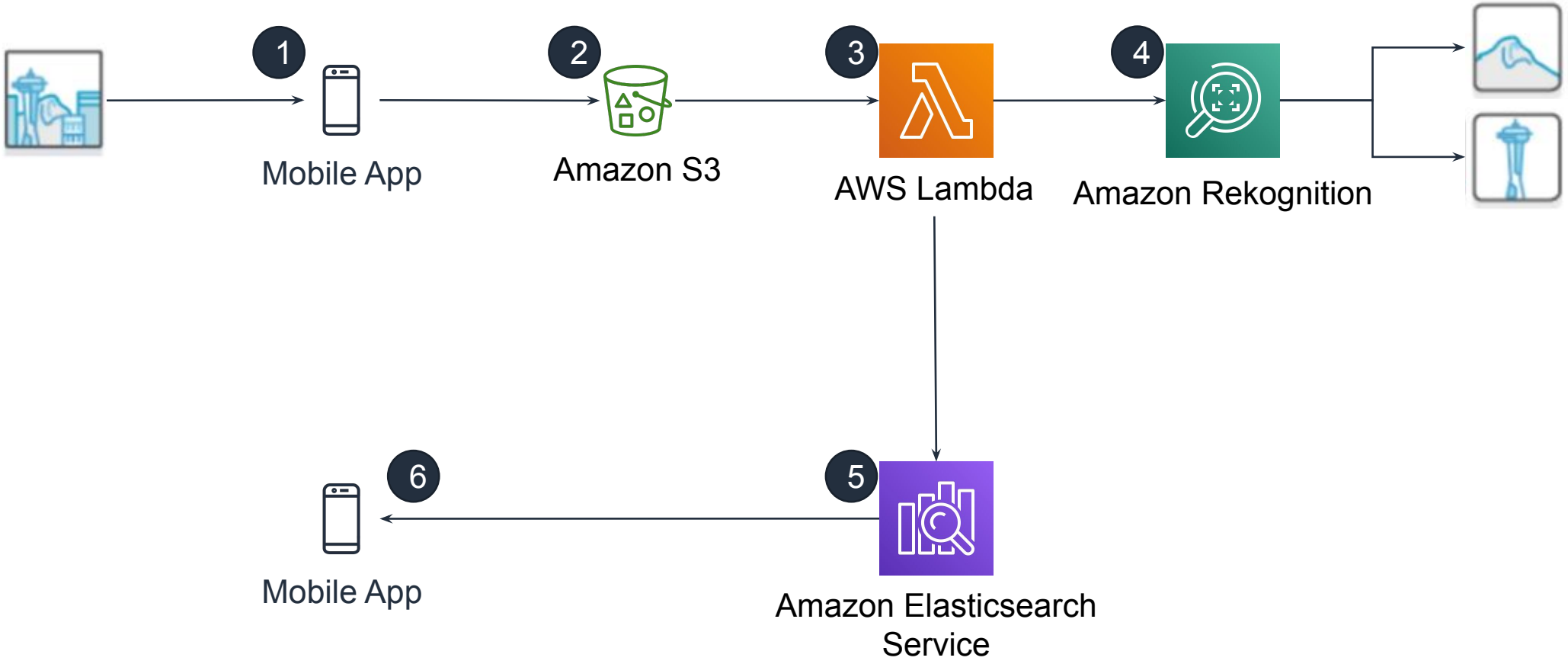
Is highly scalable and continuously learns

Integrates with other AWS services

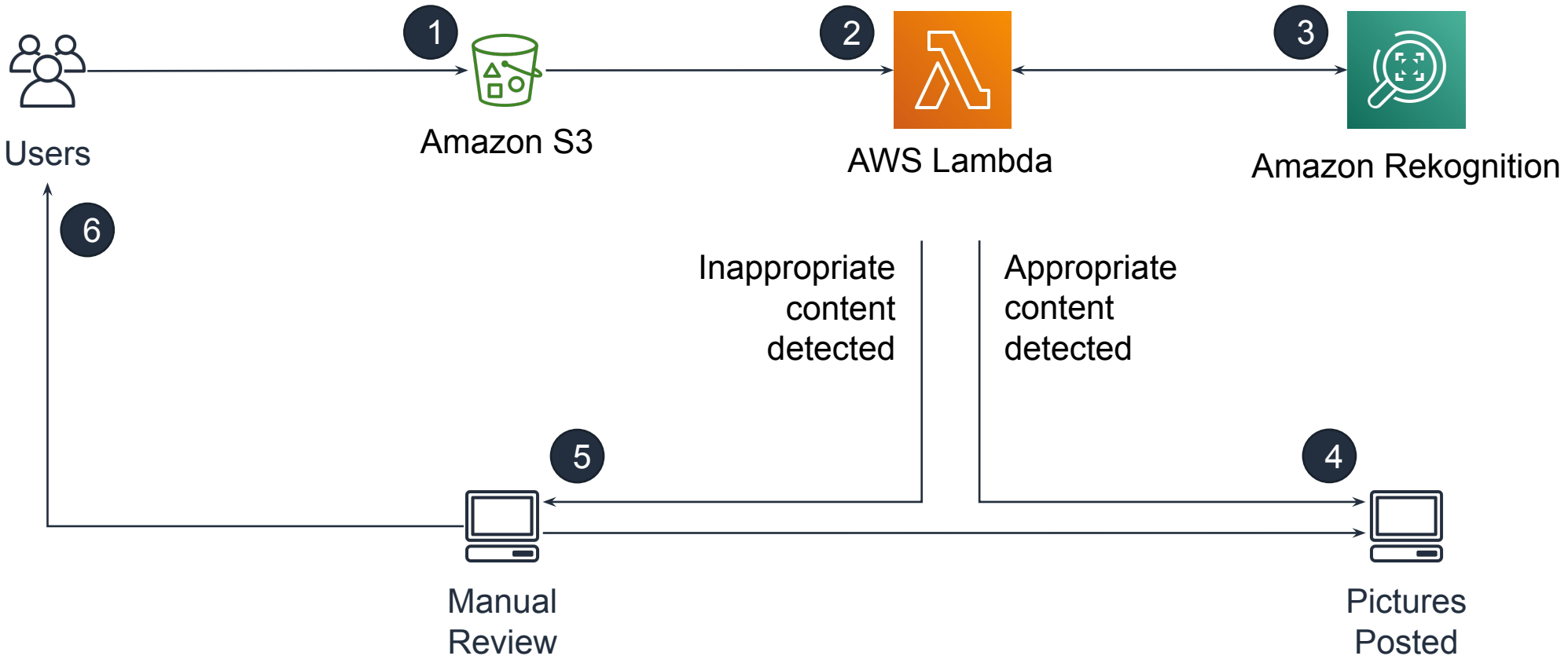
Languages supported by the Amazon Rekognition SDKs:

JavaScript	Python	PHP	.NET	Ruby	Java	Go	Node.js	C++

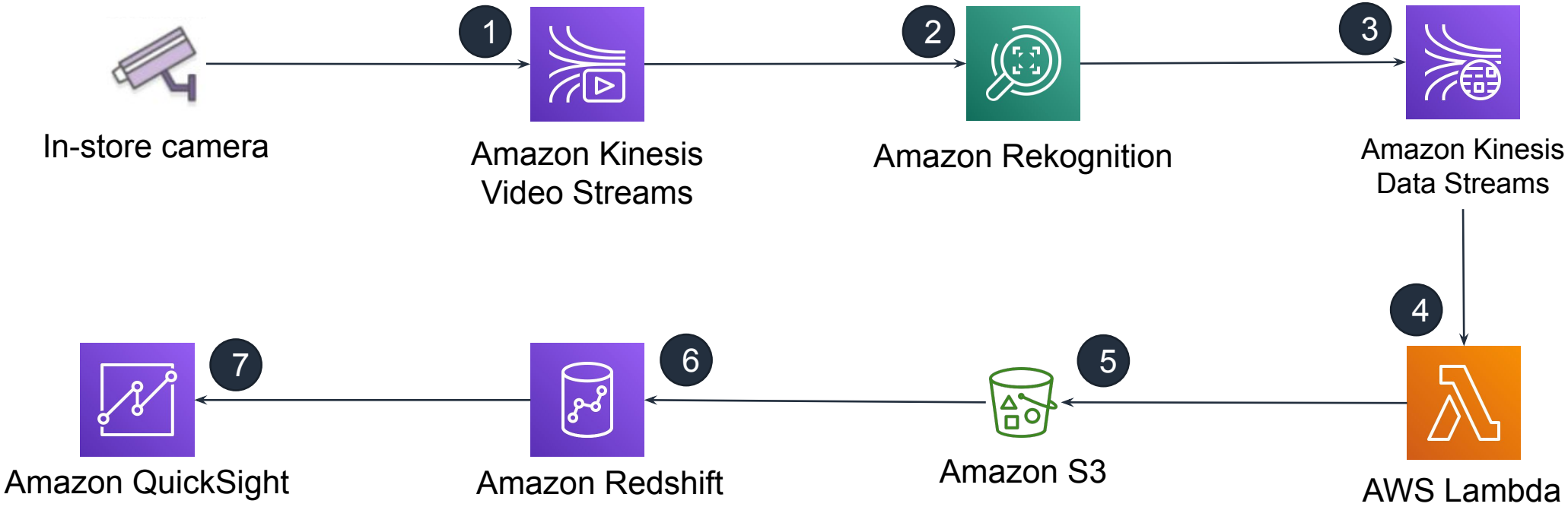
Use case 1: Searchable image library



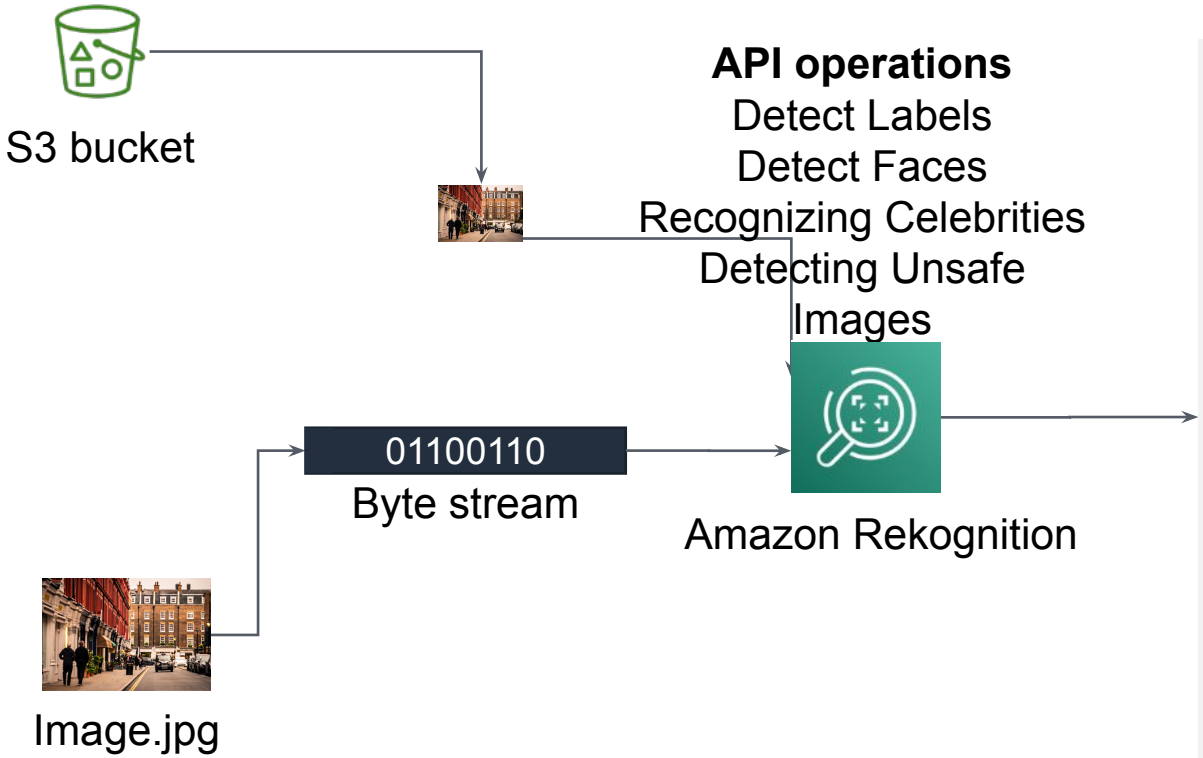
Use case 2: Image moderation



Use case 3: Sentiment analysis



Using Amazon Rekognition



JSON results


```
"Name": "Car",
   "Confidence": 96.9476089477539,
   "Instances": [
    {
            "BoundingBox": {
               "Width": 0.07903634756803513,
               "Height": 0.086999185383319 85,
               "Left": 0.7790974974632263,
               "Top": 0.6977462768554688
            },
            "Confidence": 96.9476089477539
    }
```

Image analysis: Object and scene detection

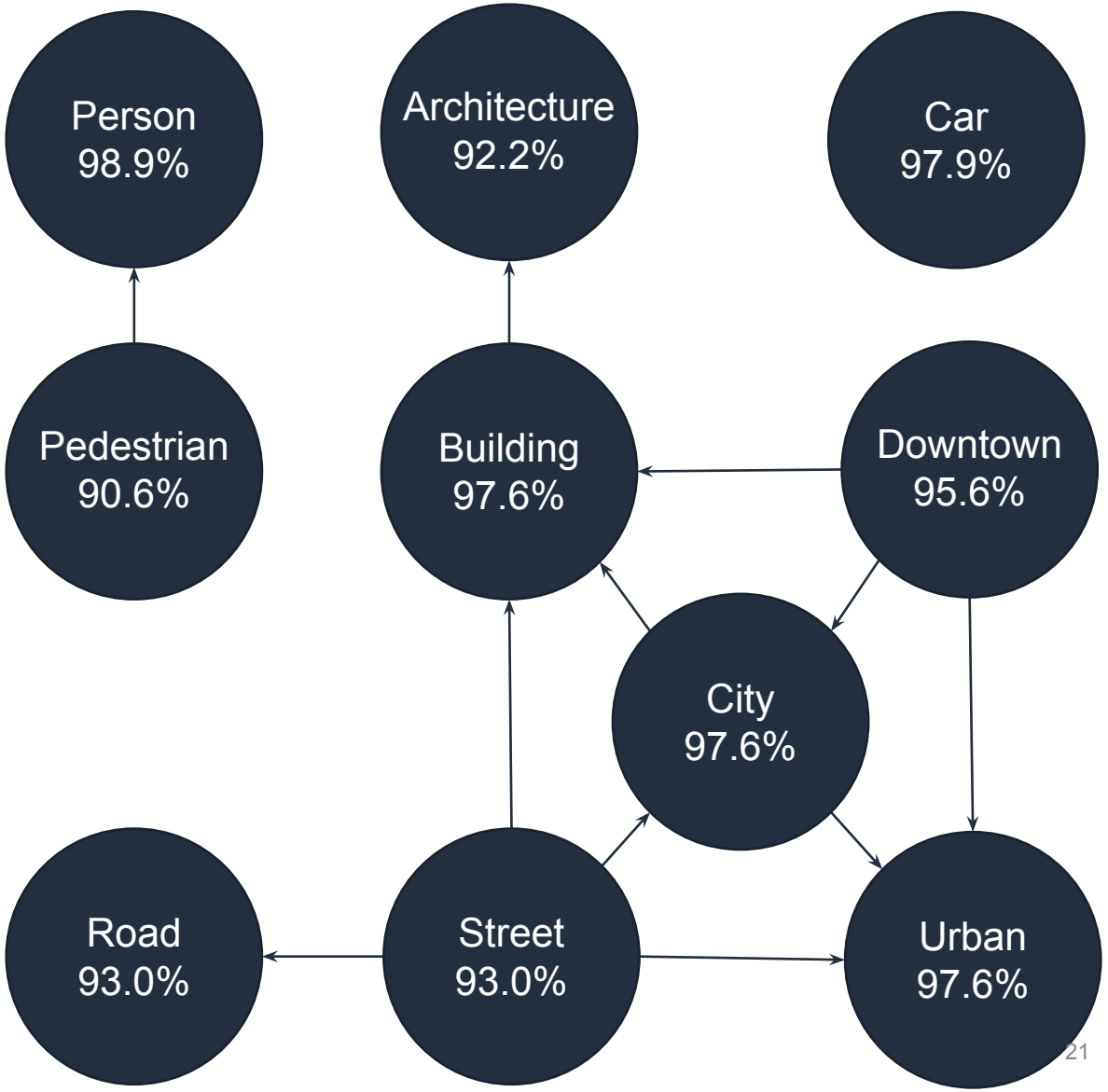
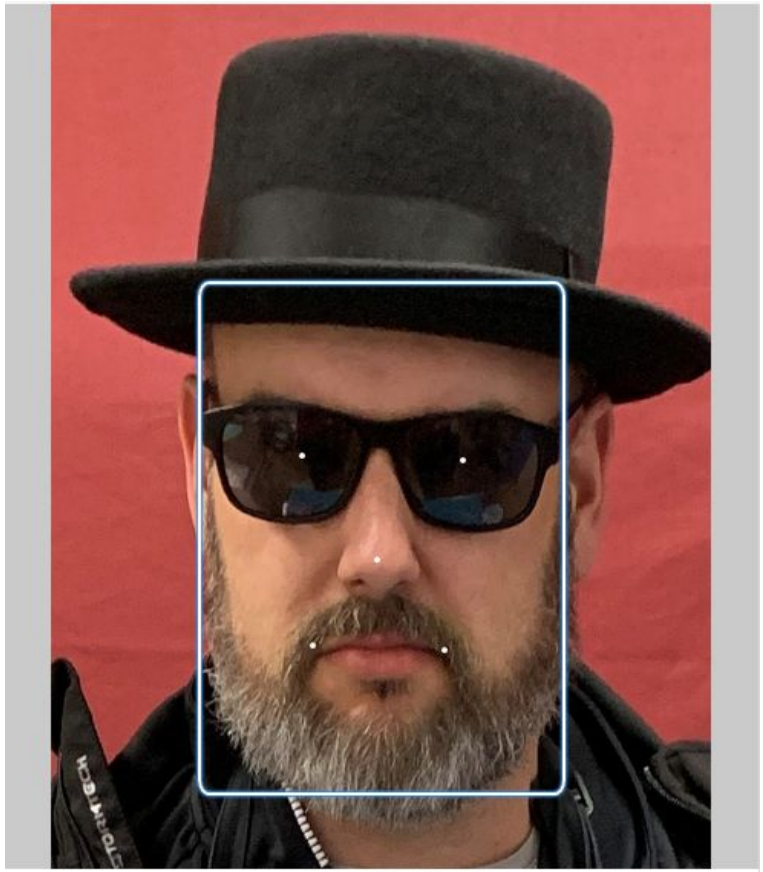


Image analysis: Object and scene detection


```
"Name": "Car",
   "Confidence": 96.9476089477539,
   "Instances": [
    {
            "BoundingBox": {
               "Width": 0.07903634756803513,
               "Height": 0.08699918538331985,
               "Left": 0.7790974974632263,
               "Top": 0.6977462768554688
            },
            "Confidence": 96.9476089477539
    }
```

Face detection

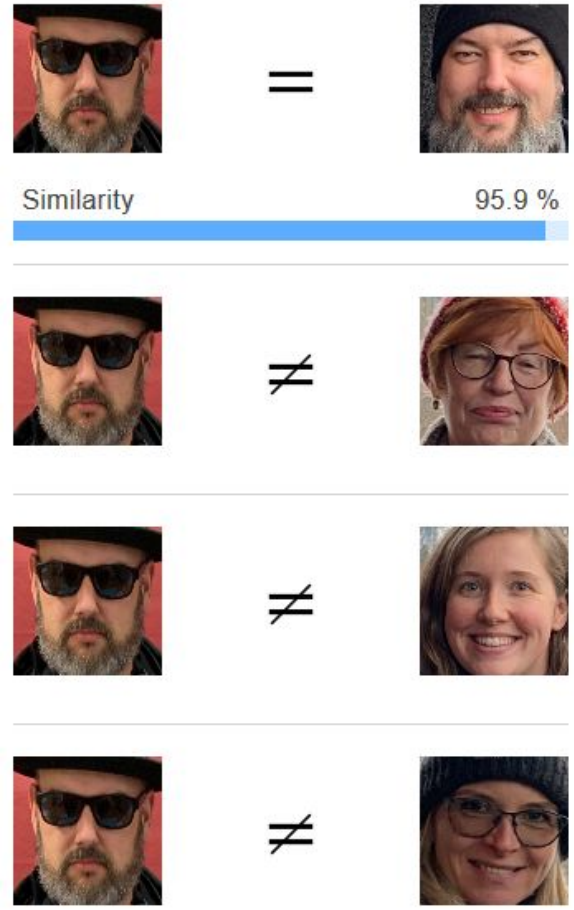
- Bounding box
- Attributes
- Emotions
- Facial landmarks
- Quality
- Pose
- Confidence score



looks like a face	99.9 %
appears to be male	99.5 %
age range	32 - 48 years old
not smiling	99.8 %
appears to be calm	99.4 %
wearing glasses	99.5 %
wearing sunglasses	96.7 %
eyes are open	99.9 %
mouth is closed	98.7 %
does not have a mustache	75.1 %
has a beard	96.6 %

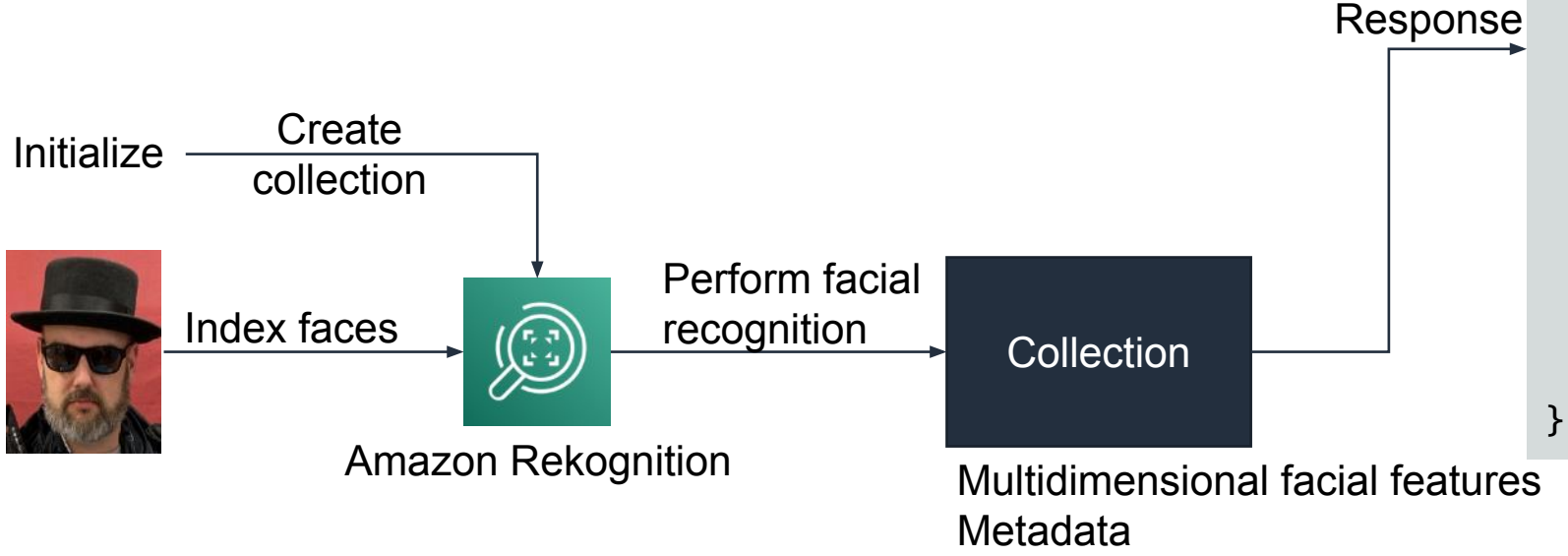
Using facial recognition

- Compare *source* with *target*
- Face match –
  - Bounding box and confidence
  - Similarity score
  - Facial landmark locations
- Source face information –
  - Bounding box and confidence
  - Facial landmarks
- Unmatched face –
  - Bounding box and confidence
  - Facial landmarks



Searching for known faces

- Must train the model
- Use images with target faces
- Perform facial recognition
- Store facial metadata




```
{
   "FaceModelVersion": "string",
   "Faces": [
      {
         "BoundingBox": {
            "Height": number,
            "Left": number,
            "Top": number,
            "Width": number
         },
         "Confidence": number,
         "ExternalImageId": "string",
         "FaceId": "string",
         "ImageId": "string"
      }
   ],
   "NextToken": "string"
}
```

Searching for Known Faces

SearchFacesByImage

Amazon Rekognition

Collection

Multidimensional facial features
Metadata

```
Response
{ "FaceMatches": [ {
    "Face": {
        "BoundingBox": { "Height": 0.0633333027, "Left": 0.171851992, "Top": 0.73666697,
"Width": 0.1106169968 },
        "Confidence": 100,
        "ExternalImageId": "input.jpg",
        "FaceId": "578e2e1b-d0b0-493c-aa39-ba476a421a34",
        "ImageId": "9ba38e68-35b6-5509-9d2e-fcffa75d1653" },
        "Similarity": 99.9764175415039 } ],
"FaceModelVersion": "3.0",
"SearchedFaceBoundingBox": { "Height": 0.063333332, "Left": 0.171851858, "Top":
0.73666667, "Width": 0.11061728 },
"SearchedFaceConfidence": 99.99999237060547 }
```

Guidelines

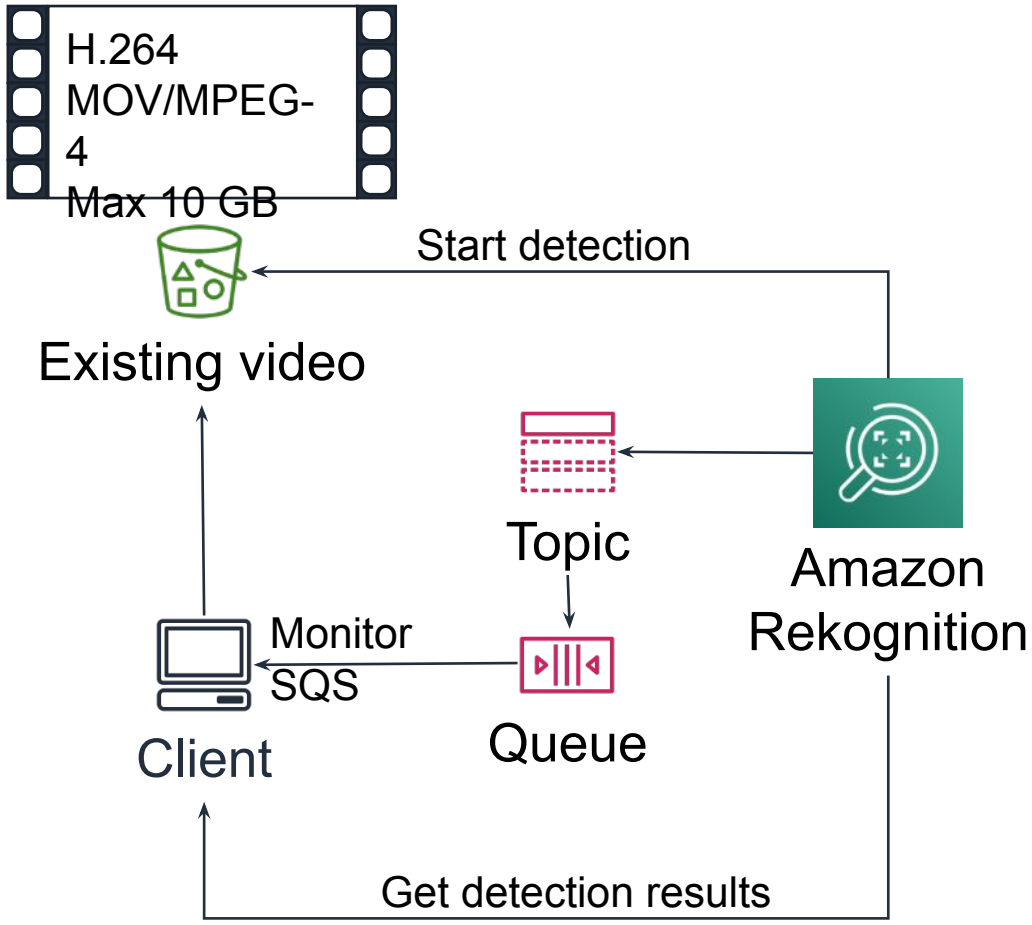
- Facial detection
  - Bounding box, attributes, emotions, landmarks, quality, pose
  - Confidence score
- Detection is based on image data
  - Gender based on image, not identity
  - Emotions are inferred from physical appearance
- You can use the appropriate confidence score for your use case
- Facial recognition should never be used in a way that violates an individual's rights, including the right to privacy, or makes autonomous decisions for scenarios that require analysis by a human*

Demonstration: Introducing Amazon Rekognition

Working with stored videos

Process:

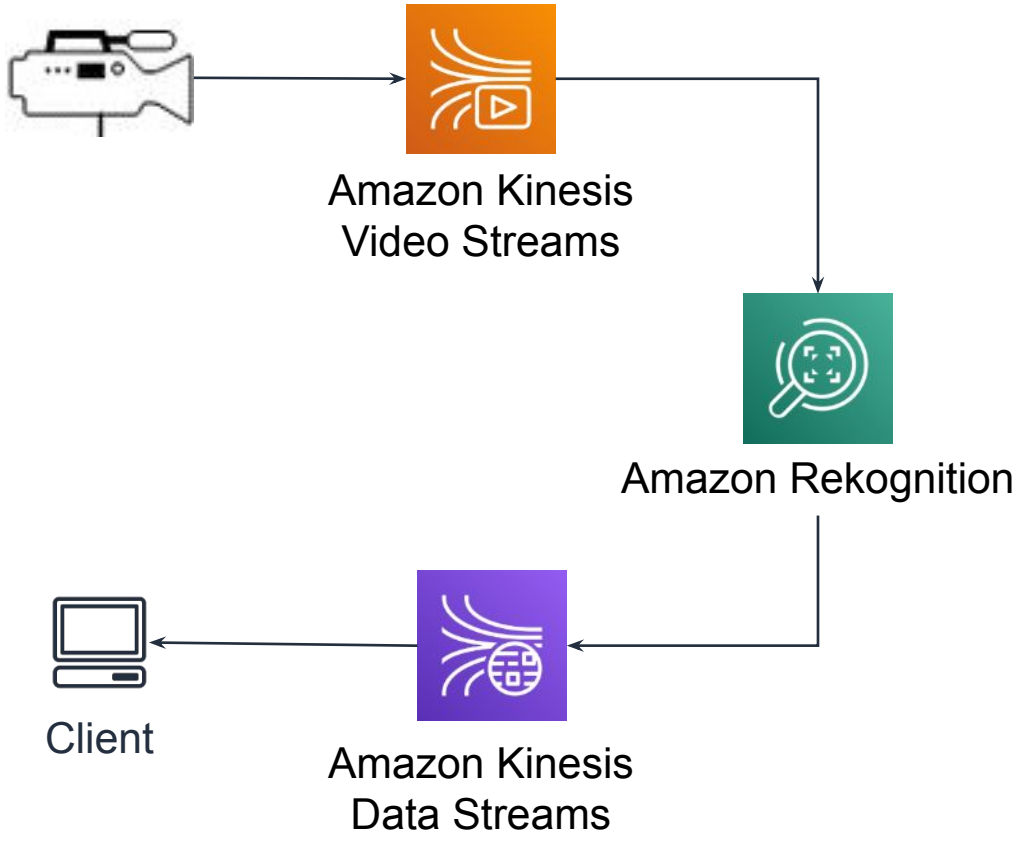
- Start detection
  - People, faces, labels, celebrities, text, inappropriate content
- Monitor Amazon Simple Queue Service (Amazon SQS) queue for completion
- Get detection results



Working with streaming videos

Application process:

1. Stream video to Amazon Kinesis Video Streams
2. Connect Amazon Rekognition Video stream processor
3. Read analysis from the Amazon Kinesis data stream



Section 2 Key Takeaways

- You can use Amazon Rekognition for image and video analysis that uses proven, highly scalable, deep learning technology that doesn't require machine learning expertise
- Provides image and video detection of faces, sentiment, text, unsafe content, and library search
- Is integrated into other AWS services

Module 5: Introducing Computer Vision

Section 3: Preparing custom datasets for computer vision

aws academy

Video analysis, labels, Ben, and Metric

AWS DeepLens Powered Cat Flap

Models need training on domain

Wood	97%
Canvas	88.7%
Text	84.7%
Envelope	71.7%
Greeting Card	70.4%
Mail	70.4%
Hardwood	60.6%
Plywood	60.4%
Paper	57.7%
Advertisement	55.8%
Poster	55.8%
Art	55.6%

Amazon Rekognition Custom Labels

- Use for –
  - Search for logos
  - Identify products
  - Identify machine parts
  - Distinguish between healthy and infected plants
- Almost all vision solutions start with an existing model
- Benefits –
  - Simplified data labeling
  - Automated machine learning
  - Simplified model evaluation, inference, and feedback

Custom labeling process

1: Collect images

2: Create training dataset

3: Create test dataset

4: Train the model

5: Evaluate

6: Use Model

Step 1: Collect images

1: Collect images | 2: Create training dataset | 3: Create test dataset | 4: Train the model | 5: Evaluate | 6: Use Model

- Typically use a few hundred images
- Build domain-specific models
- Use 10 PNG or JPEG images per label
- Use images similar to the images that you want to detect

Step 2: Create training dataset

1: Collect images | 2: Create training dataset | 3: Create test dataset | 4: Train the model | 5: Evaluate | 6: Use Model

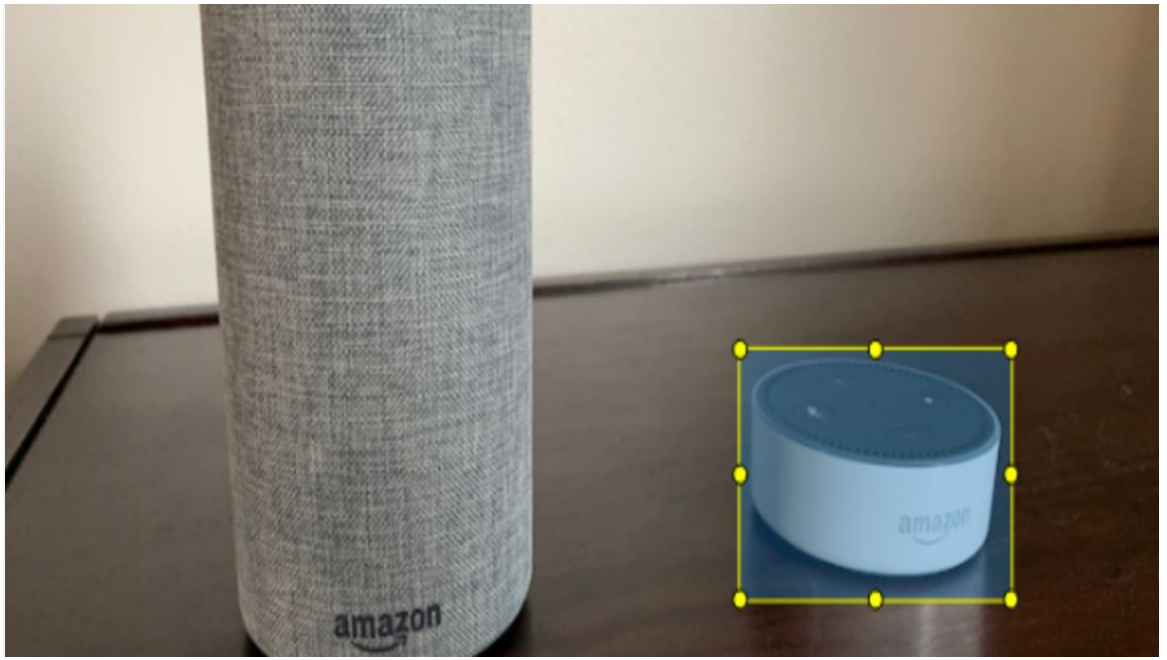
- Dataset: Data about images, labels, and bounding box
- Create at least two labels
- Label the images by using the console or Amazon SageMaker Ground Truth

Image-level versus object-level labels

Scenes and concepts

Label:
beach

Objects with bounding boxes

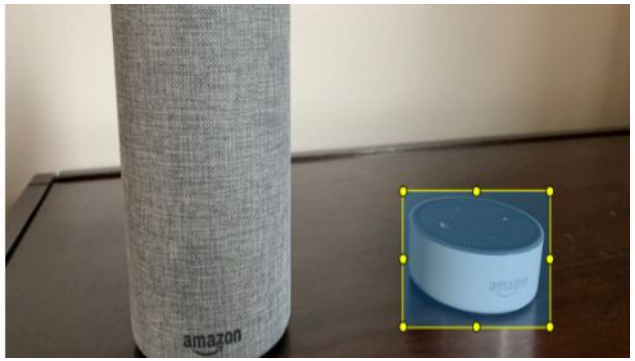


Label: *Echo Dot*

Images need labels

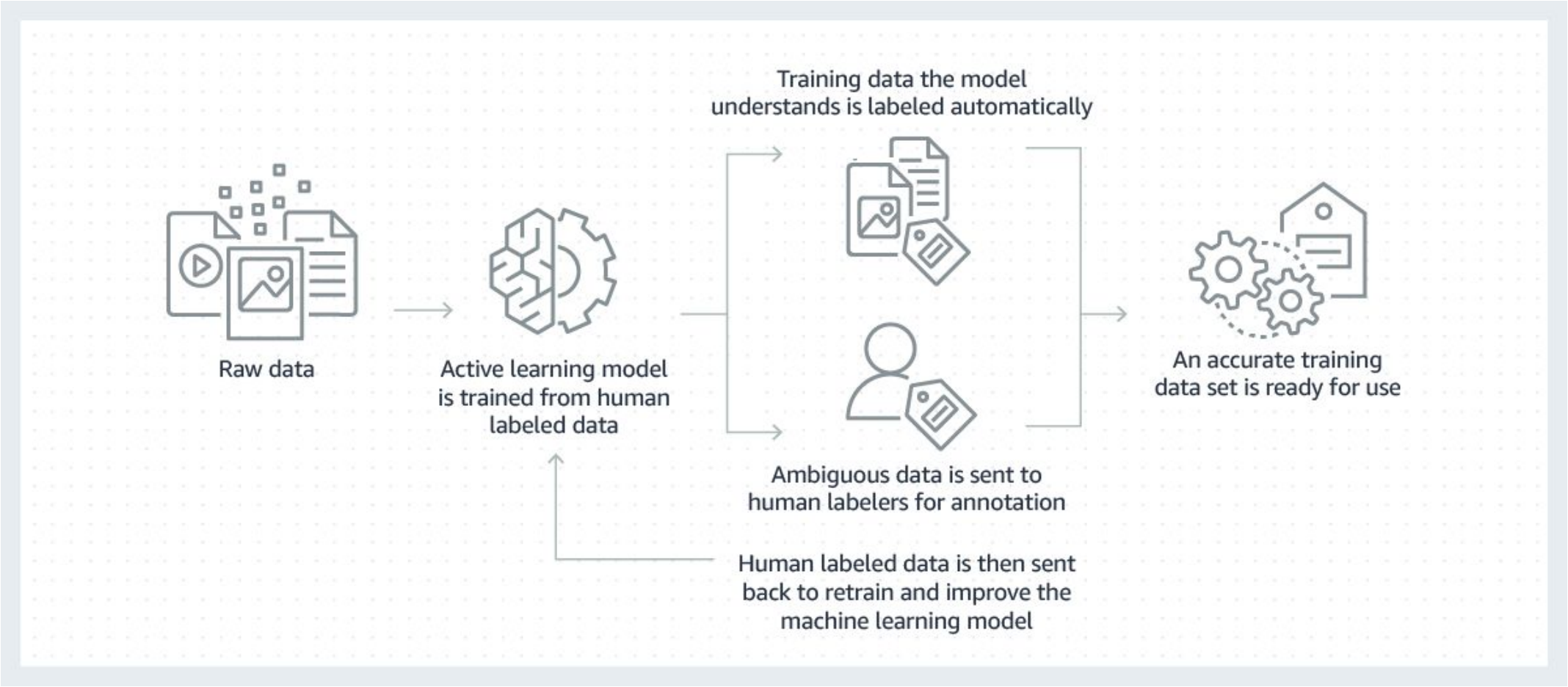
```
{
    "source-ref": "s3://b/…/img_2783.jpg",
    "small_beach": 1,
    "small_beach-metadata": {
       "confidence": 1,
       "job-name": "labeling-job/small_beach",
       "class-name": “Beach",
       "human-annotated": "yes",
       "creation-date": "2020-03-02T20:42:03.525Z",
       "type": "groundtruth/image-classification"
    }
}
```

Objects need bounding boxes




```
"small_BB": { "annotations": [
      {  "left": 186,"top": 125,"width": 346,"height": 580,
         "annotationType": "bounding-box", "id": 5,"class_id": 0},
      {  "left": 732,"top": 445,"width": 270,"height": 192,
         "annotationType": "bounding-box", "id": 6,"class_id": 1}],
   "image_size": [
      {  "width": 1152,"height": 864,"depth": 3 }]},
"small_BB-metadata": {
   "job-name": "labeling-job/small_BB",
   "class-map": { "0": "Echo", "1": "Echo dot" },
   "human-annotated": "yes",
   "objects": [ {"confidence": 1}, {"confidence": 1}],
   "creation-date": "2020-03-02T20:42:25.930Z",
   "type": "groundtruth/object-detection"
```

Amazon SageMaker Ground Truth



Step 3: Create test dataset

1: Collect images

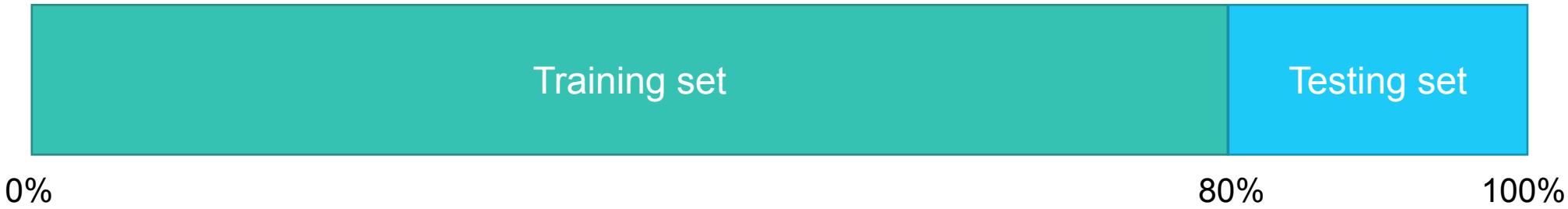
2: Create training dataset

3: Create test dataset

4: Train the model

5: Evaluate

6: Use Model



Step 4: Train the model

1: Collect images

2: Create training dataset

3: Create test dataset

4: Train the model

5: Evaluate

6: Use Model

Step 5: Evaluate - Metrics

1: Collect Images | 2: Create training dataset | 3: Create test dataset | 4: Train the model | 5: Evaluate | 6: Use Model

- Evaluate model performance
- Metrics
  - Precision
  - Recall
  - Overall model performance

Confusion Matrix

Predict		Actual: Cat	Actual: Not a Cat
	Cat	TP	FP
	Not a Cat	FN	TN

Step 5: Evaluate – Improve

1: Collect Images | 2: Create training dataset | 3: Create test dataset | 4: Train the model | 5: Evaluate | 6: Use Model

- Data
- Reducing false positives (better precision)
  - Adjust the confidence threshold to improve precision
  - Add additional classes as labels for training
- Reducing false negatives (better recall)
  - Lower the confidence threshold to improve recall
  - Use better data or more precise classes (labels) for training

Step 6: Use the model

1: Collect Images | 2: Create training dataset | 3: Create test dataset | 4: Train the model | 5: Evaluate | 6: Use Model

```
aws rekognition detect-custom-labels --project-version-arn "model_arn"\
   --image '{"S3Object":{"Bucket":"bucket","Name":"image"}}'\
   --min-confidence 70
```

Returns array of custom labels:

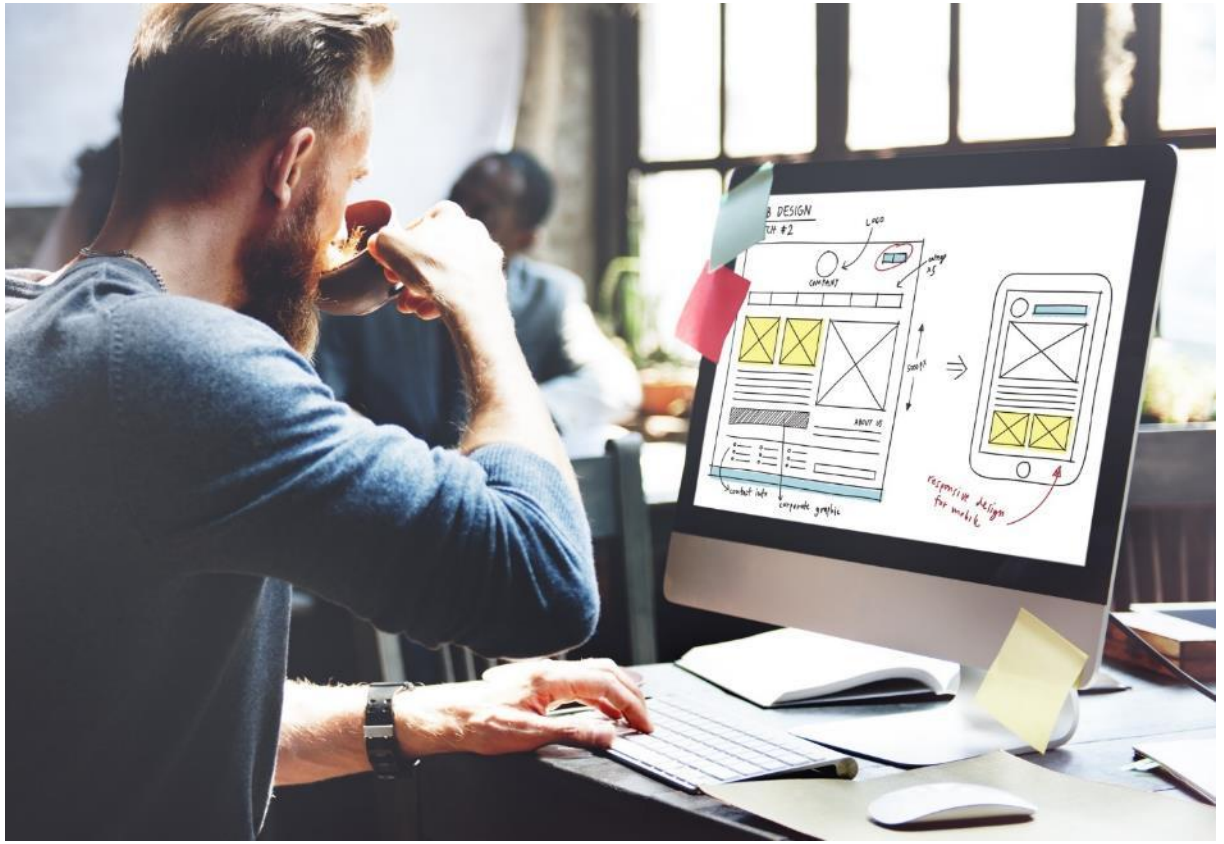
- Label
- Bounding box for objects
- Confidence

Demonstration: Labeling Images with Amazon SageMaker Ground Truth

Section 3 key takeaways

- Models must be trained for the specific domain that you want to analyze
- You can set custom labeling for the specific business case
  - Custom labeling workflow
- You must label images and create bounding boxes for objects
- You can use Amazon SageMaker Ground Truth to build training datasets for your models

Module 5 – Guided Lab: Facial Recognition



Module 5: Introducing Computer Vision

Module wrap-up

aws academy

Module summary

In summary, in this module, you learned how to:

- Describe use cases for computer vision
- Describe the AWS managed machine learning (ML) services for image and video analysis
- List the steps required to prepare a custom dataset for object detection
- Describe how Amazon SageMaker Ground Truth can be used to prepare a custom dataset
- Use Amazon Rekognition to perform facial detection

Complete the knowledge check

Additional resources

- What is Amazon Rekognition
- Welcoming Amazon Rekognition Video: Deep-Learning Based Video Recognition
- Classify a Large Number of Images with Amazon Rekognition and AWS Batch

Thank you

aws academy