

AI-3004: Azure で コンピューター ビジョン ソリューションを開発する

全6モジュール

本資料の内容は以下のMicrosoft Learnコンテンツに準拠しています。



**Azure でコンピューター ビジョン ソリ
ューションを開発する**

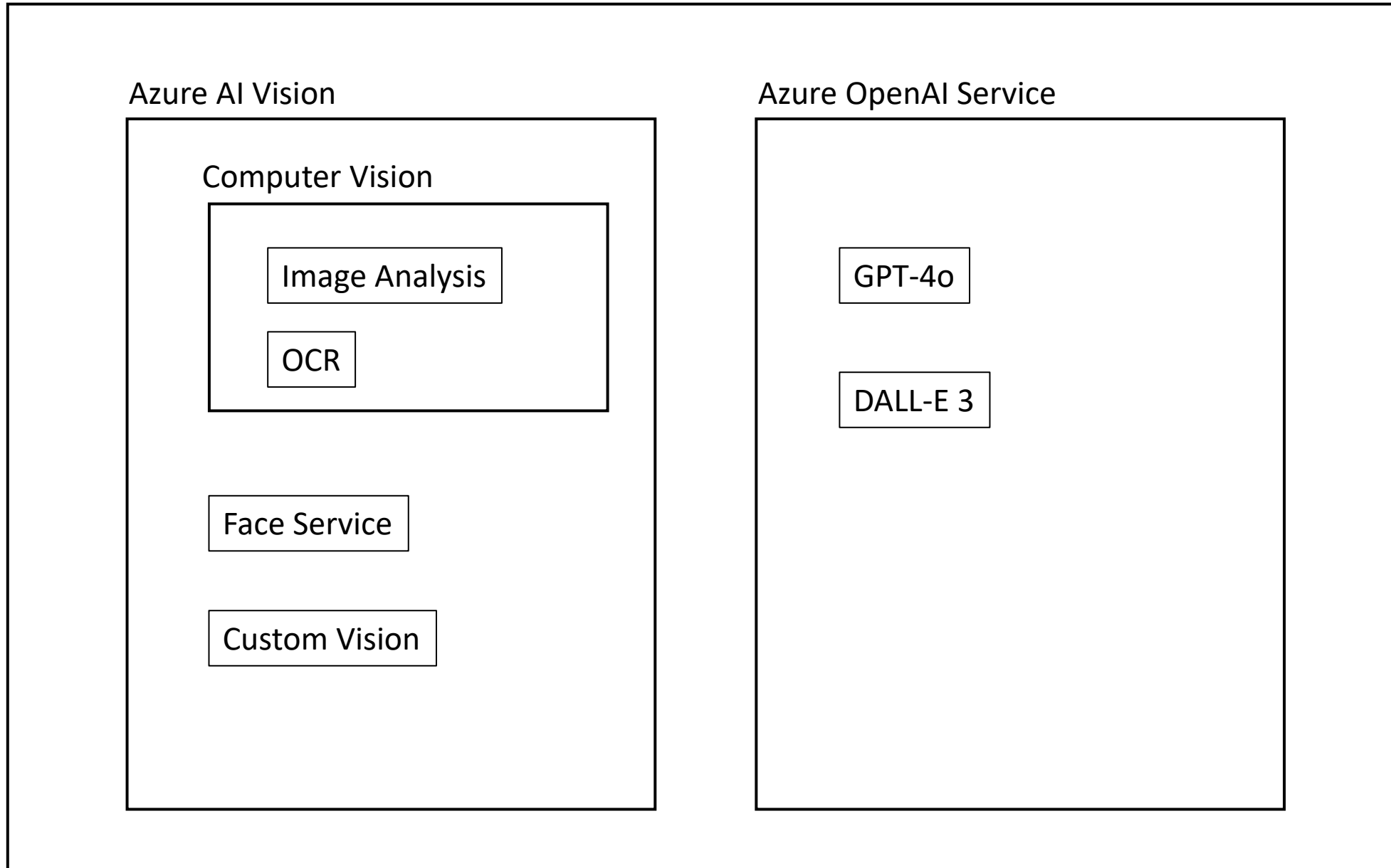
このコースでは・・・

- Azure AI Services (旧 Azure Cognitive Services) に含まれる画像処理系の機能について学習します
- 具体的には「画像分析」「画像からのテキストの読み取り」「画像からの顔の情報の検出」「画像の分類」「画像についての質問への回答」「画像の生成」について学習します
- ラボ（演習）ではこれらの機能を使用するPythonアプリを開発して動作を確認します

AI-3004 もくじ

- モジュール1 画像を分析する
 - ラボ1 Azure AI Vision > Computer Vision > Image Analysis
- モジュール2 画像内のテキストを読む
 - ラボ2 Azure AI Vision > Computer Vision > OCR
- モジュール3 顔を検出して分析する
 - ラボ3 Azure AI Vision > Face Service
- モジュール4 画像を分類する
 - ラボ4 Azure AI Vision > Custom Vision
- モジュール5 ビジョン対応のチャットアプリを開発する
 - ラボ5 Azure OpenAI Service > GPT-4oモデル
- モジュール6 AIで画像を生成する
 - ラボ6 Azure OpenAI Service > DALL-Eモデル

Azure AI Services

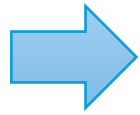


モジュール1 画像を分析する

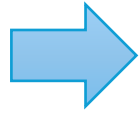
- 画像の分析とは？
- 画像の分析のためのAzureリソース
- ラボ1
- まとめ

画像の分析とは？

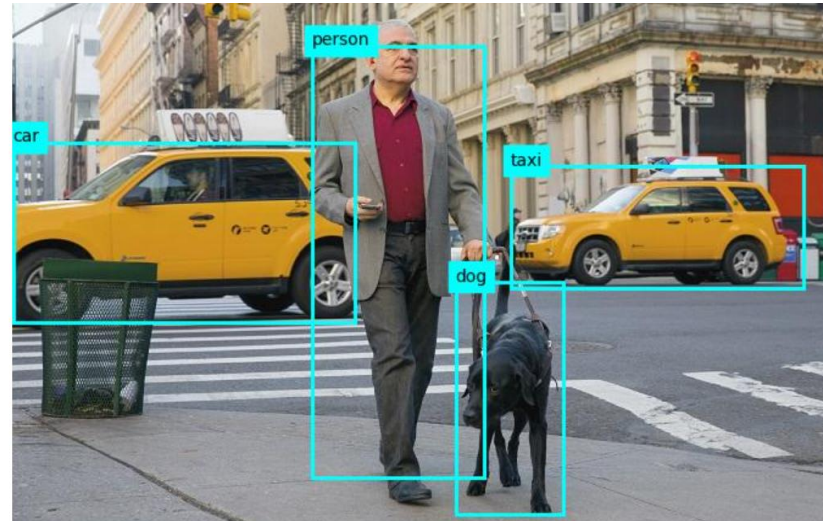
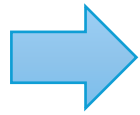
- 画像から説明文を作る、画像にタグ付けする（画像の特徴を単語化する）、画像に含まれるオブジェクトや人物を検出する、といった処理を行う



a man walking a dog on a leash on a street
(路上でリードに繋いだ犬を散歩させている男性)



outdoor, building, street, person, dog



使用するサービス・機能

- Azure AI Vision > Computer Vision > Image Analysis
- Azure portalで「Computer Vision」リソースを作成する
- 「Computer Vision」のImage Analysis APIを呼び出す

ラボ1 解説

- (別紙)

まとめ

- Azure AI Visionの「Computer Vision」の「Image Analysis」（画像分析）を使用して、画像の分析を行いました
- 画像から説明文を作る
- 画像にタグ付けする（画像の特徴を単語化する）
- 画像に含まれるオブジェクトや人物を検出する

モジュール2 画像内のテキストを読む

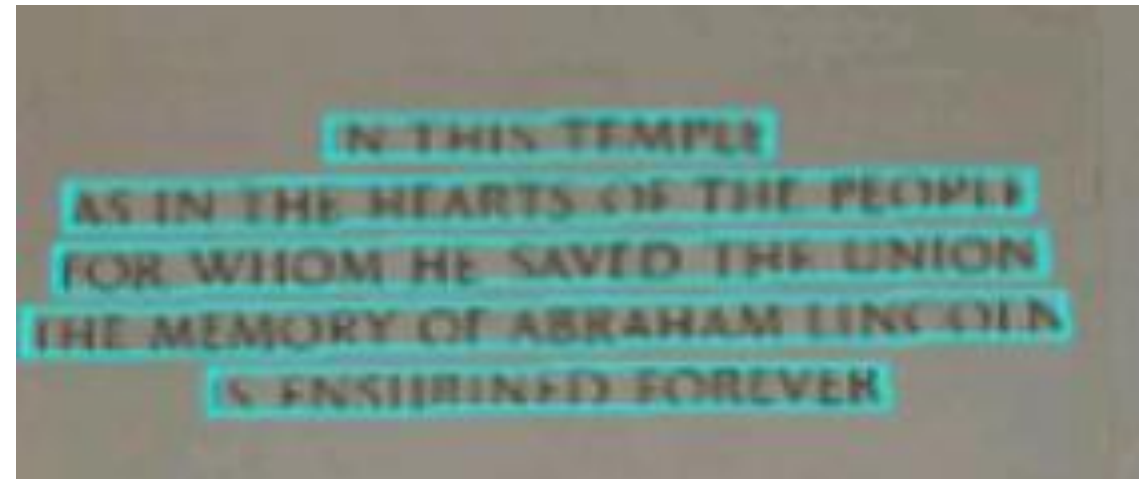
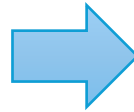
- OCRとは？
- 使用するサービス・機能
- ラボ2
- まとめ

OCRとは？

- Optical Character Recognition（光学式文字認識）
- 画像に含まれている文字情報を取り出す
- 日本語を含む多数の言語に対応



"IN THIS TEMPLE . . ."



使用するサービス・機能

- Azure AI Visionの「Computer Vision」に含まれる「OCR」（光学式文字認識）を使用
- Azure portalで、Computer Vision リソースを作成
 - これはラボ1で使ったものと同じです
- Computer Vision の OCR APIを呼び出す

ラボ2 解説

- (別紙)

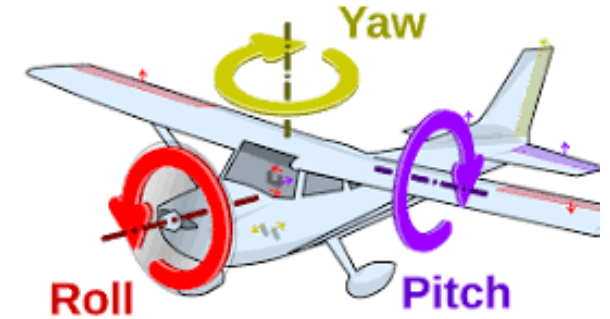
まとめ

- Azure AI Visionの「Computer Vision」の「OCR」（光学式文字認識）を使用して、画像から文字情報を抽出しました
- 画像に含まれる文字列を抽出できる
- 行単位や単語単位での位置情報も得られる

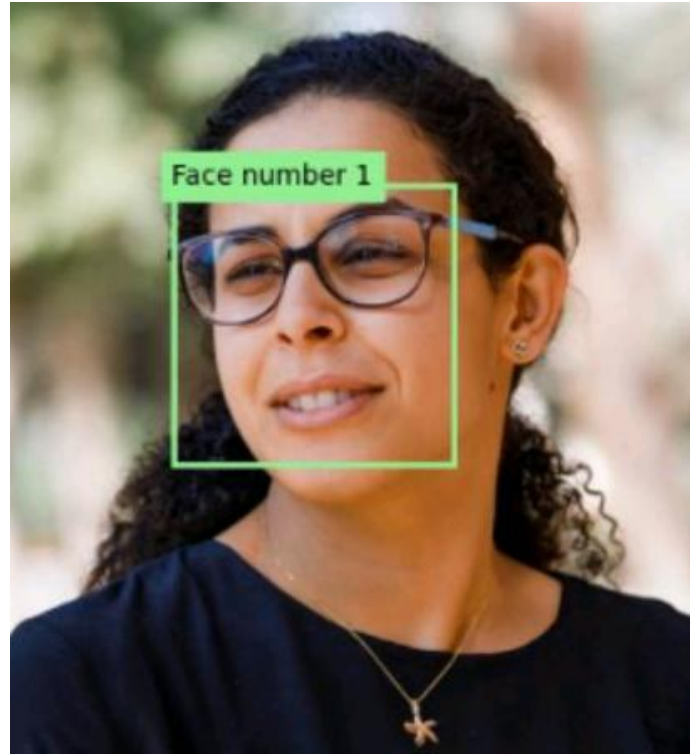
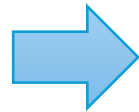
モジュール3 顔の検出と分析

- 「顔の検出と分析」とは？
- 使用するサービス・機能
- ラボ3
- まとめ

「顔の検出と分析」とは？

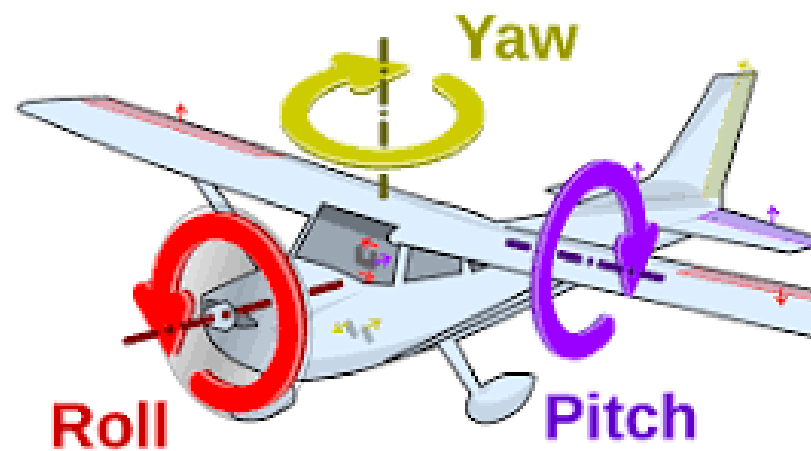


- 「Azure AI Face サービス」
- 画像に含まれている人の顔の検出を行い、位置の四角形の座標、顔の向きなどの情報を返す



```
Head Pose (Yaw): -20.9
Head Pose (Pitch): 3.7
Head Pose (Roll): -7.7
Forehead occluded?: False
Eye occluded?: False
Mouth occluded?: False
Accessories:
- AccessoryType.GLASSES
```


Head Pose (Yaw): -20.9
Head Pose (Pitch): 3.7
Head Pose (Roll): -7.7
Forehead occluded?: False
Eye occluded?: False
Mouth occluded?: False
Accessories:
- AccessoryType.GLASSES

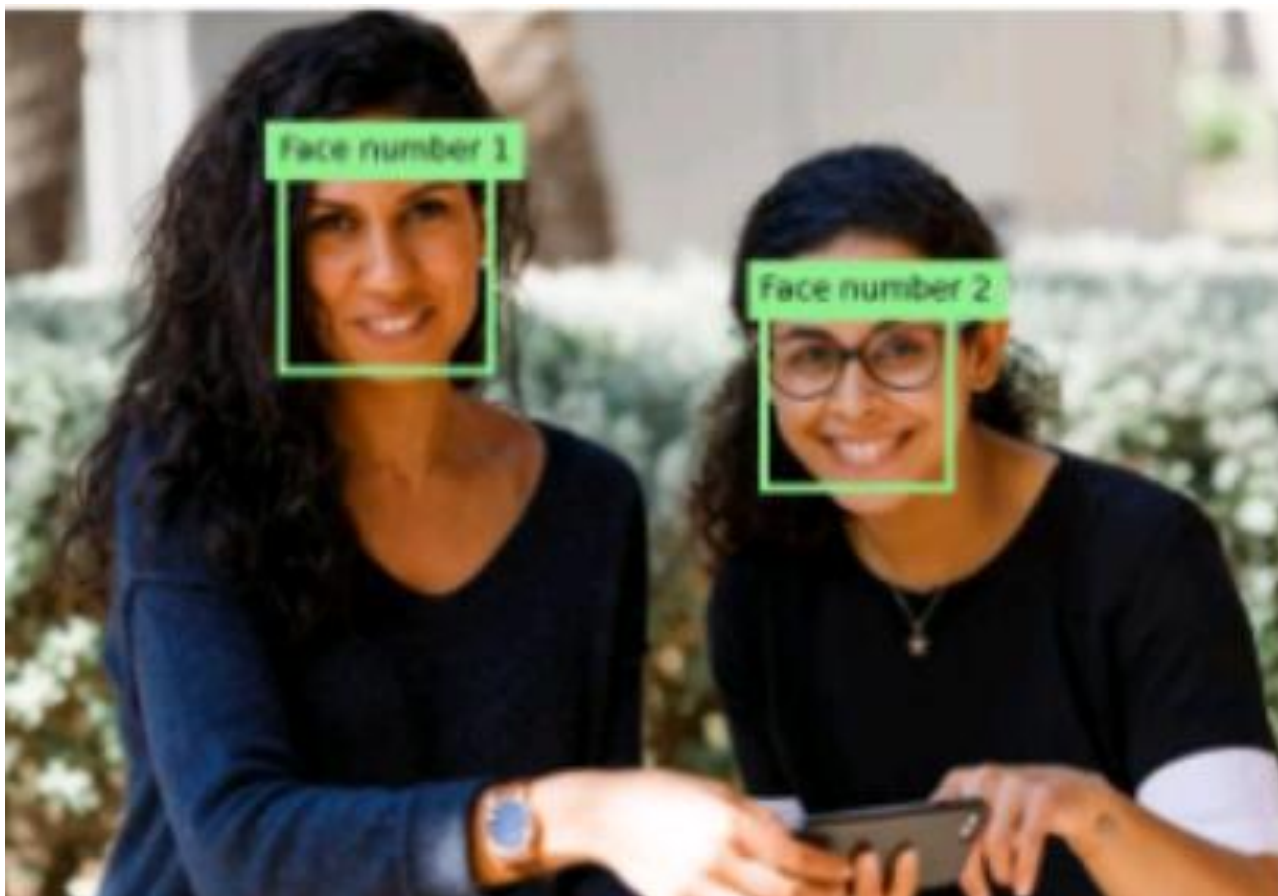


[ローリング - Wikipedia](#)

forehead occluded = 額が覆われている
eye occluded = 目が覆われている
mouth occluded = 口が覆われている

※帽子、サングラス、マスクなどで目などの部位が隠されているかどうか、ということ

複数の顔を同時に処理できる



2 faces detected.

Face number 1

- Head Pose (Yaw): -5.2
- Head Pose (Pitch): -10.0
- Head Pose (Roll): -8.1
- Forehead occluded?: False
- Eye occluded?: False
- Mouth occluded?: False
- Accessories:

Annotating faces in image...

Results saved in detected_faces.jpg

Face number 2

- Head Pose (Yaw): -8.3
- Head Pose (Pitch): -9.2
- Head Pose (Roll): -6.3
- Forehead occluded?: False
- Eye occluded?: False
- Mouth occluded?: False
- Accessories:
 - AccessoryType.GLASSES

使用するサービス・機能

- Azure AI Visionの「Face Service」に含まれる「Face Detection」(顔検出)を使用
- Azure portalで、Face Serviceリソースを作成
 - ※Azure portal上では「Face API」とも表示される
- Face Service の Face Detection APIを呼び出す

ラボ3 解説

- (別紙)

まとめ

- Azure AI Visionの「Face Service」に含まれる「Face Detection」(顔検出) を使用して、画像から顔の情報を取得できる
- 位置の四角形の座標、顔の向き (ローリング) などの情報を返す
- 画像に複数の顔が含まれている場合はすべての顔の情報を返す

モジュール4 画像を分類する

- 「画像の分類」とは？
- 使用するサービス・機能
- ラボ4

「画像の分類」とは？

- Azure AI Visionの「Custom Vision」を使用
- あらかじめ多数の画像を準備してモデルをトレーニングする
- トレーニング済みのモデルに新しい画像を与えると、画像の分類が行われる

ラボ4 解説

- (別紙)

モジュール5ビジョン対応のチャットアプリを開発する

- 「Azure OpenAI Service」の「GPT-4o」モデルを使用する
- このモデルはテキストだけではなく画像や音声の入出力に対応している
 - 「マルチモーダル」モデル、と呼ばれる
- ラボ5: 画像を入力し、それについて質問ができるチャットアプリを開発する

モジュール6 AIで画像を生成する

- 「Azure OpenAI Service」の「DALL-E」（ダリ）モデルを使用する
- このモデルはプロンプト（テキスト）に沿って画像を生成できる
- ラボ6 Azure OpenAI Service > DALL-Eモデル