

画像分析とは

[アーティクル] • 2024/09/03

Azure AI Vision Image Analysis サービスでは、画像からさまざまな視覚的特徴を抽出できます。たとえば、成人向けコンテンツが画像に含まれているかどうかを判断したり、特定のブランドや物体、人の顔を検出したりすることができます。

現在一般提供中の最新バージョンの Image Analysis 4.0 には、同期 OCR や人物検出などの新機能があります。今後はこのバージョンを使うことをお勧めします。

画像分析は、クライアント ライブラリ SDK を通じて、または [REST API](#) を直接呼び出すことで使用できます。使用を開始するには、[クイックスタート](#)に従ってください。

クイックスタート

または、Vision Studio を使用して、ブラウザーですばやく簡単に画像分析の機能を試すことができます。

Vision Studio を試す

このドキュメントには、次のような記事が記載されています。

- [クイックスタート](#)は、サービスの呼び出しと結果の取得を短時間で行えるようにする、ステップバイステップの手順です。
- [攻略ガイド](#)には、より具体的またはカスタマイズした方法でサービスを使用するための手順が記載されています。
- [概念の記事](#)では、サービスの機能と特長について詳しく説明します。
- [チュートリアル](#)はより長文のガイドであり、より広範なビジネス ソリューションの 1 コンポーネントとしてこのサービスを使用する方法を示すものです。

より構造化されたアプローチについては、画像分析のトレーニング モジュールに従ってください。

- [Azure AI Vision サービスを使用して画像を分析する](#)

Image Analysis のバージョン

📌 重要

要件に最も適した Image Analysis API バージョンを選択してください。

 テーブルを展開する

バージョン	使用できる機能	推奨
バージョン 4.0	テキストの読み取り、キャプション、高密度キャプション、タグ、物体検出、カスタム画像分類/物体検出、人物、スマートトリミング	より優れたモデル。バージョン 4.0 が目的のユース ケースをサポートしている場合は、バージョン 4.0 を使用してください。
バージョン 3.2	タグ、物体、説明、ブランド、顔、画像の種類、配色、ランドマーク、有名人、成人向けコンテンツ、スマートトリミング	幅広い機能。バージョン 4.0 が目的のユース ケースをまだサポートしていない場合は、バージョン 3.2 を使用してください

Image Analysis 4.0 API が目的のユース ケースをサポートしている場合、4.0 API を使用することをお勧めします。4.0 が目的のユース ケースをまだサポートしていない場合、バージョン 3.2 を使用してください。

また、画像キャプションを実行する場合で、Vision リソースがサポートされている Azure リージョン外にある場合は、バージョン 3.2 を使用する必要があります。画像分析 4.0 の画像キャプション機能は、特定の Azure リージョンでのみサポートされています。バージョン 3.2 の画像キャプションは、すべての Azure AI Vision リージョンで使用できます。[利用可能なリージョン](#)に関するページを参照してください。

Analyze Image (画像を分析する)

画像を分析し、その視覚的特徴や性質に関する分析情報を提示できます。この表内の機能はすべて、Analyze Image API によって提供されます。[クイックスタート](#)に従って始めてください。

 テーブルを展開する

名前	説明	概念ページ
モデルのカスタマイズ (v4.0 プレビューのみ)	画像分類または物体検出を行うためのカスタム モデルを作成してトレーニングできます。任意の画像を取り込み、カスタム タグでラベルを付けると、Image Analysis は目的のユース ケース用にカスタマイズされたモデルをトレーニングします。	モデルのカスタマイズ

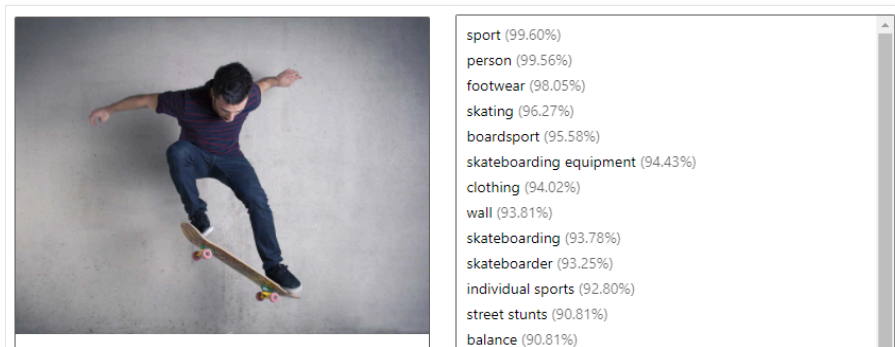
名前	説明	概念ページ
画像からテキストを読み取る (v4.0 のみ)	Image Analysis のバージョン 4.0 プレビューでは、画像から読み取り可能なテキストを抽出できます。非同期の Computer Vision 3.2 Read API と比較すると、新しいバージョンでは、パフォーマンスが向上した統合同期 API で使い慣れた Read OCR エンジンが提供されます。これにより、1 回の API 呼び出しで OCR とその他の分析情報を簡単に取得できます。	画像の OCR
画像内の人物を検出する (v4.0 のみ)	Image Analysis のバージョン 4.0 では、画像に表示された人物を検出できます。検出された各人物の境界ボックスの座標が、信頼度スコアと共に返されます。	人物検出
画像キャプションを生成する	<p>完全な文を使用して、人間が判読できる言語で画像のキャプションを生成します。Computer Vision のアルゴリズムにより、画像内で識別された物体に基づいてキャプションが生成されます。</p> <p>バージョン 4.0 の画像キャプション モデルは、高度な実装であり、広い範囲の入力画像を処理します。特定の地理的リージョンでのみ使用可能です。 利用可能なリージョンに関するページを参照してください。</p> <p>バージョン 4.0 では、画像内にある個々の物体の詳細なキャプションを生成する高密度キャプションも使用できます。API により、画像内で見つかった各物体の境界ボックスの座標 (ピクセル単位) とキャプションが返されます。この機能を使用して、画像の個別部分の説明を生成できます。</p> <div>  <div> Cows grazing in a field </div> </div>	画像キャプションを生成する (v3.2) (v4.0)
物体を検出する	オブジェクトの検出はタグ付けに似ていますが、API で返されるのは、各タグが適用された境界ボックスの座標です。たとえば犬や猫、人物が画像に含まれている場合、検出操作によって、それらのオブジェクトが、画像における対応する座標と共に一覧表示されます。この機能を使用して、画像内のオブジェクト間のリレーションシップをさらに処理できます。画像内に同じタグの複数のインスタンスが存在する場合はそれを知ることができます。	物体を検出する (v3.2) (v4.0)



視覚的特徴のタグ付け

数千個の認識可能なオブジェクト、生物、風景、および動作のセットから、画像内の視覚的な特徴を識別してタグ付けします。タグが不明確な場合や、常識的でない場合は、API 応答により、タグのコンテキストを明確にするためのヒントが示されます。タグ付けの対象は、前景の人物などの被写体に限らず、背景 (屋内または屋外)、家具、道具、植物、動物、アクセサリ、ガジェットなども含まれます。

視覚的特徴のタグ付け
(v3.2)
(v4.0)



関心領域/スマートクロップを取得する

画像の内容を分析し、指定された縦横比に一致する "関心領域" の座標を返します。Computer Vision では領域の境界ボックスの座標が返されるため、呼び出し元のアプリケーションで必要に応じて元の画像を変更できます。

バージョン 4.0 のスマート トリミング モデルは、高度な実装であり、広い範囲の入力画像を処理します。特定の地理的リージョンでのみ使用可能です。 [利用可能なリージョン](#)に関するページを参照してください。

サムネイルを生成する
(v3.2)
(v4.0 プレビュー)

ブランドの検出 (v3.2 のみ)

数千点ものグローバル ログのデータベースから、画像または動画に含まれる商業ブランドを識別します。この機能は、たとえば、ソーシャル メディアで最も人気のあるブランドや、メディアのプロダクトプレイスメントの中で最も普及しているブランドを検出する目的で使用できます。

ブランドを検出する

画像の分類 (v3.2 のみ)

親/子で引き継がれる階層を備えた [カテゴリの分類](#) を使用して、イメージ全体を識別してタグ付けします。カテゴリは単独で、または新しいタグ付けモデルと共に使用できます。

イメージの分類

名前	説明	概念ページ
	現時点では、イメージのタグ付けと分類でサポートされている言語は、英語のみです。	
顔の検出 (v3.2 のみ)	<p>イメージ内の人物の顔を検出して、検出されたそれぞれの顔に関する情報を提示します。 Azure AI Vision は検出された各顔の座標、四角い枠、性別、および年齢を返します。</p> <p>これらの目的で専用の Face API を使用することもできます。 顔識別や姿勢検出など、より詳細な分析に使用できます。</p>	顔を検出する
画像の種類の検出 (v3.2 のみ)	イメージが線による描画かクリップ アートのようにになっているかなど、イメージの性質を検出します。	イメージの種類の検出
ドメイン固有のコンテンツの検出 (v3.2 のみ)	ドメイン モデルを使用して、有名人やランドマークなど、イメージ内のドメイン固有のコンテンツを検出して識別します。 たとえば、画像に人物が含まれている場合、Azure AI Vision では、有名人用のドメイン モデルを使用して、画像内で検出された人物が既知の有名人と一致するかどうかを判断できます。	ドメイン固有のコンテンツの検出
配色の検出 (v3.2 のみ)	イメージ内にある色の使用状況を分析します。 Azure AI Vision では、画像が白黒かカラーかを特定し、カラー画像の場合は、主要な色やアクセントになる色を識別することができます。	配色の検出
画像内のコンテンツの調整 (v3.2 のみ)	Azure AI Vision を使用すると、画像内の成人向けコンテンツを検出し、さまざまな分類の信頼度スコアを返すことができます。 コンテンツをフラグ設定するためのしきい値は、自分の都合に合わせて、スライディング スケールで設定することができます。	成人向けコンテンツを検出する

💡 ヒント

画像分析のテキスト読み取りおよび物体検出機能は、[Azure OpenAI](#) サービスを通じて利用できます。 **GPT-4 Turbo with Vision** モデルを使用すると、共有する画像を分析できる AI アシスタントとチャットできます。また、Vision Enhancement オプションでは、画像分析を使用して画像に関する詳細 (読み取り可能なテキストとオブジェクトの位置) を AI アシスタントに提供します。 詳細については、[GPT-4 Turbo with Vision のクイックスタート](#)に関する記事を参照してください。

Product Recognition (v4.0 プレビューのみ)

Product Recognition API を使用すると、小売店の棚の写真を分析できます。製品の有無を検出し、境界ボックス座標を取得できます。モデルのカスタマイズと組み合わせで使用して、特定の製品を識別するようにモデルをトレーニングします。Product Recognition の結果を店舗のプラノグラム ドキュメントと比較することもできます。

Product Recognition

マルチモーダル埋め込み (v4.0 のみ)

マルチモーダル埋め込み API を使うと、画像とテキスト クエリの "ベクトル化" が可能になります。それらは、多次元ベクトル空間内の座標に画像を変換します。その後は、受信したテキスト クエリをベクトルに変換することもでき、セマンティックの近さに基づいて画像をテキストと照合できます。これにより、ユーザーはテキストを使って一連の画像を検索できます。画像タグや他のメタデータを使う必要はありません。セマンティックの近さにより、多くの場合、検索でより良い結果が得られます。

2024-02-01 API には、102 個の言語でのテキスト検索をサポートする多言語モデルが含まれています。元の英語のみのモデルは引き続き使用できますが、同じ検索インデックス内で新しいモデルと組み合わせることはできません。英語のみのモデルを使用してテキストと画像をベクター化した場合、これらのベクターは多言語テキストと画像のベクターと互換性がありません。

これらの API は、特定の地理的リージョンでのみ使用できます。 [利用可能なリージョン](#)に関するページを参照してください。

マルチモーダル埋め込み

背景の削除 (v4.0 プレビューのみ)

Image Analysis 4.0 (プレビュー) では、画像の背景を削除できます。この機能では、背景を透明にして検出された前景オブジェクトの画像を出力すること、または検出された前景オブジェクトの不透明度を示すグレースケール アルファ マット画像を出力することができます。

背景の削除

 テーブルを展開する

元の画像	背景が削除された状態	アルファ マット
		

サービスの制限

入力の要件

Version 4.0

画像分析は、次の要件を満たす画像で動作します。

- 画像は JPEG、PNG、GIF、BMP、WEBP、ICO、TIFF、または MPO 形式で表示する必要があります
- 画像のファイル サイズは、20 メガバイト (MB) 未満である必要があります
- 画像のディメンションは、50 x 50 ピクセルより大きく、16,000 x 16,000 ピクセル未満である必要があります

💡 ヒント

マルチモーダル埋め込みの入力要件は異なります。そのリストについては、[マルチモーダル埋め込み](#)に関するページを参照してください

言語のサポート

さまざまな言語で、さまざまな画像分析機能を使用できます。[言語のサポート](#)に関するページを参照してください。

利用可能なリージョン

画像分析 API を使用するには、サポートされているリージョンで Azure AI Vision リソースを作成する必要があります。画像分析機能は次のリージョンで使用できます。

[🔗 テーブルを展開する](#)

リージョン	Analyze Image (画像を分析する) (4.0 キャプションを除く)	Analyze Image (画像を分析する) (4.0 キャプションを含む)	Product Recognition	マルチモーダル埋め込み	背景の削除
米国東部	✓	✓	✓	✓	✓
米国西部	✓	✓		✓	✓
米国西部 2	✓		✓	✓	
フランス中部	✓	✓		✓	✓
北ヨーロッパ	✓	✓		✓	✓
西ヨーロッパ	✓	✓		✓	✓
スウェーデン中部	✓			✓	
スイス北部	✓			✓	
オーストラリア東部	✓			✓	
東南アジア	✓	✓		✓	✓
東アジア	✓	✓			
韓国中部	✓	✓		✓	✓
東日本	✓			✓	

データのプライバシーとセキュリティ

Azure AI サービス全般に言えることですが、Azure AI Vision サービスを使用する開発者は、顧客データに関する Microsoft のポリシーに留意する必要があります。詳細については、Microsoft Trust Center の [Azure AI サービス ページ](#) を参照してください。

次のステップ

好みの開発言語と API バージョンのクイックスタート ガイドに従って、画像分析の使用を開始します。

- [クイックスタート \(v4.0\): Vision REST API またはクライアント ライブラリ](#)
- [クイック スタート \(v3.2\): Vision REST API またはクライアント ライブラリ](#)

フィードバック

このページはお役に立ちましたか?

 Yes

 いいえ

[製品フィードバックの提供](#)  | [Microsoft Q&A でヘルプを表示する](#)