Vision API

Analyze an Image

**Current**

86 categories

* Thumbnail
* OCR
* Adult

**Build//**

**Objects:** keep **86 categories divided into 15 top concepts each with several sub-concepts.**

Visual taxonomy feature used to categorize content in an image based on a list of 86 concepts, ranging from broad to specific. For the full taxonomy in text format, please go [here](https://www.projectoxford.ai/images/bright/vision/examples/86categories.txt).

Any relationship between 86 categories and new tags, hints and/or descriptions?

**Category:** People (person is not a category currently)

**2000 Tags:** Example: person, man, indoor, Satya Nadella (recognizes celebrities), monitor, zinnia (in pot) {hint: flower}

**Hint**: (instead of taxonomy) for example {Cello} Hint: musical instrument or {Zinnia} Hint: flower

**Description:** “...”

**Confidence level (0.0 -1.0)**

**Image example:** Satya Nadella giving a talk

**Description:** “Satya Nadella on a stage giving a speech.”

Plus, look at Chris Thrasher doc and my own rewrite of “Analyze and Image”

Need to redo the photos and their description

Overview

By uploading an image or specifying an image URL, Microsoft Computer Vision algorithms can analyze visual content in different ways based on inputs and user choices. As something new with the release of version 1.0, Vision API offers image “tagging” based on more than 2000 recognizable subjects. Where tagging may create ambiguity or eccentricity, “hints” serve to clarify the context of the object, animal, person or scene. A collection of content tags forms the foundation for an image “description” displayed in human readable language; this in turn may be used to create auto-generated “captions”. Indexing your photos has never been easier.

## Analyze an image through Vision API’s new tagging and caption generator

## Summary

After uploading an image or specifying an image URL, Vision API’s algorithms output a number of tags based on the objects identified in the image. Tagging is not limited to the main subject, such as a person in the foreground, but also includes the setting (indoor or outdoor), furniture, tools, plants, animals, accessories, gadgets etc. Based on these tags the algorithm will generate a caption of what it “locates” in the image.

## Two Options

Caption generation is provided through two user interfaces.

1. Option One

**Description:** An addition to the analyze operation returns one image caption per picture.

1. Option Two

**Evaluation:** The evaluation operation allows for returning n-best results for n!=1. In other words, when no definitive caption can be obtained, a user can select to get a number of captions back with a confidence score for each.

You can still use the 86-category classification algorithm either in combination with or instead of the new features. Tagging and caption generation is separate from the 86-category classification system.

## Custom Models

A method for providing users with new custom image-understanding models, for example of dogs, plants, skateboarding etc.

1. Option One

**Scoped Analysis:** Analyze only a given model by invoking an HTTP POST call as follows.

POST /vision/v1.0/models/{model}/analyze

For this option, no other query parameter {visualFeatures, details} is valid. The output will be an array of tags.

Example:

|  |
| --- |
| { "details": [   { "name": "golden retriever", "score": 0.98},  { "name": "labrador retriever", "score": 0.78}]} |

Additionally, a method for listing models will now be required.

GET /vision/v1.0/models

Return type will be a new JSON type:

|  |
| --- |
| { "models": [  { "name": "celebrities", "categories": [ "people\_" ] },  { "name": "dogs", "categories": [ "animal\_dog" ] },  { "name": "plants", "categories": [ "plant\_" ] } ]} |

The categories field is a list of one or more categories from the 86-category taxonomy. Note also that categories ending in an underscore will match that category and its children (for example, people\_ as well as people\_group, for th ecelebrity model). The current taxonomy can be found [here](https://www.projectoxford.ai/images/bright/vision/examples/86categories.txt).

1. Option Two

### **Enhanced Analysis:** Analyze to provide additional details related to categories.

For applications where users want to get generic image analysis *in addition* to details from one or more domain-specific models as mentioned above. A new query parameter to be used with the models looks as follows.

POST /vision/v1.0/analyze?details=[all|{model1},[{model2},…]]

When this method is invoked, the 86-category taxonomy classifier is called first. If any of the categories match that of known/matching models, a second pass of classifier invocations will follow. For example, if details=all, or details include 'dogs', we will call the dog classifier *after* the 86-category classifier is called *and* the result includes object\_animal\_dog. This will increase latency for users interested in dogs, but these users will always have the option of directly invoking the dog classifier as shown in option 1.

All v1 query parameters will behave the same in this case. If visualFeatures=categories is *not* specified, it will be implicitly enabled.