Using Cognitive Services

Eng Teong Cheah
MVP Visual Studio &
Development Technologies

Agenda

Cognitive Services overview
Processing language
Processing image and video

Cognitive Services overview

What is cognitive service?

Microsoft Cognitive Services (formerly Project Oxford) are set of APIs, SDKs and services available to developers to make their applications more intelligent, engaging and discoverable.

Microsoft Cognitive Services expands o Microsoft's evolving portfolio of machine learning APIs and enables developers to easily add intelligent features –such as emotion and video detection; facial, speech and vision recognition; and speech and language understanding – into their applications.

What is cognitive service?

Our vision is for more computing experiences and enhanced productivity aided by systems that increasingly can see, hear, speak, understand and even begin to reason.

Processing language

Bing Spell Check API

The Bing Spell Check API lets you perform contextual grammar and spell checking.

Bing has developed a web-based spell-checker that leverages machine learning and statistical machine translation to dynamically train a constantly evolving and highly contextual algorithm.

Bing Spell Check API

The spell-checker is based on a massive corpus of web searches and documents.

Linguistic Analysis API

These APIs provide access to natural language processing (NLP) tools that identify the structure text.

The current release provides 3 types of analysis:

- Sentence separation and tokenization
- Part-of-speech tagging
- Constituency parsing

Text Analysis

Text Analytics API is a cloud-based service that provides advanced natural language processing over raw text, and includes 3 main functions:

- Sentiment analysis
- Key phrase extraction
- Language detection

Translator

Microsoft Translator Text API is a cloud-based machine translation service.

With this API you can translate text in near real-time from any app or service through a simple REST API call.

The API uses the most modern neural machine translation technology, as well as offering statical machine translation technology.

Web Language Model API

A REST-based cloud service providing state-of-the-art tools for natural language processing.

Using this API, your application can leverage the power of big data through language models trained on web-scale corpora collected by Bing in the EN-US market.

Web Language Model API

The Web LM REST API support 4 lookup operations:

- Joint (log 10) probability of a sequence of words.
- Conditional (log 10) probability of one word given a sequence of preceding words.
- List of words (completions) most likely to follow a given sequence of words.
- Word breaking of strings that contain no spaces.

Processing image & video

Face API

A cloud-based service that provides the most advanced face algorithms.

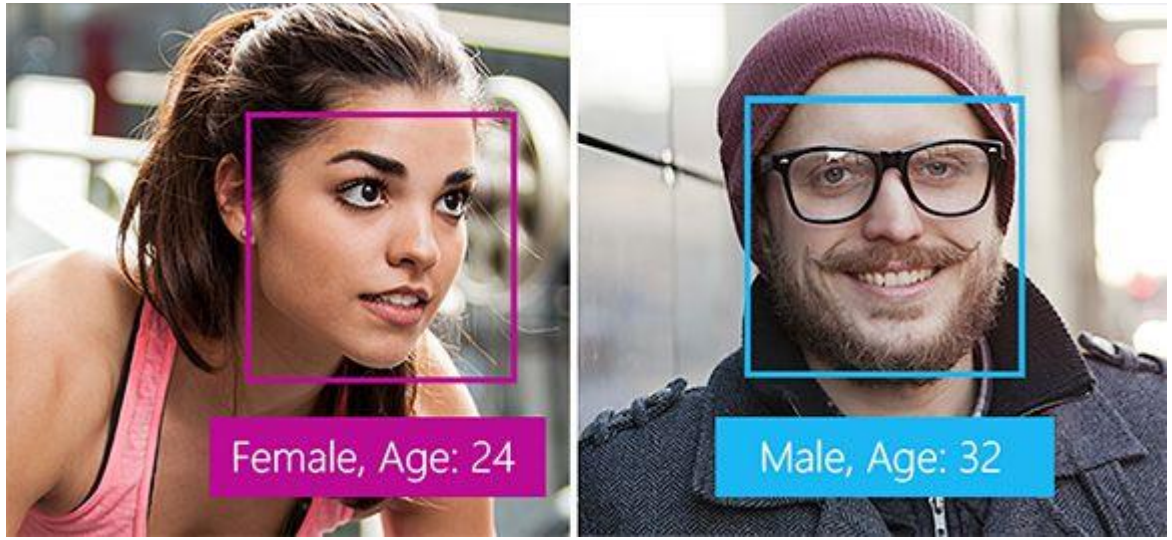
Face API has 2 main functions:

- Face detection with attributes
- Face recognition

Face Detection

Face API detects up to 64 human faces with high precision face location in an image. And the image can be specified by file in bytes or valid URL.

Face Detection



Face rectangle (left, top, width, and height) indicating the face location in the image is returned along with each detected face.

Optionally, face detection extracts a series of face-related attributes such as pose, gender, age, head pose, facial hair and glasses.

Face Verification

Face API verification performs an authentication against 2 detected faces or authentication from one detected face to one person project.

Finding Similar Face

Given a target detected face and a set of candidate faces to search with, the service finds a small set of faces that look most similar to the target face.

Two working modes, **matchFace** and **macthPerson** are supported.

Face Grouping

Given one set of unknown faces, face grouping API automatically divides them into several groups based on similarity.

Each group is a disjointed proper subset of the original unknown face set, and contains similar faces.

And all the faces in the same group can be considered to belong the same person object.

Emotion API

The Emotion API beta takes an image as an input, and returns the confidence across a set of emotions for each face in the image, as well as bounding box for the face, from the Face API.

The emotions detected are happiness, sadness, surprise, anger, fear, contempt, disgust or neutral.

These emotions are communicated cross-culturally and universally via the same basic facial expressions, where are identified by Emotion API.

Emotion API

Interpreting Results:

In interpreting results from the Emotion API, the emotion detected should be interpreted as the emotion with the highest score, as scores are normalized to sum to one.

Users may choose to set a higher confidence threshold within their application, depending on their needs.

What is Content Moderator?

Content moderation is the process of monitoring user-generated content on online and social media websites, chat and messaging platforms, enterprise environments, gaming platforms, and peer communication platforms.

The goal is to track, flag, assess, and filter out offensive and unwanted content that creates risk for your organization.

Moderated content might include text, images, and videos.

Computer Vision

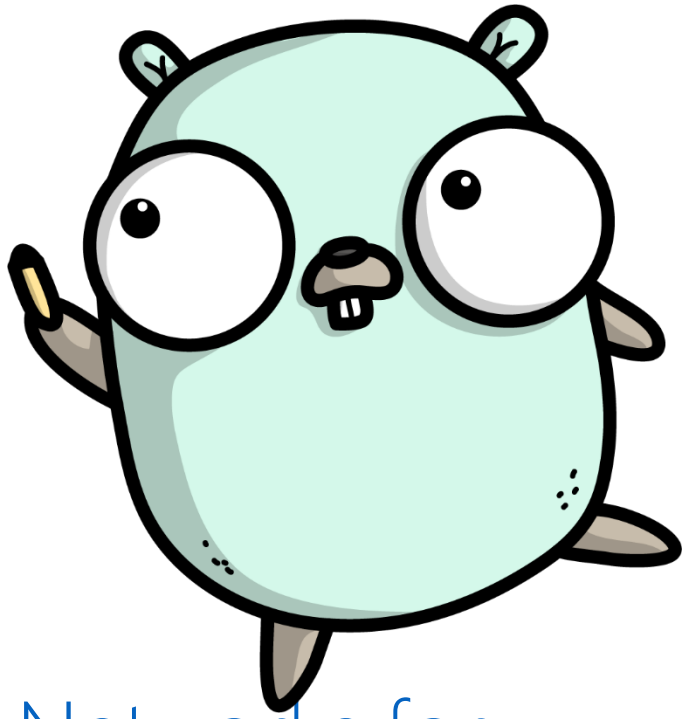
The cloud-based Computer Vision API provides developers with access to advanced algorithms for processing images and returning information.

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.

Demo

Build a face detection application

Resources



TutorialsPoint

Microsoft Docs

Lecture Collection | Convolutional Neural Networks for Visual Recognition(Spring 2017)

Python Numpy Tutorial

Image Credits: @ashleymcnamara

Thank you

Eng Teong Cheah
Microsoft MVP Visual Studio & Development Technologies
Twitter: @walkercet
Github: https://github.com/ceteongvanness
Blog: https://ceteongvanness.wordpress.com/
Youtube: http://bit.ly/etyoutubechannel