Final Project

Computer Vision API

Beri, Shailesh

Deep Azure@McKesson

Dr. Zoran B. Djordjević

Introduction

- Finding actionable information from images, generating captions, identifying objects in images and interpreting text in images has been a important breakthrough in Cognitive services
- Azure 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
- Various Applications in this space range from 1) Identification of faces or celebrities in images, 2) Blocking images with adult content, 3) Text In Image recognition, 4) Optical character recognition (OCR) and several hundreds more

Problem Statement

- Need a Computer Vision Reusable Feature which performs Computer Image Analysis, Identifies objects in the image and Describes these objects in a single function call

Description of Data and URL of Data Source

- Data used for the project are Image URL(s) from Internet
- To support the key objective of qualitative analysis of the Computer Vision API a variety of images were obtained from the Internet and were tested to verify results
- Sample URL(s)
  - https://ausopen.com/sites/default/files/201801/28/o_federer_f_rla_28012018_42.jpg
  - https://ausopen.com/sites/default/files/201801/28/o_federer_f_rla_28012018_35.jpg
  - https://cdn.cnn.com/cnnnext/dam/assets/180202172405-01-week-in-politics-0204-restricted-super-169.jpg
  - https://oxfordportal.blob.core.windows.net/vision/Analysis/3.jpg
  - https://c1.staticflickr.com/7/6013/5918998899_3051a519f6_b.jpg
  - https://thumbs.gfycat.com/MeanIlliterateBlackbear-max-1mb.gif

Description of Hardware

- Local
  - Windows 10 Home; Dell Laptop
  - Windows 7 Professional; Dell Laptop
- Azure Cloud Services
  - Computer Vision API

Description of Software

- Development Tools
  - Programming Language – Python 3.6
  - Platform – Anaconda – Jupyter
- **API(s) used for project**
  - Image Analysis (URL) 'https://{}.api.cognitive.microsoft.com/vision/v1.0/analyze'.format(_region)
  - Image Description (URL) 'https://{}.api.cognitive.microsoft.com/vision/v1.0/describe?%s'.format(_region)
- Python Packages
- Image Analysis – Numpy, Requests, CV2,
- Image Description Packages - http.client, urllib.request, urllib.parse, urllib.error, base64
- Image Rendering Packages –Matplotlib.pyplot
- Other Packages –Time, Operator

Installation and Configuration Steps

- Create new resource group in “West US” region
- Sign-Up for Computer Vision API in Azure Portal under the resource group
  - Select Computer Vision API under AI + Cognitive Services
- Select correct pricing tier
- Capture Subscription keys
- Install opencv-python using PIP in cmd
- Call API(s) by running Python program
- Check Metrics in Azure Portal for the API and check calls made vs. successful calls

Example Image Data File

https://photos.smugmug.com/02Sports-2/Warriors/Warriors-stun-Thunder-in-Game-/i-4DDRNhm/0/a0e14f53/L/SJM-WARRIORS-0529-148-L.jpg


```
Processing Azure Cognitive Services Computer Vision with  SDK for Python
Enter URL of Image : https://photos.smugmug.com/02Sports-2/Warriors/Warriors-stun-Thunder-in-Game-/i-4DDRNhm/0/a0e14f
53/L/SJM-WARRIORS-0529-148-L.jpg
```

Visualization of Results

```
Image Analysis

{'categories': [{'name': 'people_crowd', 'score': 0.546875}], 'color': {'dominantColorForeground': 'Black', 'dominant
ColorBackground': 'Black', 'dominantColors': ['Black', 'Brown'], 'accentColor': '9F642C', 'isBwImg': False}, 'request
Id': 'd73383a5-7d4e-462c-bf0d-c5ad92fb7031', 'metadata': {'height': 534, 'width': 800, 'format': 'Jpeg'}}


Image Description

{'description': {'tags': ['person', 'sport', 'game', 'basketball', 'court', 'player', 'playing', 'ball', 'man', 'grou
p', 'holding', 'standing', 'female', 'young', 'people', 'walking', 'wearing', 'woman', 'crowd', 'soccer', 'street'],
'captions': [{'text': 'a group of men playing a game of basketball', 'confidence': 0.840567729389681}, {'text': 'a gr
oup of people playing a game of basketball', 'confidence': 0.839567729389681}, {'text': 'a group of people playing ba
sketball on a court', 'confidence': 0.838567729389681}, {'text': 'a group of young men playing a game of basketball',
'confidence': 0.670308292944782}, {'text': 'a group of people playing a basketball game', 'confidence': 0.66930829294
4782}, {'text': 'a group of men playing a basketball game', 'confidence': 0.668308292944782}, {'text': 'a group of me
n playing a game of basketball on a court', 'confidence': 0.667308292944782}, {'text': 'a group of people standing on
a basketball court', 'confidence': 0.666308292944782}, {'text': 'a group of people on a basketball court', 'confidenc
e': 0.665308292944782}, {'text': 'a group of basketball players on the court', 'confidence': 0.664308292944782}]}, 'r
equestId': '60f5cd50-1e85-4026-b65d-434d6c17e211', 'metadata': {'height': 534, 'width': 800, 'format': 'Jpeg'}}
 Image Processing Completed
```

Data Samples (Image URLs)

- https://ausopen.com/sites/default/files/201801/28/o_federer_f_rla_28012018_42.jpg

- https://ausopen.com/sites/default/files/201801/28/o_federer_f_rla_28012018_35.jpg

- https://cdn.cnn.com/cnnnext/dam/assets/180202172405-01-week-in-politics-0204-restricted-super-169.jpg

- https://oxfordportal.blob.core.windows.net/vision/Analysis/3.jpg

Summary

- **Lessons Learned**
  - The Computer Vison API(s) are a powerful tool to get 360 degree view of an image, its characteristics, persons, objects and texts in the image
  - Azure Computer API capability is good and flexible. There is room for improvement by enhancing the API through machine learning as more and more images are analyzed and the image knowledge base grows
- **Pros:**
  - A single Computer Vision API service to be deployed in Azure to access all the functions by making calls to different underlying API(s)
  - API(s) return all information back in JSON format which can be directly stored in database
  - Multiple Image formats are supported (JPEG, GIF, BMP. PNG)
  - Option to perform enhanced image analysis from 86 category taxonomy
- **Cons**
  - Image size is restricted to 4 MB
  - More AI and Machine learning capabilities should be provided

YouTube URLs, GitHub URL, Last Page

- Two minute (short):
  - https://youtu.be/s7AtUIe5kks

- GitHub Repository with all artifacts:
  - https://github.com/shaileshberi/Azure-Computer-Vision