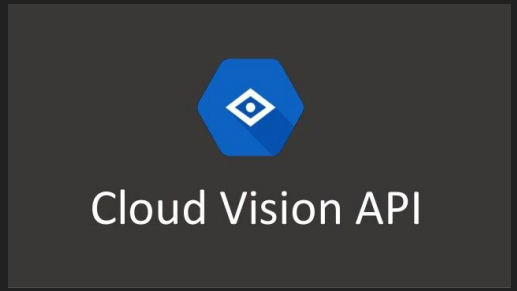
Image Processing via the Cloud

CIS 4010 - Team 3

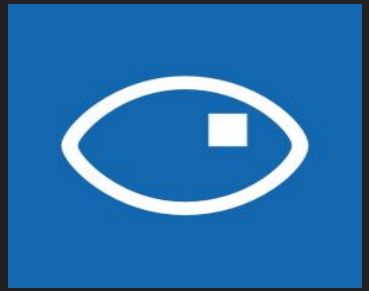
Christian Cornelis (0939357)
Mohammadamin Sheikhtaheri (0930853)
Samuel Tracz (0927451)

Project Statement

- “Compare the accuracy of the labelling systems implemented on 3 major pre-trained image processing APIs - Google’s Vision API, Azure’s Computer Vision API, and Amazon’s Rekognition API.”



https://miro.medium.com/max/638/1*f-9oMvg-zUL7t-Pv6cRV6w.jpeg



https://connectoricons-prod.azureedge.net/cognitiveservicescomputervision/icon_1.0.1308.1899.png

https://miro.medium.com/max/600/1*ozYrdn7kUvdkQ2Eh6PARkg.png

Christian

Contributions (Why should you care?)

- Investigated the limitations of labelling images via cloud AI services
  - Is it ready to be used in real world applications?

- Compared and contrasted the correctness of labelling abilities using images outside of the normal scope of images that are likely to have been used to train the algorithms with

Christian

Contributions (Implementation)

- Used the Python SDKs for each cloud service to create different scripts that would run the image processing algorithms on a set of images
- Compared the ease of which each of these services could be set up and tested
- Compared the response types of each of these services

Cloud Aspects - “tutorial” of cloud services used

- Utilized the following Cloud image processing APIs:
  - Amazon Rekognition
  - Azure Computer Vision
  - Google Vision

- Requirements:
  - Amazon Rekognition: AWS account
  - Azure Computer Vision: Azure account, and connection keys
  - Google Vision: Google account, enabling the API

Mohammadamin

Major Accomplishments (1)

- Found limits that some or all the pretrained algorithms failed to recognize and label correctly
- Discovered that the algorithms have a hard time recognizing objects/actions that look similar
  - Jellyfish vs mushroom
  - Falling vs jumping
- Learned that there is more than one category of image recognition
  - Face detection
  - Text recognition
  - Object recognition
  - Labeling
  - Caption creation

Mohammadamin

Major Accomplishments (2)

- Found that each of these services is extremely-easy to interact with
  - This means that setup is not a barrier to further-testing the limitations of these services

- Discovered that images containing multiple objects are difficult to classify
  - Multiple labels tend to be assigned with high confidence
  - Many of these labels do not pertain to the context of the image
  - Each service tended to display this issue
- Discovered that Rekognition has a clear advantage in labelling with its ‘Parents’ aspect of the response from the API

Samuel

Future Work

- Expand scope - not just labels
  - All three of these services also provide object-detection results when they're used, which would be an interesting area to focus on alongside the labelling capabilities
- Google vision docs (https://cloud.google.com/vision/docs)
- Amazon Rekognition docs (https://docs.aws.amazon.com/rekognition/)
- Azure Computer Vision docs (https://docs.microsoft.com/en-us/azure/cognitive-services/computer-vision/)

Samuel

If we had to do this again… (1-2)

- Spend more time attempting to find images that resulted in false-positives in any of the three image processing services tested in this project.
- This would increase the confidence in our analyses, and perhaps allow us to gain a deeper understanding of these "black box" algorithms that provide little-to-no reasoning for the results that they generate. In doing so, this may also reveal more categories of images that these algorithms do not label properly, or at all.

Samuel

Demo