<https://cloudacademy.com/blog/google-vision-api-image-analysis/>

Title of the Article: Build powerful applications that see and understand the content of images the Google Vision API

Name of Journal, Date of Publication: Build powerful applications that see and understand the content of images the Google Vision API, Jan 28th, 2016

Primary Author(s): Alex Casalboni

Summary of the background:

Google Vision API includes machine learning models and image processing algorithms to make it possible. Google Vision API contains six features.

Summary of the objectives:

The six features include label detection, text detection, face detection, landmark detection, logo detection, and safe search detection. Label Detection executes “Image Content Analysis on entire image and provides relevant labels.” Text Detection performs Optical Character Recognition and provides text. Face Detection can detect face. Landmark detection can detect geographic landmarks while logo detection can detect company logos. These features can all be used once using the API but the speed can be slower.

Summary of the discussion:

Label detection adds semantics to images or video streams by providing labels for any image. These labels that are provided are selected among thousands of categories by Google, which requires image classification, analysis, understanding, and reasoning. The detection is performed by image as a whole. Even though a multiple set of labels can be extracted on a single object, it might lead to lower-quality labels.

Optical Character Recognition requires high-resolution images and very clear algorithm, in which Google’s API returns string and bounding box.

The Face detection literally looks for the face inside an image. In order to capture people’s face, facial landmarks extraction, 3D analysis, skin texture analysis, and many other techniques are required using continuous practices and data.

What did you learn from article? What is most pertinent for your research?

From this article, I learned some special technique that Google Vision API can do, including detecting people’s face when given a picture.

What questions do you have about the content that remain unanswered?