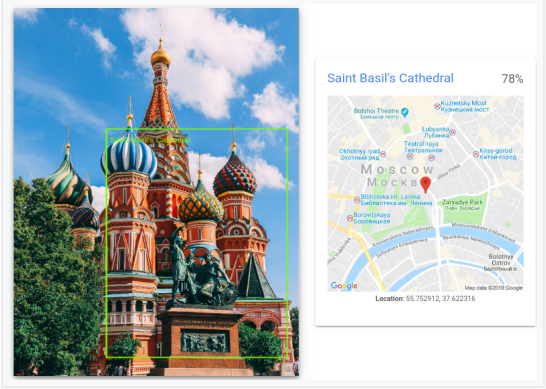
Features list

Cloud Vision API currently allows you to use the following features:

All feature types	
Face detection (https://cloud.google.com/vision/docs/detecting-faces) 1 (#fn1) 	• Locates faces with bounding polygons, and identifies specific facial "landmarks" such as eyes, ears, nose, mouth, etc. along with their corresponding confidence values. • Returns likelihood ratings for emotion (joy, sorrow, anger, surprise) and general image properties (underexposed, blurred, headwear present). • Likelihoods ratings are expressed as 6 different values: `UNKNOWN`, `VERY_UNLIKELY`, `UNLIKELY`, `POSSIBLE`, `LIKELY`, or `VERY_LIKELY`.
Landmark detection (https://cloud.google.com/vision/docs/detecting-landmarks) 2 (#fn2) 	• Provides the name of the landmark, a confidence score and a bounding box in the image for the landmark. • Gives coordinates for the detected entity.

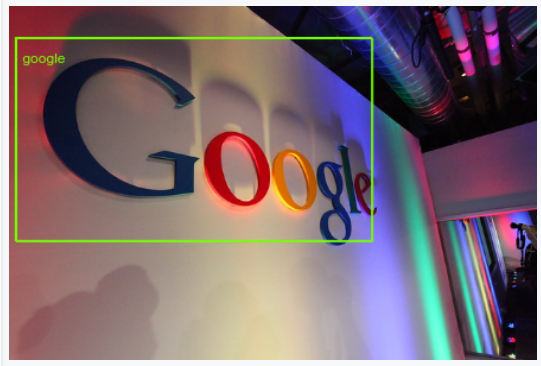
All feature types

Logo detection

(https://cloud.google.com/vision/docs/detecting-logos)

3 (#fn3)

- Provides a textual description of the entity identified, a confidence score, and a bounding polygon for the logo in the file.

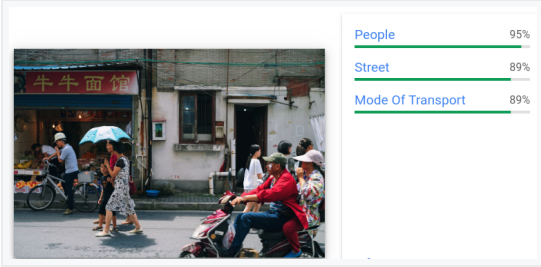


Label detection

(https://cloud.google.com/vision/docs/labels) 4 (#fn4)

- Provides generalized labels for an image.
- For each label returns a textual description, confidence score, and topicality rating.

★ Need customized labeling? With AutoML Vision (https://cloud.google.com/vision/automl/docs/) you can create a custom machine learning model for your specific image labeling use case.



All feature types

Text detection

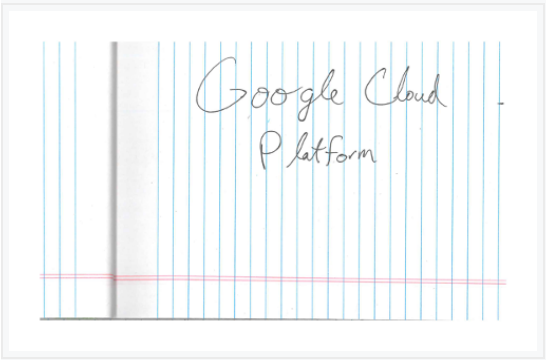
(https://cloud.google.com/vision/docs/ocr)

- Optical character recognition (OCR) for an image; text recognition and conversion to machine-coded text. Identifies and extracts UTF-8 text in an image.
- **Images**: Optimized for ***sparse*** areas of text within a larger image.
- **Response**: Returns both a list of words identifed with text, bounding boxes, and confidence scores (`textAnnotations`), as well as the structural hierarchy for the OCR detected text (`fullTextAnnotation`).
  - Hierarchy of extracted text structure:
    - TextAnnotation -> Page -> Block -> Paragraph -> Word -> Symbol.
    - Each structural component from Page on may further have their own properties such as detected languages, breaks, etc.
- **Languages supported**: Works with currently supported, mapped, and experimental (https://cloud.google.com/vision/docs/languages) languages.
- Feature enum value: `TEXT_DETECTION`.

All feature types

Document text detection (dense text / handwriting)

(https://cloud.google.com/vision/docs/pdf)



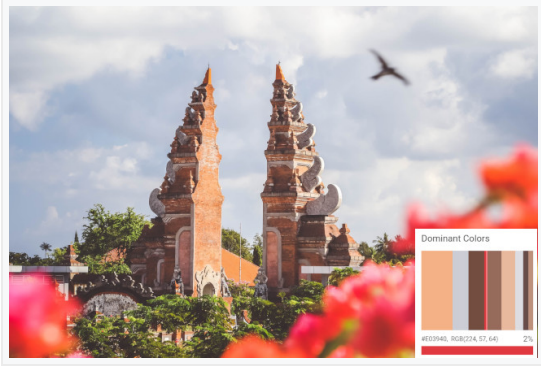
- Optical character recognition (OCR) for a file (PDF/TIFF) or dense text image; dense text recognition and conversion to machine-coded text.
- **Files**: Optimized for document files (PDF/TIFF).
- **Images**: Optimized for ***dense*** areas of text in an image (images that are documents), and images that contain handwriting.
- **Response**: Returns the structural hierarchy for the OCR detected text (**`fullTextAnnotation`**).
  - Hierarchy of extracted text structure:
    - TextAnnotation -> Page -> Block -> Paragraph -> Word -> Symbol.
    - Each structural component from Page on may further have their own properties such as detected languages, breaks, etc.
- **Languages supported**: Works with currently supported, mapped, and experimental (https://cloud.google.com/vision/docs/languages) languages.
- Feature enum value: **`DOCUMENT_TEXT_DETECTION`**.
  - Takes precedence when both **`DOCUMENT_TEXT_DETECTION`** and **`TEXT_DETECTION`** are requested.

All feature types

Image properties

(https://cloud.google.com/vision/docs/detecting-properties)

5 (#fn5)



- Returns dominant colors in an image.
- Each color is represented in the RGBA color space, has a confidence score, and displays the fraction of pixels occupied by the color [0, 1].

Object localization

(https://cloud.google.com/vision/docs/object-localizer)

6 (#fn6)

- Provides general label and bounding box annotations for multiple objects recognized in a single image.
- For each object detected the following elements are returned: a textual description, a confidence score, and normalized vertices [0,1] for the bounding polygon around the object.

★ Need customized object detection? With AutoML Vision Object Detection (https://cloud.google.com/vision/automl/object-detection/docs/) you can create a custom machine learning model for your specific image object detection use case.

All feature types

Crop hint detection

(https://cloud.google.com/vision/docs/detecting-crop-hints)

7 (#fn7)

- Provides a bounding polygon for the cropped image, a confidence score, and an importance fraction of this salient region with respect to the original image for each request.
- You can provide up to 16 image ratio values (width:height) for a single image.

All feature types

Web entities and pages

(https://cloud.google.com/vision/docs/detecting-web)

8 (#fn8)

Image credit: Quinten de Graaf, Unsplash.

Category	Responses
Web entities	• ***entityId***: /m/02p7_j8, ***score***: 1.3225499, ***description***: Carnival in Rio de Janeiro • ***entityId***: /m/06gmr, ***score***: 1.1684971, ***description***: Rio de Janeiro • ***entityId***: /m/04cx88, ***score***: 1.05945, ***description***: Brazilian Carnival ...
Full matching images	• ***url***: https://1000lugaresparair.files.wordpress.com/2017/11/quinten-de-graaf-278848.jpg • ***url***: https://freewalkingtourrotterdam.com/wp-content/uploads/2017/07/quinten-de-graaf-278848.jpg ...
Partial matching images	• ***url***: https://www.linnanneito.fi/wp-content/uploads/sambakarnevaali-riossa.jpg • ***url***: https://static.airhelp.com/wp-content/uploads/2019/02/26105557/two-women-in-carnival-costumes.jpg ...
Pages with matching images	• ***url***: https://travelnoire.com/best-carnival-celebrations-around-the-world/, ***pageTitle***: Best \u003cb\u003eCarnival\u003c/b\u003e Celebrations Around The World - Travel Noire, ***fullMatchingImages***: [{***url***: https://travelnoire.com/wp-content/uploads/2019/02/quinten-de-graaf-278848-unsplash.jpg}] • ***url***: https://bespokebrazil.com/rio-carnival-2019/, ***pageTitle***: Visit \u003cb\u003eRio Carnival 2019\u003c/b\u003e with the Brazil Specialists - Bespoke Brazil, ***partialMatchingImages***: [{***url***: https://bespoke-brazil-2018-bespokebrazil.netdna-ssl.com/wp-content/uploads/2019/01/Carnival-1.jpg}] ...
Visually similar images	• ***url***: https://www.brazilbookers.com/_images/photos/rio-carnival-images/rio-carnival-2016-carnival-date.jpg • ***url***: https://image.redbull.com/rbcom/010/2017-02-08/1331843859949_3/0100/0/1/watch-rio-carnival-2017-live-on-red-bull-tv.jpg ...
Best guess labels	rio carnival 2019 dancers

- Provides a series of related Web content to an image.
- Returns the following information:
  - **Web entities**: Inferred entities (labels/descriptions) from similar images on the Web.
  - **Full matching images**: A list of URLs for fully matching images of any size on the Internet.
  - **Partial matching images**: A list of URLs for images that share key-point features, such as a cropped version of the original image.
  - **Pages with matching images**: A list of Webpages (identified by page URL, page title, matching image URL) with an image that satisfies the conditions described above.
  - **Visually similar images**: A list of URLs for images that share some features with the original image.
  - **Best guess label**: A best guess as to the topic of the requested image inferred from similar images on the Internet.

Explicit content detection (Safe Search)

(https://cloud.google.com/vision/docs/detecting-safe-search)

- Provides likelihood ratings for the following explicit content catgories: **`adult`**, **`spoof`**, **`medical`**, **`violence`**, and **`racy`**.
- Likelihoods ratings are expressed as 6 different values: **`UNKNOWN`**, **`VERY_UNLIKELY`**, **`UNLIKELY`**, **`POSSIBLE`**, **`LIKELY`**, or **`VERY_LIKELY`**.

1. *Image credit*: Himanshu Singh Gurjar (https://unsplash.com/photos/iSi02D_Qx_w) on Unsplash (https://unsplash.com/) (*annotations added*). ↩ (#ref1)

2. *Image credit*: Nikolay Vorobyev (https://unsplash.com/photos/jaH3QF46gAY) on Unsplash (https://unsplash.com/) (*annotations added*). ↩ (#ref2)

3. *Image credit*: Robert Scoble (https://www.flickr.com/photos/scobleizer/4249731778/) (CC BY 2.0 (https://creativecommons.org/licenses/by/2.0/), *annotation added*). ↩ (#ref3)

4. *Image credit*: Alex Knight (https://unsplash.com/photos/wtwUptVqrKU) on Unsplash (https://unsplash.com/) . ↩ (#ref4)

5. *Image credit*: Jeremy Bishop (https://unsplash.com/photos/QUwLZNchflk) on Unsplash (https://unsplash.com/). ↩ (#ref5)

6. *Image credit:* Bogdan Dada (https://unsplash.com/photos/J9cBJjlpYKU) on Unsplash (https://unsplash.com/) (*annotations added*). ↩ (#ref6)

7. *Image credit*: Yasmin Dangor (https://images.unsplash.com/photo-1526398018-f9ee0723f228?ixlib=rb-1.2.1&auto=format&fit=crop&w=1400&q=80) on Unsplash (https://unsplash.com/) (*original and cropped image shown*). ↩ (#ref7)

8. *Image credit*: Quinten de Graaf (https://unsplash.com/photos/KB0Ipylp7dc) on Unsplash (https://unsplash.com/). ↩ (#ref8)

Except as otherwise noted, the content of this page is licensed under the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/), and code samples are licensed under the Apache 2.0 License (https://www.apache.org/licenses/LICENSE-2.0). For details, see our Site Policies (https://developers.google.com/terms/site-policies). Java is a registered trademark of Oracle and/or its affiliates.

Last updated November 6, 2019.