

Memespector GUI by Jason Chao

# Memespector GUI

Graphical User Interface Client for Computer Vision APIs

### Settings

<input checked="" type="checkbox"/> <b>Google Vision</b>	Features	7 selected	Credential file	~/Documents/Memespector-GUI-macOS/google-cloud-token.json <a href="#">Browse...</a>		
<input checked="" type="checkbox"/> <b>Microsoft Azure</b>	Features	7 selected	Endpoint	ognitiveservices.azure.com/	Subscription key	abcdef0123456789
<input checked="" type="checkbox"/> <b>Clarifai</b>	Model	General	API key	abcdef0123456789		
<input checked="" type="checkbox"/> <b>Open Source</b>	Model	VGG16 (ImageNet)	Endpoint	https://europe-west1-digital-methods-resources.cloudfunctions.net/classify_imag		

### Image Sources

Add [image files on this computer](#) [a folder containing images on this computer](#) [a text file containing image locations](#) [images on the web](#)

### Output

JSON file (full results)	/Users/User/Documents/Memespector-GUI-macOS/cv-apis-date_time.json	<a href="#">Browse...</a>
CSV file (simplified results)	/Users/User/Documents/Memespector-GUI-macOS/cv-apis-date_time.json	<a href="#">Browse...</a>

### Invocation

[Invoke APIs](#)

[User Manual](#) [About](#)

Jason Chao  
SFB 1187 Medien der Kooperation  
Universität Siegen

## Why AI?



Writing/coding explicit rules to classify cats and dogs is very complicated.

# What do machine learning models do?



Label: dog



Label: cat

1. Train a model



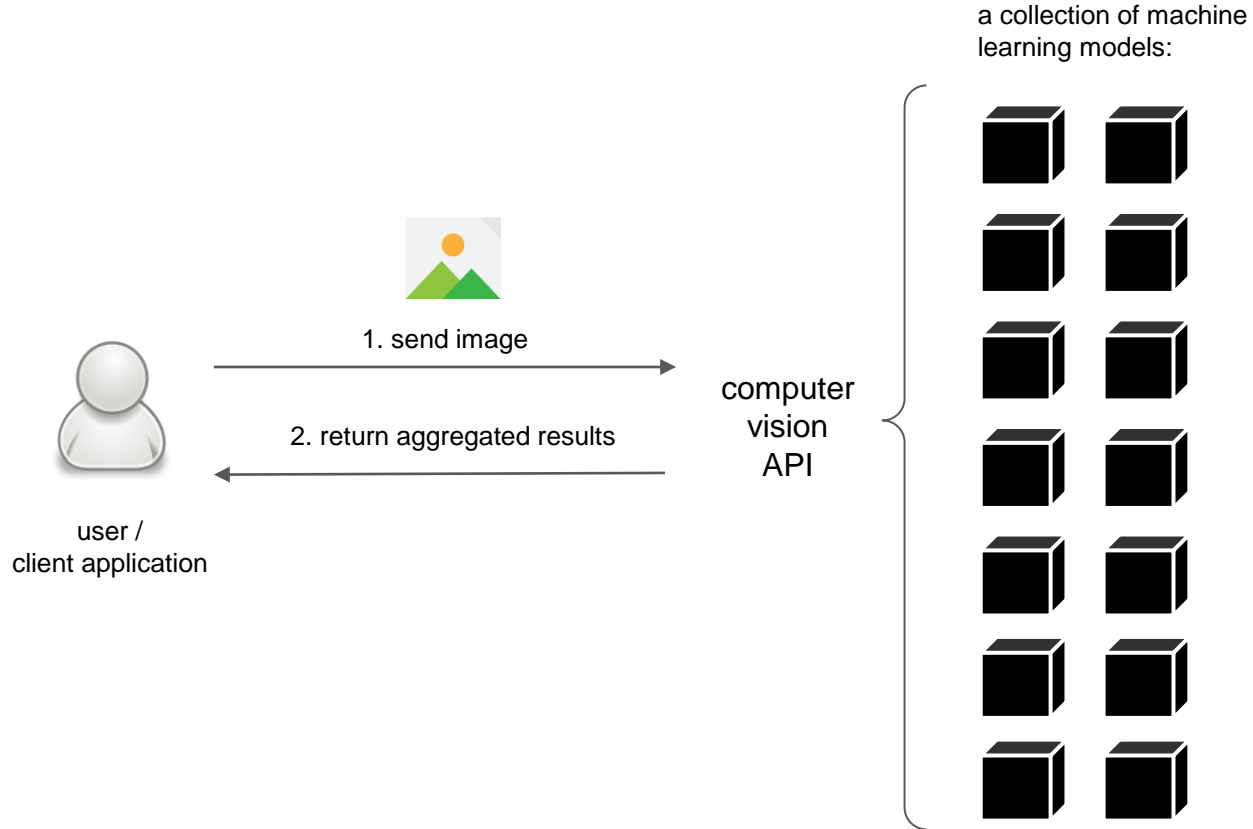
2. Input an unseen picture



3. Output the result of classification

Label: cat  
Confidence: 0.85

# What is a (commercial) computer vision API?



## What can Google Vision API tell us about an image? (Label and text detection)



Photo credit: PBS

### Labels:

- Sky
- Building
- Crowd

### Objects:

- Building
- Person
- Footwear

### Text:

- DON'T SHOOT OUR KIDS

What can Google Vision API tell us about an image? (Face detection)



### Face

- Joy: Very likely
- Sorrow: Very unlikely
- Anger: Very unlikely
- Surprise: Very unlikely
- Exposed: Very unlikely
- Blurred: Very unlikely
- Headwear: Very unlikely



WARNING

Some may find the following images disturbing.

Viewer discretion is advised.

## What can Google Vision API tell us about an image? (Web detection)



### Web entities:

- Derek Chauvin
- George Floyd
- Trial of Derek Chauvin
- Sentence

### Full matching image URL:

- [kiwifarms.net/...](#)
- [the-sun.com/...](#)
- ...

### Pages with full matching image:

- [wikipedia.org/...](#)
- [nbcnews.com/...](#)
- ...



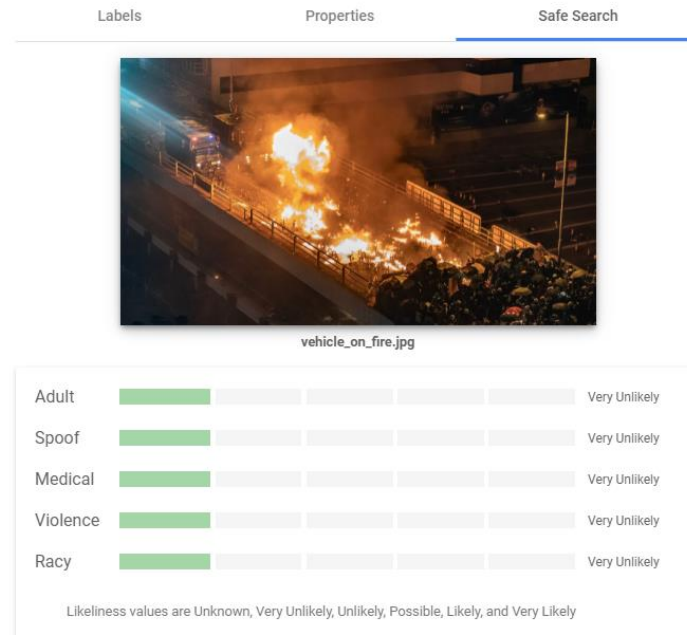
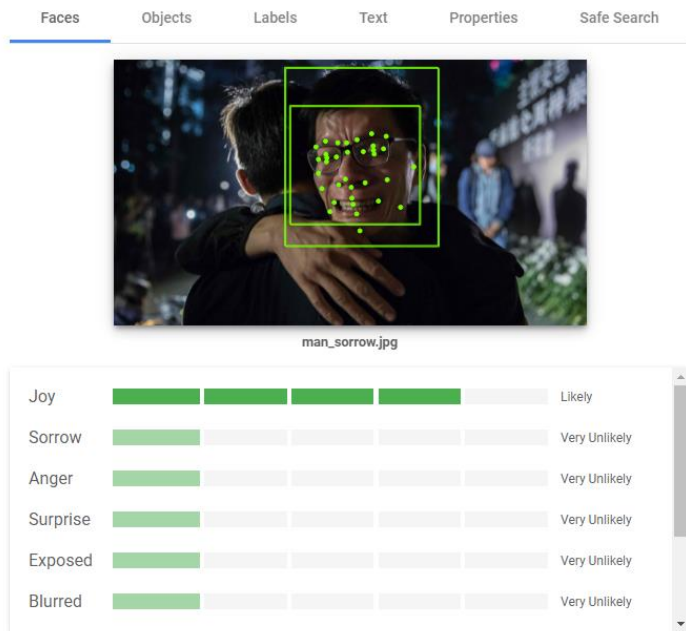
What can Google Vision API tell us about an image? (Safety detection)



Safety:

- Adult: Unlikely
- Spoof: Unlikely
- Medical: Unlikely
- Violence: Very likely
- Racy: Possible

# Caution: AI is not perfect



# Try the drag-and-drop demo of Google Vision API yourself

<https://cloud.google.com/vision/docs/drag-and-drop>

The screenshot shows the Google Cloud Vision API documentation page. The browser address bar displays "cloud.google.com/vision/docs/drag-and-drop". The page title is "Try it!". Below the title, there is a "Send feedback" button. The main content area contains the following text:

Use the application below to return image annotations for your image file. Click the **Show JSON** button to view the raw response.

- Maximum file size is 4MB.
- Your browser must have JavaScript enabled.

A blue note box contains the following text:

★ **Note:** Vision API offers two [feature types](#) for text detection (also called optical character recognition, or OCR). The text detection feature used in this demo is [DOCUMENT\\_TEXT\\_DETECTION](#). Results from this feature may differ from results returned from a [TEXT\\_DETECTION](#) feature request.

Below the note is a "Try the API" section with a dashed box containing the text "Upload your image" and a placeholder image of a mountain range.

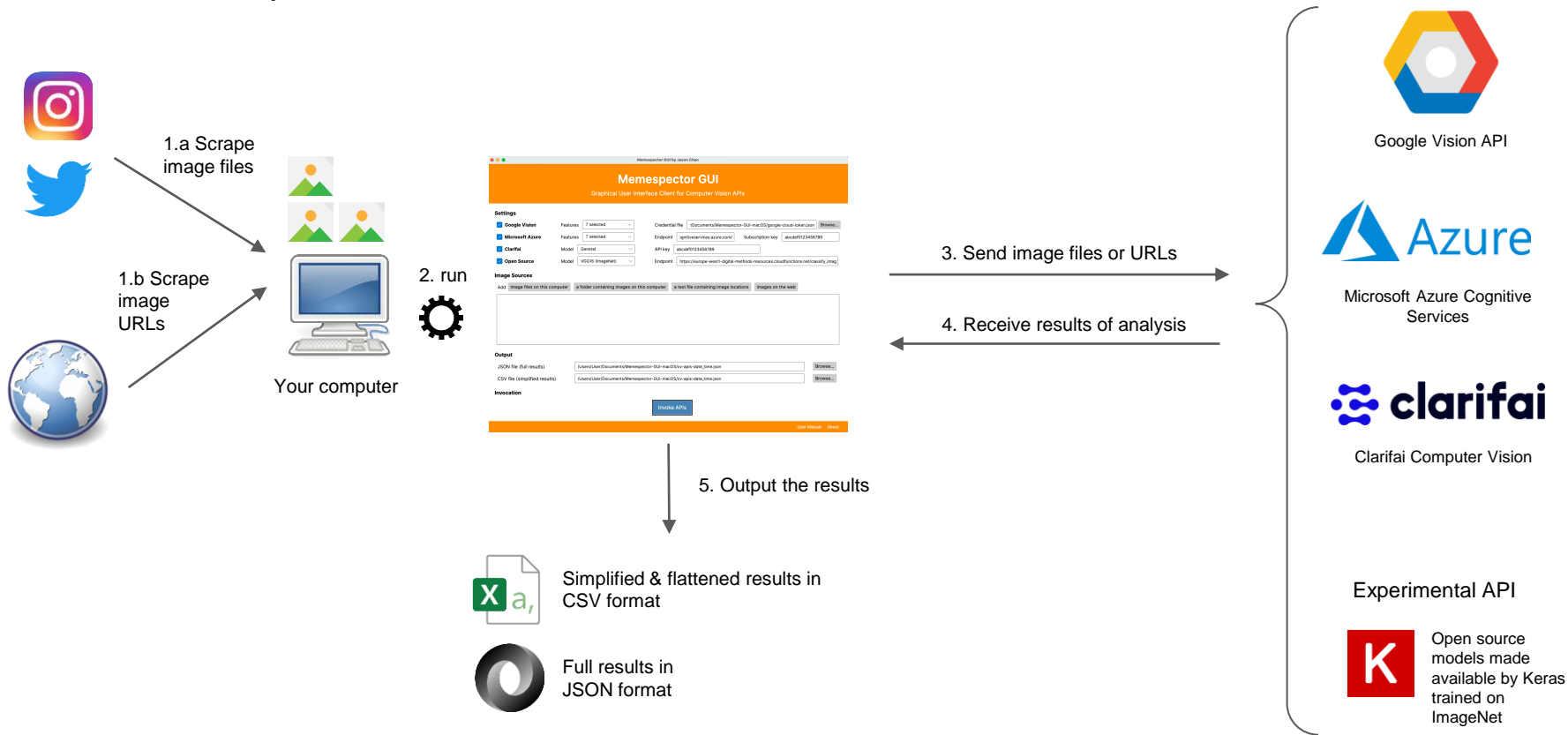
The left sidebar contains a navigation menu with the following items:

- Vision API
- Product overview
- Features list
- Try it!
- Quickstarts
- All Quickstarts
- Set up the Vision API
- Using client libraries
- Using the command line
- Using API explorer
- Samples
- All Vision API code samples
- All code samples for all products
- How-to Guides
- All How-to guides
- Before you begin
- Optical character recognition (OCR)
- Detect crop hints
- Detect faces
- Detect image properties
- Detect labels
- Detect landmarks
- Detect logos
- Detect multiple objects
- Detect explicit content (SafeSearch)
- Detect Web entities and pages
- Batch feature detection
- Using Vision with Spring framework
- Using Vision with Spring framework
- Base64 encode

The screenshot shows the "Try the API" demo interface. The page title is "Try the API". The interface has a navigation bar with the following tabs: "Faces", "Objects", "Labels", "Text", "Properties", and "Safe Search". The "Objects" tab is selected. Below the navigation bar is a large image of a conference room with several people sitting at tables. The image is annotated with green bounding boxes around various objects. Below the image is the filename "smart2020.jpg". To the right of the image is a list of detected objects with their confidence scores:

Object	Confidence
Person	90%
Person	87%
Chair	84%
Person	80%
Chair	77%
Person	76%
Person	71%
Person	70%

# What is Memespector-GUI?



Google Vision	Microsoft Azure Cognitive Services	Clarifai Vision	Open source model (ImageNet)
<ul style="list-style-type: none"> <li>● <b>Safety</b> - adult, violent and racist elements</li> <li>● <b>Face</b> - emotional expressions of faces</li> <li>● <b>Label</b> - generalised labels defined by Google</li> <li>● <b>Web</b> - web entities (inferred descriptions from similar images on the web), similar images, full/partial matching images, visually similar images and web pages with matching images</li> <li>● <b>Text</b> - the text recognised</li> <li>● <b>Landmark</b> - well-known or prominent sites</li> <li>● <b>Logo</b> - logos of popular products</li> </ul>	<ul style="list-style-type: none"> <li>● <b>Adult</b> - explicitly sexual, sexually suggestive and blood/gore</li> <li>● <b>Face</b> - human faces with age and gender</li> <li>● <b>Objects</b> - objects or living things with bounding box coordinates</li> <li>● <b>Tags</b> - recognisable objects, living beings, scenery and actions</li> <li>● <b>Categories</b> - 86-category taxonomy</li> <li>● <b>Description</b> - a human-readable sentence that describes the image's contents</li> <li>● <b>Brands</b> - logos of brands in consumer electronics, clothing and more</li> </ul>	<ul style="list-style-type: none"> <li>● <b>General</b> - concepts including objects, themes, moods and more</li> <li>● <b>Apparel</b> - fashion-related concepts</li> <li>● <b>Celebrity</b> - recognised celebrities</li> <li>● <b>Color</b> - dominant colours present</li> <li>● <b>Food</b> - food items</li> <li>● <b>Moderation</b> - gore, drugs, explicit nudity or suggestive nudity</li> <li>● <b>NSFW</b> - nudity</li> <li>● <b>Textures and Patterns</b> - common textures (feathers, woodgrain), unique/fresh concepts (petrified wood, glacial ice) and overarching descriptive concepts (veined, metallic)</li> <li>● <b>Travel</b> - specific features of residential, hotel and travel-related properties</li> </ul>	<p>Labels from the <a href="#">ImageNet</a> dataset</p> <p><i>Caution: The API that serves open source pre-trained computer vision models is experimental. It does not offer the same level of performance as the commercial APIs. The default endpoint in Memespector-GUI is for evaluation purpose only.</i></p>

<https://github.com/jason-chao/memespector-gui/>