

Emotion Detection API Comparison

The three main API's (Google Vision, IBM Watson Visual Recognition and Microsoft Cognitive Services) and an alternative (Affectiva API) were compared using the below topics.

- Presentation & First Impressions
- Features accomplished
- Documentation & Learning curve
- Price of use
- Community adoption

Google Vision API

1 – Presentation & First Impressions

The API presentation page brings attention citing topics like powerful machine learning models, object and face detection, improper content detection and language automatic identification. The “Try it yourself” example is well presented and provides fast evaluation and nice visualization. Another good point is to have everything in one page format such as feature description, example of use, users comment's and pricing table.

2 – Features accomplished

Google Vision API provides different categories of features such as face analysis, label content, web parsing content, color information and safe searching. Each of them are composed by some features and the presentation allow us to try an example and check the results. Below are some considerations:

- Faces tab provides information about face emotion detection, his confidence and head pose features.
- Labels tab provides useful content and categorization such as gender, professional and dress object identification.
- Web tab brings information parsed from websites the contains the original image or similar ones.
- Properties tab has information about color domination and aspect ratios.
- Safe Search brings categorized features related to some specific fields.
- JSON tab show us a brief example of format to read in case of developing.

3 – Documentation & Learning curve

The API follows default Google 2017 documentation and provides quickstart materials, tutorials and code samples of use. The content of both are easy to understand in terms of basic concepts and call requests but has some difficult steps for a beginner programmer.

The server-side and Google Cloud configuration content with a lot of authentication is always considered a hard work for first users of Google services. These steps could be more understandable to users in order to decrease the configuration time taken.

4 – Price of use

Google products are usually based on number of query amounts and this API is not different. A free limit of 1000 requests is available per month in order to permit students and initial company users to make use of first features and stay tuned about the services. Therefore, with an amount greater than 1000, the user/company is charged for each next 1000 requests.

The price range for companies are inferior if they have a lot of request use since units more than 5 million per month are charged by an inferior rate. In overall, the price is not so high if we think of financial business application returns.

5 – Community Adoption

These services are innovative and brings new features for application developers and consumers. The main blog websites about image recognition and technology in general cites Google Vision API and its advantages.

Analysing the Google Trends by the “Google Vision API” term we can see that since 2016 it seems to be increasing the interests in its products and services. We can also notice that Asian countries are in the top of the list, maybe could be related to the high number of users looking for Computer Vision and Robotic tasks resolution.

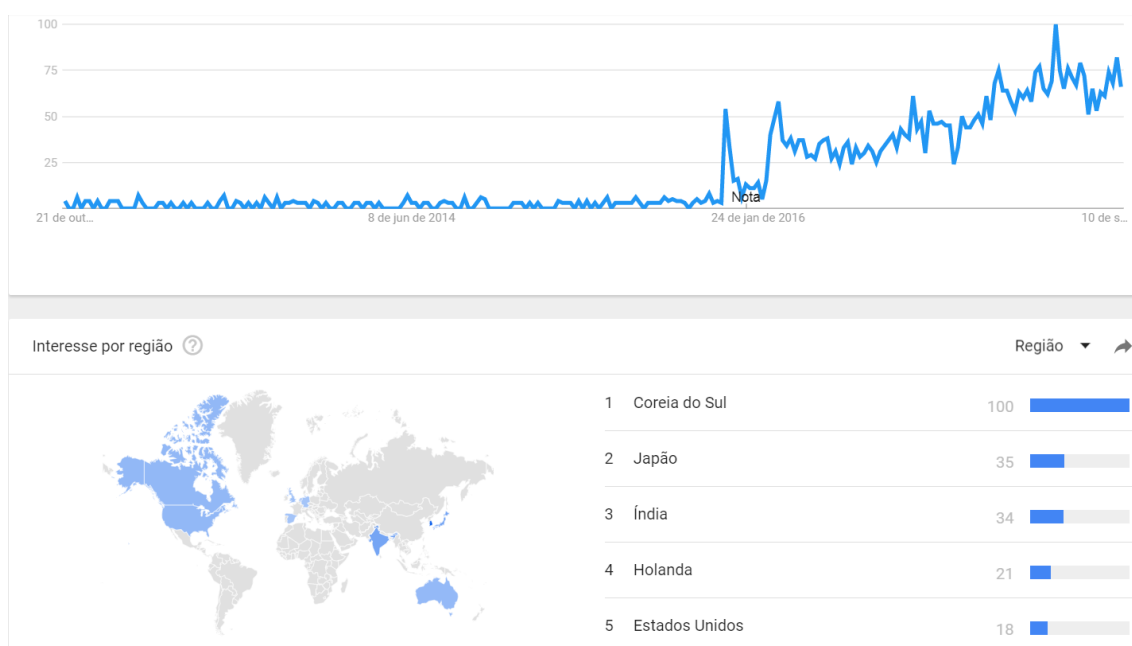


Figura 1 - Google Trend for "Google Vision API" term

Microsoft Cognitive Services API

1 – Presentation & First Impressions

The home page cites various services included in this product not limited to emotion such as video processing, face detection, content moderator and other services. All of them have examples and an appropriate menu to check more information. The “Try it” examples are easy to understand and visualize, with the possibility to upload your own image.

The emotion API page demonstrates with a lot of informative content, what features the service provides, the useful links to documentation, SDK, pricing and a stack overflow community. The first impressions trying the examples were satisfied since it could predict the right emotion with fast evaluation providing the JSON response which is great for a developer to start being familiar with the service.

2 – Features accomplished

Microsoft Emotion API provides face detection and emotion categorization scores for multiple persons. The classes covered by the service are anger, contempt, disgust, fear, happiness, neutral, sadness and surprise.

The service allows video emotion recognition with a good visualization of resulted emotions that are the same cited for image classification. The tested results had fast evaluation through a sample video by extracting the frames and sending them to the API. Some examples provided on GitHub from Microsoft also helped to understand the development concepts.

In respect of further metrics, the API can lead to 1 video per minute limited to 100MB and 10 image transactions per second with 4MB image maximum size. The facial landmarks detection is limited to 27 for a number of faces per image up to 64.

3 – Documentation & Learning curve

The documentation provides quick-starts in different programming languages (C#, Java, JavaScript, PHP, Python, Ruby) and tutorials in some of them (C# and Python). The content of both bring the developer direct to code with no introductory concepts and configuration since it assumes that the developer also has familiarity with JSON requests and authentication configuration.

The server-side is based on Azure, which is an already adopted cloud service and has many tutorials on the internet. So, the authentication and Cloud configuration for better performance could be solved in few minutes even for first users.

There is also an API reference explaining image and video parameters in respect of request and response. It also provides extra information about the types of errors and examples using Android SDK and Windows SDK.

4 – Price of use

The Emotion API pricing page details the 3 categories of use which are Free, Basic and Standard. In the free option, the user can have 30000 image transactions per month and 300 video uploads (limited to 1 per minute). The basic option increases the image processing rate to 10 image transactions per second with a cost of \$0.10 per 1000 transactions. The most advanced option keeps the image transaction rate and increases the video uploads up to 3000 by a cost of \$0.25 per 1000 transactions.

A beginner to intermediate developer of Visual Recognition and Deep Learning fields could work with free option in order to test their system and get familiar with the service. However, a company or an advanced user will need to pay an extra amount, choosing the basic option for image applications and standard option for video applications.

5 – Community Adoption

Microsoft services are already adopted by the developers community due to its variety and efficiency. The Emotion API is one of the most cited tools for emotion recognition using both image and video, being compared with personal projects made by university labs. Forums, blogs and technology websites have detailed content and are always trying to post news about it.

Analysing the Google Trends by the “Microsoft Emotion API” term we can see that since 2016 is increasing the interest in this service but very instable. However, this fact is due to the API was recently launched and the academic papers is still focus on scientific works using open source code and methodology for emotion evaluation.



Figura 2 - Google Trends for "Microsoft Emotion API" term

IBM Watson Visual Recognition API

1 – Presentation & First Impressions

The home page explain simple topics that the API may be useful like organize image libraries, understand an individual image, recognize food, detect faces and create custom classifier. The “Try the service” examples are so simple and provides the possibility to upload your own image. However, the evaluation is not so fast and the result is not so impressive.

Another point that needs improvement is the fact that the documentation, pricing and feature description do not appear in the main page.

The first impressions trying the examples could not be considered amazing since it brings only simple evaluation and categorization. The simplicity of the page does not call the user to interact to the tool sample.

2 – Features accomplished

IBM Watson Visual Recognition API provides some classes related to the image. The presentation allow us to visualize an example and check the results, turning into classification of categories that the image can be included and a score representing the reliability of evaluation. The result also contains the type hierarchy that could be useful to show sub-categories.

A second option is to train your own classifier or choose the specific image classes included under Creative Commons 2.0 general license (which means that they could be used for projects without legal problems). These included image collections also contains interesting sub-classes giving ideas of the powerful tools that could be built using the API.

3 – Documentation & Learning curve

The documentation provides different ways to learn how to build your own classifier or change the existed ones. It presents options such as step by step tutorials, video materials, webinars and blog posts. The content of both are very easy to understand for a beginner even those topics related about credentials, POST calls and JSON response format.

The server-side is based on IBM Bluemix, which leads all the IBM Watson services and has an easy adoption. The authentication configuration content is explained with details which is good for first users. There is also an API reference that takes care of resume the main parameters, attributions and some programming languages examples such as Java, Python and Node.js.

4 – Price of use

This Watson product is free for simple use with less than 5000 training images and 250 image classifications per day. An intermediate developer of Visual Recognition and Deep Learning fields usually will need more than that for personal projects, specially a bigger amount of training data. So, the free price option only covers those who only want to test the API and decide if will be a good choice for future projects.

A standard price option is also available for people that need an extra amount of training space and image classification. Their prices change according to the type of use such as image classification, face detection and custom classifier training. Again, for an expert user or company that needs a full application with custom model, visual recognition system and uses sensitive data, there is a premium plan but the price is not specified and needs the contact with the IBM sellers.

5 – Community Adoption

The IBM Watson services are frequently appearing in social media, forums and technology websites. The Visual Recognition was created to expand their field into object detection, image and object classification, in order to compete with the other tools available. However, this API still does not have advanced features that contributes mainly to emotion recognition system because it only provides image categorization.

Analysing the Google Trends by the “IBM Watson Visual Recognition” term we can see that only in the end of 2016 is increasing the interest in this service but very instable. However, checking for the general term “IBM Watson” show us that their services have searches for a high number of people confirming the potential of the services to compete in cognitive field.



Figura 3 - Google Trend for "IBM Watson Visual Recognition" term

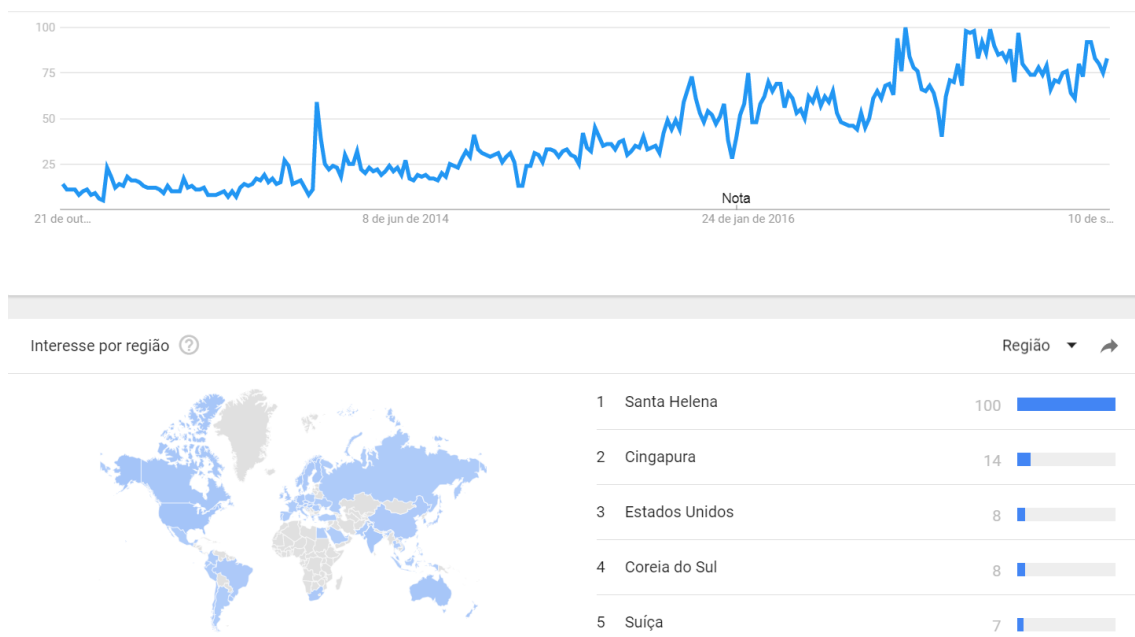


Figura 4 - Google Trends for "IBM Watson" term

Affectiva API

1 – Presentation & First Impressions

The home page demonstrates what it is possible to build with the Emotion API, the SDK support platforms and general links for further information. The “Try it” option is present in the page in different formats as well as a mention for Android and IOS applications.

In terms of presentation and possibilities, Affectiva brings the user to dive into emotion recognition world since the “Demo” section has examples like mapping emotions while watching Youtube videos, check how the market tool for online media works and a funny GIF search service using user’s emotion. The first impressions trying the services shows that Affectiva is using the most advanced techniques for emotion recognition and the visualization of results is better than the main services described previously.

2 – Features accomplished

The SDK allows developers to make use of different platforms and get various insights based on emotion and other features like gender recognition, ethnicity, age and even use of glasses. Some platforms included are iOS, Android, Unity, Windows, Linux and MacOS, which we can make use of video or photo files during streaming and offline.

There are many emotions available and also a brief description on how to detect each of them using expressions descriptors based on AU’s. Moreover, valence and engagement could be detected using the API also making use of the EMFACS model developed by Friesen and Ekman.

The tool gives results of the emotions joy, anger, disgust, surprise, fear, sadness and contempt. An interesting fact is that the action units cited are easily accessed and could be used as metrics to understand the classifier. The three API’s analysed previously did not show these features and only provided the final results.

3 – Documentation & Learning curve

The documentation provides different examples to learn how to build your own classifier on each platform. It presents alternatives using camera stream, video file, video frame stream or photo files. The content is intuitive for a beginner specially those developers which never worked with emotion recognition applications.

The API service is described and the credentials for authentication configuration needs to be approved by the company. There is also a reference of job names and data formats supported that can help developers to understand the features available.

4 – Price of use

Affective product is free for simple use and the main page did not bring further information about pricing. It seems that a user could use it for free until reach an amount of

\$1000. However, no information was found in order to describe how these services are calculated and a complete price ranging table is not provided.

So, the free price option only covers those who only want to test the API and decide if will be a good choice for future projects. The price is based on user’s use case and I consider that for an expert user or company that needs a full application, Affectiva is a competitive company to ask for a budget.

5 – Community Adoption

Affectiva has some of the best scientists of this field as its employers and the company is frequently cited in news, technology blogs and papers. Their services brings much more detailed results and extra information to the user compared to other API’s available. However, this API still does not have consolidated pricing conditions and their cloud services seems to be limited.

Analysing the Google Trend term “Affectiva”, it points to an unstable interest in the last 5 years in different countries. However, the company and their founders are increasingly in conferences trying to collect more users to the services. The background of Affectiva API and SDK are more related to academy studies and it seems that now they are trying to offer their emotion services as business demand.

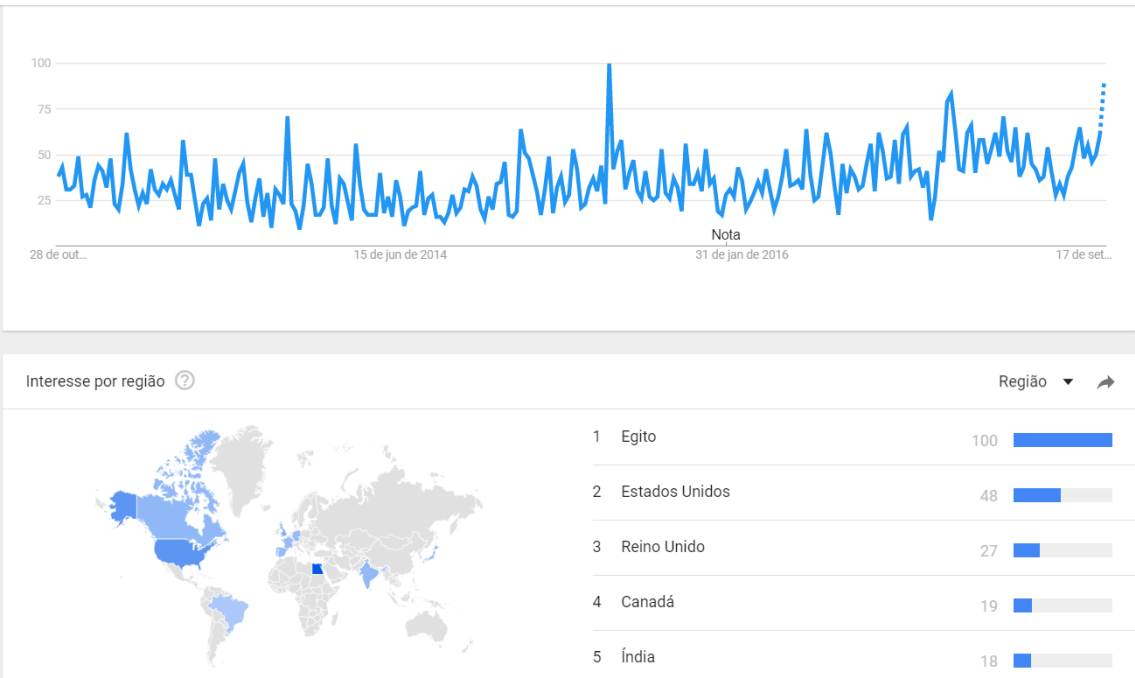


Figura 5 - Google Trends for "Affectiva" term