

The Rapid Rise of Computer Vision Technology

Computer vision is just one of the applications for machine learning experiencing rapid growth and increasing popularity in recent years. Machine learning technologies that are on the rise include [predictive analytics](#), [natural language processing](#), sentiment analysis, graph databases and [graph analysis](#). Wikipedia [defines](#) computer vision as:

... a field that includes methods for acquiring, processing, analyzing, and understanding images and, in general, high-dimensional data from the real world in order to produce numerical or symbolic information, e.g., in the forms of decisions. ... Sub-domains of computer vision include scene reconstruction, event detection, video tracking, object recognition, learning, indexing, motion estimation, and image restoration.

While it seems that computer vision is used primarily for building image-recognition platforms and applications, computer vision technology can be used for a variety of use cases across many industries. In a recent *ProgrammableWeb* [article](#), CamFind API CEO Dominik Mazur provides specific examples such as object recognition, search and e-commerce, surveillance and security, astronomy and outer space applications, and panoramic photography.

Computer vision has been in the news quite often in recent months. In February, Google [launched](#) Project Tango Smartphone, a project that uses

3-D sensing, machine learning, computer vision and other technology to achieve the goal of "giving mobile devices a human-scale understanding of space and motion." Project Tango devices are equipped with customized hardware and software that tracks the motion of the device in 3-D and at the same time creates a map of the environment. According to the Project Tango [website](#), the sensors in Project Tango devices "allow the device to make over a quarter million 3-D measurements every second, updating its position and orientation in real time, combining that data into a single 3-D model of the space around you."

Earlier this month, Google [acquired Jetpac](#), a startup that has created a mobile application that uses computer vision and other machine learning technologies to extract and analyze data from public Instagram photos. The data generated from the photos is used to create city guides. At the time of this writing, the [Jetpac City Guides app](#), which is available only on the iPhone, is a visual guide for more than 6,000 cities worldwide. On Aug. 15, Berlin-based company EyeEm [announced](#) the acquisition of Sight.io, a computer vision and machine learning technologies startup that has developed a platform featuring automated photo managing and processing capabilities.

Even the Walt Disney Co. has a division focused on the research and development of advanced technologies, which includes computer graphics, video processing, computer vision, data mining, machine learning and wireless communications. [Disney Research](#) consists of research laboratories located in several cities around the world, and its goal is to "positively impact profitability of The Walt Disney Company by inventing technologies and making discoveries that are novel on a global scale."

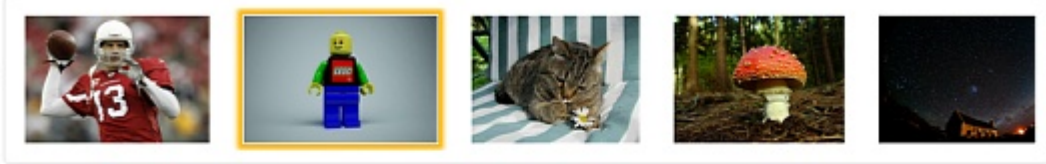
Below are a few examples of companies using computer vision technology, most of which specialize in image-recognition platforms and applications. A

few of these companies also provide platforms that feature natural language processing capabilities. These companies were chosen to show a sampling of the market and also because they provide APIs.

AlchemyAPI

Use one of our images
select an image from our gallery

Use your own image
point us at an image



Enter your own API Key

[Click here to learn more about Image Tagging.](#)


Visual

JSON

API

Tag	Score
lego	0.999864

[Contact us to learn more about Visual Search.](#)



The AlchemyVision API [demo](#) shows how the platform can be used for image tagging and visual search.

[AlchemyAPI](#) specializes in advanced text and image analysis, and the company provides a cloud-based platform that features advanced NLP and image-recognition functionality. Using [AlchemyAPI's API](#), developers can incorporate sentiment analysis, keyword extraction, entity extraction, image


tagging and other NLP and image-recognition features into their applications.

Back in May, *ProgrammableWeb* [reported](#) that AlchemyAPI had launched a product line called AlchemyVision, which features a new computer vision API. In March, the company launched [Taxonomy and Sentiment Analysis APIs](#), and in a *ProgrammableWeb* [article](#) covering the news, AlchemyAPI CEO Elliot Turner hinted that the company plans to debut in the near future APIs that use deep-learning based technology.

Diffbot

TRY THE API WITH ANY ARTICLE OR BLOG POST:

Analysis of <http://www.programmableweb.com/news/five-great-gaming-apis-developers-to-play/brief/2014/01/22/five-great-gaming-apis-developers-to-play-brief-2014-01-22-163400-gmt>

TITLE	Five Great Gaming APIs For Developers To Play With
AUTHOR	Janet Wagner
DATE	Wed, 22 Jan 2014 16:34:00 GMT
IMAGE	

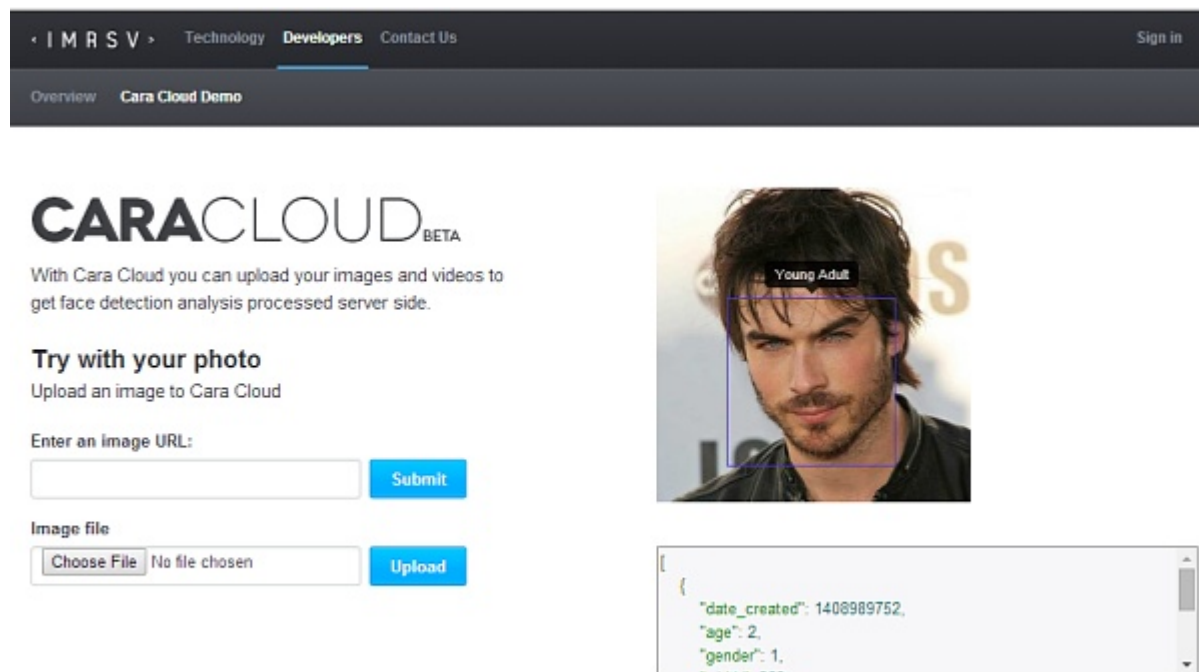
Diffbot provides a [demo](#) for its Article API and Product API.

[Diffbot](#) is a cloud-based data mining platform that uses computer vision, machine learning, NLP and other technologies to understand and extract

data from Web pages, including Web pages from news article sites, e-commerce sites and image galleries. Diffbot provides developers with several APIs that can be used to extract data from Web pages, which can then be used to build data-driven applications. Diffbot developer products include a [Frontpage API](#), [Page Classifier API](#), [Article API](#), [Product API](#) and other data-extraction tools.

In August 2013, *ProgrammableWeb* [reported](#) that Diffbot had released its Product API, which the company had been developing for two years. The Product API is capable of extracting from e-commerce Web pages data such as product title, model number, SKU, regular price, sales price and image. Earlier this year, the company [released](#) API client libraries for more than 35 programming languages.

IMRSV



The screenshot shows the IMRSV website with a navigation bar containing links for Technology, Developers, and Contact Us, along with a Sign in button. Below the navigation bar, there is a section for 'Overview' and 'Cara Cloud Demo'. The main content area features the 'CARACLOUD BETA' logo and a description: 'With Cara Cloud you can upload your images and videos to get face detection analysis processed server side.' Below this, there is a 'Try with your photo' section with a prompt to 'Upload an image to Cara Cloud'. This section includes a text input field for 'Enter an image URL:' with a 'Submit' button, and a file upload section with a 'Choose File' button, a 'No file chosen' status, and an 'Upload' button. To the right of the upload section, there is a photo of a man with a beard and a blue bounding box around his face, with the label 'Young Adult' above it. Below the photo, there is a code block showing JSON data:

```
{  "date_created": 1408989752,  "age": 2,  "gender": 1,  "height": 180}
```

The Cara Cloud [demo](#) shows a very basic example of facial-detection analysis from an uploaded photo. Photo from [Google Images](#).

[IMRSV](#), formerly Immersive Labs, specializes in technology that is capable of analyzing facial expressions via any webcam. Using computer vision, the IMRSV Cara platform can measure facial reactions such as smiles, frowns and surprise.

The [IMRSV Cara API](#) allows programmatic access to the platform, which developers can use to incorporate facial-detection analysis functionality into their applications. Applications can upload images and videos to the platform, which will then return facial-detection analysis information. The API is particularly useful for embedded, wearable and research analytics. A [demo](#) on the IMRSV website demonstrates how the facial-detection analysis works for both images and video.

Khronos OpenVX

The screenshot shows the Khronos Group website with the following structure:

- Header:** KHRONOS GROUP logo, tagline "CONNECTING SOFTWARE TO SILICON", and a navigation bar with links: Developers, Conformance, Membership, News, Events, Forums.
- Navigation:** Home, Developers, Books.
- QUICK LINKS:** Overview, Registry, Forums, Bug Tracking, Presentations, Ref Cards, Books, Jobs.
- Reference Cards Section:** A grid of 12 cards, each representing a different specification:
 - OpenCL:** 2.0 | 1.2 | 1.1 | 1.0 | BUY 2.0
 - OpenGL:** 4.5 | 4.4 | 4.3 | 4.2 | 4.1 | 4.0 | 3.3 | BUY 4.5
 - OpenGL ES:** 3.1 | 3.0 | 2.0 | BUY 3.1
 - WebGL:** 1.0 | BUY 1.0
 - COLLADA:** 1.4 | BUY 1.4
 - OpenSL ES:** 1.1
 - OpenMAX AL:** 1.0
 - OpenVG:** 1.1
 - EGL:** 1.4
 - OpenWF Composition:** 1.0
 - OpenWF Display:** 1.0
 - Khronos:** KHRONOS GROUP logo, with a link "Download Latest Cards".

The Khronos Group focuses on creating open standards for parallel computing. Open standards established by the group include OpenGL and WebGL. Image credit: [Khronos Group](#).

The [Khronos Group](#) is a nonprofit consortium of media-centric technology companies that focuses on developing Khronos API specifications and creating open standard APIs to "enable the authoring and acceleration of graphics, vision, sensor processing and dynamic media on a wide variety of platforms and devices." The Khronos Group is responsible for the creation and development of a variety of open standards, including OpenGL, WebGL, OpenCL and WebCL. In May 2013, *ProgrammableWeb* [reported](#) that the group was beginning work on the development of an open API for advanced control of mobile and embedded cameras and sensors.

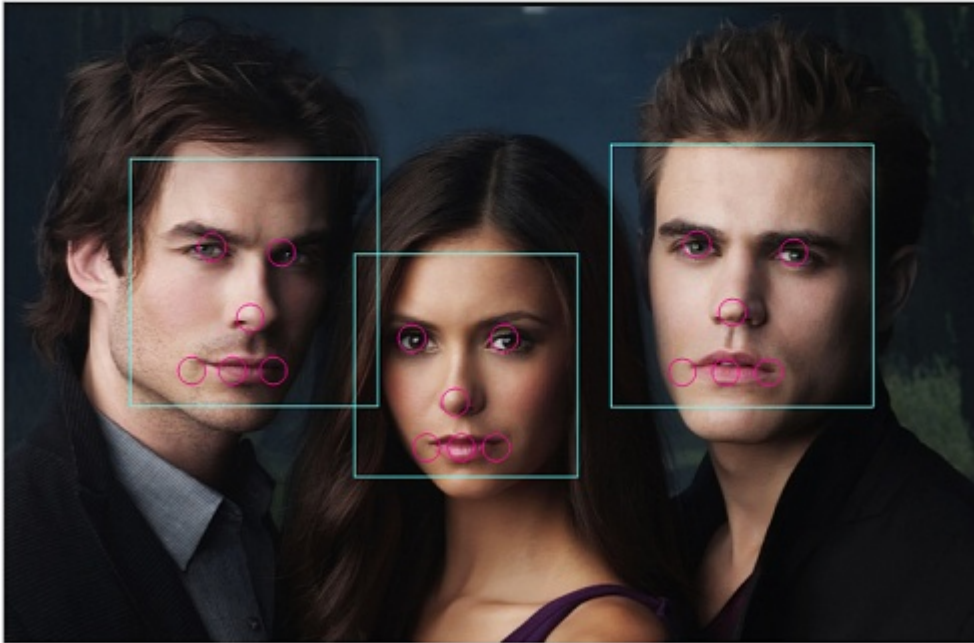
In November of last year, the Khronos Group [released](#) to the public the OpenVX 1.0 provisional specification, an open standard for enabling computer vision algorithms specifically for use cases such as face, body and gesture tracking, automatic driver assistance systems, and object and scene reconstruction. The OpenVX specification features the [OpenVX Hardware Acceleration API](#) designed for computer vision applications and libraries.

Lambda Labs

Raw Results:

```
{
  "status": "success",
  "images": {
    "url": "http://www.nebocodepro.net/images/vampire-series.jpg"
  },
  "photos": [
    {
      "url": "http://www.nebocodepro.net/images/vampire-series.jpg",
      "width": 719,
      "height": 1024
    }
  ]
}
```

Results:



Demos for Lambda Labs facial-recognition and facial-detection technology can be found on the company's [website](#). Photo from [Google Images](#).

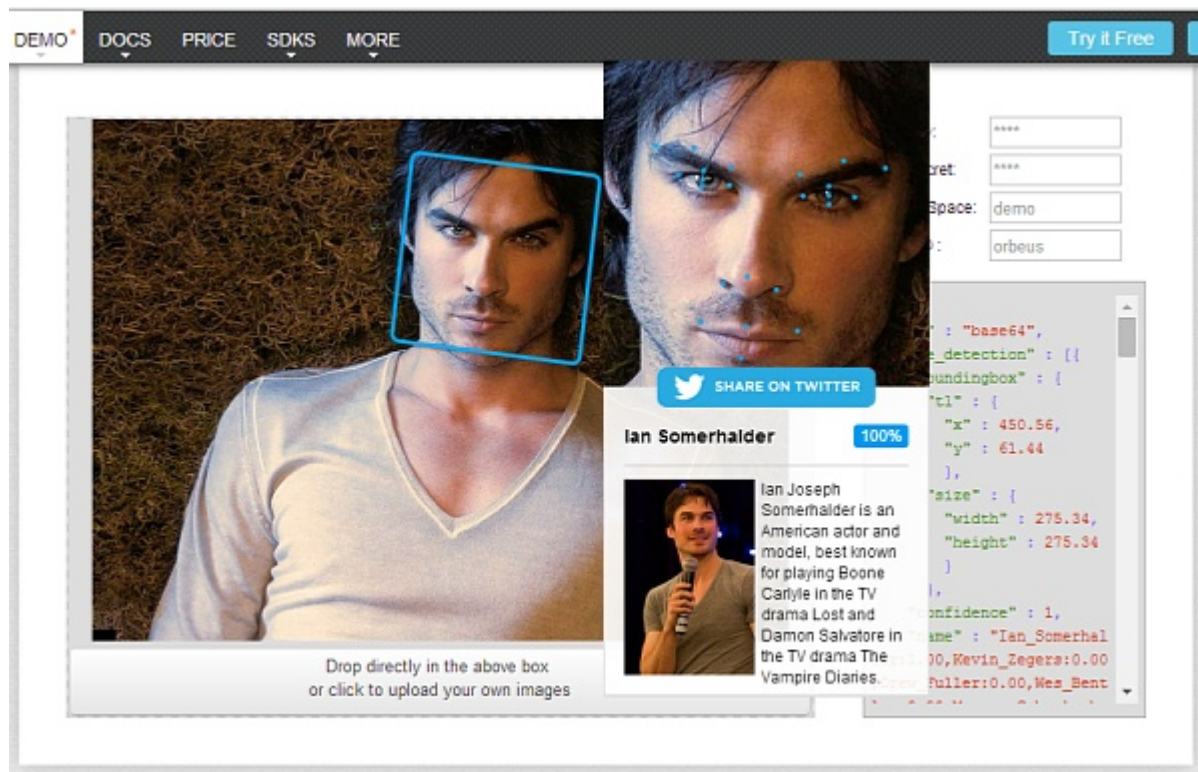
[Lambda Labs](#) is a computer vision and artificial intelligence technology company based in San Francisco. The company launched the beta version of its open source Face Recognition API back in September 2012. The [Lambda Labs Face Recognition API](#) provides facial recognition, facial detection, gender classification and other facial attributes from photos. The company is also [developing](#) a Google Glass facial-recognition application despite Google's strict policy regarding facial-recognition and voice-print use cases.

Last year, *ProgrammableWeb*'s Ajay Ohri [interviewed](#) Lambda Labs founder Stephen Balaban, who said that Lambda Labs facial-recognition technology

is being used primarily by ad agencies and Web and mobile application developers. Balaban also told *ProgrammableWeb* the following:

The massive centralization of personal information by both enterprise and government is making citizens ask important questions about what privacy means going forward. I've discussed with others in the industry about standardized protocols that may help individuals maintain control over their privacy. Think Robot Exclusion Standard (robots.txt) for face-recognition systems.

Orbeus



Celebrity recognition is a new Rekognition platform feature; a [demo](#) is available. Photo from [Google Images](#).

[Orbeus](#) is a computer vision company that has developed an integrated visual recognition engine capable of detecting, recognizing and analyzing

faces, scenes and objects all together in photos. Orbeus provides developers with access to the cloud-based Rekognition open API platform, which makes it possible to incorporate computer vision functionality into third-party applications. The [Orbeus Rekognition API](#) features facial- and concept-recognition capabilities; it is able to detect and recognize faces as well as recognize scenes, landmarks and other objects.

Last year, the company launched visual recognition APIs for Google Glass, which do not violate Google's policy regarding facial-recognition use cases. The APIs for Google Glass are not capable of detecting a person's identity. In a TechCrunch [article](#) covering the launch of the Orbeus Rekognition for Glass APIs, Orbeus CEO Ning Xu is quoted as saying, "Our API actually offers different face detection, face reading and scene understanding, so we're not just a facial-recognition company. But even without facial recognition, we can do a lot of things with your face without revealing your identity."

Conclusion

Computer vision is just the tip of the iceberg when it comes to the advances being made in the field of machine learning technology. Some [technology experts](#) believe data science, which includes machine learning, represents "a larger potential disruption than the industrial revolution." Whether this belief is proved to be true or false is only a matter of time.