

[Regístrate gratis](#)[Suscríbete](#)[Lee La Vanguardia en](#)[Iniciar sesión](#)

Un investigador catalán de Google participa en un avance clave para la descripción natural de imágenes

Oriol Vinyals ha creado, junto con sus compañeros de Google Brain Project, un algoritmo capaz de describir imágenes con el lenguaje propio de los humanos

Innovación | 04/12/2014 - 19:30h | Última actualización: 05/12/2014 - 09:56h



La finalidad del software es dotar a las máquinas de la capacidad de mejorar los procesos de búsqueda y catalogación de información, en este caso de imágenes Google

JUAN MANUEL GARCÍA CAMPOS | [Sigue a este autor en Twitter](#)
Barcelona

Google publicó hace unos días en Research Blog los [resultados de una investigación](#) sobre un sistema de aprendizaje automático capaz de describir con un lenguaje natural una escena representada en una fotografía. El algoritmo desarrollado por un grupo de investigadores de la compañía de Mountain View describe en una frase de unas quince palabras, y con una precisión muy similar a la capacidad humana, la escena que muestra una imagen determinada, más allá de identificar los objetos que aparecen en ella.

Así, el software es capaz de asignar a la siguiente imagen un subtítulo que la describe tan bien como lo haría una persona: 'A group of young people playing a game of frisbee' (Un grupo de jóvenes jugando una partida de frisbee).



El equipo que ha desarrollado esta investigación forma parte de **Google Brain**, un proyecto de Google dedicado al desarrollo de técnicas de "aprendizaje profundo" en disciplinas como el '**Machine Learning**', una rama de la inteligencia artificial que "entrena" a los ordenadores en el reconocimiento y aprendizaje del habla humana. La finalidad de este programa es dotar a las máquinas de la capacidad de mejorar los procesos de búsqueda y catalogación de información, en este caso de

imágenes.

Uno de los científicos que firma el 'paper' que detalla esta investigación es el sabadellense de 31 años **Oriol Vinyals**. Oriol estudió Ingeniería de Telecomunicaciones en la UPC de Barcelona, titulación que "dobló" con la de Matemáticas a través del CFIS (Centre de Formació Interdisciplinària Superior), un centro de excelencia de la UPC que incorpora cada año a un número limitado de estudiantes de alto nivel para cursar estudios interdisciplinarios. Tras realizar el proyecto de final de carrera en la CMU (Carnegie Mellon University, en Pittsburgh, Estados Unidos), uno de los más destacados centros de investigación superior en el área de ciencias de la computación y robótica, Vinyals encadenó una estancia de 9 meses en la Universidad de Berkeley (California) para especializarse en Machine Learning, un master de un año en Ciencias de la Computación en San Diego y un doctorado de cuatro años (2009-2013), otra vez en Berkeley. Durante todo este tiempo, dedicó los meses de verano a realizar estancias ('internships') de tres meses en Microsoft y Google, donde participó en diversas investigaciones sobre su especialidad.

Este [impresionante curriculum](#) le sirvió para que las principales tecnológicas y compañías financieras estadounidenses se lo rifaran al finalizar sus estudios: "En Estados Unidos hay mucha demanda de perfiles especializados en Machine Learning. Recibí muchas propuestas", reconoce. Escogió la de Google "porque me permitía rodearme de los mejores científicos del mundo y porque me daban total libertad para hacer lo que me gusta: proyectos de investigación pura, sin limitaciones de recursos ni fechas de entrega predeterminadas".

El resultado de esta última investigación sobre la construcción de una "descripción natural de la imágenes" supone un avance gigantesco en el campo del reconocimiento del lenguaje humano. El estudio trabaja con una base de datos común para investigadores de diversos centros (un equipo de [Stanford acaba de liberar el código de su propio algoritmo](#)). El objetivo es desarrollar una métrica capaz de evaluar cómo funcionan las frases resultantes del algoritmo desarrollado por los investigadores para describir imágenes en relación a las descripciones hechas por humanos.

Así, por ejemplo, ante una imagen que muestra tres tipos distintos de pizza sobre el fogón de una cocina, la descripción de un humano ("Three different types of pizza on top of a stove") y la que hace el algoritmo desarrollado por los investigadores ("Two pizzas sitting on top of a stove top oven") difieren en pocos detalles.



"El problema -explica Vinyals "es que todas las descripciones hechas por humanos son válidas, pero no son necesariamente exactas". Según la métrica desarrollada por este equipo de Google, sobre una puntuación de 100, la exactitud de las descripciones de los humanos es de 70 puntos, mientras que su algoritmo alcanza los 65 puntos, por 55 el de Stanford: "Funciona mucho mejor de lo que esperábamos -admite Oriol- en especial teniendo en consideración que la base de datos de imágenes sobre la que trabajamos es relativamente pequeña".

La aplicación más obvia de esta investigación es la mejora del buscador de imágenes de Google, pero puede ser útil para múltiples campos. "Por ejemplo"-comenta Vinyals-, para el desarrollo de aplicaciones móviles que ofrezcan una descripción de las fotografías que se muestran de un 'smartphone', de modo que los discapacitados visuales puedan saber qué imagen se expone en la pantalla del móvil, aunque no la puedan ver". El joven científico catalán desarrolla esta idea en una reciente [entrevista](#) que le ha concedido a la BBC.

En definitiva, esta investigación desarrollada por los cerebros más selectos de Google concuerda a la perfección con la misión

de la compañía fundada por Larry Page y Sergey Brin: “Organizar la información del mundo y hacerla útil y accesible para todos”.

Los próximos pasos del proyecto - avanza Oriol- serán “seguir entrenando al sistema con el fin de mejorar el algoritmo, dotando a los modelos de una base de datos mucho más amplia” -poniendo en práctica técnicas de big data- y, a medio plazo, “aplicar el algoritmo a alguno de los servicios de Google (probablemente Google Images) o crear algún producto en el que se pueda sacar provecho de este descubrimiento (por ejemplo, una aplicación para Android, el sistema operativo móvil de Google)”

Sigue a [Juan Manuel García Campos](#) en [Twitter](#)

Otras noticias



El 'bunga bunga' llega a Nueva York



¿Qué emociones buscas cuando lees?



Google cambia su algoritmo de búsqueda para favorecer a las webs adaptadas al...



Facebook actualiza su algoritmo para priorizar a los amigos sobre las marcas

En otras webs

recomendado por [Normas de participación](#)

5 Comentarios








Regístrate

6 conectados



Seguir		Compartir en	Comentar como
--------	--	--------------	---------------

Recientes | Antiguos

-  **NeektheGreek** 06/12/2014
Ni pies ni cabeza , o sea que para el desarrollo de tan pauperrima aplicación para la vida hay recursos ilimitados i sin necesidad de tiempo para conseguir resultados, pero para las investigaciones dirigidas a combatir enfermedades mortales, hay que mojarse.....no entiendo nada.
[Me gusta](#) [Responder](#)
-  **Enric** 04/12/2014
Que raro que sea catalán normalmente cuando alguien hace algo bien es español y cuando esta mal es catalán.
2   [Me gusta](#) [Responder](#)
-  **charley** 05/12/2014
[@Enric](#) Pero estas en La Vanguardia , donde los pedos catalanes huelen mejor que los demás españoles. ¿Te extraña ?
[Me gusta](#) [Responder](#)
-  **sheldon** 04/12/2014
Felicidades al gobierno español por prohibir el acceso a los miles de expedientes de pacientes para que un sistema parecido encuentre patrones comunes para luchar contra el cáncer...
[Me gusta](#) [Responder](#)
-  **Whiskey** 05/12/2014
Sheldon, tienes razón. Pero todo lo que quieras tocar en el campo médico es de una sensibilidad terrible. Exiaten dificultades incluso en alojar sistemas y documentos en la nube por la situacion de los servidores. la ley es la ley
[Me gusta](#) [Responder](#)

Powered by Livefyre