OPEN AI reveals most stunning AI system ever

12 April 2022, 02:00

Pascal Pollet

The development of AI technology is going at high speed. A new system shows how images can be generated from text.

Open AI was founded in 2015 as a research organisation, supported by Elon Musk and some other investors. Open AI had already caused quite a stir with the development of language model GPT-3 (Generative Pre-trained Transformer 3). GPT-3 is a deep-learning model containing 175 billion parameters and can be used to answer questions, generate text and translate texts. Building on its language model knowledge, OpenAI launched a new, stunning AI system in early April 2022: DALL.E2.

DALL.E2 is able to generate images from text. For example, when users enter "Astronaut riding a horse", the system will generate an image of an astronaut sitting neatly on a horse. The system was created by training a deep-learning model based on images and their text description. DALL.E2 goes even further than the systems that generate deep fake images: it also seems to understand the relationships between objects, which allows it – for the first time ever- to make meaningful combinations between different concepts (such as astronaut, riding, horse) in an image.



(*Source: OpenAI*)

In addition to generating new images, the system can also be used to edit existing images. For example, when you have a picture of a monkey, you can tell the system to make the monkey pay taxes, which will trigger the system to transform the picture accordingly. DALL.E2 can also mark parts of an image, which can then be filled in with another image based only on a textual

description of the adjustment of your choice.

The following video illustrates the possibilities of DALL.E2.

Accept marketing-cookies to watch this video. For FireFox users, disable the ‘Enhanced tracking protection’ of your browser to view this video.

DALL.E2 generates images in two steps: first, DALL.E2 uses language model CLIP, which can connect images with textual descriptions. DALL.E2 starts generating a raw interim solution, which according to CLIPP contains the main image features of the textual description. Second, DALL-E 2 applies a diffusion model to improve the interim solution, until the image fits the description according to CLIP. Diffusion models are “image improvers” that are trained by adding random pixels to images, after which the model is taught to restore the original image.

Source: OpenAI

Authors



Pascal Pollet