OPEN AI dévoile le système d'IA le plus étonnant de tous les temps

12 avril 2022, 02:00

Pascal Pollet

La technologie de l'IA se développe à la vitesse de l'éclair. Un nouveau système montre comment des images peuvent être générées à partir d'un texte.

Fondée en 2015 en tant qu'organisme de recherche, avec le soutien d'Elon Musk et de quelques autres investisseurs, Open AI n'en est pas à son coup d'essai : la société est en effet à la base du développement du modèle de langage GPT-3 (Generative Pre-trained Transformer 3), un modèle d'apprentissage profond comportant 175 milliards de paramètres. Il peut être mis en œuvre pour répondre à des questions, générer du texte et traduire des textes. Début avril 2022, s'appuyant sur sa connaissance des modèles de langage, OpenAI a lancé un nouveau système d'IA étonnant : DALL.E2.

DALL.E2 est capable de générer des images à partir d'une description textuelle. Exemple : en saisissant « Astronaute à cheval », l'utilisateur se voit proposer une image d'un astronaute correctement assis sur un cheval. Le système a été créé par l'entraînement d'un modèle d'apprentissage profond sur la base d'images et de leur description textuelle. DALLE.E2 va beaucoup loin que les systèmes qui génèrent de fausses images profondes, étant donné que le système semble également comprendre les relations entre les objets. Cela permet, pour la toute première fois, de combiner différents concepts (astronaute, équitation, cheval) de manière significative dans une image.



(*Source: OpenAI*)

En plus de générer de nouvelles images, le système peut également être mis en œuvre pour éditer des images existantes. Partant de l'image d'un singe, par exemple, on peut demander au système de lui faire payer des impôts, après quoi le système transforme l'image en conséquence. DALL.E2 permet aussi de marquer des parties d'une image et de les remplir par une autre image sur la seule base d'une description textuelle de la modification souhaitée.

La vidéo suivante illustre les possibilités de DALL.E2.

Acceptez les cookies marketing pour regarder cette vidéo. Pour les utilisateurs de FireFox, désactivez la « protection renforcée contre le suivi » de votre navigateur pour visionner cette vidéo.

DALL.E2 génère les images en deux étapes. Tout d'abord, DALL.E2 utilise le modèle de langage CLIP, qui est capable de relier des images et des descriptions textuelles. DALL.E2 commence par générer une solution intermédiaire qui, selon CLIP, contient les principales propriétés d'image de la description textuelle. DALL.E2 applique ensuite un modèle de diffusion pour améliorer la solution intermédiaire jusqu'à ce que l'image corresponde parfaitement à la description selon CLIP. Les modèles de diffusion sont des "améliorateurs d'image" qui sont formés en ajoutant des pixels aléatoires aux images, puis en apprenant au modèle à générer à nouveau l'image originale.

Source: OpenAI

Auteurs



Pascal Pollet