

QUAND L'APPRENTISSAGE PROFOND FAIT DE L'ART PSYCHÉDELIQUE

L'apprentissage profond est performant pour la reconnaissance d'objets dans une image, mais il est en réalité assez difficile de comprendre ce qui se passe au cours de l'analyse à chaque niveau du réseau de neurones. La société Google a développé le logiciel DeepDream qui utilise l'apprentissage profond. Les chercheurs peuvent analyser le résultat à différentes étapes du processus en demandant au logiciel de faire ressortir ce qu'il « voit » dans l'image. Ainsi, telles des paréidolies – l'illusion de voir, par exemple, un chien dans un nuage qui en a vaguement la forme –, le logiciel, qui s'est, par exemple, entraîné à reconnaître des images d'animaux, verra des oiseaux et des rongeurs dans une photographie d'un champ de maïs. Chacun peut créer ses propres images psychédéliques sur le site Deepdreamgenerator.com.



réduit. Dans un tel réseau, la plupart des minima locaux correspondent à un niveau d'apprentissage quasi équivalent à celui du minimum global.

Bien que les problèmes théoriques de l'optimisation soient, en principe, résolubles, la construction de grands réseaux comportant plus de deux ou trois couches avait souvent échoué. À partir de 2005, les recherches financées par l'ICRA ont commencé à porter leurs fruits et les obstacles ont été surmontés progressivement. En 2006, nous sommes parvenus à entraîner des réseaux neuronaux plus profonds en utilisant une technique qui opérait couche par couche.

Du bruit pour un meilleur apprentissage

Par la suite, en 2011, nous avons trouvé un moyen d'entraîner des réseaux encore plus profonds (avec davantage de couches de neurones virtuels) en modifiant les opérations effectuées par chacun des neurones, ce qui les rapprochait davantage des neurones biologiques dans leur façon de traiter l'information. Nous avons aussi découvert qu'en injectant du bruit aléatoire dans les signaux transmis entre les neurones au cours de l'apprentissage – comme ce qui a cours dans le cerveau –, les réseaux apprenaient mieux à identifier correctement une image ou un son.

Par ailleurs, deux facteurs essentiels ont contribué au succès des techniques d'apprentissage profond. Le premier est l'augmentation d'un facteur 10 de la puissance de calcul des ordinateurs grâce aux processeurs dédiés au traitement d'image (conçus initialement pour les jeux vidéo !), ce qui a permis d'entraîner des réseaux plus grands en un temps raisonnable. Le second facteur est que l'on avait désormais accès à d'énormes bases de données étiquetées, avec lesquelles les algorithmes d'apprentissage ont pu s'exercer à reconnaître un « chat », par exemple, dans des images qui comportaient un chat parmi d'autres éléments.

Une autre raison aux récents succès de l'apprentissage profond est que le réseau est capable d'apprendre à effectuer un ensemble de calculs qui construisent ou analysent pas à pas une image, un son ou un autre type de donnée. Avec une profondeur suffisante, ces dispositifs excellent dans beaucoup de tâches de reconnaissance visuelle ou auditive. Des travaux théoriques et expérimentaux récents ont d'ailleurs montré qu'il est impossible d'effectuer efficacement certaines de ces opérations mathématiques sans un réseau assez profond.

Mais que se passe-t-il au cœur du réseau de neurones profond ? Chaque couche transforme son entrée et produit une sortie qui est envoyée à la couche

suivante (voir l'encadré page 45). Les premières couches se concentrent sur les détails aux échelles les plus petites, puis les couches suivantes agrandissent les échelles considérées. Plus les couches sont profondes, plus elles vont représenter des concepts abstraits. Par exemple, les premières couches pourraient isoler des éléments caractéristiques d'une chaise et l'image de la chaise pourrait émerger du traitement de neurones d'une couche plus profonde, même si le concept de « chaise » ne faisait pas partie des étiquettes de catégorie sur lesquelles le réseau s'est entraîné. Et le concept de chaise pourrait lui-même n'être qu'une étape intermédiaire vers la création d'un concept encore plus abstrait au niveau d'une couche plus profonde, que le réseau pourrait catégoriser comme une « scène de bureau ».

Jusqu'à récemment, les réseaux de neurones artificiels se distinguaient en grande partie par leur capacité à effectuer des tâches telles que la reconnaissance de motifs dans des images statiques. Mais d'autres types de réseaux neuronaux affichent des résultats intéressants sur des événements dynamiques. Les « réseaux neuronaux récurrents », par exemple, ont démontré leur capacité à effectuer une séquence de calculs pour traiter la parole ou de la vidéo. Les données séquentielles sont constituées d'unités (phonèmes ou mots dans le cas de la parole) qui se suivent. La

façon dont les réseaux neuronaux récurrents traitent leurs entrées présente une certaine ressemblance avec la façon de fonctionner du cerveau. Les signaux qui se propagent parmi les neurones changent constamment à mesure que les entrées sensorielles sont traitées.

Un traitement dynamique

Les réseaux récurrents parviennent à prédire quel sera le mot suivant dans une phrase et peuvent ainsi produire une nouvelle séquence de mots, l'un après l'autre. Ils peuvent aussi s'atteler à des tâches plus complexes. Après avoir « lu » tous les mots d'une phrase, le réseau est à même de deviner le sens de la phrase entière. Un réseau récurrent distinct peut alors utiliser le traitement sémantique du premier réseau pour traduire la phrase dans une autre langue.

La recherche sur les réseaux neuronaux récurrents a connu sa propre période de stagnation à la fin des années 1990 et au début des années 2000. Mes propres travaux théoriques suggéraient qu'ils rencontreraient

des difficultés à apprendre à récupérer de l'information du passé lointain, à savoir les premiers éléments de la séquence en cours de traitement (un peu comme si vous essayiez de réciter les premières phrases d'un livre alors que vous venez d'atteindre la dernière page). Mais certains de ces problèmes ont été résolus grâce à des techniques consistant à stocker l'information de façon à ce qu'elle persiste durablement. Ces réseaux neuronaux utilisent la mémoire temporaire (mémoire cache) d'un ordinateur pour traiter des informations multiples et dispersées, comme des idées contenues dans différentes phrases réparties au fil d'un document.

Le retour en force des réseaux neuronaux à apprentissage profond après le long sommeil de l'intelligence artificielle n'est pas uniquement un triomphe technique. C'est aussi une leçon de sociologie des sciences. En particulier, il souligne la nécessité de soutenir des idées qui font fi du *statu quo* technologique et d'encourager la tenue d'un portefeuille de recherches diversifié, laissant de la place à des champs provisoirement passés de mode. ■

BIBLIOGRAPHIE

Y. LeCun et al., **Deep learning**, *Nature*, vol. 521, pp. 436-444, 2015.

Y. Bengio et al., **Representation learning: A review and new perspectives**, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35(8), pp. 1798-1828, 2013.

A. Krizhevsky et al., **ImageNet classification with deep convolutional neural networks**, *26th Annual Conference on Neural Information Processing Systems (NIPS 2012)*, 2012.

SUR LE WEB

Cours de Yann LeCun au Collège de France, 2015-2016 : <https://www.college-de-france.fr/site/yann-lecun/index.htm>

PARTEZ EN CROISIÈRE DANS LES MERS DU MÉSOZOÏQUE

avec le 2^e hors-série *Espèces*

En kiosque mi-juillet 2016



À la lumière des dernières découvertes, les plus grands spécialistes vous proposent d'aller à la rencontre des contemporains marins des dinosaures. Découvrez la faune extraordinaire qui a régné pendant 185 millions d'années sur les océans du Trias, Jurassique et Crétacé.

Sous la direction de Nathalie Bardet et Peggy Vincent (MNHN), reconstitutions de Camille Renversade.

Revue trimestrielle dédiée
aux sciences de la Vie et de la Terre

www.especies.org

84 pages – 8,90 €

Vente par correspondance www.bopresse.fr/especies.htm

Tél. 05 65 81 54 86 / contact@bopresse.fr

Également disponible par abonnement, en PDF
et application smartphone/tablette

[f revue.Especies](https://www.facebook.com/revue.Especies)

[e http://especies-revue.blogspot.fr](http://especies-revue.blogspot.fr)