

sur les exemples connus (il suffit d'une mémoire de capacité suffisante). Mais cela n'a pas grand intérêt. Même si on accumule des millions d'images de coucher de soleil, le nombre d'images possibles d'une telle scène est infini. Ainsi, l'objectif de l'apprentissage est d'être capable de donner de bonnes réponses sur de nouveaux exemples, inconnus du système, en généralisant les exemples connus. Le bon niveau de généralisation dépend du contexte. Ainsi, on peut vouloir reconnaître la notion d'arbre en général, mais on peut aussi vouloir distinguer les chênes des hêtres, ou vouloir reconnaître le hêtre d'un jardin donné...

Une structure faite de couches multiples

Un tel algorithme doit aussi s'appuyer sur certaines hypothèses relatives aux données et à ce que pourrait être une solution possible à un problème donné. Par exemple, le logiciel doit intégrer comme principe que si des données d'entrée d'une fonction particulière sont semblables, leur sortie ne devrait pas être radicalement différente : modifier quelques pixels sur une image de chat ne devrait pas transformer l'animal en chien.

De telles hypothèses faites sur des images sont utilisées par les réseaux de neurones dits convolutifs. Ces programmes sont à l'origine du renouveau de l'intelligence artificielle. Les réseaux neuronaux convolutifs utilisés dans l'apprentissage profond comprennent de nombreuses couches de neurones organisées de telle manière que le logiciel sera robuste vis-à-vis des changements dans l'objet qu'il tente d'analyser. Il sera capable de reconnaître l'élément même s'il a un peu bougé par exemple. Ainsi, un réseau bien entraîné sera capable de reconnaître un visage que les photographies représentent sous des angles variés.

La configuration d'un réseau convolutif s'inspire de la structure en couches multiples du cortex visuel, c'est-à-dire la partie de notre cerveau qui reçoit les signaux des yeux. Ces nombreuses couches de neurones virtuels justifient le qualificatif de « profond » donné à ces réseaux et confèrent à ces derniers une meilleure capacité à appréhender le monde environnant.

L'apprentissage profond est devenu envisageable il y a une dizaine d'années grâce à certaines innovations, alors que



LA MACHINE GAGNE AU JEU DE GO.

En mars 2016, le logiciel AlphaGo a affronté le joueur coréen, Lee Sedol, considéré comme l'un des meilleurs du monde (ci-dessus au milieu, avec Sergey Brin, cofondateur de Google, et Demis Hassabis, cofondateur de Google DeepMind qui a développé AlphaGo). Fondé sur l'apprentissage profond, AlphaGo a battu l'humain quatre manches à une (ci-contre, la seule partie gagnée par Lee Sedol).



l'intérêt pour l'intelligence artificielle était au plus bas. Un organisme canadien financé par le gouvernement et par des fonds privés, l'ICRA (Institut canadien de recherches avancées), a contribué à ranimer la flamme en soutenant un programme dirigé par Geoffrey Hinton, de l'université de Toronto, et dont faisaient partie Yann LeCun de l'université de New York et du centre de recherche de Facebook à Paris, Andrew Ng de l'université de Stanford, Bruno Olshausen de l'université de Californie à Berkeley, moi-même et plusieurs autres. En raison du scepticisme ambiant, il était difficile à cette époque de publier des articles sur le sujet et même de persuader des étudiants de faire leur thèse dans ce domaine. Mais nous étions convaincus qu'il fallait persévérer dans cette voie.

La réticence vis-à-vis de l'intelligence artificielle découlait initialement de l'idée que la création de réseaux neuronaux était sans espoir à cause de la difficulté à optimiser leur performance pour apprendre efficacement.

L'optimisation est une branche des mathématiques qui vise à trouver la meilleure combinaison de paramètres, ici les poids synaptiques des neurones virtuels, pour atteindre un certain objectif. Lorsque la relation qui lie les paramètres et l'objectif

est assez simple – plus précisément lorsque l'objectif est une « fonction convexe » des paramètres – alors on peut améliorer progressivement les paramètres et l'on parle d'optimisation convexe. La procédure d'entraînement est répétée jusqu'à ce que les paramètres s'approchent aussi près que possible des valeurs qui produisent le meilleur résultat. On cherche en fait un minimum global, à savoir le jeu de paramètres qui permet d'atteindre la valeur la plus basse (et la meilleure) de l'écart à l'optimum.

Des obstacles surmontés peu à peu

En général, la situation n'est pas aussi simple, la fonction reliant les paramètres à l'objectif qu'ils atteignent n'étant pas convexe. Or l'optimisation non convexe représente un défi bien plus difficile. De nombreux chercheurs pensaient qu'il n'était pas possible de le relever. L'algorithme d'apprentissage pourrait en effet se retrouver coincé dans un minimum local, d'où il ne pourrait sortir s'il ajuste les paramètres par petites touches.

Or mes collègues et moi avons montré que si le réseau de neurones est assez grand, le problème des minima locaux est fortement