



INTELLIGENCE ARTIFICIELLE

UNE INTELLIGENCE ARTIFICIELLE QUI RECONNAIT LES PAYSAGES DU MONDE ENTIER !



Photo CC-BY-NC by steele12



(a)



Photo CC-BY-NC by edwin.11



(b)



Photo CC-BY-NC by jonathank



(c)

Le type de photo présenté au logiciel (gauche) et la localisation probable trouvée par celui-ci (droite). Pour la tour Eiffel, aucun doute possible, pour le paysage de crique, deux lieux possibles, pour la photo de plage, des dizaines d'origines possibles (Tobias Weyand et al., Google)

Google a encore frappé. Des chercheurs de son laboratoire *Computer Vision & Machine Learning* ont conçu un "réseau de neurones profond" qui reconnaît sur des photos le lieu où elles ont été prises, qu'il s'agisse d'un paysage, d'une rue ou d'un monument. Ce n'est pas une reconnaissance à 100 %, loin de là, mais ses capacités dépassent celles des humains.

Déjà connue pour son aptitude à reconnaître et catégoriser les images (visages, objets, etc.), cette technologie prouve encore une fois son excellence dans les tâches "cognitives" visuelles et illustre la mainmise des géants du net dans le domaine de l'Intelligence artificielle.

UNE INTELLIGENCE ARTIFICIELLE BIO-INSPIRÉE...

Les réseaux de neurones artificiels dits "[profonds](#)", car ils contiennent plusieurs couches de neurones, sont à la mode

depuis le milieu des années 2000 surtout grâce à l'entrée des grandes firmes du Web comme Google et Facebook dans le domaine de recherche en IA.

Inspirés du fonctionnement du cortex visuel des animaux, ces algorithmes nous épatent déjà par leur capacité à reconnaître les [visages](#) ou [les scènes](#), ou encore à créer des ["concepts" visuels](#) abstraits représentant de grandes catégories comme celle des "chats", des "humains", des "chaises", etc.

... QUI APPREND À RECONNAÎTRE LE MONDE

Leur secret : ils réussissent à "filtrer", de couche en couche, les occurrences visuelles les plus significatives d'une image, ce grâce à une phase d'apprentissage préalable surveillé par un opérateur, qui leur sert de base de repère.

Ainsi, si l'on veut qu'ils sachent reconnaître des chats, il suffit de leur en faire voir des millions en leur indiquant qu'il s'agit d'un chat (apprentissage) : ils apprennent alors à reconnaître les similitudes entre ce qu'ils ont appris et l'image d'un chat qu'ils n'auraient pas vu précédemment. C'est de la statistique et du calcul de probabilités, mais sur un volume de données immense.

LA PHASE D'APPRENTISSAGE : 91 MILLIONS DE PHOTOS

[Dans la cas présent](#), durant sa phase d'apprentissage l'algorithme des chercheurs – nommé PlaNet – a vu défiler quelque 91 millions d'images de différents lieux de la Terre (photos de rues, d'immeubles, de paysages, etc.) contenues dans différentes bases de données du Web.



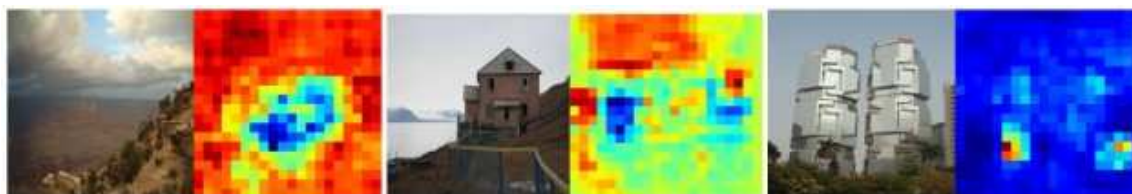
Quelques paysages soumis au logiciel. En jaune, leur localisation réelle, en vert la localisation supposée donnée par des humains, en bleu la localisation calculée par le logiciel (Tobias Weyand et al., Google)

Bien sûr, ces images étaient accompagnées des informations de localisation afin que le réseau associe les traits communs entre plusieurs images à une localisation unique – par exemple, des images des toits de Paris ont des caractéristiques communes que le logiciel "extraît" par filtrage et associe à la localisation de Paris.

LES RÉSULTATS DES TESTS

Puis les chercheurs ont testé sur 34 millions d'images non utilisées pour l'apprentissage la capacité du logiciel à les "rapprocher" statistiquement de ce qu'il avait appris, en le corrigeant quand il se trompait.

Enfin, ils sont allés chercher 2,3 millions d'images dans la base de donnée Flickr pour mesurer leur compétence. Résultat : PlaNet est capable de donner la ville d'origine pour 3,6 % d'images représentant des rues et 10.1 % d'images représentant des vues plus larges.



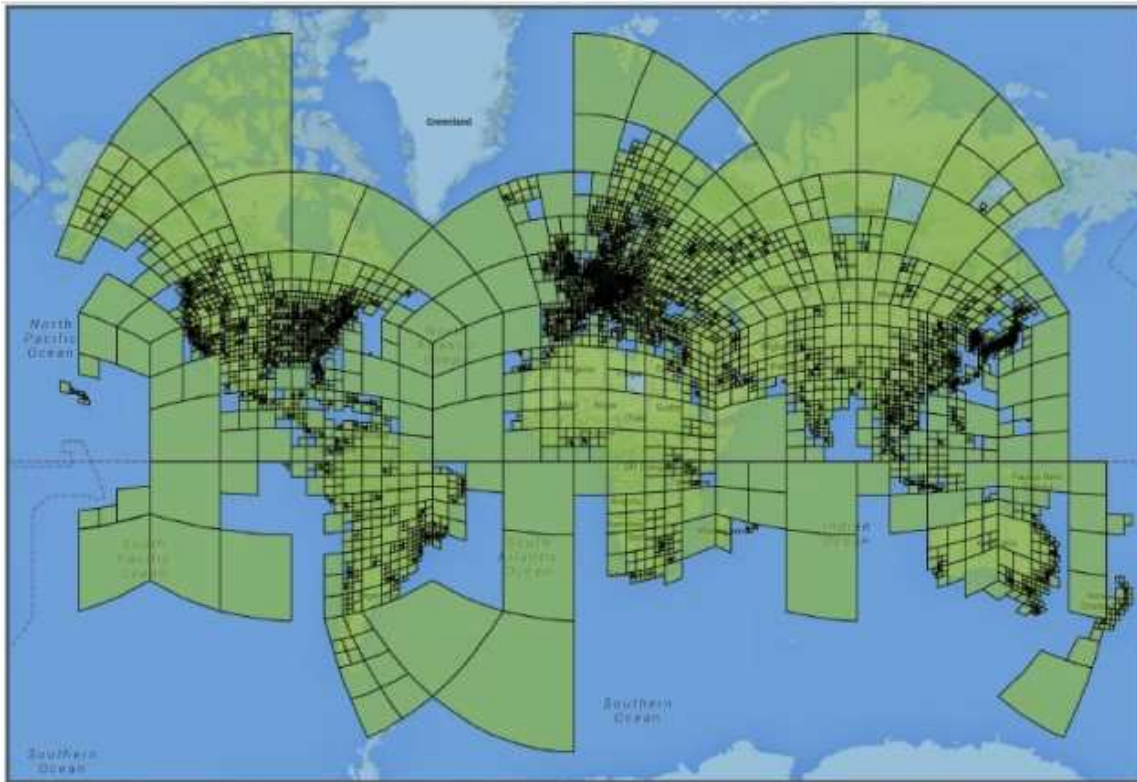
Représentation du calcul de probabilités du logiciel : en bleu, le maximum de correspondance entre l'image et une localisation calculée par le logiciel, en rouge les éléments de l'image que le logiciel ne reconnaît pas comme appartenant à la localisation calculée (minimum de probabilité). Grand Canyon, Norvège, Shanghai (T. Weyand et al., Google).

Mais s'il se trompe souvent de ville, il peut néanmoins resituer le pays d'origine de 28,4 % des images et le continent d'origine de 48 % d'entre elles. Cela semble peu, mais mis au défi par 10 humains ayant beaucoup voyagé, PlaNet les a battu 56 % des fois.

LE MONDE SUBDIVISÉ EN 26 000 SECTEURS

Bien sûr, l'entraînement du logiciel n'a pas été aléatoire : les photos de villes et des lieux habités ou très fréquentés étant

favorisé par rapport à celles de grands espaces déserts et sous-représentés en termes de photos circulant dans le Web (les Océans et les Pôles ayant été carrément ignorés).



Les quelque 26 000 secteurs que les chercheurs ont déterminé pour découper le monde (T. Weyand et al., Google).

Concrètement, les chercheurs ont divisé le monde en 26 263 secteurs de tailles différentes : les grandes villes contenant plusieurs de ces secteurs, tandis que les espaces peu peuplés ou fréquentés n'en contenant qu'un seul (voir carte). Pour chaque secteur, le nombre de photos montrés au logiciel durant l'entraînement était le même.

UNE INTELLIGENCE AVANT TOUT STATISTIQUE

Ce résultat montre encore une fois que l'approche massivement statistique dans le domaine de l'IA est très efficace, surtout depuis l'invention des réseaux de neurones profonds qui apprennent à filtrer et isoler des schémas récurrents à l'aide d'un entraînement sollicitant des milliards de données.

Et il n'est donc pas étonnant que les géants du Web, propriétaires de la plupart de ces données et disposant de moyens de calcul extraordinaires, soient aujourd'hui à la pointe de ce type de recherche.

Román Ikonicoff

>Lire aussi:

- [Victoire historique de l'intelligence artificielle au jeu de go](#)
- [Voyez les hallucinations d'une Intelligence artificielle !](#)
- [une intelligence artificielle résout seule une énigme biologique](#)

> Lire également dans le site des Grandes Archives de Science & Vie :

- [Moteurs de recherche : voici qu'ils légendent les images](#) – S&V n°1169 – 2015. Grâce à la fusion de plusieurs technologies IA, les chercheurs de chez Google ont conçu un système qui reconnaît et légende les images.



- [Reconnaissance faciale : le triomphe de l'intelligence artificielle](#) – S&V n°1162 – 2014 – Le domaine de l'Intelligence artificielle progresse à très grande vitesse grâce aux ressources fournies par les réseaux numériques. L'exemple avec la reconnaissance automatique des visages.



- [Robots : leur intelligence dépasse déjà la nôtre](#) – S&V n°1166 – 2014 – Peu à peu et sans grandes vagues, l'intelligence artificielle et la robotique sont sorties de l'échec relatif des années 1980 pour finir par devenir des acteurs essentiels dans nos activités. Ils nous dépassent déjà.



VOUS AIMEREZ AUSSI

contenus sponsorisés Outbrain

AILLEURS SUR LE WEB

contenus sponsorisés Outbrain

.....

.....

.....

.....

.....

A VOIR SUR SCIENCE&VIE

Vidéo : la plus grande éponge jamais observée

Reconnaître un visage humain n’a rien d’exceptionnel, même les poissons y arrivent

Le ciel du mois de juin 2016

Qui va gagner les Jeux olympiques de Rio ? Un modèle informatique aurait la réponse

Quand nous lisons sur un écran, nous pensons autrement !

Un arc-en-ciel a-t-il forcément sept couleurs ?

Recommandé par

