

L' Intelligence Artificielle expliquée

L'intelligence artificielle, thème abondamment utilisé en science-fiction et plus récemment à la mode dans les "unes" des journaux, marie fantasmes de toute puissance des machines et peurs des dérives. Dans cet article, **nous faisons le point sur l'intelligence artificielle en traitement d'image, les possibilités offertes par la technologie et ses limites quant à la photographie.**

Les origines de l'intelligence artificielle

Avant d'être une anticipation de la science-fiction, l'intelligence artificielle est un fantasme d'un des pères de l'informatique, Alan Turing, dans un article scientifique publié en 1950. Contrairement aux algorithmes conventionnels qui exécutent une succession d'opérations programmées par un être humain, l'intelligence artificielle repose sur le principe du système informatique apprenant par lui-même. On code donc un mécanisme d'apprentissage et on fournit au système des données pour qu'il s'initie à faire certaines opérations.

Petit point sur les appellations...

Au cours de sa longue évolution, le vocabulaire de l'intelligence artificielle s'est diversifié. Vous pouvez donc entendre parler de manière équivalente d'intelligence artificielle ou de *machine learning* (apprentissage automatique). Le terme de *deep learning* (apprentissage profond) est aussi très utilisé et correspond à une catégorie de réseaux de neurones (*neural network*). Précisons également qu'une grande partie du vocabulaire de l'intelligence artificielle s'exprime dans la langue de Shakespeare, bien que nous ayons souhaité utiliser le français autant que possible.

Pourquoi l'intelligence artificielle est devenue une réalité ?

L'intelligence artificielle est un sujet de recherches depuis près de 70 ans et

pourtant le domaine ne commence à avoir des applications pratiques en photographie qu'aujourd'hui. Dans les faits, le point bloquant des systèmes d'intelligence artificielle est le besoin de données. Il en faut beaucoup, si possible de qualité, avec des annotations dans certains cas... Pour des applications photographiques, on parle au bas mot en centaines de milliers d'images pour entraîner un algorithme.

Où trouver toute cette information ? Sur Internet, bien sûr ! Et c'est là l'explication principale du boom que connaît l'intelligence artificielle depuis quelques années, en grande partie dû à la présence d'une grande quantité de données en ligne. D'ailleurs, et ça n'est pas très étonnant, les champions de la recherche en apprentissage automatique sont les grands groupes Google, Facebook, Microsoft...

La puissance de calcul nécessaire pour l'entraînement de l'algorithme est aussi importante. L'utilisation de cartes graphiques et le calcul parallèle ont permis une accélération des temps d'entraînement, ce qui a autorisé d'envisager des architectures de réseaux d'apprentissage plus complexes (le fameux apprentissage profond/*deep learning*) et des traitements de problèmes utilisant des images de plus grande taille.

L'intelligence artificielle en traitement des images

Comment fonctionne l'intelligence artificielle ?

L'intelligence artificielle repose sur l'utilisation de réseaux de neurones. Les neurones en question sont en fait des fonctions mathématiques qui s'inspirent lointainement des neurones humains.

Comment fonctionne un neurone ?

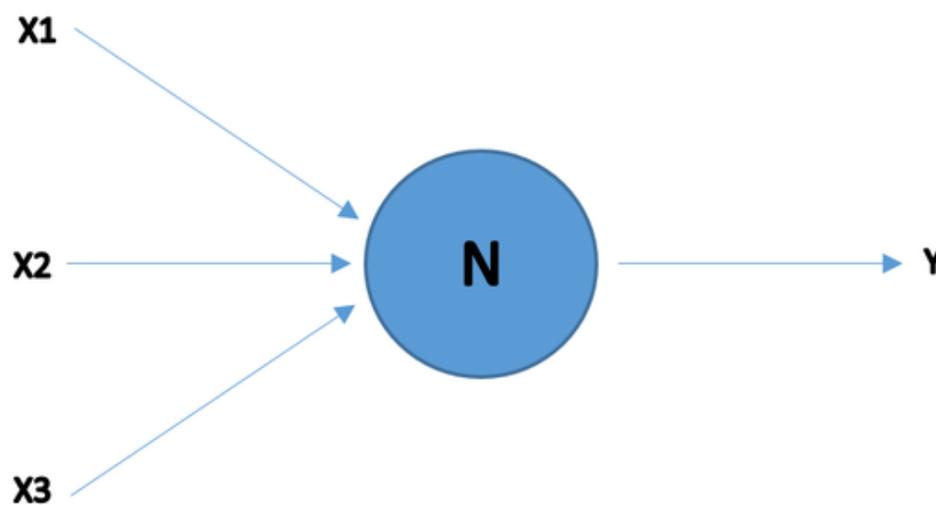
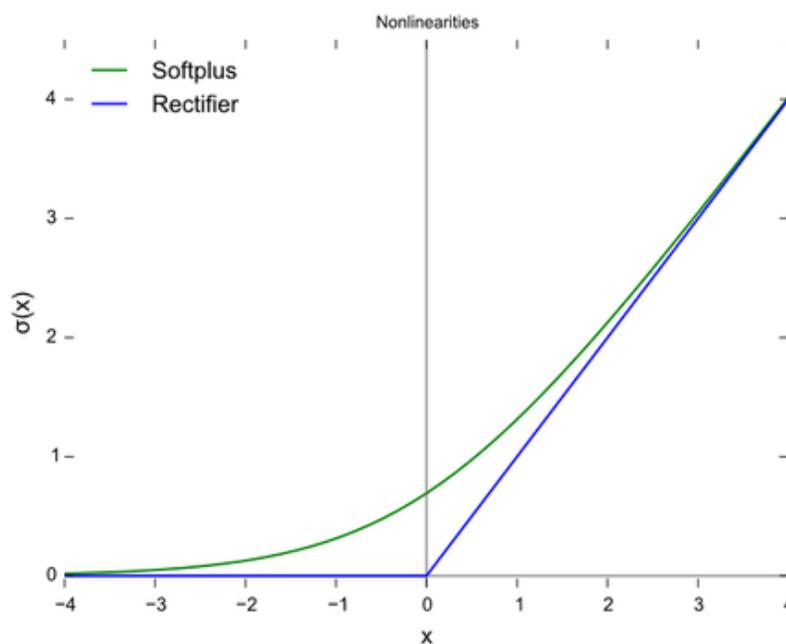


Schéma d'un neurone.

Un neurone N prend en entrée une série de valeurs X_1, X_2, \dots, X_N , et va calculer une moyenne pondérée à partir de ces valeurs. $V = P_1.X_1 + P_2.X_2 + \dots + P_N.X_N$ avec P_1, P_2, \dots, P_N des poids positifs de somme unitaire. Les poids P_1, P_2, \dots, P_N sont autant de paramètres qui seront ajustés pendant la phase d'entraînement.

Le neurone possède une fonction d'activation qui va permettre de faire passer ce résultat en sortie en fonction de la valeur. Ici, par exemple, vous pouvez voir la fonction d'activation *ReLU* (il en existe des dizaines).



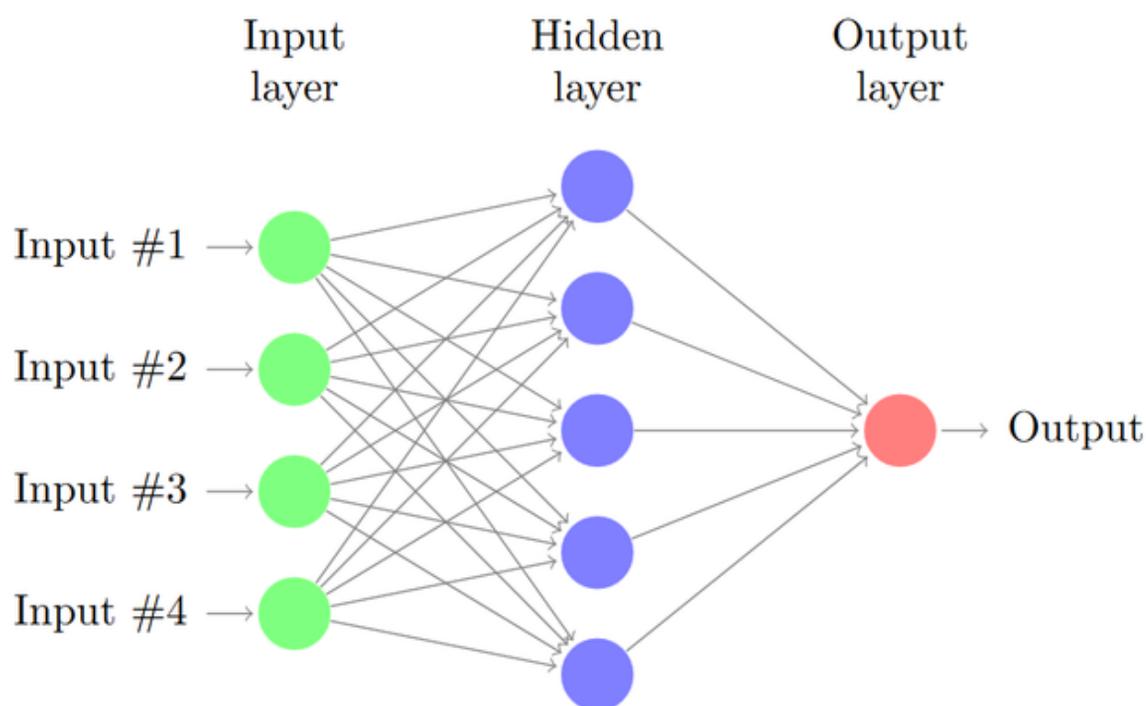
Fonction d'activation d'un neurone.

En sortie de ce neurone, on a donc une nouvelle valeur Y que l'on peut, par exemple, passer à un autre neurone.

Comment fonctionne un réseau ?

Un réseau de neurones est un ensemble de neurones individuels liés entre eux pour former un système organisé capable de prendre une décision. En pratique, il y a beaucoup d'architectures différentes parce que les besoins sont différents et nous nous contenterons de présenter un schéma général.

Les réseaux sont organisés en couches (*layers*) qui regroupent des neurones en parallèle. Chaque couche suivante récupère les sorties de la couche précédente en entrée pour ses neurones. Et ainsi de suite...



Réseau de neurones comportant plusieurs couches.

On peut empiler les couches de neurones à l'infini *a priori*. Seulement, plus il y a de couches de neurones, plus il y a de paramètres à optimiser dans la phase d'entraînement. Et plus il y a de paramètres à optimiser, plus il faut d'images d'entraînement pour bien faire.

L'entraînement d'un réseau de neurones consiste à faire varier progressivement l'ensemble des paramètres des neurones pour une grande quantité d'images et/ou de données. Il faut définir une métrique qui permettra de savoir si un nouvel ensemble de paramètres est meilleur qu'un autre – c'est un processus d'optimisation. Le processus d'entraînement peut facilement prendre plusieurs jours sur des cartes graphiques performantes.

Pour de grands réseaux de neurones, le risque d'instabilité du

comportement est élevé. On parle d'*overfitting* ou de surapprentissage. Des stratégies existent pour améliorer la stabilité, telle la technique de *dropout* (abandon, en français) qui consiste à ignorer aléatoirement certains neurones pendant l'entraînement. Cela dit, l'instabilité reste l'un des principaux défis des concepteurs de réseaux de neurones.

Comment utilise-t-on un réseau ensuite ?

Une fois que le réseau a été entraîné à l'aide de milliers, voire de millions de données, on enregistre des tableaux de coefficients qui correspondent aux paramètres optimaux des différents neurones. Ces paramètres peuvent ensuite être utilisés pour une nouvelle donnée. Chez l'utilisateur, un processeur (idéalement graphique) va simplement servir à appliquer les opérations décrites par les neurones. Le processus est relativement rapide pour des images de taille raisonnable. Dans le cas de systèmes embarqués, des [systèmes dédiés](#) à l'apprentissage profond commencent à apparaître pour proposer une architecture électronique optimisée.

En vidéo

Pour ceux qui souhaiteraient une présentation vulgarisée de qualité pour mieux comprendre, vous pouvez regarder cette vidéo de David Louapre ([Science étonnante](#)).

La photographie numérique se prête-t-elle à l'usage de l'intelligence artificielle ?

La photographie numérique est compatible avec l'usage de l'intelligence artificielle (c'est numérique, les techniques de filtrage de convolution sont déjà bien connues et peuvent être exécutées rapidement). Le seul souci avec les photographies contemporaines réside dans leur taille. Un grand nombre de papiers de recherche travaillent pour des raisons d'efficacité sur des images de petite taille (typiquement 100 x 100 px). Or, les solutions algorithmiques à 100 x 100 px ne sont pas toujours utilisables pour des images de plusieurs dizaines de mégapixels.

Quelles applications en photographie ?

Classement d'images

Notation d'images

La notation d'image est un Graal pour de nombreux sites comme Flickr,

Facebook ou Instagram. L'idée est de pouvoir détecter automatiquement les “belles” photographies pour pouvoir les mettre en avant. On peut donc entraîner une machine à noter les images.

Les résultats sont assez prometteurs : on peut citer le projet NIMA ([Neural Image Assessment](#)), une technologie de Google (AR ranks photos on their technical and aesthetic quality cf DPReview). Vous pouvez aussi tester par vous-même les performances obtenues par [Everypixel](#). Pour autant, l'esthétique est toujours une notion discutable et elle ne fait pas tout dans une photo, encore moins son prix. L'image *99cents* d'Andreas Gursky, vendue plus de 3 millions de dollars, n'obtient ainsi qu'une note esthétique de 0,2 % chez Everypixel...

Organisation intelligente avec détection des sujets

De nombreux logiciels, dont Adobe Lightroom, intègrent des fonctionnalités de classement, mais aussi des étiquettes pour permettre de chercher *a posteriori* toutes les photos liées à un même endroit, une même personne, un même type de photos, etc., de sorte qu'en cherchant “Photos de paysage en Camargue au coucher du soleil”, vous puissiez rapidement remettre la main sur la photo souhaitée.

La recherche en reconnaissance d'objet permet déjà de faire de nombreuses choses : reconnaître le contenu de la scène dans la plupart des cas ([Google Caption Generator](#) ou encore la [présentation TED de Fei Fei Li](#)).

Du côté des portraits, des méthodes de reconnaissance et de classification des émotions existent. Vous pouvez, par exemple, tester la solution proposée par [Microsoft Azure](#).

Redimensionnement et super-résolution

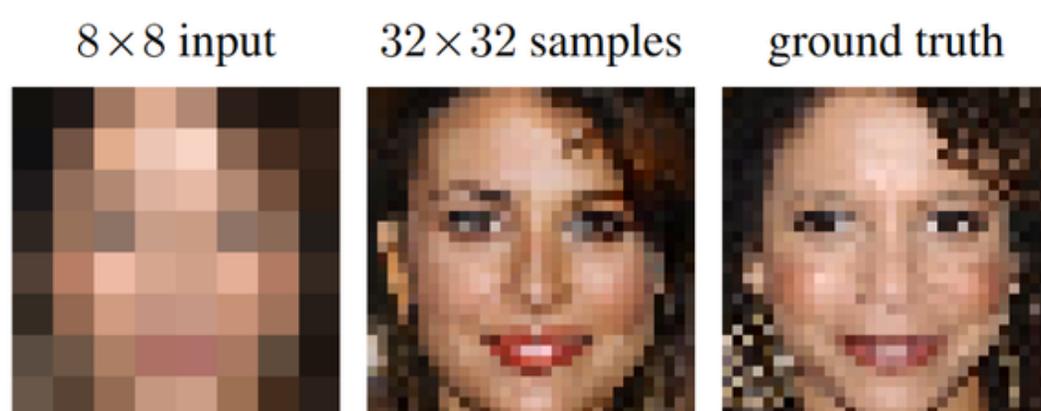
Entendons-nous bien : il n'y a pas de super-résolution possible à partir d'une seule image et d'un éclairage classique. La super-résolution nécessite en théorie une série d'images (vidéos, par exemple) ou bien un éclairage très particulier (microscopie à fluorescence). Dans le cas de la

photographie, le capteur a enregistré certaines informations et aucun traitement numérique ne pourra faire apparaître de la résolution qui n’existait pas. Au mieux – et c’est ce sur quoi repose la super-résolution à apprentissage profond –, on peut demander à un algorithme de “rêver” les détails qui manquent, quitte à ce que ces détails soient partiellement faux.

Pour donner une analogie, pensez au portrait-robot réalisé par les policiers à partir d’une description relativement vague du suspect. Les policiers n’ont jamais vu le visage décrit, mais ils peuvent lui donner un aspect réaliste. En pratique, l’apprentissage profond fonctionne bien pour effectuer des agrandissements d’images. L’image ci-dessous est tirée d’un [article scientifique de 2014](#) et montre que sur cet exemple la technique offre de meilleurs résultats que l’interpolation bicubique.

Super-résolution SRCNN.

Dans le cas particulier des portraits, Google a proposé un [algorithme](#) pour augmenter la résolution des images. Il est assez ironique de parler de super-résolution pour des images de 32 x 32 px, mais il y a effectivement une meilleure résolution que sur l’image fournie en entrée (8 x 8) avec un résultat assez réaliste.



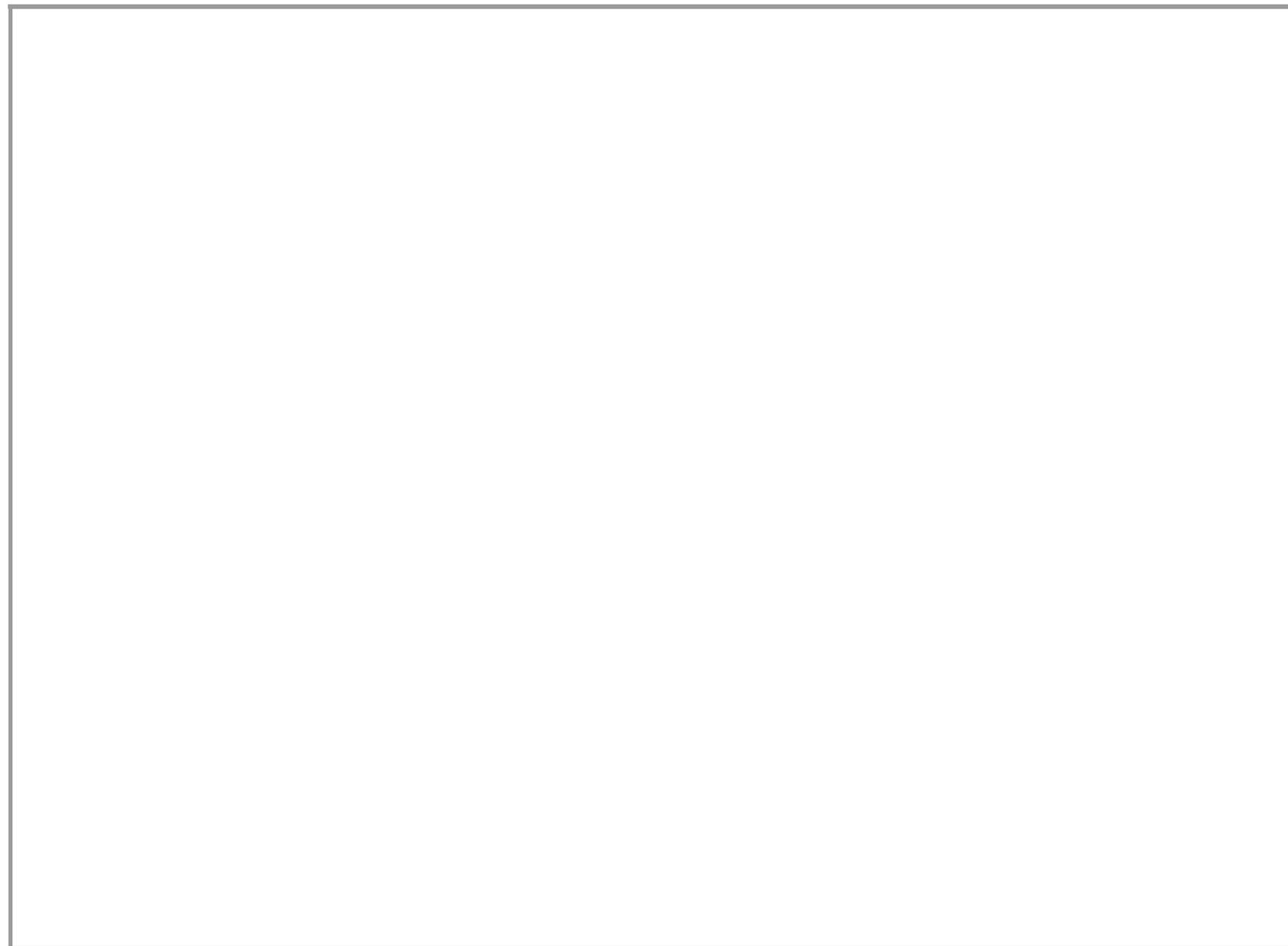
Super-résolution appliquée au portrait – Recherche Google.

Colorisation des images en noir et blanc

La colorisation des images attire aussi une grande quantité de recherches (par exemple, cet [article de 2016](#)), bien que la problématique soit moins

utile au photographe.

Il semblerait qu'Adobe s'intéresse également au sujet et il ne serait pas surprenant de voir apparaître une option de ce genre dans une version de Photoshop à venir. Vous pouvez jeter un œil à la présentation du projet Scribbler.



Débruitage et dématricage

Le débruitage d'image est un sujet de recherche toujours actif. Des chercheurs de Disney ont, par exemple, œuvré à débruiter des images en utilisant l'apprentissage profond. Les détails sont disponibles dans leur [article publié en 2017](#).

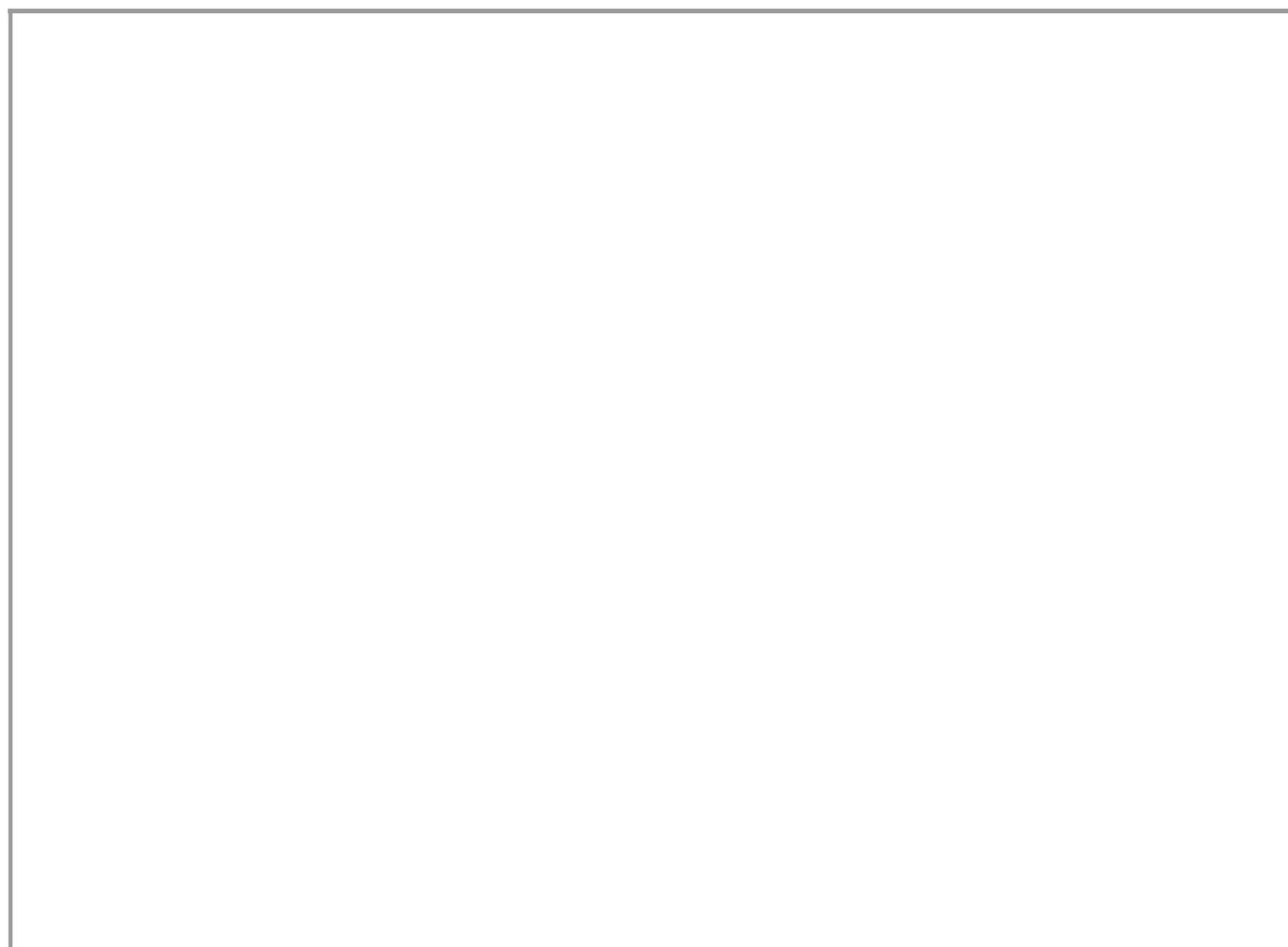
Débruitage par réseau de neurones – Disney.

Plus généralement, le débruitage est souvent effectué en même temps que le dématricage. L'idéal étant de réaliser les deux processus d'un coup avec l'aide de réseaux de neurones. Vous pouvez trouver un [exemple de](#)

[recherche](#) sur le sujet fait par le MIT. L'intérêt des réseaux de neurones pour ce problème n'est pas forcément évident pour l'instant, mais les résultats pourraient s'améliorer dans le futur.

Masque de sélection et inpainting

Les utilisateurs de Photoshop sont probablement déjà au courant d'une nouvelle technique de sélection de sujet utilisant de l'intelligence artificielle. L'idée consiste à proposer une alternative à la baguette magique pour permettre une sélection rapide et performante de sujets dotés de grandes dimensions.



La sélection de sujet pourrait être très utile [en vidéo](#), notamment dans un but d'incrustation. En photographie, l'option peut servir pour remplacer un objet non désiré, possiblement par une technique d'inpainting. L'inpainting connaît d'ailleurs quelques améliorations liées à l'apprentissage profond (voir par exemple cet [article](#)).

Estimation de la profondeur et bokeh

L'estimation de la profondeur est un objectif prioritaire de la recherche en réseaux de neurones, à cause notamment de ses applications en [conduite autonome](#). En tant qu'observateur, vous pourriez à partir d'une photographie imaginer avec une bonne précision la distance à tous les objets de la scène. Il devrait donc être possible de créer un réseau de neurones qui fasse de même à partir d'une image ou d'une vidéo (pour le moment, ce n'est pas vraiment chose faite).

En photographie, obtenir une carte de profondeur, c'est pouvoir jouer sur la profondeur de champ. Une option très prisée des concepteurs de smartphones par le biais des modes "portrait". Google utilise ainsi de l'apprentissage profond pour ses [téléphones Pixel 2](#).

Réduction de profondeur de champ obtenue par réseau de neurones – Google Pixel 2.

Retouches automatiques

L'ajustement de la luminosité et du contraste peut être fastidieux et quelques entreprises travaillent à produire des systèmes de traitement par lot utilisant l'intelligence artificielle. C'est le cas particulier du logiciel [PhotoLemur](#) qui propose de procéder à une série d'améliorations automatiquement et sans assistance humaine. Les retouches automatiques sont aussi très prisées des concepteurs de smartphones qui, souvent de manière très excessive, promettent ainsi de [rivaliser avec les appareils photo reflex](#).

Dernière initiative du domaine, des chercheurs de Google se sont amusés à créer un algorithme pour retoucher les images de Google Street View et leur donner un [aspect plus artistique](#). Jugez plutôt !

Google Street View s'essaye à la photographie d'art...

Les limites...

On peut faire à peu près tout avec de l'apprentissage profond. L'ensemble des problématiques du traitement d'image classique peut être revisité à la sauce intelligence artificielle. Mais est-ce que tout mérite pour autant de passer par de l'apprentissage profond ?

Certains problèmes sont déjà résolus

Ce qui compte en traitement d'image pour les utilisateurs, c'est le résultat image, quels que soient la méthode et accessoirement le temps de traitement. Et il y a des problèmes qu'on sait très bien résoudre sans en passer par un apprentissage sur des millions d'images (dématriçage, ajustement d'histogramme, détection de visage, correction des yeux rouges...). La partie d'entraînement de l'algorithme est gourmande en ressources et prend du temps. À l'usage, les réseaux de neurones peuvent être rapides, mais rien ne justifie d'utiliser de l'intelligence artificielle pour toutes les opérations.

Cas en coin

Le comportement de l'intelligence artificielle en dehors des cas d'entraînement peut être très aléatoire (*overfitting* ou surapprentissage), d'autant plus que le nombre de neurones est élevé. On peut difficilement prévoir comment un réseau de neurones va se comporter devant une image vraiment différente de la base de données utilisée pour son entraînement. Et ce gâchement pourrait bien faire n'importe quoi...

Le risque d'un détournement marketing

L'appellation "Intelligence Artificielle" est très utilisée, à tort et à travers. Nous voudrions conclure en rappelant que ce n'est qu'un type d'algorithmes parmi d'autres. Le fait qu'un logiciel utilise de l'intelligence artificielle ne garantit pas qu'il produise des bons résultats. L'appellation ne doit pas être dévoyée dans un but marketing pour faire croire au consommateur qu'intelligence artificielle est synonyme de meilleures performances, il s'agit simplement d'une autre manière de résoudre les problèmes de traitement d'image.

Pour aller plus loin

Tout public

→ [Conférence Futurs de l'Intelligence Artificielle](#) – École Normale Supérieure 2016

→ [Article : Comment le *deep learning* révolutionne l'intelligence artificielle ?](#) - Le Monde 2015

Avancé

→ [Conférences de Yann LeCun](#) pour le Collège de France