



# Queste de savoir

Quand l'IA médicale pose des questions  
éthiques

---

19 février 2021



# Table des matières

1.	Décortiquons l'article . . . . .	1
1.1.	Un peu de dermato . . . . .	1
1.2.	L'apprentissage automatique . . . . .	4

En février 2017 est paru un article qui a fait grand bruit dans la communauté scientifique, à la fois du côté de la médecine et de l'informatique. Publié dans *Nature*<sup>1</sup>, l'article est baptisé *Demartologist-level classification of skin cancer with deep neural networks*. Il est disponible gratuitement sur [le site de l'auteur principal](#) [↗](#) pour ceux que ça intéresse (et assez simple à lire).

La publication explique qu'en utilisant un système d'apprentissage automatique (*machine learning* en anglais), ses auteurs sont parvenus à réaliser des diagnostics de cancer de la peau à partir de photos de lésions dermatologiques avec une précision comparable à celle de dermatologues expérimentés. Ces premiers tests se sont faits sur des cas cliniques (une photo accompagnée d'un diagnostic) pour construire la base de référence de l'algorithme puis valider son efficacité.

Tout semblait aller pour le mieux dans le meilleur des mondes alors les concepteurs ont décidé de passer à une phase de test grandeur nature avec de vrais patients, mais ils ont alors été confrontés à un problème éthique qu'ils ont jugé majeur et qui les a conduits à interrompre au moins temporairement le développement et le test de cet outil et ce qui a mis un sacré coup de frein à la recherche en apprentissage automatique dans le domaine médical.

## 1. Décortiquons l'article

Il ne va pas être question d'analyser l'article pour en vérifier la validité (le travail serait long et dépasserait le simple cadre de ce billet), mais plutôt de comprendre les concepts principaux qui permettent de saisir l'essentiel du propos.

### 1.1. Un peu de dermato

La **dermatologie** est la discipline médicale qui s'intéresse aux affections de la **peau**, des **muqueuses** (un tissu de surface en continuité avec la peau qui tapisse l'intérieur du corps) et des **phanères** (ongles et poils dont cheveux). Les dermatologues s'intéressent également à la **vénérologie**, qui est la spécialité dédiée aux infections sexuellement transmissibles (ou maladies vénériennes), ce qui est moins important pour notre propos.

---

1. Pour ceux qui ne connaissent pas, il s'agit probablement du journal scientifique le plus prestigieux du monde.

## 1. Décortiquons l'article

### 1.1.1. La peau

La peau est un organe de surface qui tapisse la quasi-totalité de l'extérieur de notre corps (à l'exception des extrémités des doigts recouvertes par les ongles et des yeux, recouverts par la cornée). C'est un organe très exposé aux agressions externes mais c'est également un organe très accessible à l'examen puisqu'il suffit de découvrir le patient pour y avoir un accès direct.

La peau a 3 rôles principaux:

- **protéger l'organisme des agressions extérieures** par deux mécanismes, la barrière physique qui arrête la majorité des pathogènes et une concentration de cellules immunitaires pour commencer à combattre ceux qui parviendraient à passer malgré tout.
- **communiquer à l'organisme** (majoritairement au cerveau) **des informations sur le milieu extérieur**. Les nerfs relaient les informations captées par les récepteurs de la peau telles que le toucher, la température, la douleur ou encore la proprioception<sup>2</sup>.
- **réguler la température de l'organisme**. Ça passe par deux mécanismes, la contraction (quand il fait froid, pour garder la chaleur à l'intérieur) ou la dilatation (quand il fait chaud, pour favoriser l'évacuation de la chaleur vers l'extérieur) des petits vaisseaux sanguins qui parcourent la peau et la production de sueur par les glandes sudoripares (qui permet d'évacuer la chaleur).

Elle est composée de 3 couches (de l'extérieur vers l'intérieur):

- l'**épiderme** constitué de cellules (les kératinocytes) serrées les unes contre les autres pour former une barrière physique,
- le **derme**, constitué de vaisseaux sanguins, de nerfs avec leurs récepteurs, de cellules immunitaires, de follicules pileux, de glandes sudoripares et sébacées...
- l'**hypoderme**, une couche essentiellement composée de graisse (pour amortir les chocs et tenir chaud) qui laisse passer les différentes structures (vasculaires et nerveuses) vers le derme.

### 1.1.2. Les lésions de la peau

La dermatologie est probablement l'une des spécialités médicales les plus riches qui soient<sup>3</sup>. Il y a des *milliers* de diagnostics possibles pour un simple bouton sur la peau en fonction de la couleur, de la forme, du mode d'apparition, ~~de la phase de la lune, de la position de mercure,~~ etc.

Même en excluant les lésions traumatiques (plaies et brûlures), le spectre des possibles en dermatologie est atrocement large: ça va de la simple **ecchymose** (le bleu de base) au **purpura fulminans** (des petites taches rouges avec des caractéristiques spécifiques qui signent une atteinte gravissime du système vasculaire, le plus souvent dû à une méningite à méningocoque) en passant par le **zona** (une éruption cutanée qui suit le trajet d'un nerf, dû à la réactivation d'une infection dormante au virus de la varicelle), l'**acné**, l'**érythème** (une rougeur)... Bref, la clinique dermato est loin d'être une sinécure.

Parmi ces lésions, il en existe toutefois des plus graves que d'autres. Si personne ne peut nier que le purpura est indicateur d'un problème le plus souvent sérieux, il s'agit d'un problème d'apparition rapide et dont le traitement a tout intérêt à être très rapide lui aussi. Il existe au contraire des lésions qui se développent lentement ce qui les rend d'office plus susceptibles d'être

## 1. Décortiquons l'article

surveillées par les patients eux-mêmes et légitime donc le développement d'un outil qui leur faciliterait la vie.

Parmi ces lésions, les plus graves sont probablement les cancers de la peau.

### Les cancers

Un cancer pour faire simple, c'est un ensemble de cellules qui pour des raisons très diverses vont acquérir plusieurs mutations génétiques qui vont perturber gravement leur fonctionnement et les conduire à proliférer et former des masses que l'on appelle **tumeurs**. Les tumeurs peuvent être bénignes (un kyste ou un adénome) ou bien malignes et on parle alors de cancer. Les tumeurs malignes vont perturber gravement les fonctions du tissu aux dépens duquel elles se développent et peuvent également envahir d'autres tissus, formant de nouvelles masses que l'on appelle **métastases**.

Sans traitement, les symptômes s'accumulent au fil des années et des tissus atteints, conduisant quasiment systématiquement au décès du patient à moyen ou long terme.

Il existe globalement deux types de cancers de la peau:

- le **carcinome épidermoïde cutané** (*keratinocyte carcinoma* en anglais) qui touche les kératynocytes, un des principaux types de cellules cutanées. Ce sont les plus fréquents et les moins graves.
- le **mélanome** (*melanoma* ou parfois *malignant melanoma* en anglais) qui se développe aux dépens des mélanocytes, un autre grand type cellulaire de la peau. Ils sont moins fréquents mais beaucoup plus graves.

Ces deux maladies forment des lésions particulières qui sont toutefois très ressemblantes à d'autres maladies parfaitement bénignes pour qui n'a pas l'œil averti (n'importe qui qui n'est pas dermato en gros): vous avez probablement déjà entendu dire que les grains de beauté (ou *nævi*, des tumeurs bénignes de mélanocytes) peuvent évoluer en mélanome et qu'il faut donc les surveiller.

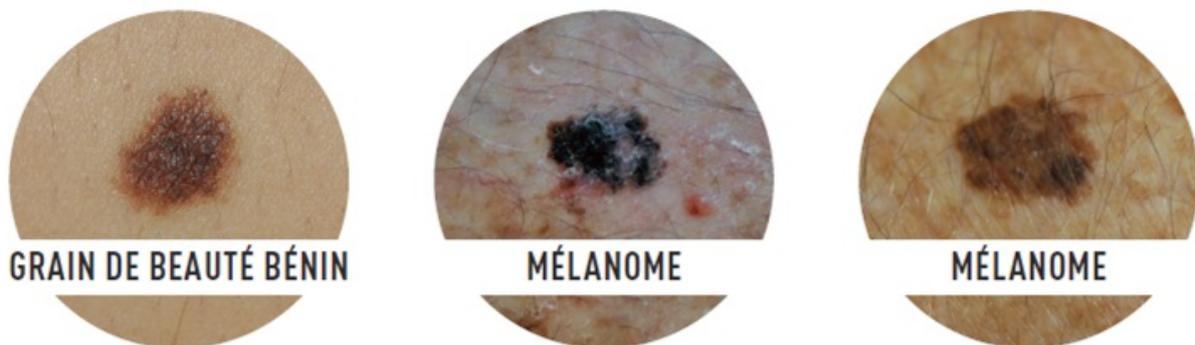


FIGURE 1.1. – Grain de beauté vs. mélanome (provient de site [vocationsante.fr](http://vocationsante.fr)), pas toujours facile de faire la différence !

Faire la différence nécessite toutefois l'avis d'un spécialiste, et c'est précisément ce que l'article auquel nous nous intéressons se propose de bousculer en permettant à chacun de faire son

## 1. Décortiquons l'article

suivi grâce à un algorithme de réseau de neurones avec une fiabilité comparable à celle d'un spécialiste.

Mais qu'est-ce que c'est un réseau de neurone? Très bonne question!

### 1.2. L'apprentissage automatique

On a tous entendu des mots très à la mode comme *Machine Learning* ou *Deep Learning*. Récemment, des algorithmes de *Deep Learning* comme AlphaZero nous ont bluffé en parvenant à battre les meilleurs joueurs de Go (un jeu de stratégie dans lequel, jusqu'à il y a peu, les joueurs professionnels restaient très nettement au-dessus des programmes informatiques) et les meilleurs programmes d'échecs (notamment Stockfish, considéré comme le moteur d'échecs le plus puissant du monde, très au-dessus des meilleurs joueurs humains).

Tous ces termes réfèrent finalement au même domaine: celui de l'**apprentissage automatique**. C'est un domaine de l'informatique (et même de l'intelligence artificielle) qui ambitionne de permettre aux ordinateurs de traiter des données tout en améliorant continuellement (et surtout automatiquement) leurs performances dans le traitement de ces données au fur et à mesure qu'ils les traitent. Autrement dit, l'algorithme devient capable de modifier certains de ses propres composants pour rendre le traitement plus efficace.

On distingue classiquement deux approches: l'**apprentissage supervisé** et l'**apprentissage non supervisé**.

Dans l'apprentissage supervisé, on donne à l'algorithme un certain nombre d'entrées (par exemple des images) et on lui programme une tâche (comme par exemple classer ces images en fonction de la forme qu'il reconnaît dessus). En plus de ça, on lui donne également des images de référence qui sont déjà classées (soit par un être humain, soit par un autre algorithme) qui lui permettent de s'entraîner et d'optimiser ses propres performances. Les classes existent donc au préalable.

Dans l'apprentissage non supervisé, l'algorithme est complètement lâché dans la nature. On lui donne des images et on le laisse se débrouiller pour les classer en fonction de catégories qu'il va devoir définir par lui-même.

L'apprentissage supervisé est beaucoup plus précis et beaucoup plus efficace puisque l'algorithme a un cadre de référence. En revanche, l'apprentissage non supervisé est beaucoup plus imaginaire et permet parfois de trouver des regroupements ou des corrélations auxquelles on n'aurait pas pensé à la base.

Bref, à chaque approche ses objectifs. Dans notre cas, nous avons des classes déjà existantes (qui sont les différentes maladies) et nous recherchons la performance du tri plutôt que la découverte de nouvelles classes. Notre choix s'oriente donc assez naturellement vers l'apprentissage supervisé.

#### 1.2.1. Les réseaux de neurones convolutifs

C'est le nom de l'algorithme choisi par les auteurs pour mener à bien leur étude, il est spécialisé dans le traitement d'images. Ils ont utilisé Google Inception V3, un réseau de neurone convolutif développé par Google. Mais avant d'expliquer comment ils l'ont utilisé, essayons de comprendre ensemble quels sont les grands principes de cet algorithme.

## 1. Décortiquons l'article

**1.2.1.1. Les réseaux de neurones** Les réseaux de neurones sont des algorithmes d'apprentissage supervisé extrêmement populaires. La théorie initiale (même si elle s'en est quelque peu éloignée depuis) s'inspire du fonctionnement des neurones biologiques.

### Les neurones biologiques

Biologiquement, pour faire simple, un neurone va recevoir des informations sous formes d'**influx électriques** en provenance d'autres neurones par le biais de structure que l'on appelle des **dendrites**. Certains de ces influx (ou **potentiels post-synaptiques**) sont **activateurs** et d'autres sont **inhibiteurs**. Le corps du neurone (ou **soma**) va additionner tous ces influx et en fonction de résultat va émettre ou non un influx électrique de sortie que l'on appelle un **potentiel d'action**. Le potentiel d'action va se propager en suivant l'**axone** du neurone (un genre de fil qui conduit l'influx) puis être communiqué aux neurones suivants par le biais de **synapses** (qui sont un genre de séparation entre l'axone de notre neurone et les dendrites des neurones suivants).

Pour un réseau de neurones informatiques, l'idée est assez similaire. L'algorithme va recevoir une entrée sous forme de données informatiques et va lui appliquer un traitement (qui est une fonction mathématique) afin de générer une sortie. La fonction qui est utilisée est soumise à un certain nombre de paramètres que l'on peut faire varier afin de modifier la sortie.

Et c'est précisément tout l'intérêt du réseau de neurones: à l'aide d'un système de validation, il va pouvoir comparer la sortie qu'il génère avec le classement de référence qu'on lui a donné (souvenez-vous qu'on est bien dans le cadre d'un apprentissage supervisé) afin d'optimiser ses paramètres pour avoir une sortie la plus proche possible de la référence.

L'algorithme de réseau de neurones le plus simple s'appelle le **perceptron**. C'est un algorithme à une couche dans lequel il y a des neurones d'entrée (un par paramètre) et des neurones de sorties (un par classe). Chaque entrée est liée à chaque sortie ainsi, pour une image donnée, le perceptron donne pour chaque classe une probabilité que l'image appartienne ou non à cette classe en fonction des paramètres considérés.

Mais il existe maintenant des réseaux de neurones à plusieurs couches qui permettent d'effectuer des traitements intermédiaires afin d'augmenter encore plus leur efficacité à classer des images. Ces couches intermédiaires forment alors ce qu'on appelle des **réseaux de neurones profonds**, d'où l'appellation de *Deep Learning*.

**1.2.1.2. Utilisation du réseau de neurones dans l'étude** Comme je vous l'ai dit, des réseaux de neurones, il en existe de plein de types différents afin de résoudre plein de problèmes différents. En l'occurrence, notre problème à nous c'est de faire du traitement d'image et pour cela, nous allons utiliser un type de réseau de neurones particulièrement adapté: le **réseau de neurones convolutif**.

Un réseau de neurones convolutif, c'est un réseau de neurones qui utilise un opérateur mathématique qui s'appelle le **produit de convolution** et qui permet entre autres d'appliquer des filtres sur les images. L'un de ces filtres est un filtre de comparaison qui permet de comparer l'image avec un ensemble de références (que l'on appelle *features* en anglais). La présence ou non de certaines de ces références permet alors de classer l'image dans une des catégories, ou,

## 1. Décortiquons l'article

pour être un peu plus précis, permet à l'algorithme de donner une probabilité d'appartenance de l'image à chaque catégorie.

Comme je vous l'ai dit un peu plus haut, l'algorithme de réseau de neurones convolutifs utilisés pour l'étude s'appelle Google Inception V3 et voici à quoi il ressemble:

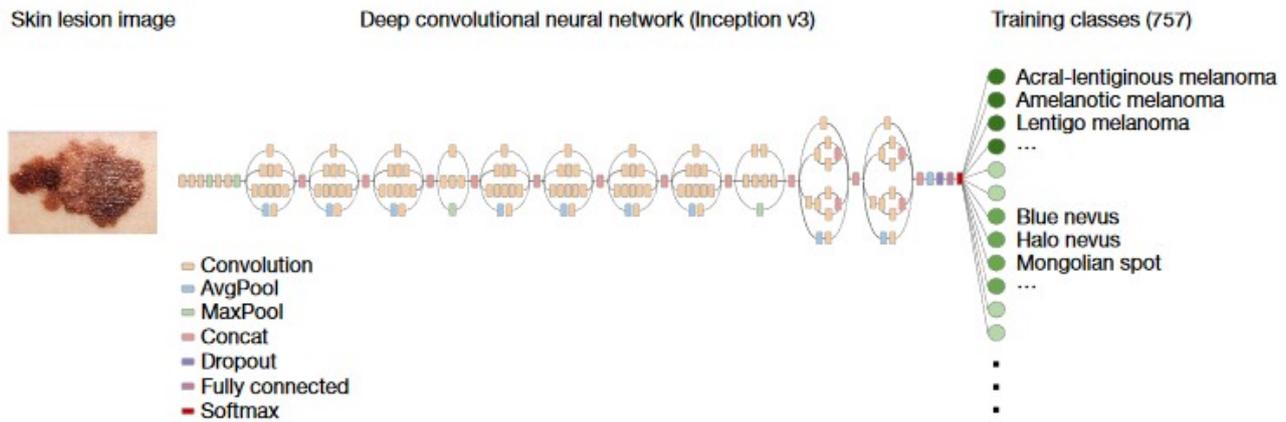


FIGURE 1.2. – Google Inception V3 (ça fait peur un peu non ?) — extrait de l'article donné en lien plus haut.

Ce réseau de neurones est constitué d'un très grand nombre de blocs, parmi lesquels<sup>4</sup>:

- des **blocs de convolution**, très nombreux, qui constituent le cœur du système (on a vu l'utilité un peu plus haut),
- des **blocs de pooling** (*AvgPool* et *MaxPool*) qui consistent à réduire la taille des images (*Avg* pour *average*, c'est-à-dire en faisant une moyenne et *max* pour *maximum* c'est-à-dire en prenant des maxima locaux),
- des **blocs de concaténation** (*Concat*) qui permettent de regrouper les résultats de plusieurs convolutions faites en parallèle et qui permet également de corriger un certain nombre de résultats par une méthode d'activation linéaire,
- un **bloc dropout** qui permet d'éliminer le produit de certaines analyses pour optimiser l'ensemble,
- les **blocs fully connected et softmax** sont chargés, à partir des différentes références, de produire pour chaque classe une probabilité que l'image à d'appartenir à cette classe. Ils sont également capables de tenir compte des itérations précédentes pour optimiser les paramètres et se rapprocher autant que possible de la cible.

L'algorithme a été pré-entraîné par Google sur 1,28 millions d'images de toutes sortes (classées dans 1000 catégories, n'ayant rien à voir avec la dermato) ce qui permet à l'algorithme d'avoir une base en matière de traitement d'images et de ne pas partir de rien, ce qui serait très coûteux en temps de calcul. Ces images permettent notamment à l'algorithme d'apprendre à tenir compte des différences d'angle, de zoom ou encore d'exposition entre les images.

En utilisant un **apprentissage par transfert** (*transfer learning* en anglais), l'équipe a réutilisé ce réseau de neurones déjà entraîné puis l'a adapté à ses propres besoin pour pouvoir analyser pas moins de 129450 images de lésions dermatologiques à classer dans 757 catégories différentes obtenant une sensibilité (la capacité à correctement classer dans une pathologie les lésions qui

## 1. Décortiquons l'article

relèvent effectivement de cette pathologie) et une spécificité (la capacité à ne pas classer dans une pathologie les lésions n'en relevant pas) proches de celles d'un dermatologue.

---

Peu après la publication de l'article, l'équipe a lancé les tests auprès de vrais patients qui prenaient en photo leurs lésions afin d'en obtenir une classification. L'objectif affiché (et même écrit dans l'article) c'était de pouvoir développer une application mobile pour permettre à chacune et chacun de faire son propre suivi simplement et de chez soi.

Cependant, un problème sujet majeur s'est très rapidement présenté: l'algorithme était extrêmement efficace avec les peaux claires et beaucoup moins efficace avec les peaux foncées (donnant de faux diagnostics, ce qui est dangereux). La communauté scientifique a avancé deux explications à ce phénomène:

- de manière générale, le diagnostic dermatologique des peaux foncées est beaucoup plus difficile que celui des peaux claires car les lésions sont elles-mêmes généralement foncées et sont donc beaucoup plus difficiles à observer,
- il y a un biais lié à la base d'images utilisées pour entraîner l'algorithme. Celles-ci ont été collectées par plusieurs hôpitaux de Californie du nord, une zone où les peaux claires sont beaucoup plus présentes.

À partir de là, l'équipe avait deux choix possibles:

- continuer l'exploitation de l'algorithme en précisant qu'il ne fonctionne pas pour les peaux foncées, créant ainsi une inégalité de prise en charge basée sur la couleur de peau,
- cesser l'exploitation de l'algorithme au risque d'aboutir à un défaut de prise en charge et donc à ne pas soigner des patients qui auraient pu l'être.

L'équipe a choisi la 2<sup>ème</sup> option.

Et vous, qu'en pensez-vous? Auriez-vous fait le même choix?

---

2. C'est la capacité du corps à savoir où se situent nos membres même quand on a les yeux fermés. Fermez les yeux et essayez de toucher votre nez avec l'un de vos index et vous verrez. 🍊

3. Je me souviens d'une discussion mémorable sur Discord avec Arius qui me soutenait que c'est la spécialité la plus dure. À titre personnel, je penche pour l'immunologie, mais il y a match.

4. Pour une description plus précise de la façon dont fonctionne ce réseau de neurones, je vous invite à lire [la publication de Google à ce sujet](#) 📄 (un peu ardue).