



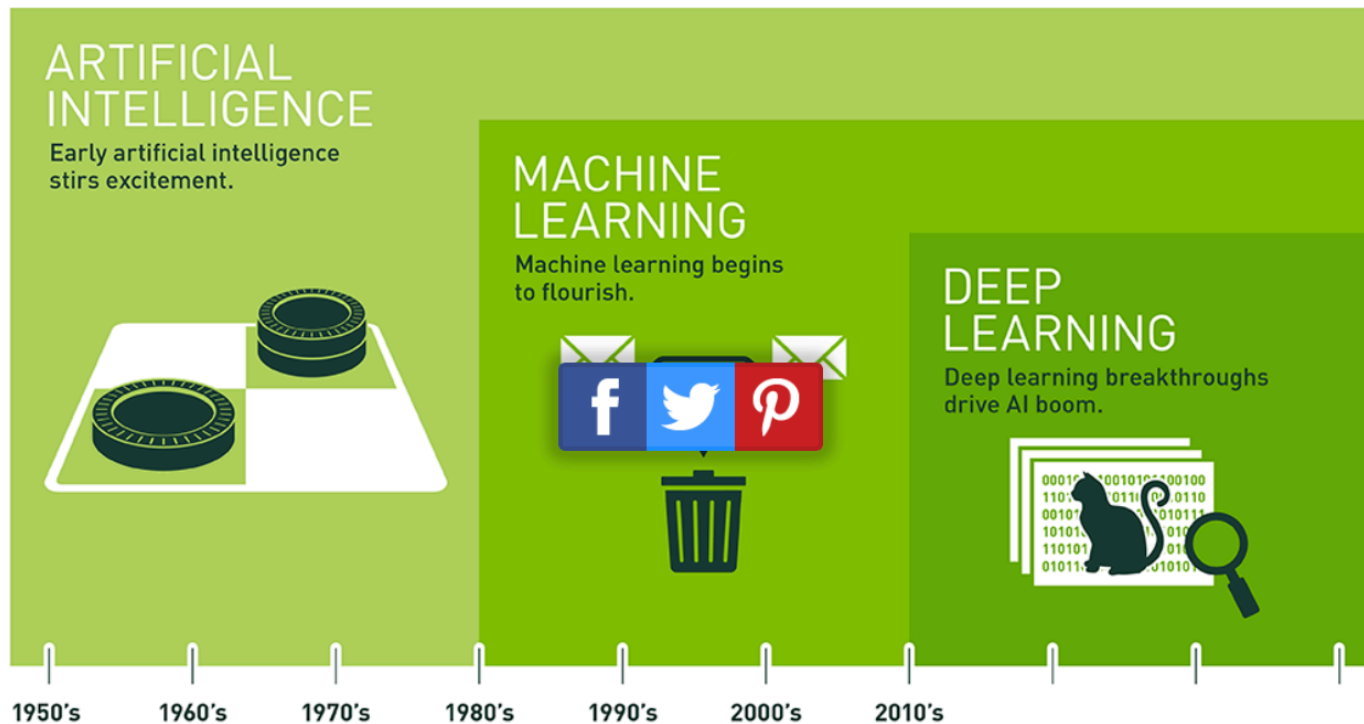
المدرسة الوطنية المتعددة التقنيات
Ecole Nationale Polytechnique

Deep Learning : principe, techniques et applications

Sid-Ahmed Berrani

Deep Learning

- Réseaux de neurones artificiels : méthodes inspirées de la biologie et du fonctionnement du cerveau humain.



<https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/>

Deep Learning

- Cela fait référence à l'intelligence artificielle.
- Plus particulièrement à l'apprentissage automatique.
- Et plus spécifiquement aux réseaux de neurones artificiels.
- Méthodes inspirées de la biologie et du fonctionnement du cerveau humain.

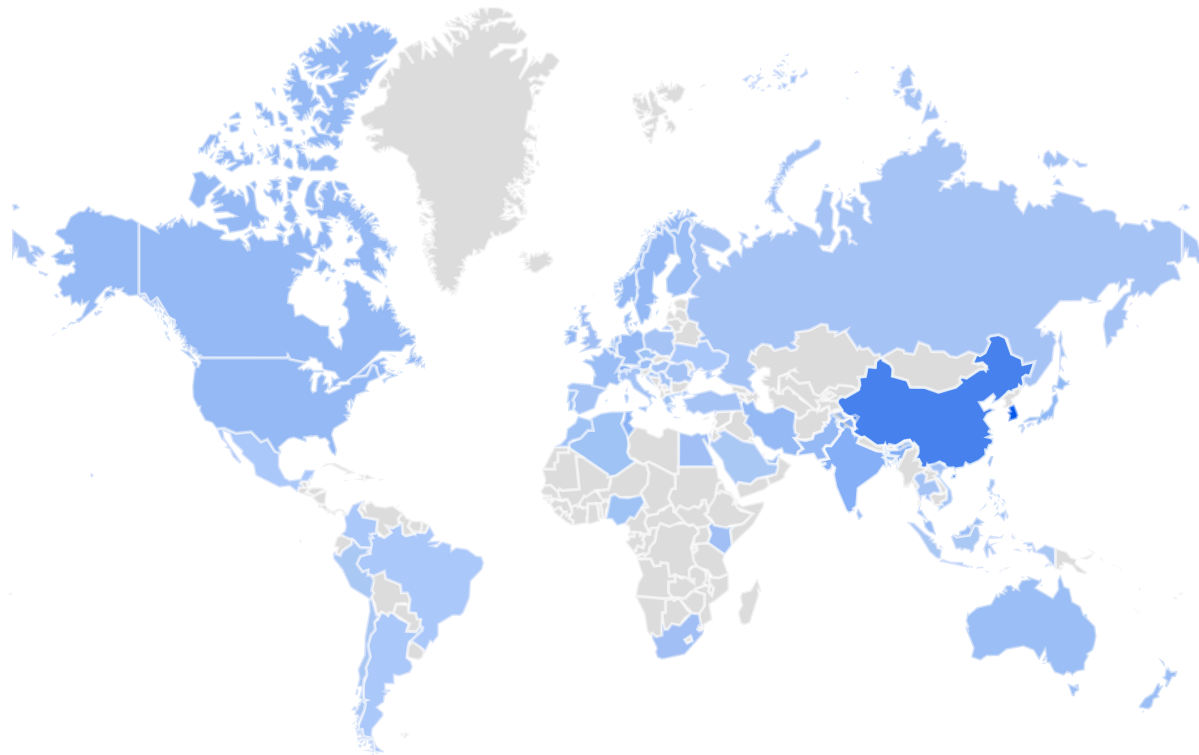
Deep Learning : une tendance

Evolution de la popularité du mot-clef « Deep learning »
sur la moteur de recherche Google



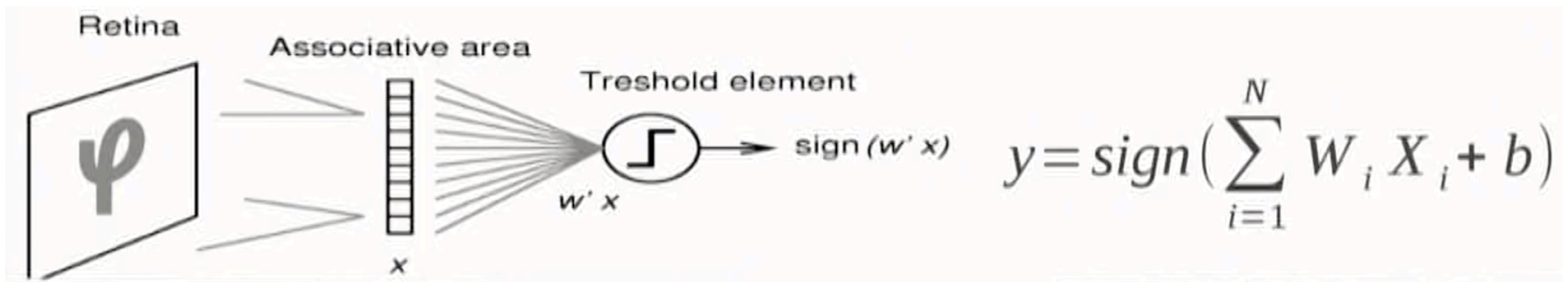
Deep learning : une tendance

Intérêt par région du mot-clef « *Deep learning* » sur la moteur de recherche Google.



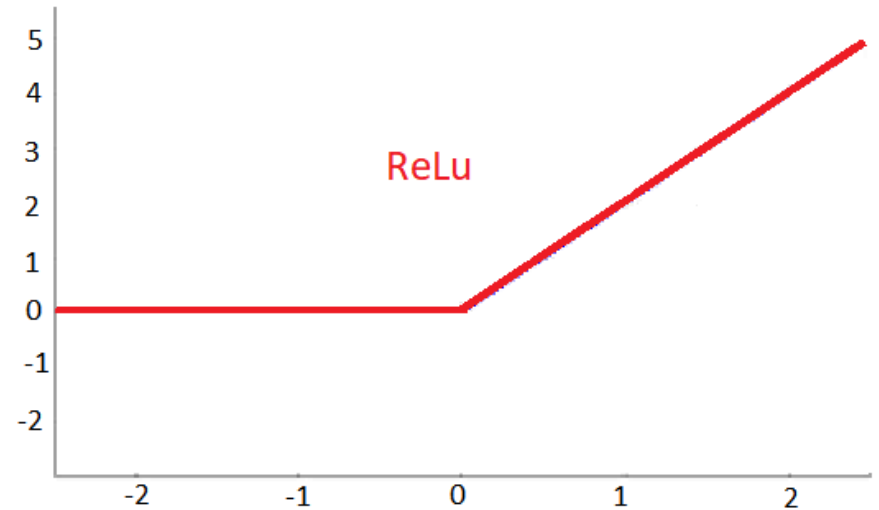
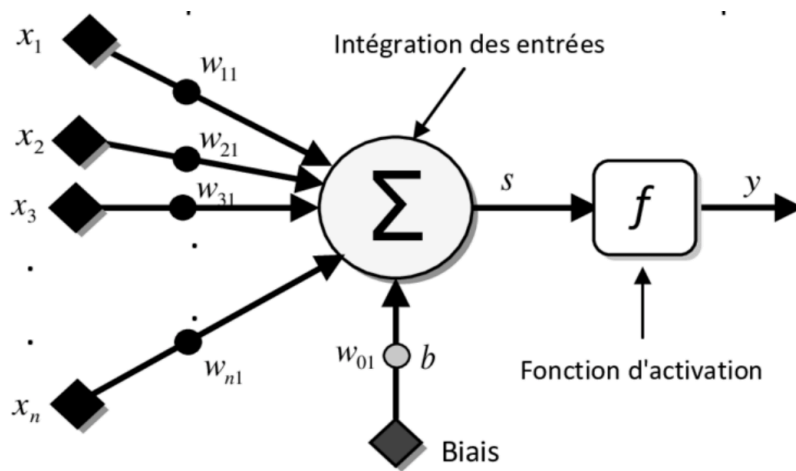
Les origines de tout cela...

- Cela a commencé, il y a très longtemps...
- Le perceptron, le début de l'apprentissage machine (Cornell 1957).
 - Un classifieur linéaire.
 - Les paramètres sont appris à partir d'exemples annotés.



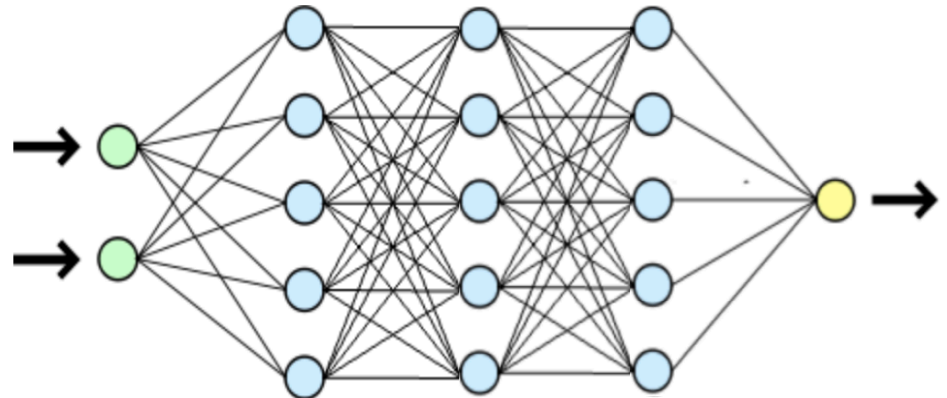
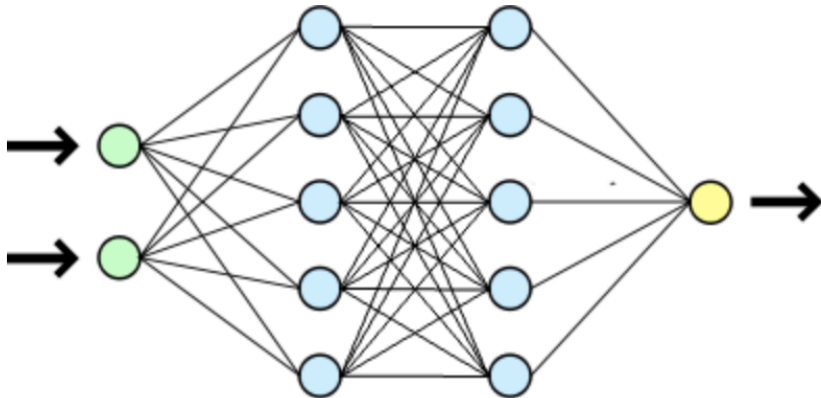
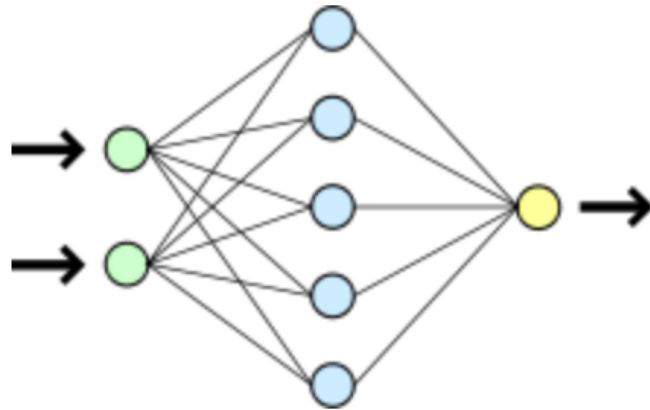
Quelques notions de base

- L'unité de base est le neurone artificiel.



Quelques notions de base

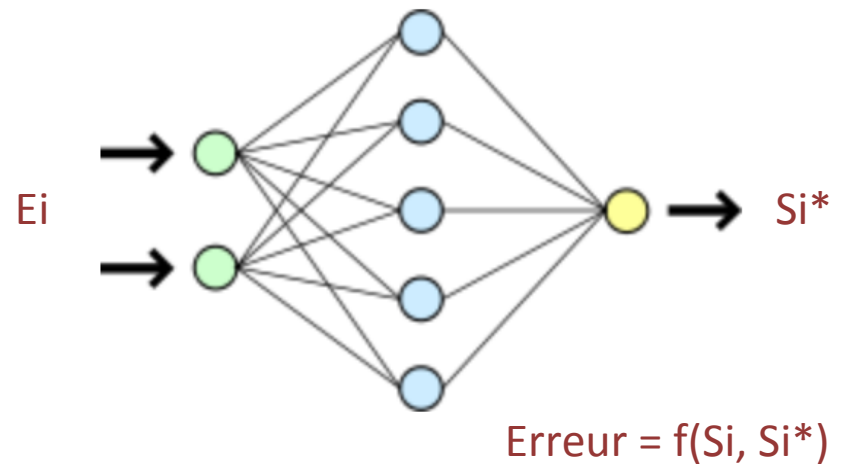
- Les neurones sont assemblés pour former un réseau de neurones...



Quelques notions de base

- L'apprentissage supervisé par rétro-propagation du gradient

Ensemble d'apprentissage :
(E1, S1) (E2, S2) ... (En, Sn)



***Adaptation des poids
du réseau***

Quelques notions de base

- La capacité de généralisation :

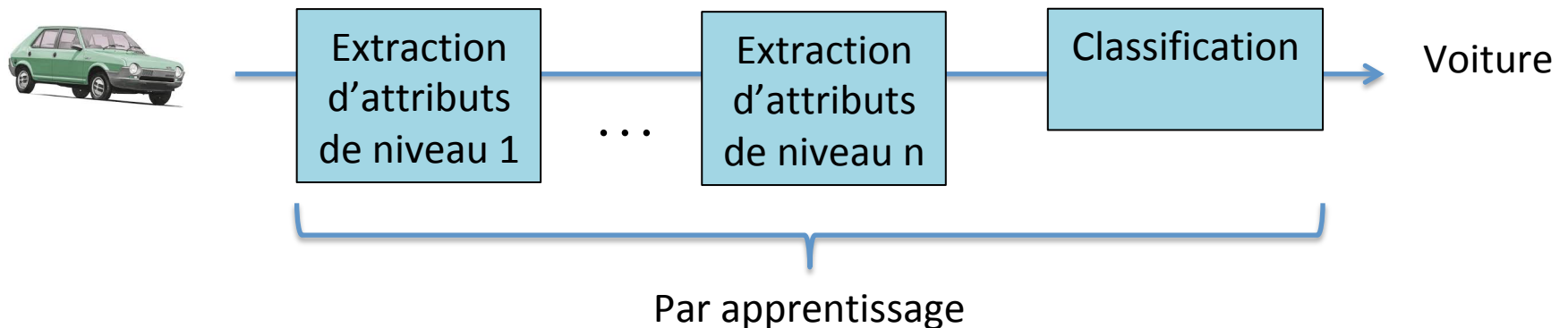


Deep Learning vs. approches traditionnelles de Machine Learning

- Approches traditionnelles



- *Deep Learning*



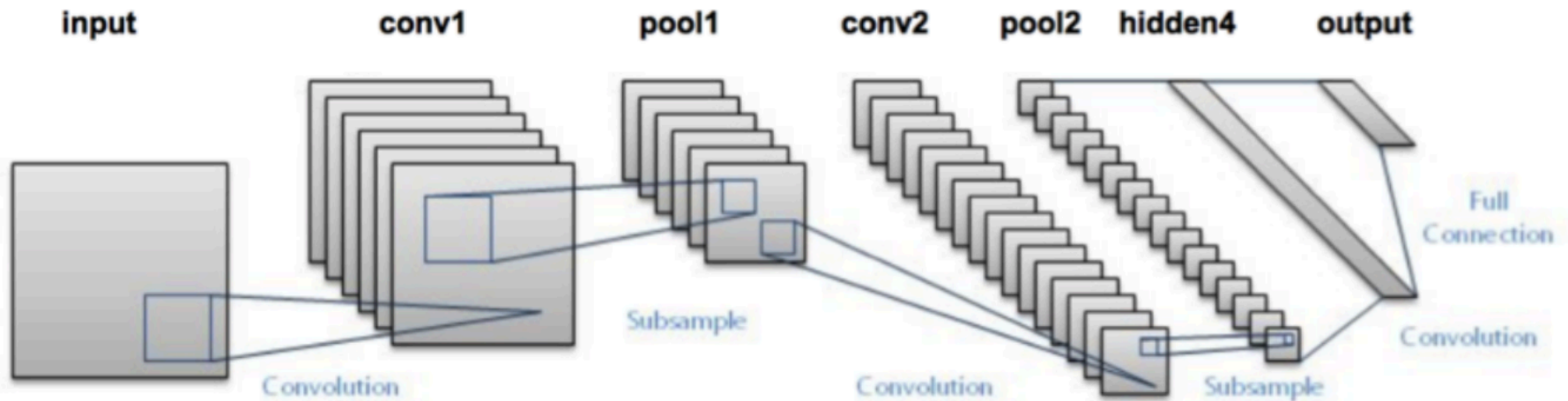
Les trois facteurs de succès du *Deep Learning*

1. La disponibilité en quantité suffisante de données d'apprentissage.
2. La disponibilité des capacités de calcul pour réaliser les apprentissages.
3. La disponibilité des « outils d'apprentissage » et des « modèles appris » en *Open Source*.

Architectures et réseaux dédiés

Les réseaux de neurones à convolution (ConvNet ou CNN)

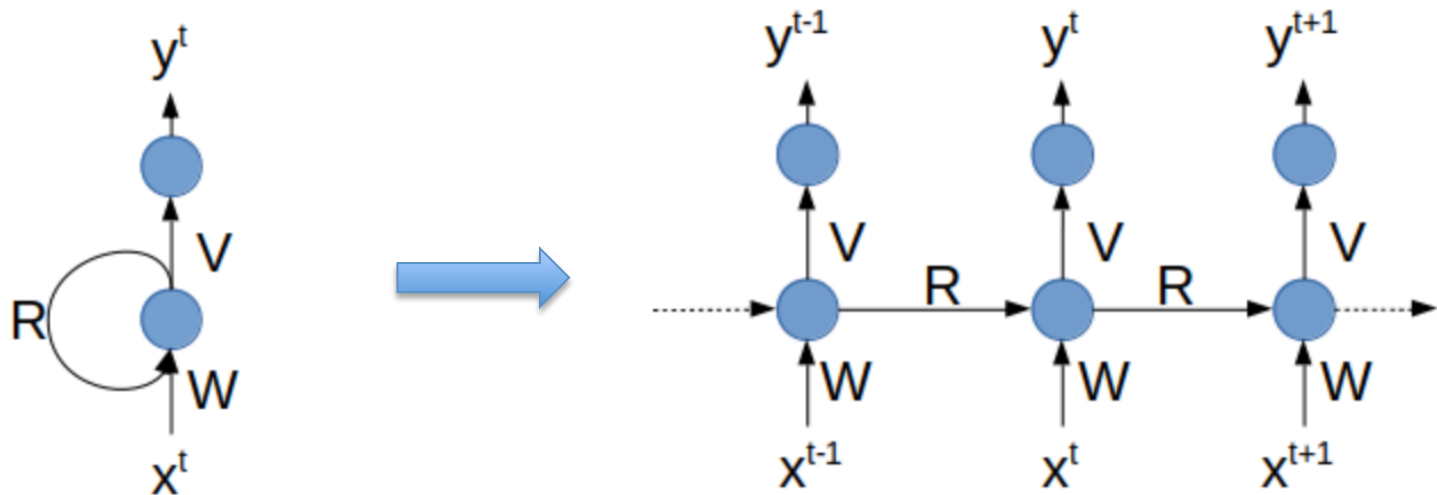
- Inspirés de ce qu'on connaît du cortex visuel.
- Adaptés à la reconnaissance d'images (signaux sous forme de tableau de nombres)



Architectures et réseaux dédiés

- **Les réseaux de neurones récurrents**

Dédiés au traitement des données séquentielles : texte, parole...



Architectures et réseaux dédiés

- **Les réseaux de neurones génératifs adversaires (GAN)**

Inventé par Ian Goodfellow, chercheur à Google en 2014.



Architectures et réseaux dédiés

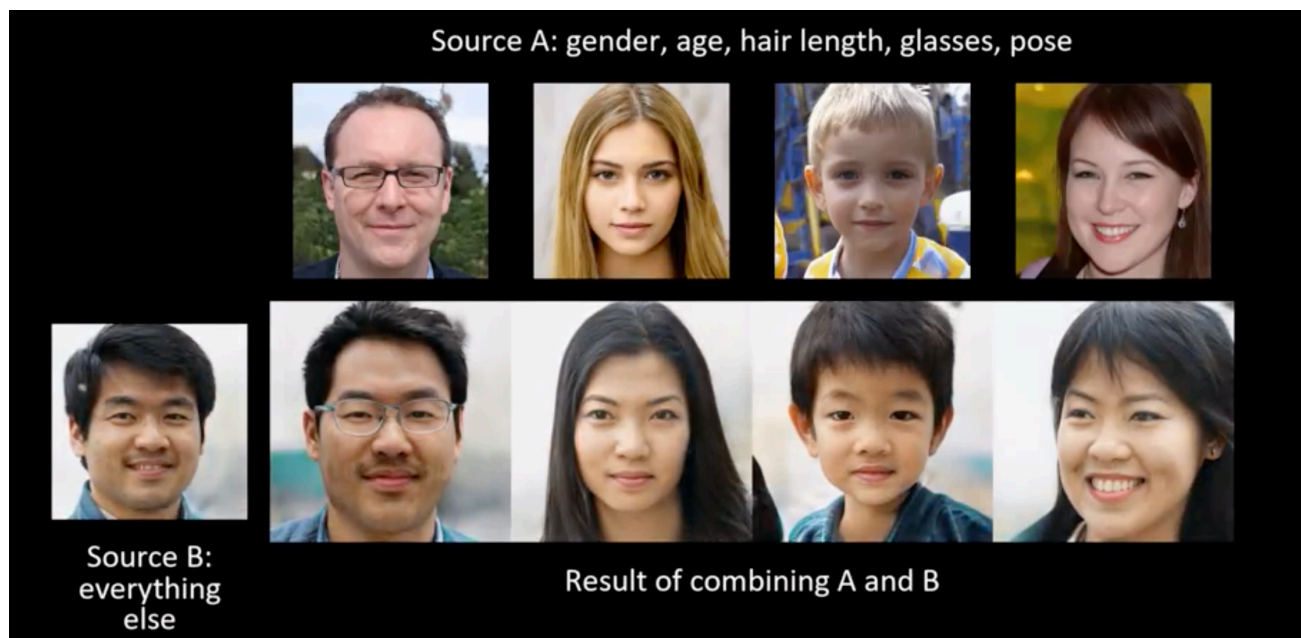
- **Les réseaux de neurones adversaires (GAN)**



Réf. : *Progressive growing of gans for improved quality, stability, and variation.*
T. Karras, T. Aila, S. Laine, J. Lehtinen. ICLR 2017

Architectures et réseaux dédiés

- Les réseaux de neurones adversaires (GAN)



Réf. : *A style-based generator architecture for generative adversarial networks*
T. Karras, S. Laine, T. Aila. CVPR 2019.

C'est bien plus qu'un *buzz word*

- En juillet 2019, Microsoft décide d'investir 1 Mds de \$ dans OpenAI.
- D'après une étude de KPMG, les entreprises devraient investir :
 - 18,3 milliards \$ en 2019.
 - 27,4 milliards \$ en 2020.
 - 232 milliards \$ en 2025.

C'est bien plus qu'un *buzz word*

En 2017, la Chine lance un plan pour devenir le leader mondial de l'IA d'ici 2030.

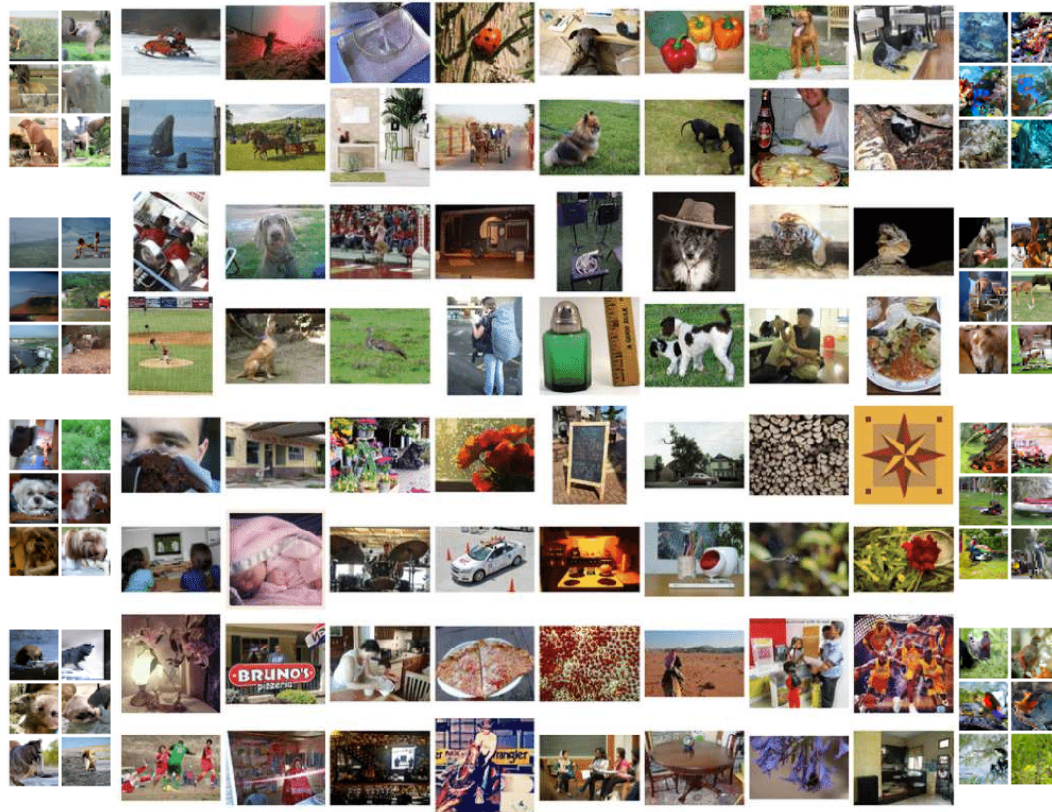
- En 2018, *SenseTime* a levé **600 millions de dollars** notamment pour financer ses recherches.
- *Baidu* s'associe au gouvernement chinois pour ouvrir un laboratoire d'intelligence artificielle.
- L'Etat fournit aux entreprises un accès facilité aux données personnelles de leurs concitoyens.
 - Les informations pour la reconnaissance faciale utilisés dans *AliPay* *proviendraient* du système national de sécurité.

Les applications du *Deep Learning*

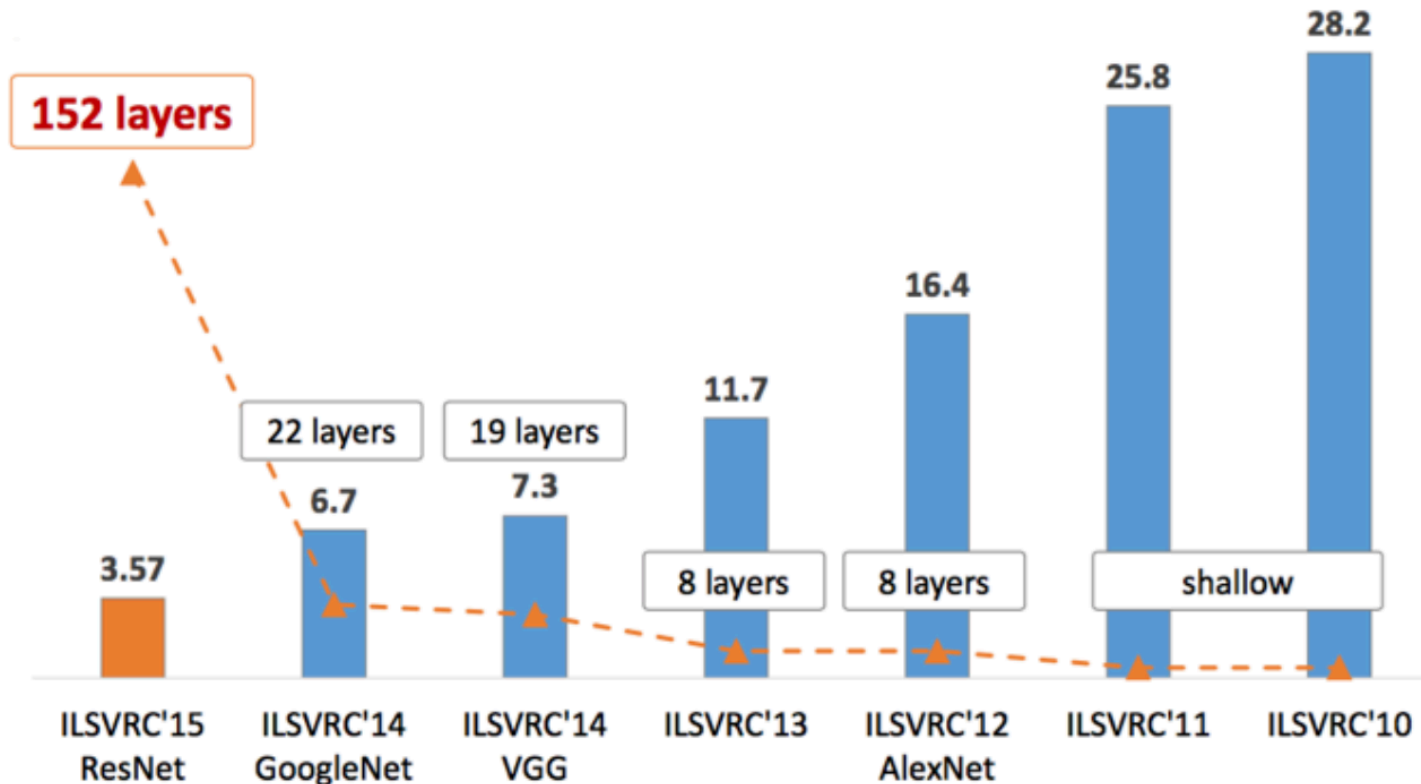
- La reconnaissance des visages.
- L'estimation de l'âge et du genre via les images de visages.
- La traduction automatique.
- La reconnaissance de la parole.
- L'aide au diagnostic médical par l'analyse des images (e.g. de tumeurs de la peau).
- *Les Deep Fake.*
- L'aide à la conduite de véhicules.
- ...

Classification des images

- La compétition internationale ILSVRC (depuis 2010).
- ImageNet : 1,2 million d'images d'apprentissage, 1000 catégories d'objets.



Classification des images



↑
Meilleur qu'un
être humaine

↑
Deep learning
60 millions de param.

La reconnaissance de visages

- En 2014, Facebook développe le système DeepFace.
- Précision obtenue par DeepFace : 97,25 %.
- Des humaines soumis aux mêmes tests répondent avec une précision de 97,53 %.
- **Réseau de neurones :**
 - 9 couches.
 - 120 millions de connections.
 - Entraîné sur 4 millions d'images réelles (chargées par les utilisateurs de Facebook) représentant +4000 personnes.
- Y. Taigman; M. Yang; M. Ranzato; L. Wolf, "DeepFace: Closing the Gap to Human-Level Performance in Face Verification", *International Conference on Computer Vision and Pattern Recognition (CVPR' 2014)*.



La reconnaissance de visages

- En 2015, Google propose Facenet.
- Précision obtenue par Facenet : 99,63 %.
- Réseau de neurones :
 - 24 couches.
 - Entraîné sur 500 millions d'images représentant 10 millions de personnes.

- F. Schroff, D. Kalenichenko, and J. Philbin. *Facenet: A unified embedding for face recognition and clustering*. *International Conference on Computer Vision and Pattern Recognition (CVPR' 2015)*.

L'identification du genre et l'estimation de l'âge



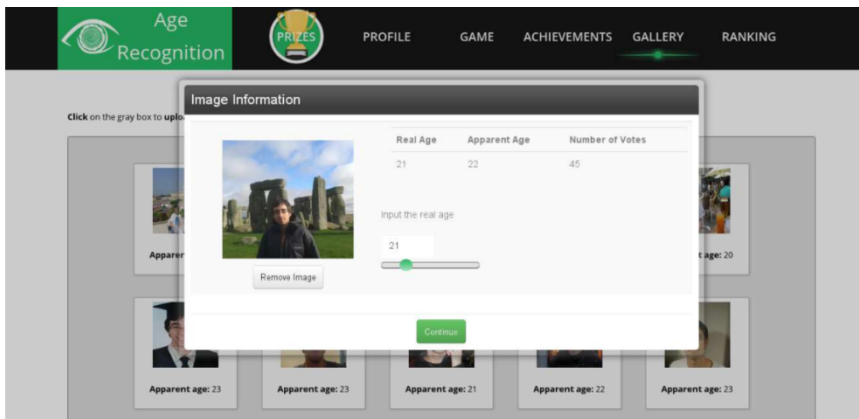
→ Un homme de 25 – 30 ans

- Annotation de photos.
- Collecte de données anonymisées.
- Marketing.
- Les applications de « soft biométrie ».
- ...

Compétition internationale ChaLearn LAP pour l'estimation de l'âge

- Organisé en marge de la conférence international CVPR 2016.

Corpus ChaLearn *Escalera et al. 2016*



5613 photos pour l'apprentissage

1978 photos pour l'évaluation

≥ 10 annotations humaines par photo

Position	Équipe	-score
1	Orange Labs	0,2411
2	palm_seu	0,3214
3	cmp+ETH	0,3361
4	WYU_CVL	0,3405
5	ITU_SiMiT	0,3668
6	Bogazici	0,3740
7	MIPAL_SNU	0,4569
8	DeepAge	0,4573

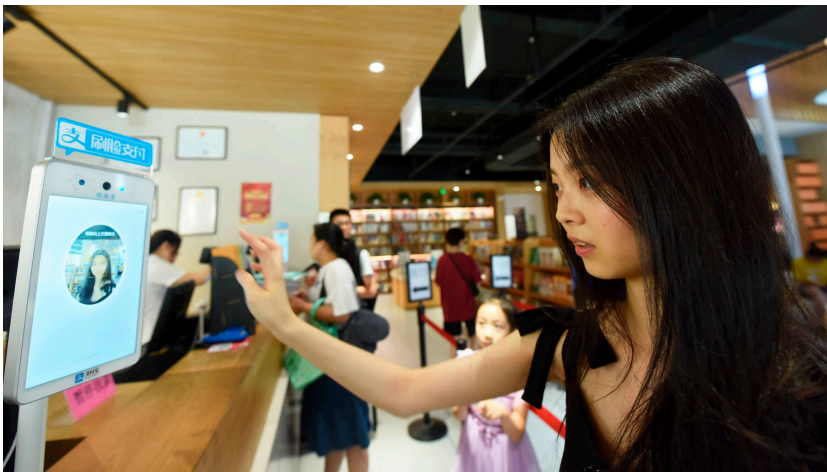
« Effective Training of Convolutional Neural Networks for Face-Based Gender and Age Prediction » par G. Antipov, M. Baccouche, **S. Berrani** et J.-L. Dugelay. **Pattern Recognition 2017**.

A quoi tout cela sert en pratique ?

- Facebook :
 - Chaque image chargée est analysée moins de 2 s après par 4 réseaux de neurones :
 1. Filtrage d'images pornographiques/violentes.
 2. Détection d'objets pour l'annotation et l'indexation.
 3. Description d'images pour les non-voyants.
 4. Reconnaissance des visages.
- Iberia (annoncé en octobre 2019) :
 - L'enregistrement des passagers avec leurs smartphones.
 - Vérification de l'identité des passagers à l'embarquement par reconnaissance faciale.
 - Destinations : Madrid-Barajas (T4) => Asturies ou Bruxelles.

A quoi tout cela sert en pratique ?

- Le visage comme porte-monnaie en Chine :



Paiement AliPay dans un librairie
(dans la ville de Hangzhou)



In September, facial-recognition technology
was introduced at ticket gates.

Paiement du ticket de métro
(dans la ville de Guangzhou)

Aujourd'hui, plus de **100 millions** de chinois sont enregistrés pour utiliser le service de paiement par reconnaissance de visages.

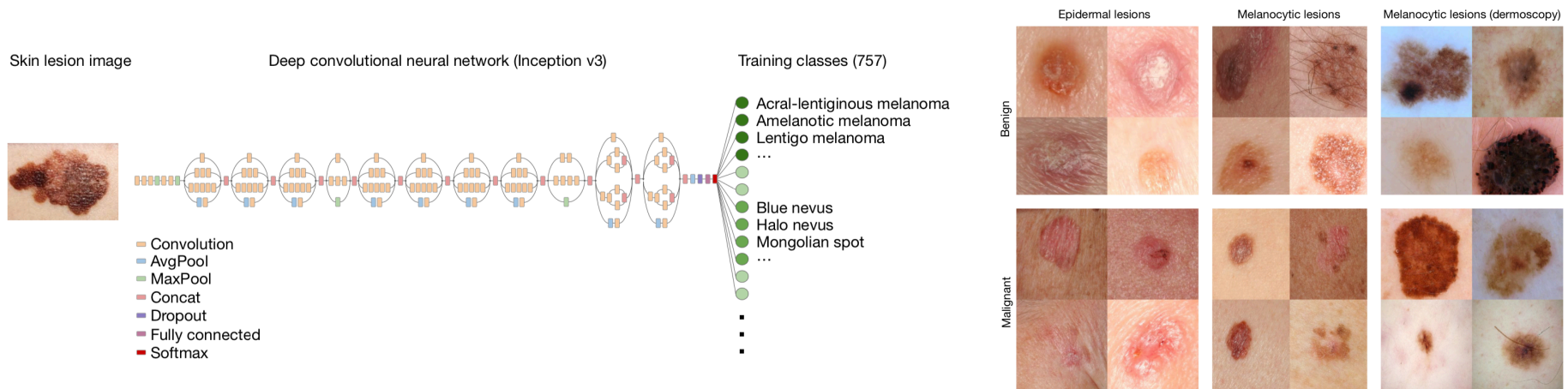
Dans le domaine médical

- **En avril 2018 :**
 - Pour la 1^{re} fois, l'agence américaine du médicament autorise une intelligence artificielle à poser un diagnostic sans être supervisée par un médecin.
 - Diagnostic de la rétinopathie diabétique.
 - IDx-DR : système qui analyse les images des yeux capturées par une caméra rétinienne.



Dans le domaine médical

- **Diagnostic des tumeurs de la peau (A. Esteva, Stanford, 2017)**
CNN, apprentissage sur **129 450 images** cliniques.
Evaluation vs. 21 dermatologues.



Réf. : *Dermatologist-level Classification of Skin Cancer with Deep Neural Networks.*
Andre Esteva, et al. *Nature Letter*, 2017.

La conduite autonome

- **Le système Autopilot de Tesla**

Tesla a commencé à en parler en 2013.

Intégré en oct. 2014 au Tesla Model S.

Développé en partenariat avec la startup Mobileye.

Tesla Model S.



La version 2.5 (août 2017) :

- 8 caméras
- 12 capteurs ultrasons
- 1 radar frontal
- GPU Nvidia Drive PX 2
- Traitement de 200 images / s.
- *Adapter la vitesse au trafic*
- *Rester dans sa voie,*
- *Changer de voie sans intervention du conducteur*
- *Bifurquer vers une autre autoroute,*
- *Sortir de celle-ci une fois la destination proche,*
- *Se garer toute seule*
- *Manœuvrer pour sortir et entrer de un garage*

=> Octobre 2019 : Tesla rachète DeepScale, une startup spécialisée dans le Deep Learning pour la conduite autonome.

Conclusion

- Le *Deep Learning* a permis de réaliser des sauts impressionnants dans les performances des techniques à base d'apprentissage.
- Il a été appliqué à des domaines très variés.
- Trois facteurs de succès : données, capacité de calcul, Open Source.
- Développements réalisés en grande partie grâce aux entreprises et aux startup.
- Le coût d'entrée est très raisonnable !
- Il reste encore beaucoup à faire, tant en recherche qu'en développement de services.