

Introduction au traitement d'images

Enseignement intégré

TS223 | Systèmes Électroniques Embarqués 2A | 2024-2025

Chapitre 1 : Introduction

Rémi Giraud

remi.giraud@enseirb-matmeca.fr

<https://remi-giraud.enseirb-matmeca.fr/>

Plan du cours

- Cours intégrés (18h)**
- **Introduction**
 - **Formation / Acquisition**
 - **Image numérique**
 - Format/Affichage/Synthèse
 - Espaces couleur caractéristiques : compression, esquisse, illusion
 - **Traitements**
 - Filtrage linéaire / non linéaire : débruitage, anonymisation
 - Détection de contours : réhaussement de contraste
 - **Transformée de Fourier**
 - Application : recouvrement fréquentiel
 - **Compression d'images**
 - Application : algorithme JPEG
 - **Transformation spatiales**

TP (10h)

?

Éval Contrôle continu + examen terminal (code + sujet)

Traitement d'images

- Culture du traitement d'images et des applications
- Comprendre le système de vision humain et ce qu'est une image couleur
- Savoir afficher de manière pertinente n'importe quel contenu 2D
- Savoir identifier des dégradations et des solutions adéquates
- Outils : changement d'espace, filtrage, Fourier, compression, ...

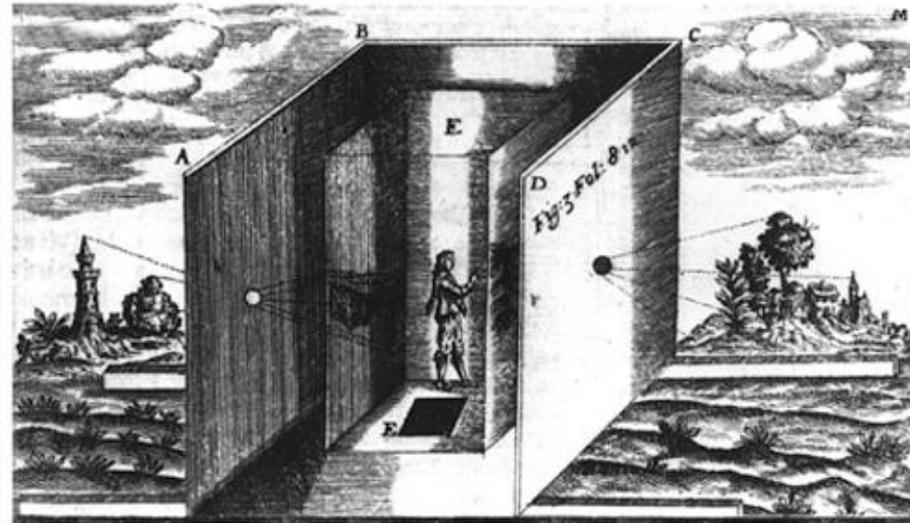
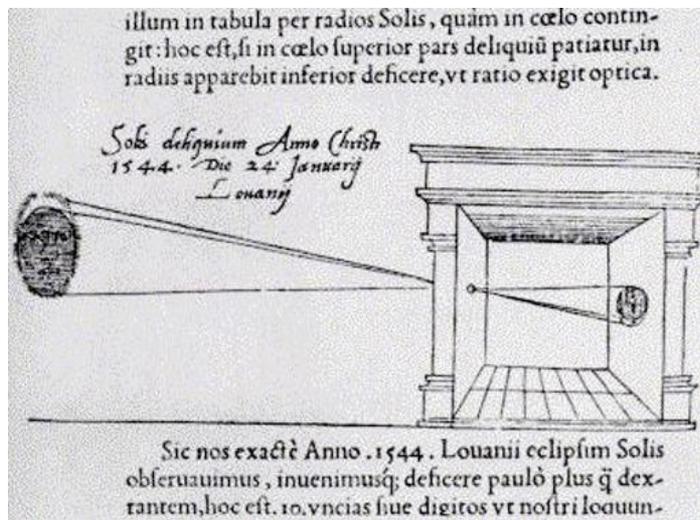
Programmation Python

- Prise en main de Spyder
- Rappels de syntaxe Python (import, indentation, boucles, fonctions, ...)
- Bibliothèques *numpy*, *matplotlib*, *scikit-image*
- Manipulation de tableaux (*np.array*, *0:L-1*, *axis=X*, *opérations vectorielles*)

L'HISTOIRE

Définition

- Image : une **interprétation visuelle** d'une fonction $f(x,y)$ où f représente la luminosité (couleur) au point (x,y) .
- Plupart des images définies sur un plan (x,y) .



Chambre noire [Kircher 1646]

La première photographie

- Première expérience réussie de fixation permanente d'une image.
(Nicéphore Niépce en 1827)



La première traversée

- Première image numérique produite en 1921 à partir d'une bande codée et transmise à travers l'Atlantique par télégraphe.
- Plusieurs semaines → quelques heures



Évolution des technologies

- Des outils de plus en plus petits et complexes



L'OMNIPRÉSENCE

Quantité de données gigantesque

- Majorité du trafic comprenant des images ou vidéos :



Exemple de Facebook

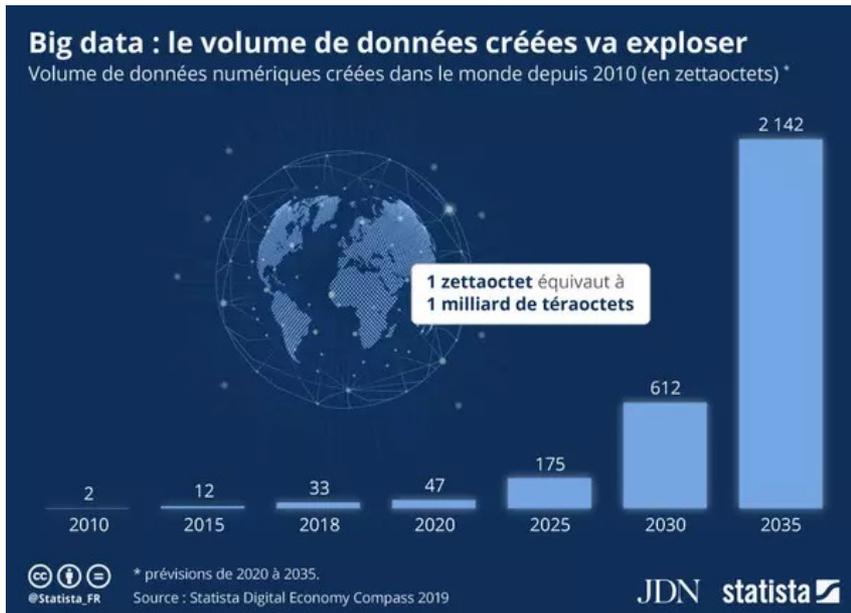
- Nombre de photos mises en ligne :
 - Par seconde : 2 263
 - Par an : 71.425 milliards



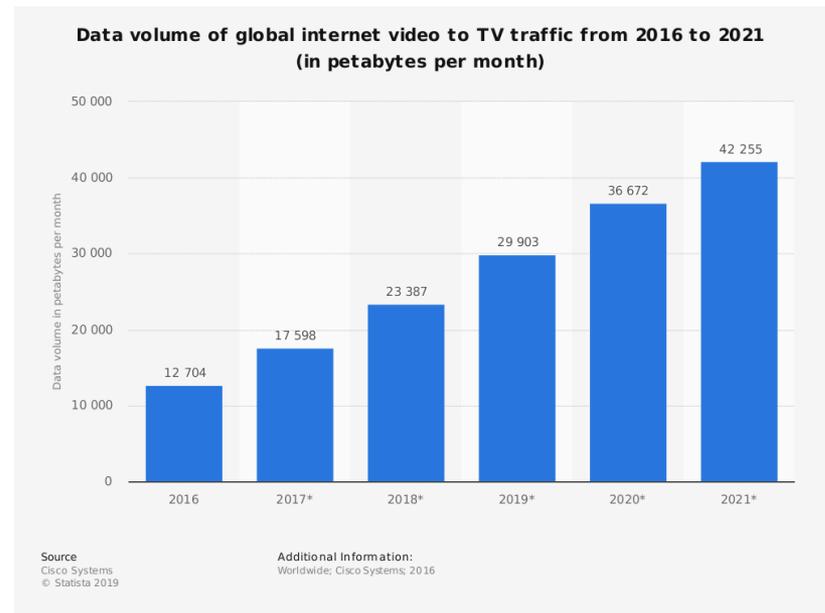
Source

Évolution exponentielle du volume de données

- Problématiques cruciales de :
 - Stockage, compression, transmission, affichage, etc.
 - Recherche de contenu

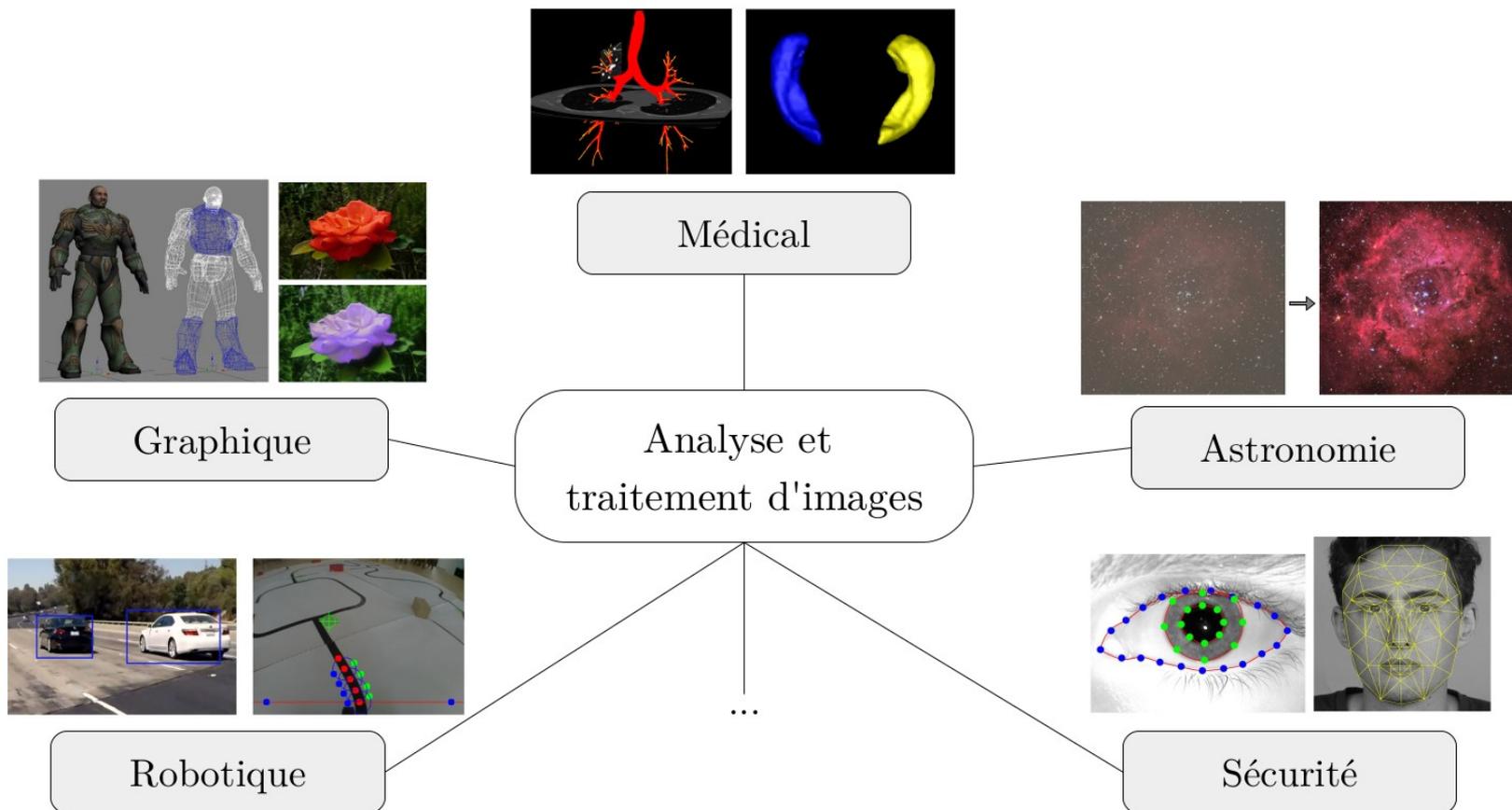


Source



Source

Pléthore de domaines d'application

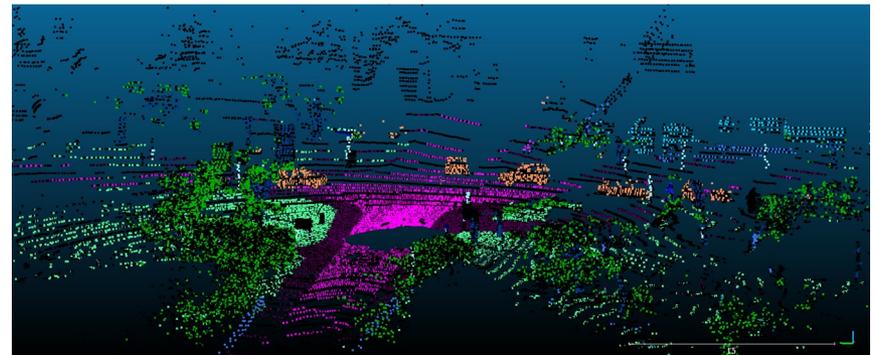
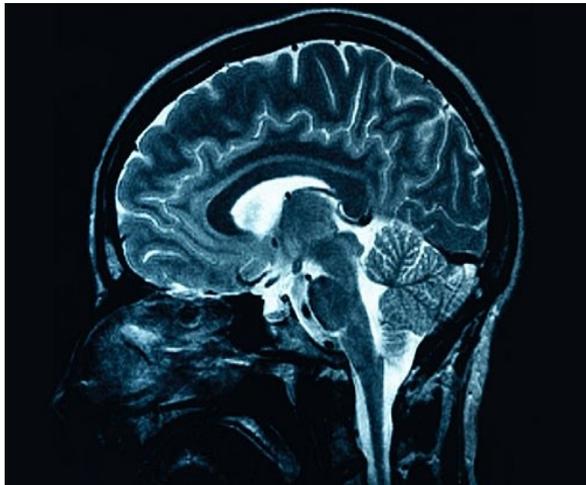
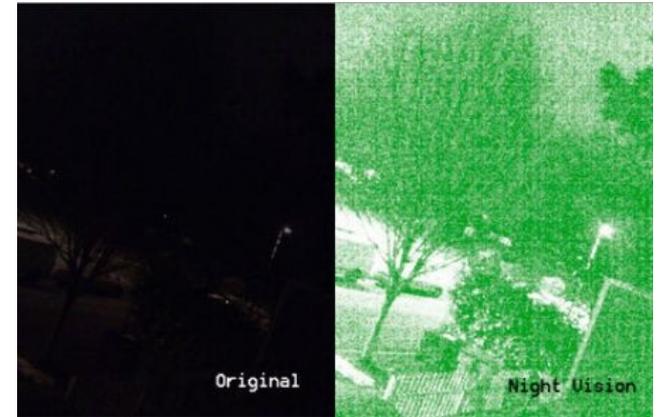
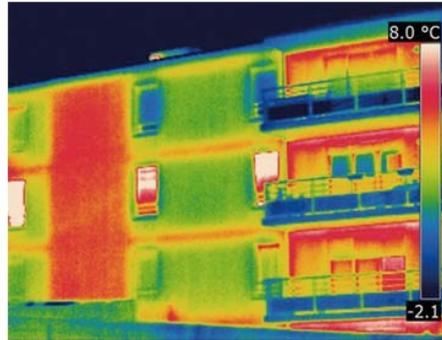


Omniprésence des données

Pléthore de dispositifs d'acquisition



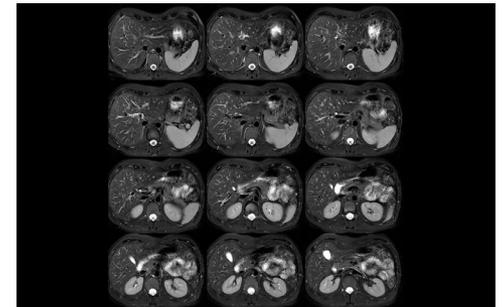
Pléthore de modalités d'acquisition



Pléthore de formats de données

- **Image (2D)**

- Appareil photographique
- IRM
- Radiographie X



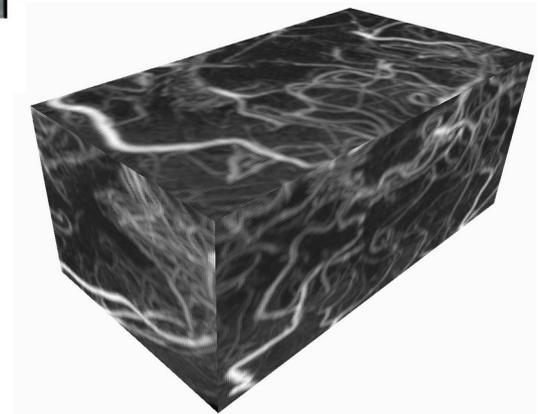
- **Vidéo (2D+t)**

- Vidéo
- Radar



- **Volumique (3D)**

- IRM
- Échographie
- Tomographies (tomodensitométrie, ...)



Nombreux formats numériques

Image couleur, multispectrale



Photo argentique en noir et blanc

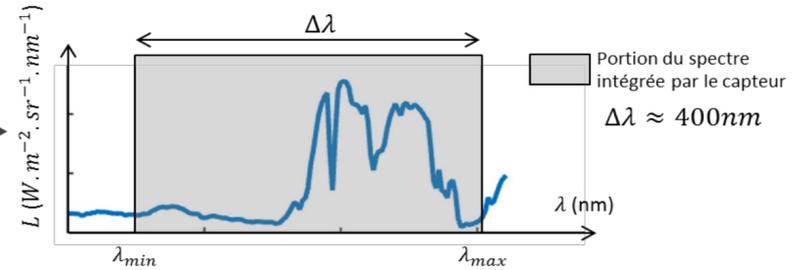


Photo numérique couleur

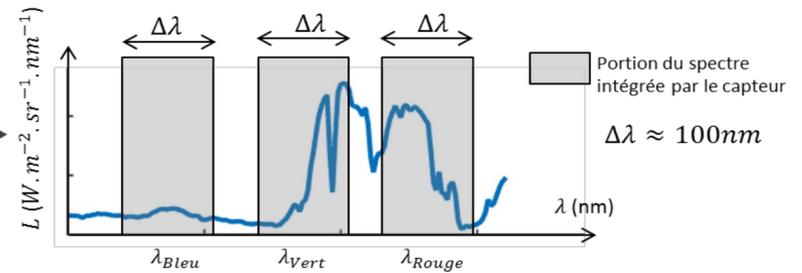
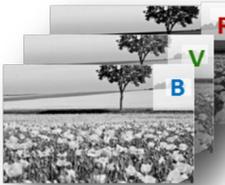
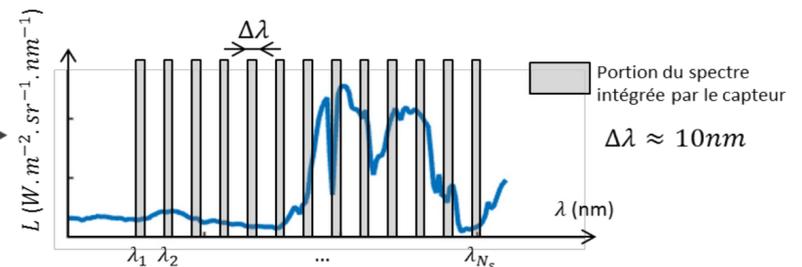
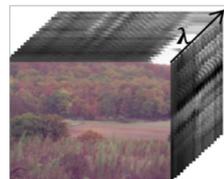


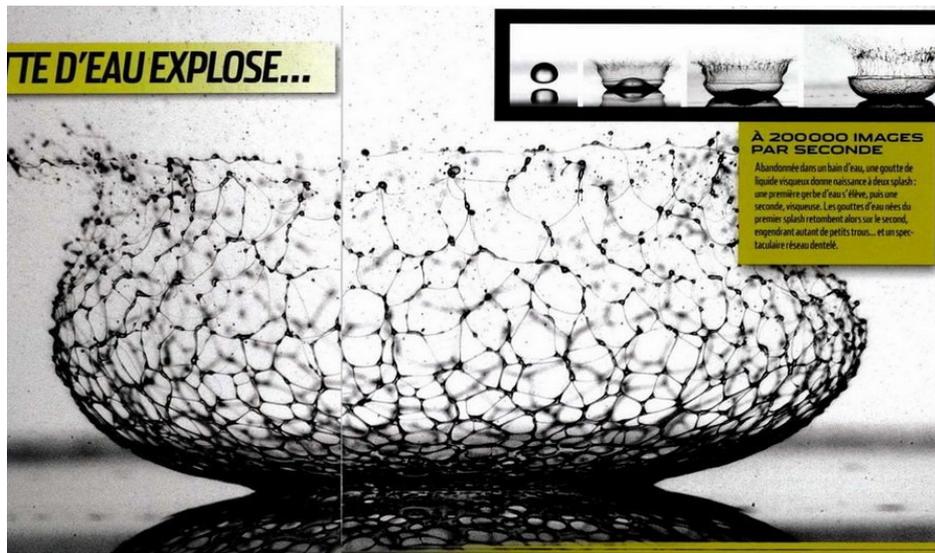
Image multispectrale



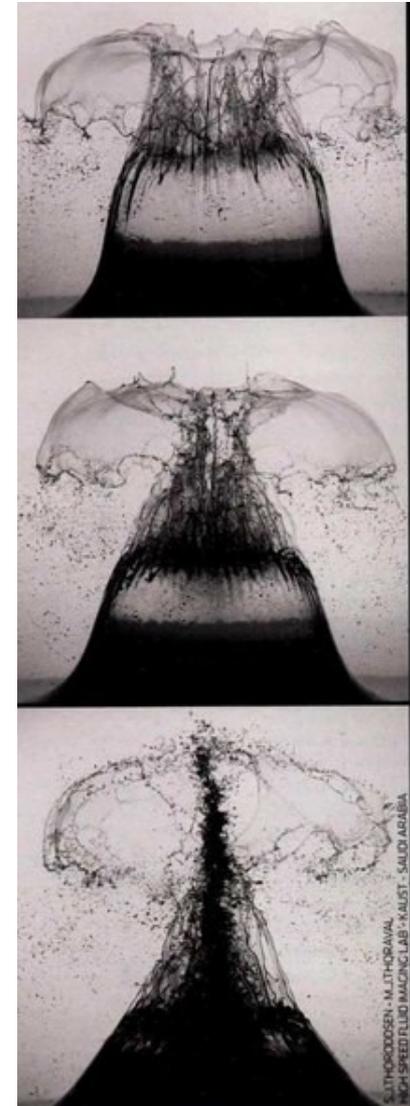
Caméras haute-fréquence

Plus rapide des caméras : 25 millions d'images/s

Peut saisir un événement aussi fugace que 40 milliardièmes de seconde !



Caméra ultra-rapide : 40.000 images par seconde



Fréquence perceptible

- L'œil humain ne peut distinguer *clairement* plus de 24 images/s
- Même si le cerveau est capable de repérer un événement de 0.01s

Standard cinéma : 24 images/s

Quelques exceptions :

Le Hobbit 48 images/s



Sorties classiques (télé, ordinateur) : $\gg 24$ images/s

Ex : Taux de réactivité de 8ms ≈ 125 images/s

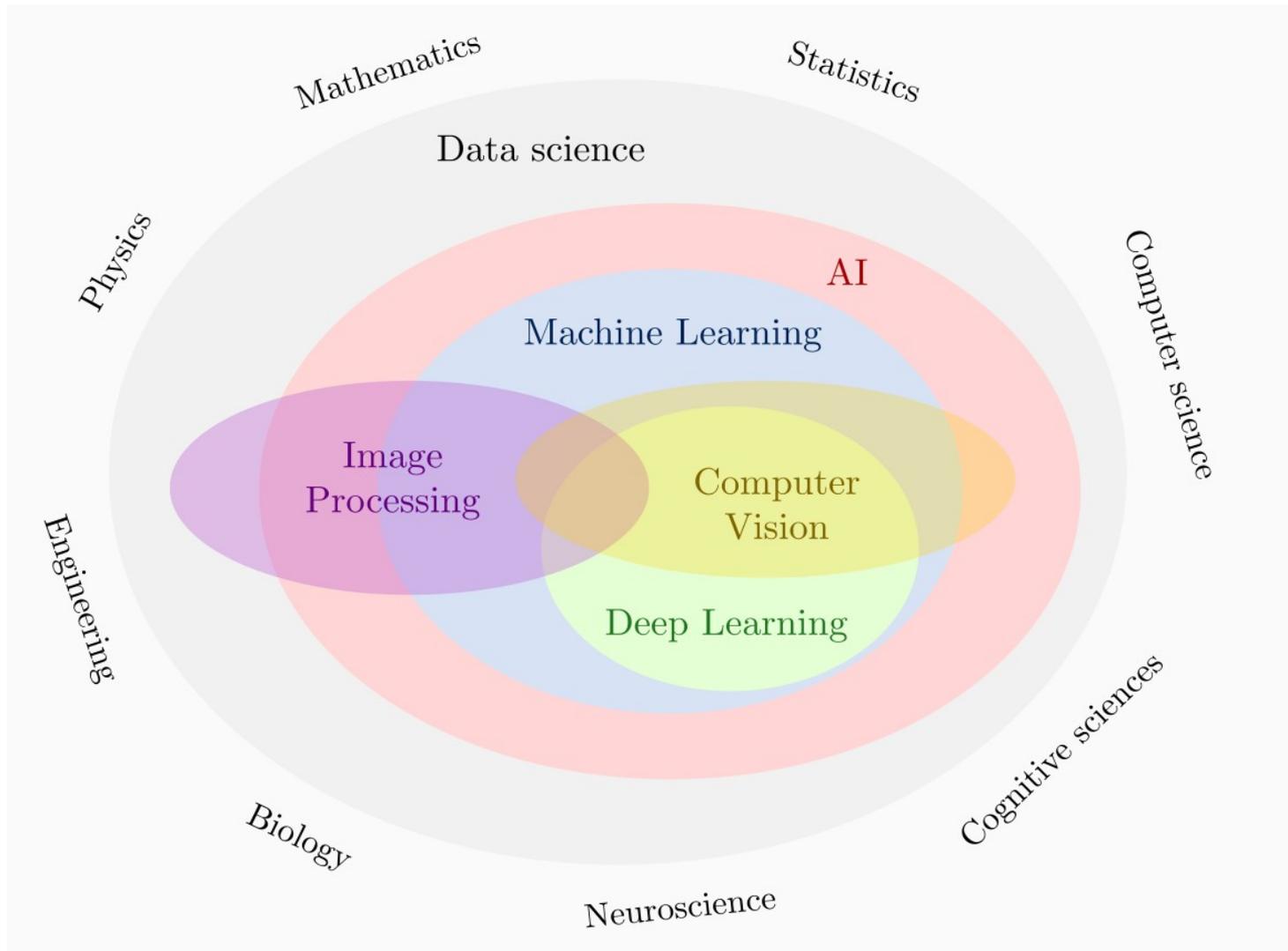
Grande réactivité, recopie de l'information

Fréquence perceptible

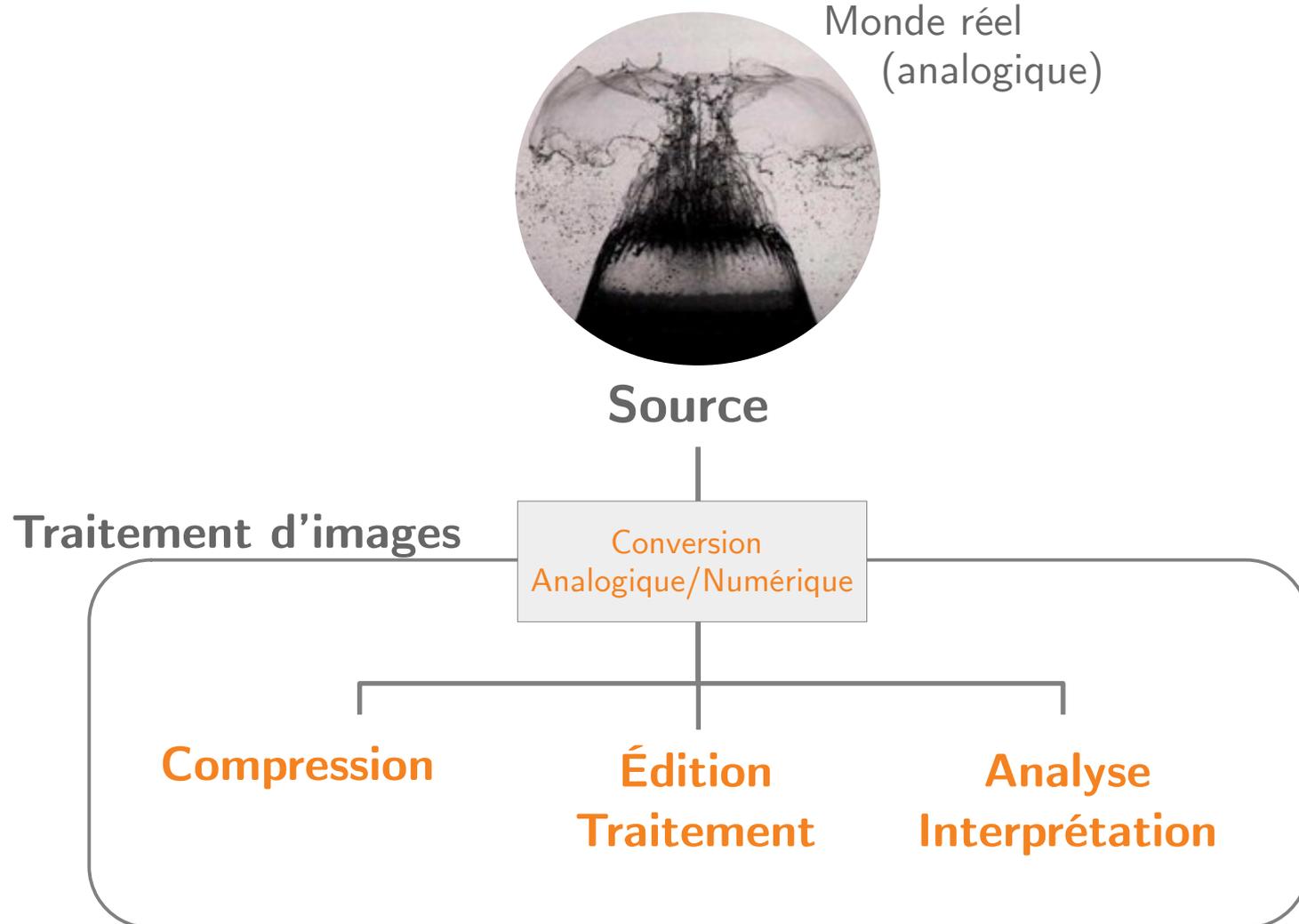


LE TRAITEMENT D'IMAGES

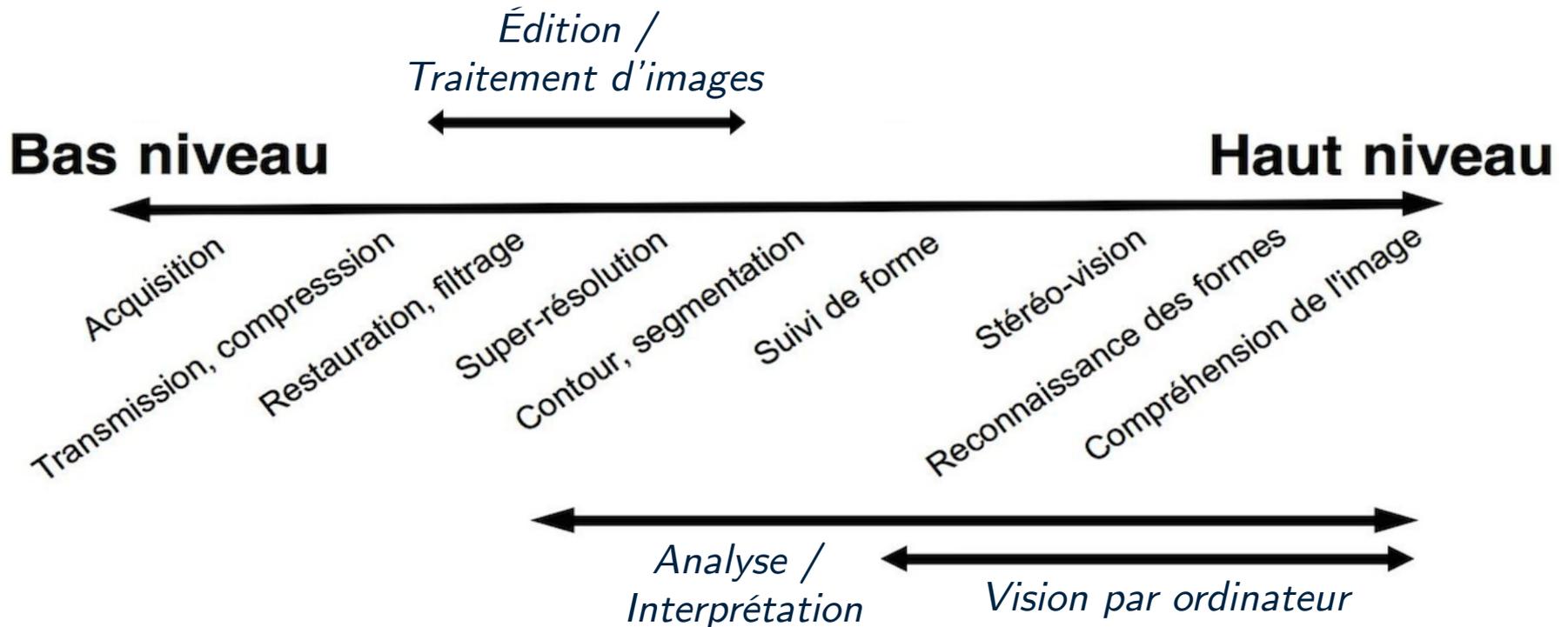
Un domaine transdisciplinaire



De l'acquisition à l'application



Vocabulaire



Édition d'images : améliorer, transformer, ... l'image (débruitage, colorisation, ...)

Analyse d'images : extraire des informations de l'image (présence d'objets, contours, ...)

Vision par ordinateur : extraire automatiquement, analyser et comprendre les informations utiles dans l'image (classification, reconnaissance, ...)

LES APPLICATIONS

Compression d'image fixe

Propriétés

- Algorithmes « *lossless* » ou « *lossy* »
- « *Lossy* » : dégradation du contenu peu *perceptible*
- Fort facteur de compression ~ 10

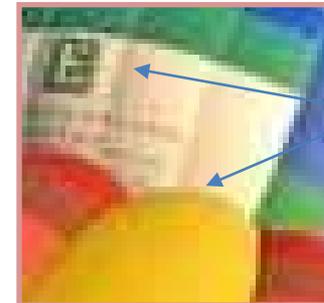
$$F = \frac{Q_{\text{initiale}}}{Q_{\text{finale}}}$$

Format	Dimensions spatiales	Taille mémoire
3 Mpixels	2048×1536	9 Mo
7 Mpixels	3072×2304	20,2 Mo
10 Mpixels	3648×2736	28,6 Mo
12 Mpixels	4000×3000	34,3 Mo

500×375, 549 ko



35,9 ko, $F \sim 15$ 17,9 ko, $F \sim 30$



dégradation
des contours
et des aplats
de couleurs

Compression de vidéo

Propriétés

- Algorithmes « *lossy* » nécessaires

Durée\Format	352×288 VCD	720×576 DVD Vidéo	1280×720 HD 720p	1920×1080 HD 1080p
1 s	7,60 Mo	31,1 Mo	69,1 Mo	156 Mo
1 min	456 Mo	1,87 Go	4,15 Go	9,33 Go
1 h	27,4 Go	112 Go	249 Go	560 Go
1 j	657 Go	2,69 To	5,97 To	13,4 To

25 image/s, format couleur = 3 canaux



Gain F ~ 100 (MPEG-4 AVC)
à 400 (MPEG-4 SVC)

Restauration de géométrie

- Utilisation de modèle de déformation de géométrie



Vue omnidirectionnelle



Reconstruction

Débruitage

Flou



Défloutage

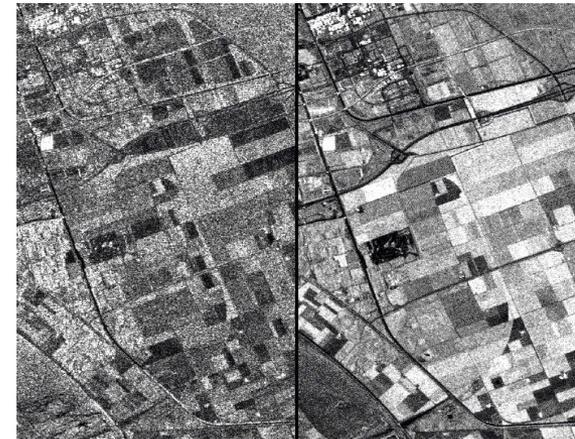
Bruit



*Trame parasite
(bruit harmonique
additif)*



*Détramage par
filtrage fréquentiel*



*Speckle
(bruit
multiplicatif)*

*Réduction par
moyennage
temporel*

Super-résolution

bicubic



SRGAN



original



Source : [Ledig et al., 2017]

Inpainting



Rayures



Texte



Inpainting



Inpainting



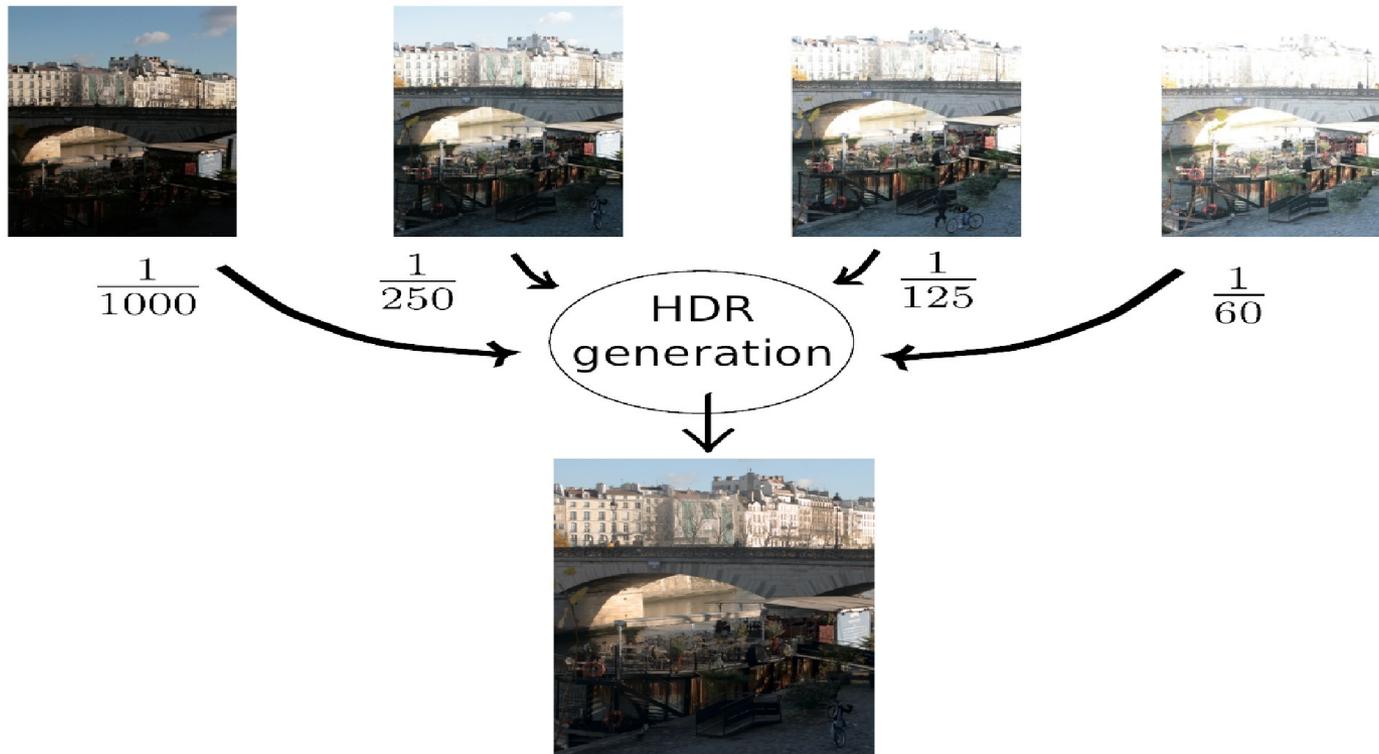
Inpainting



Source : [Gao et al., 2020]

High Dynamic Range (HDR)

- Permet d'améliorer la qualité de l'image en apportant plus de détails et de profondeur dans les zones sombres et les zones claires de l'image.
- Images produites plus proches de ce que l'œil est capable de percevoir.



High Dynamic Range

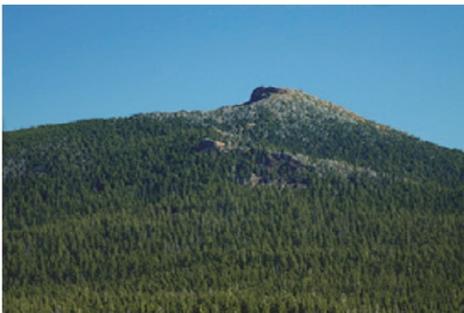


Colorisation

- Nombreuses applications :
 - Restauration de vieux films,
 - Étalonnage,
 - Applications utilisateurs, ...



Transfert de couleurs (avec référence)



Source : [Pierre et al., 2014]

Transfert de couleurs (sans référence, par apprentissage)



Source : [Zhang et al., 2016]

Transfert de couleurs (sans référence, par apprentissage)



Source : [Zhang et al. 2021]

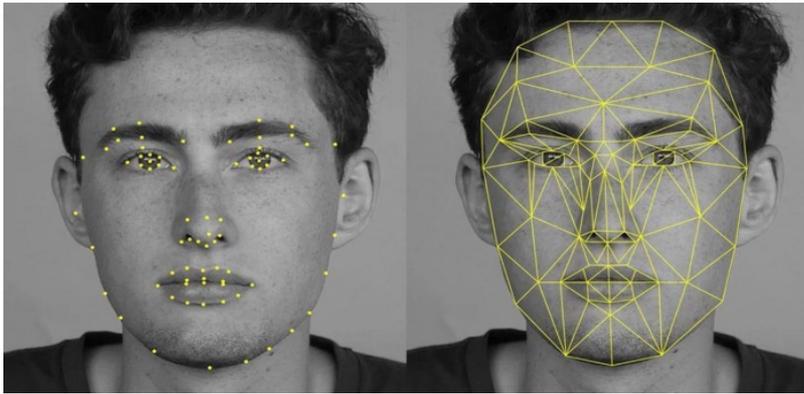
Transfert de couleurs vidéo



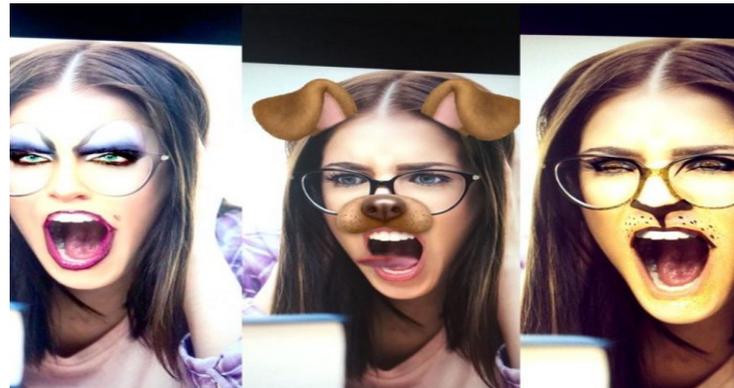
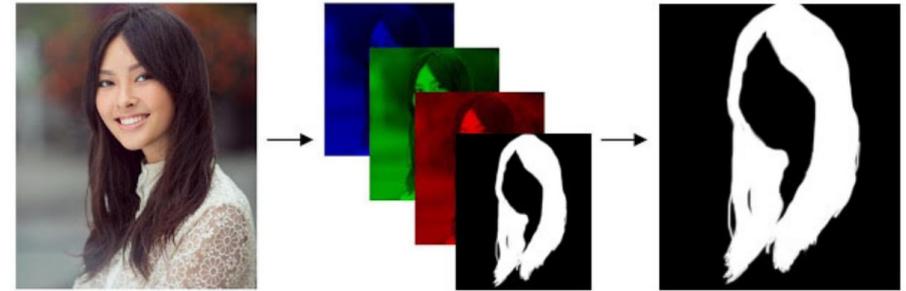
Source : [Frigo et al., 2014]

Filtres sur smartphones en temps-réel

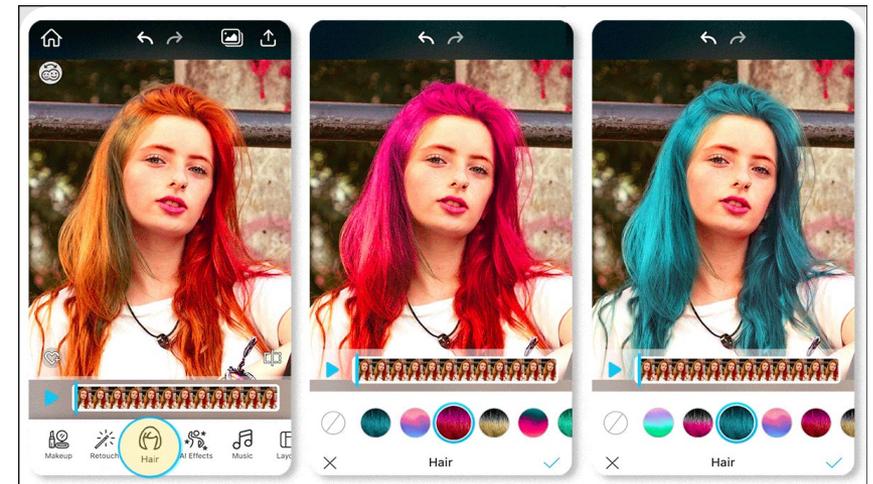
Détection des points d'intérêt du visage



Segmentation des cheveux



Projection déformée du filtre

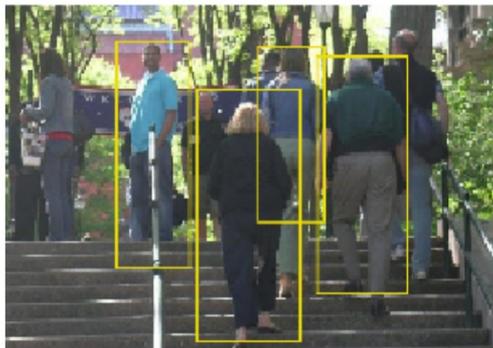


Coloration automatique

Problématiques d'analyse d'images

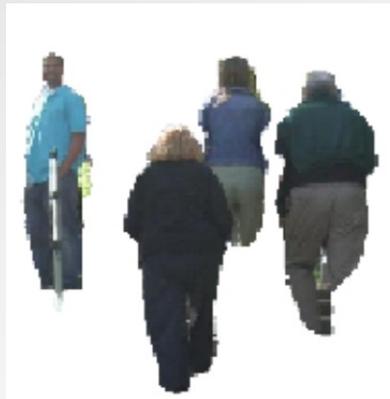
« Détection »

Présence d'un *objet*
Mouvement dans une vidéo
(*bounding box*)



« Segmentation »

Détourage précis
Mesures de volume,
contours, ...



« Reconnaissance/ Classification »

Quel objet ?

2 hommes / 2 femmes

Muriel ...
Michel ...
Jean-Marc ...
Micheline ...

Difficultés

Changement de point de vue



Illuminations différentes



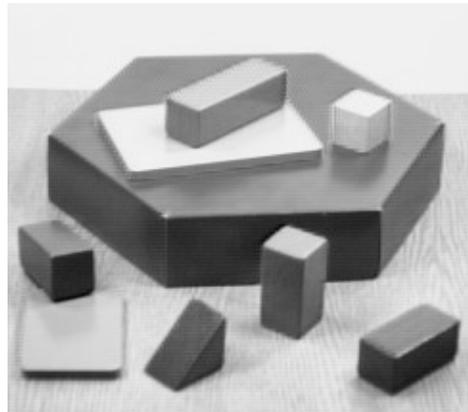
Mouvement



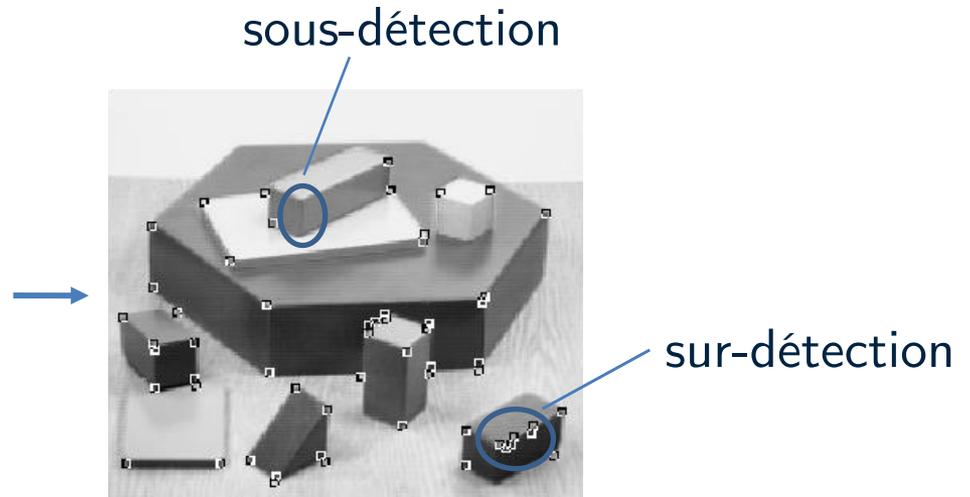
Variabilité intra-classe



Détection de coins



Scène simple



Détecteur de Harris

Granulométrie

Extraction de statistiques :

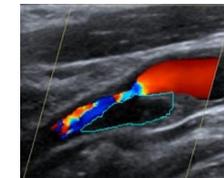
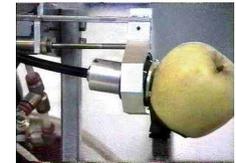
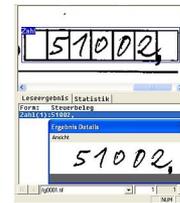
- taille moyenne,
- variance,
- répartition des tailles, ...



contour
incomplet

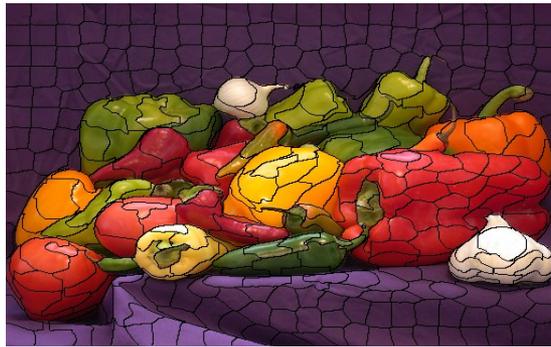
Pléthore d'applications

- Chaînes de production
 - Contrôle de posture, détection de défauts,
 - Adaptation des traitements
- Reconnaissance de texte (OCR)
 - Tri postal, contrôle radar
- Sciences de la vie et de la terre
 - Suivi de cultures, de la dynamique de catastrophes naturelles
- Sécurité/Défense
 - Détection d'objets, biométrie (empreintes digitale, vocale, rétinienne, ...)
- Commerce
 - Vente en ligne, produits dangereux
- Robotique, domotique et automobile
 - Détection et contournement d'obstacles, commande vocale
- Médecine
 - Détection d'anomalies et assistance chirurgicale



Segmentation simple

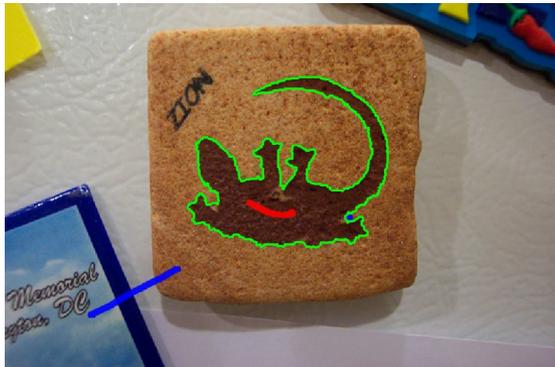
- Sur-segmentation « à l'aveugle », n'utilisant que l'information RGB de l'image
→ Réduction du nombre d'éléments à traiter



Source : [Achanta et al., 2012]



- Segmentation d'un objet par interaction utilisateur



Source : [Price et al., 2010]



Source : [Kirillov et al., 2023] (Démonstration)

Édition d'images – Incrustation (Chroma-Keying)



Source : [Sarim et al., 2009]

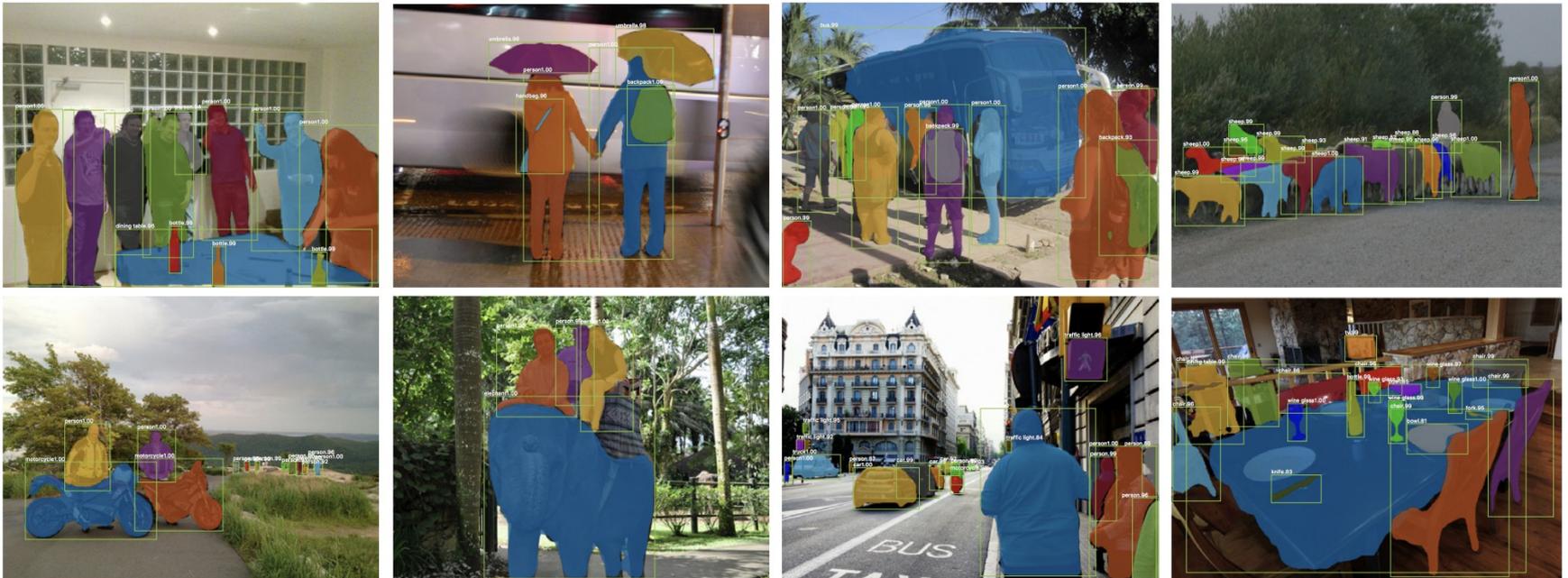


Source : [Gupta et al., 2017]

Pour vos images, essayez : <https://www.remove.bg/fr>

Segmentation sémantique / Classification pixellique

- Association d'un attribut de classe (au sens objet) à chaque pixel
- Détection des instances (différents éléments d'une même classe)



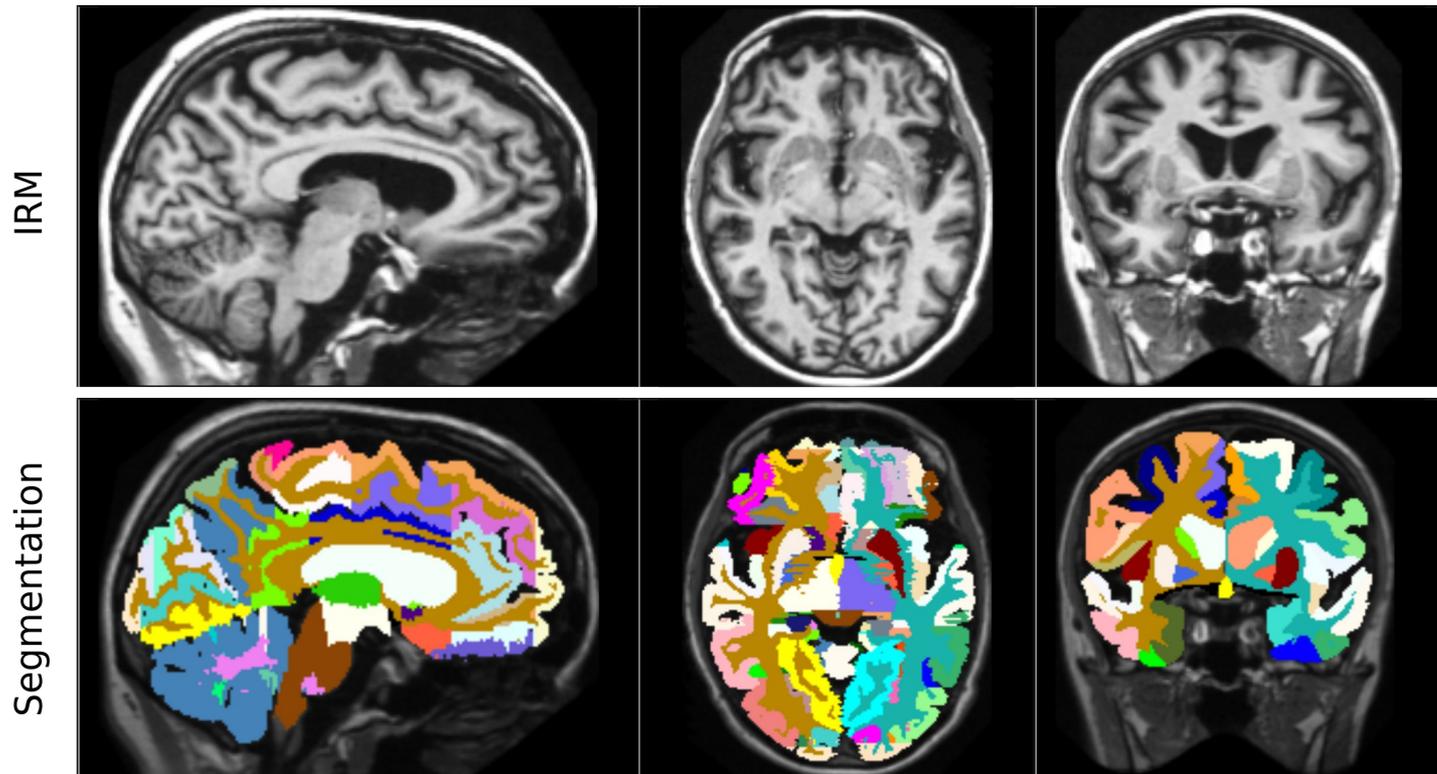
Source : [He et al., 2017]

Imagerie médicale

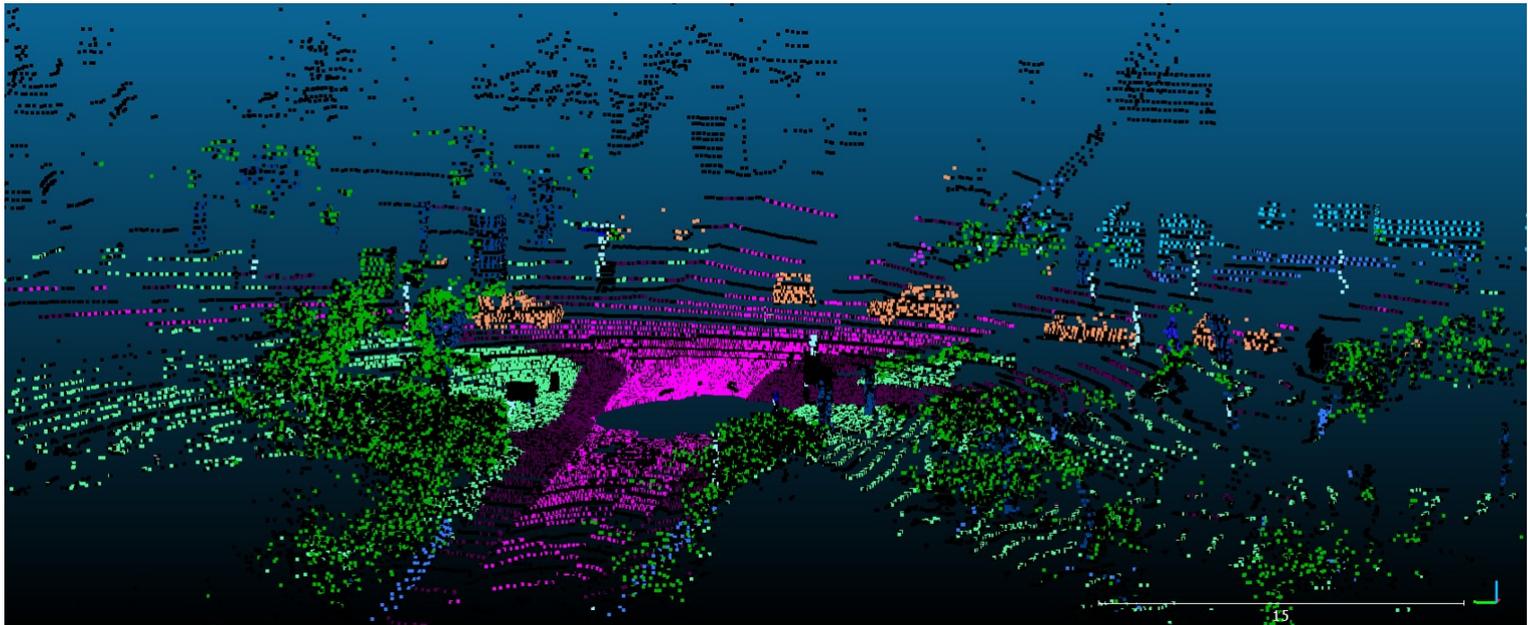
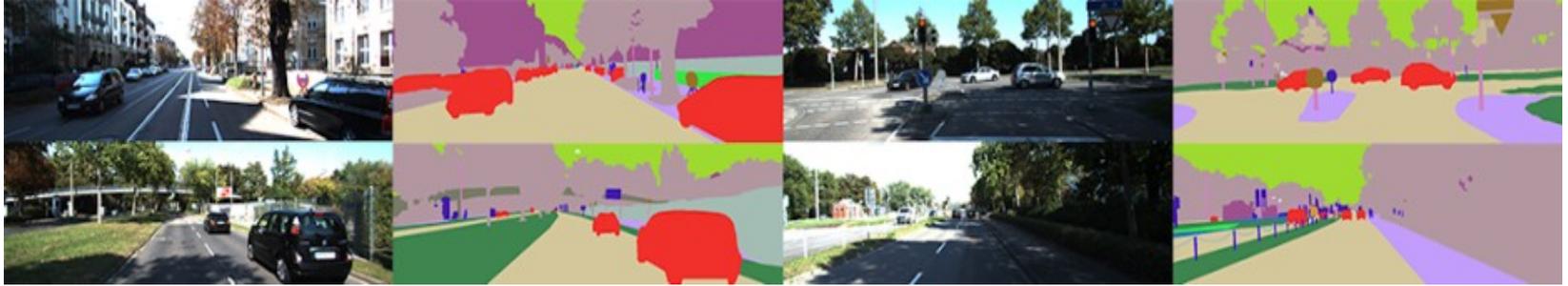
Très coûteux pour les médecins

→ Suivi des patients facilité

→ Études statistiques à grande échelle



Multi-modalités - Navigation autonome



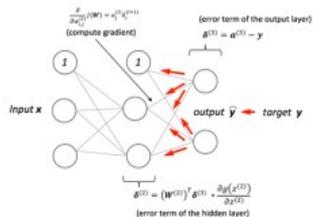
Source : [Geiger et al., 2012]

Révolution avec l'apprentissage profond

Chronologie



1958 Perceptron



1974 Backpropagation



Convolution Neural Networks for Handwritten Recognition

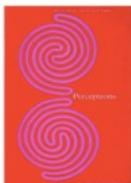
1998



Google Brain Project on 16k Cores

2012

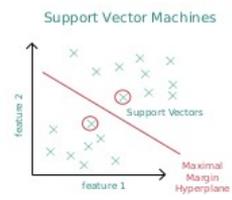
awkward silence (AI winter)



1969 Perceptrons book

~1980 Multilayer network

1995 SVM reigns

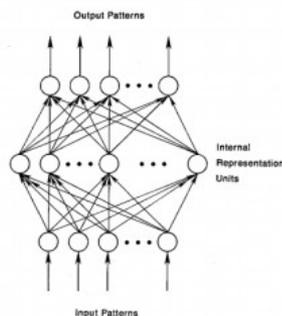
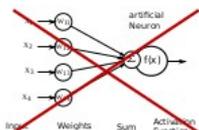


2006 Restricted Boltzmann Machine



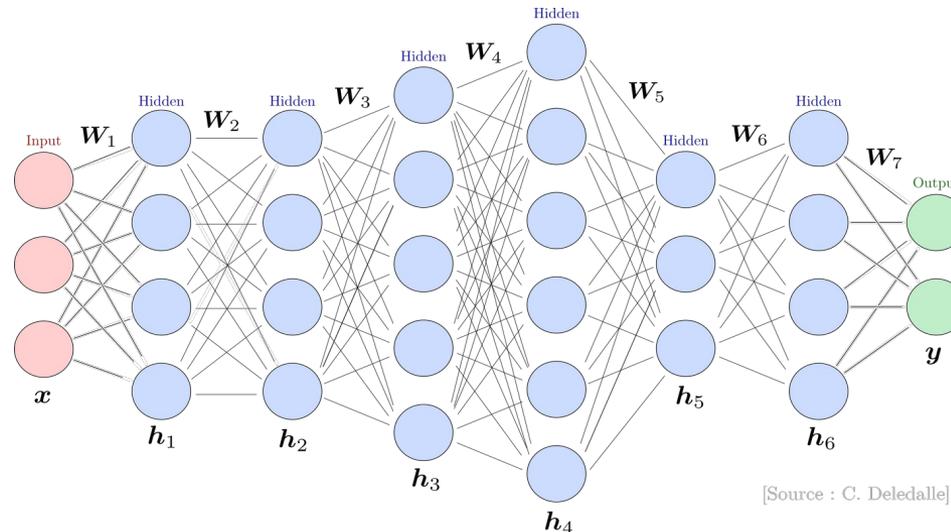
2012 AlexNet wins ImageNet IMAGENET

Perceptron criticized



Principe

- **Réseaux de neurones** : Méthode d'apprentissage "inspirée" du cerveau humain.
 - Consiste en l'inter-connexion de petites unités (fonctions paramétriques), appelées neurones.
 - Milliers/millions de paramètres de transformation initialisés aléatoirement et raffinés itérativement grâce à une large base d'images (apprentissage)
- **Apprentissage profond** : multiples couches de neurones inter-connectées
 - Rendu efficace par plusieurs optimisations et implémentation sur GPU



Lien historique avec l'image → Classification

ImageNet :

- 12 millions d'images étiquetées
- 22 000 classes
- Étiquetées par crowd-sourcing (Amazon Mechanical Turk)

ImageNet challenge (ILSVRC) :

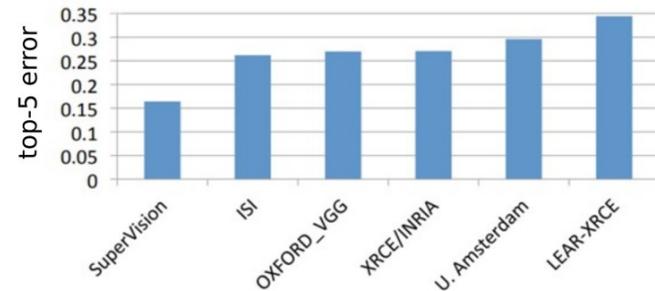
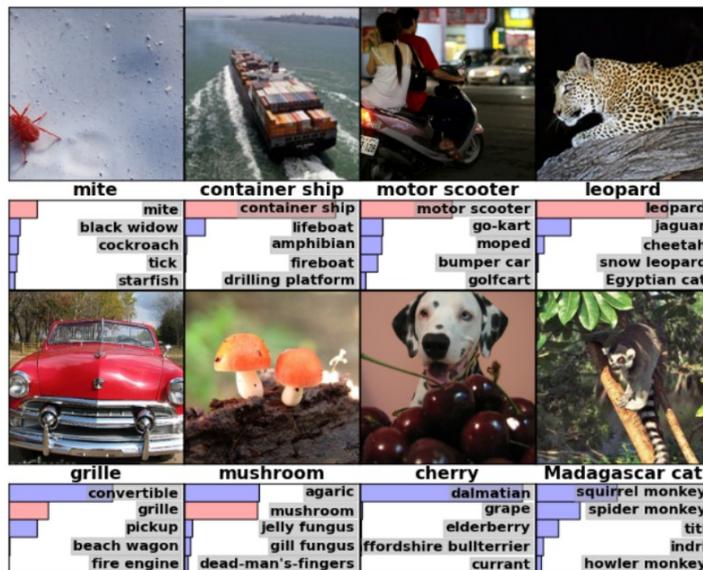
- Challenge annuel depuis 2010
- Limité à 1 000 classes
- 1.2 million d'images de taille 256×256 pour l'entraînement
- 50 000 images pour la validation 100 000 pour le test



Révolution avec l'apprentissage profond

AlexNet (Supervision) (Krizhevsky, Sutskever, Hinton, 2012)

- Large réseaux de neurones à convolution (62.3 millions de paramètres)
- 6 jours d'apprentissage sur 2 GPUs (GTX 580 3GB)
- Utilisation des nouveaux outils d'apprentissage profond (calcul sur GPU, normalisation, optimisation, augmentation de données, ...)
- 1er au challenge ILSVRC2012 avec une erreur top-5 à 16.4% (2e : 26.2%).



→ Explosion de l'apprentissage profond depuis 2010 dans **tous les domaines**

Nouvelles applications

Captioning : Écrire une phrase qui décrit l'action dans l'image



"girl in pink dress is jumping in air."



"black and white dog jumps over bar."



"young girl in pink shirt is swinging on swing."



"man in blue wetsuit is surfing on wave."



"little girl is eating piece of cake."



"baseball player is throwing ball in game."



"woman is holding bunch of bananas."



"black cat is sitting on top of suitcase."

Source : [Karpathy, et al., 2015]

Transfert de style : Peinture



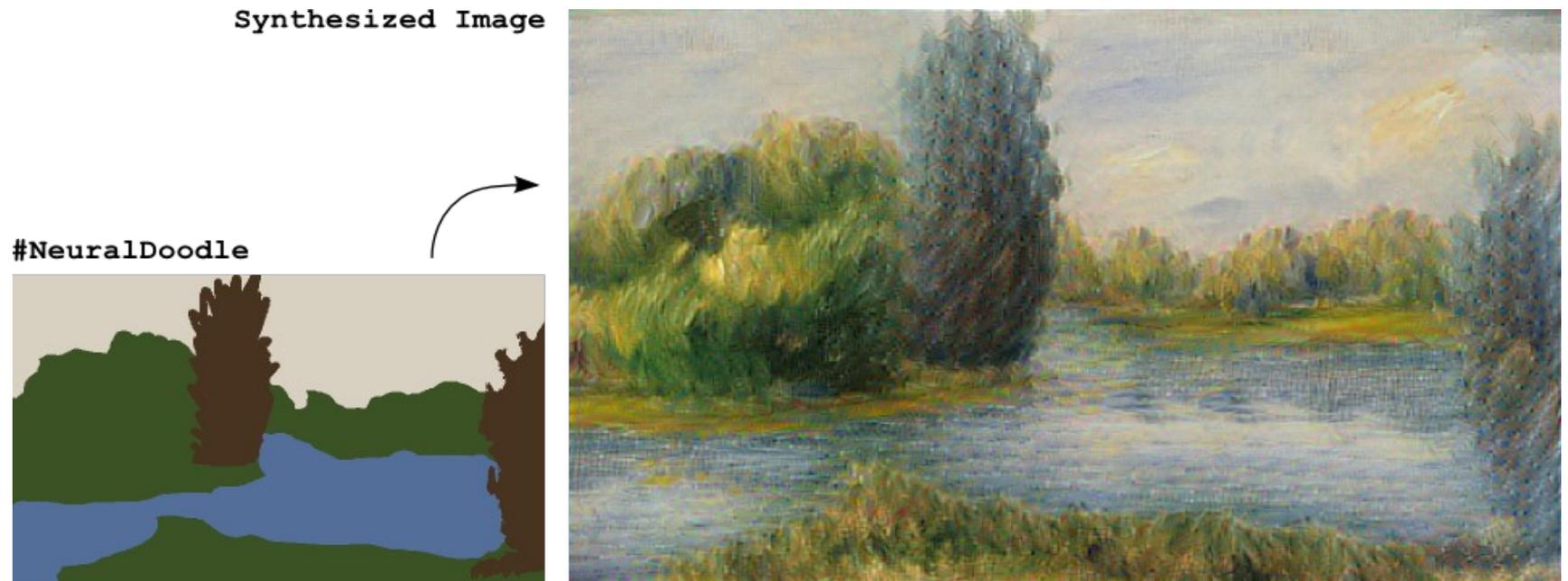
Nouvelles applications

Transfert de style : Peinture



Source : [Gatys et al., 2015]

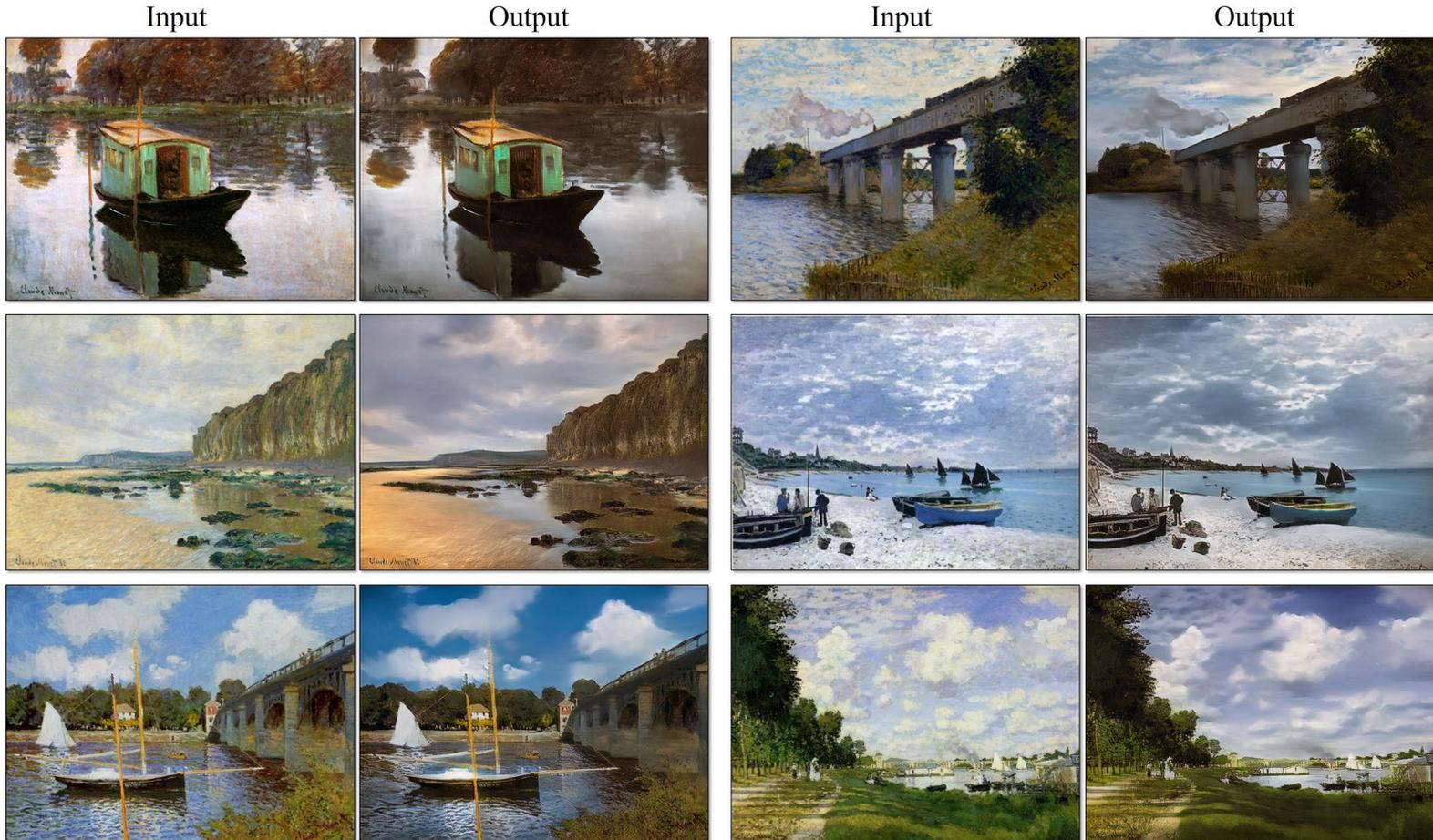
Transfert de style : Synthèse sémantique de peinture



Source : [Champandard et al., 2016]

Nouvelles applications

Transfert de style : Peinture → photo



Source : [Zhu et al., 2017]

Nouvelles applications

Transfert de style : photo → manga



Source

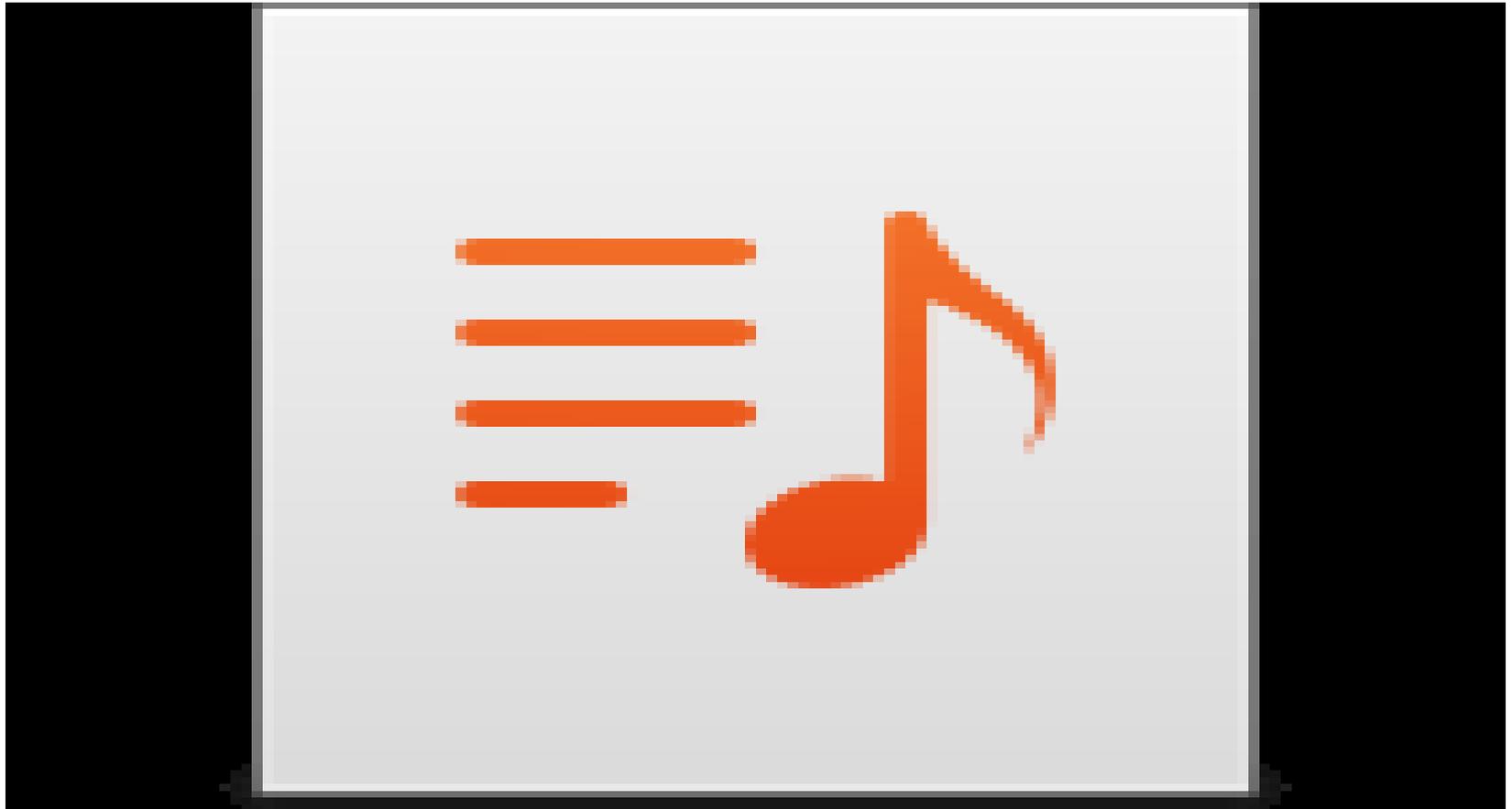
Nouvelles applications

Transfert de style : photo → manga



Source

Transfert de style : Réalité augmentée en vidéo



Source : [Ruder et al., 2016]

Nouvelles applications

Synthèse d'images : mono-classe, faible résolution



Source : [Radford et al., 2016]

Synthèse d'images : changement de classes



Source : [\[Zhu et al., 2017\]](#)

Nouvelles applications

Synthèse d'images : visages réalistes



Source : [Karras et al., 2019]

Synthèse d'images : morphing



Synthèse d'images : depuis une phrase descriptive

'A painting of a squirrel eating a burger'



'A watercolor painting of a chair that looks like an octopus'



'A shirt with the inscription: "I love generative models!"'



Source : [Rombach, et al., 2021]

Synthèse d'images : depuis une phrase descriptive

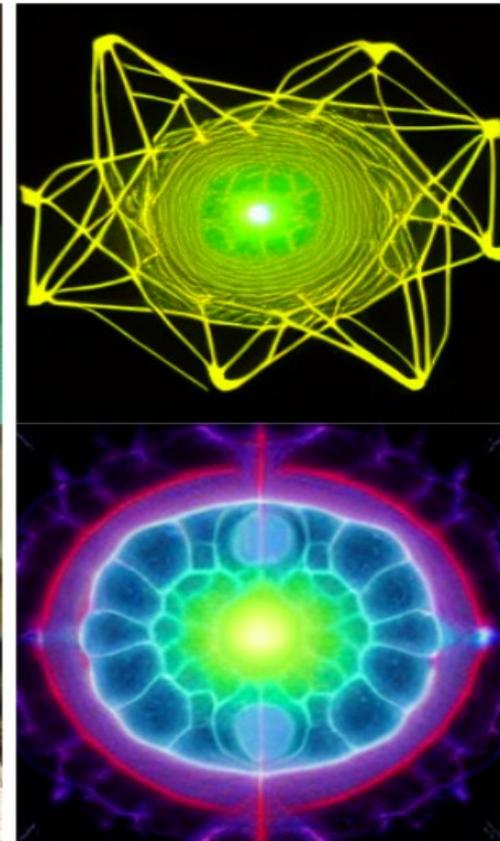
'A zombie in the style of Picasso'



'An image of an animal half mouse half octopus'



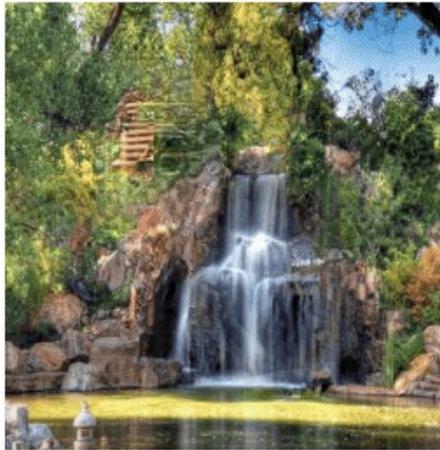
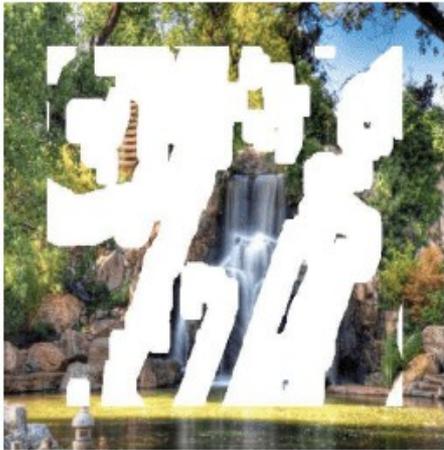
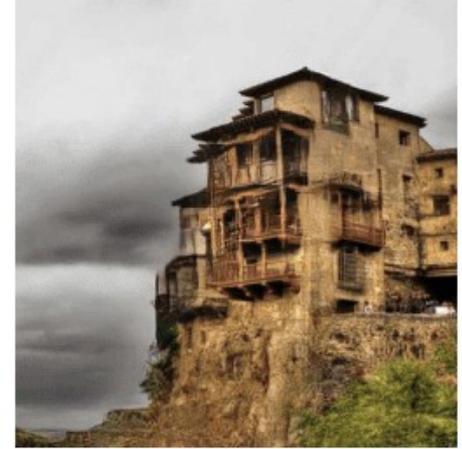
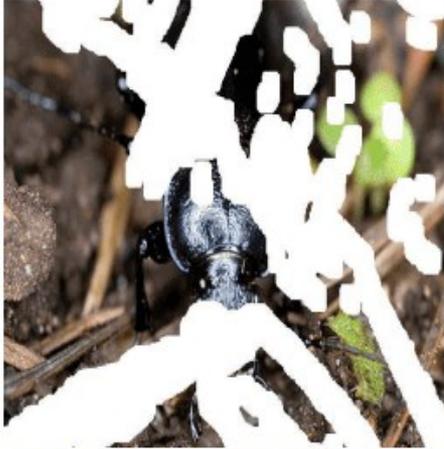
'An illustration of a slightly conscious neural network'



Source : [Rombach, et al., 2021]

Nouvelles applications

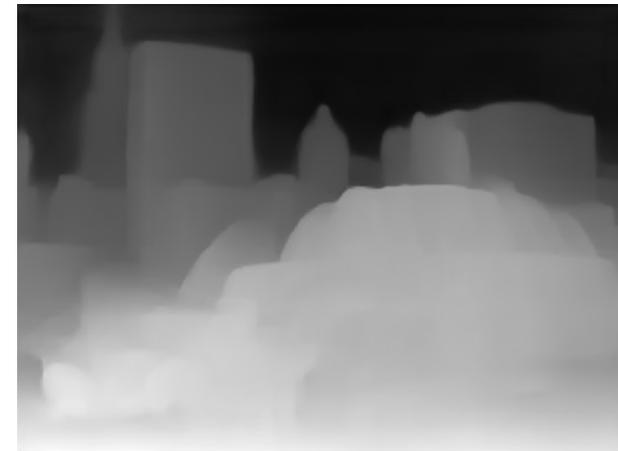
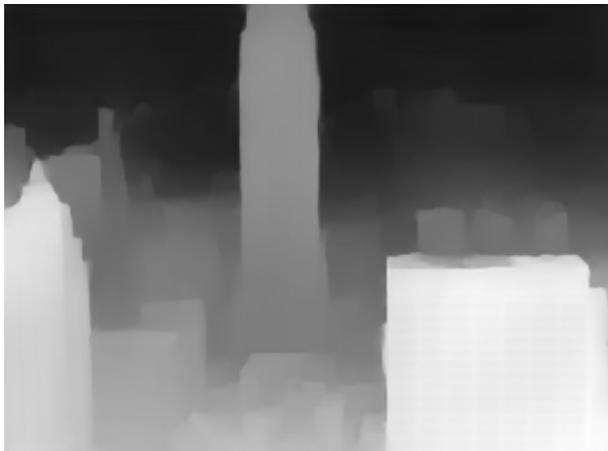
Inpainting++



Source

Nouvelles applications

Estimation de profondeur depuis une image 2D



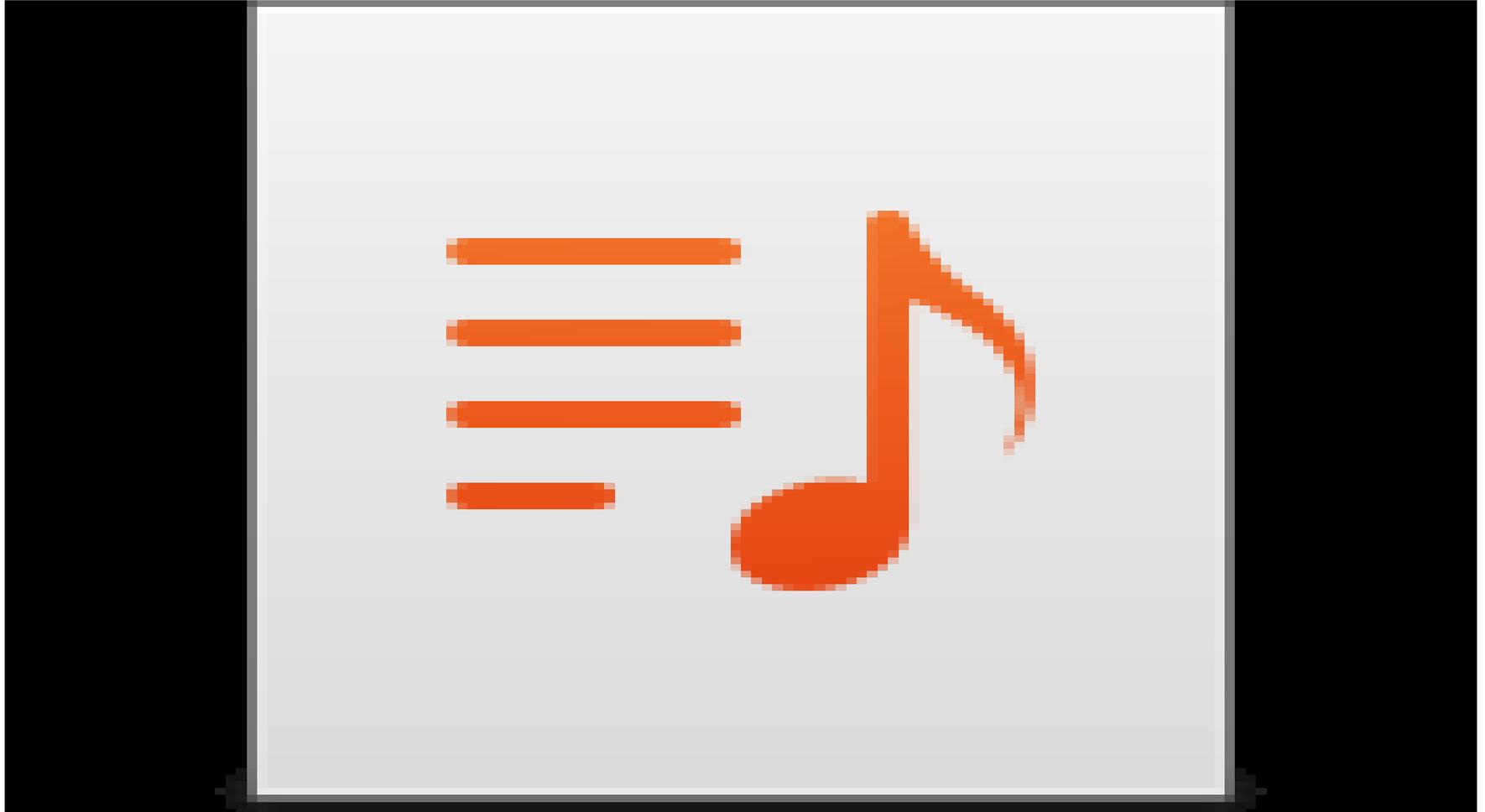
Source : [Li et al., 2018]

Simulation de 3D par inpainting de profondeur et d'image



Source : [Shih et al., 2020]

Assistance par reconnaissance d'objets en temps-réel



Source

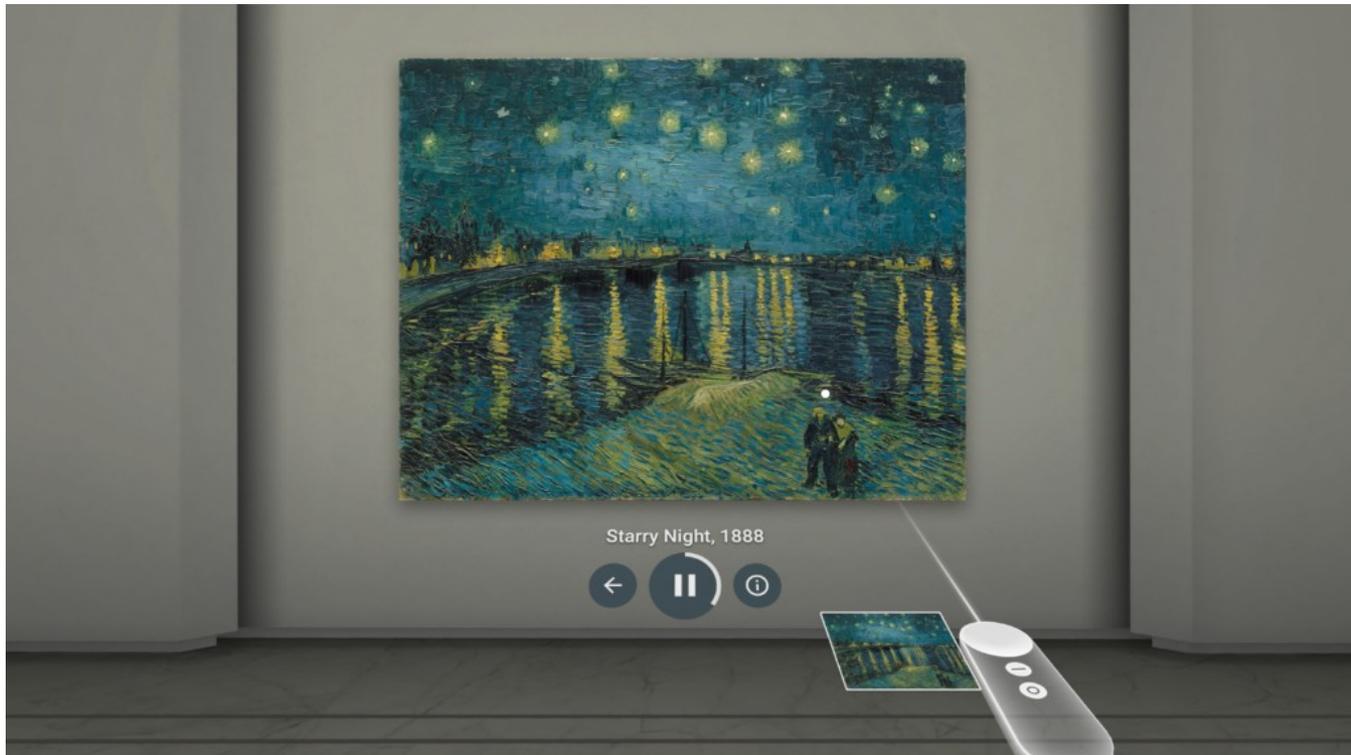
Nouvelles applications

Réalité augmentée – Traduction



Source

Réalité augmentée – Aménagement d'intérieur



Source

TOP 15 des applications gratuites de réalité augmentée

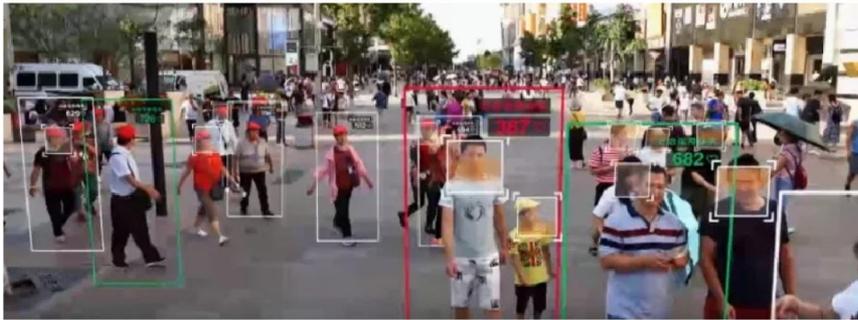
Réalité augmentée – Jeu vidéo



Nouveaux risques

Vidéo surveillance

Outil de contrôle / Abus ?



The Chinese state wants to control its citizens via a system of social scoring that punishes behavior it doesn't approve of. Image Credit: Telecoms

Navigation autonome

Responsabilité en cas d'accident ?



Tesla said autopilot was activated during a fatal Model X crash last week in California.

Génération automatique de contenus (vidéo, image, son)

Preuve de l'authenticité ?

Graphistes ?

Propriété intellectuelle ?



Vidéo de Vladimir Poutine interviewé par son double DeepFake



Image générée automatiquement par Midjourney



Générique de One Piece chanté par Johnny Hallyday

- Introduction
- **Formation / Acquisition**
- **Image numérique**
 - Format/Affichage/Synthèse
 - Espaces couleur caractéristiques : compression, esquisse, illusion
- **Traitements**
 - Filtrage linéaire / non linéaire : débruitage, anonymisation
 - Détection de contours : réhaussement de contraste
- **Transformée de Fourier**
 - Application : recouvrement fréquentiel
- **Compression d'images**
 - Application : algorithme JPEG
- **Transformation spatiales**