

Complément

Convolution et CNN

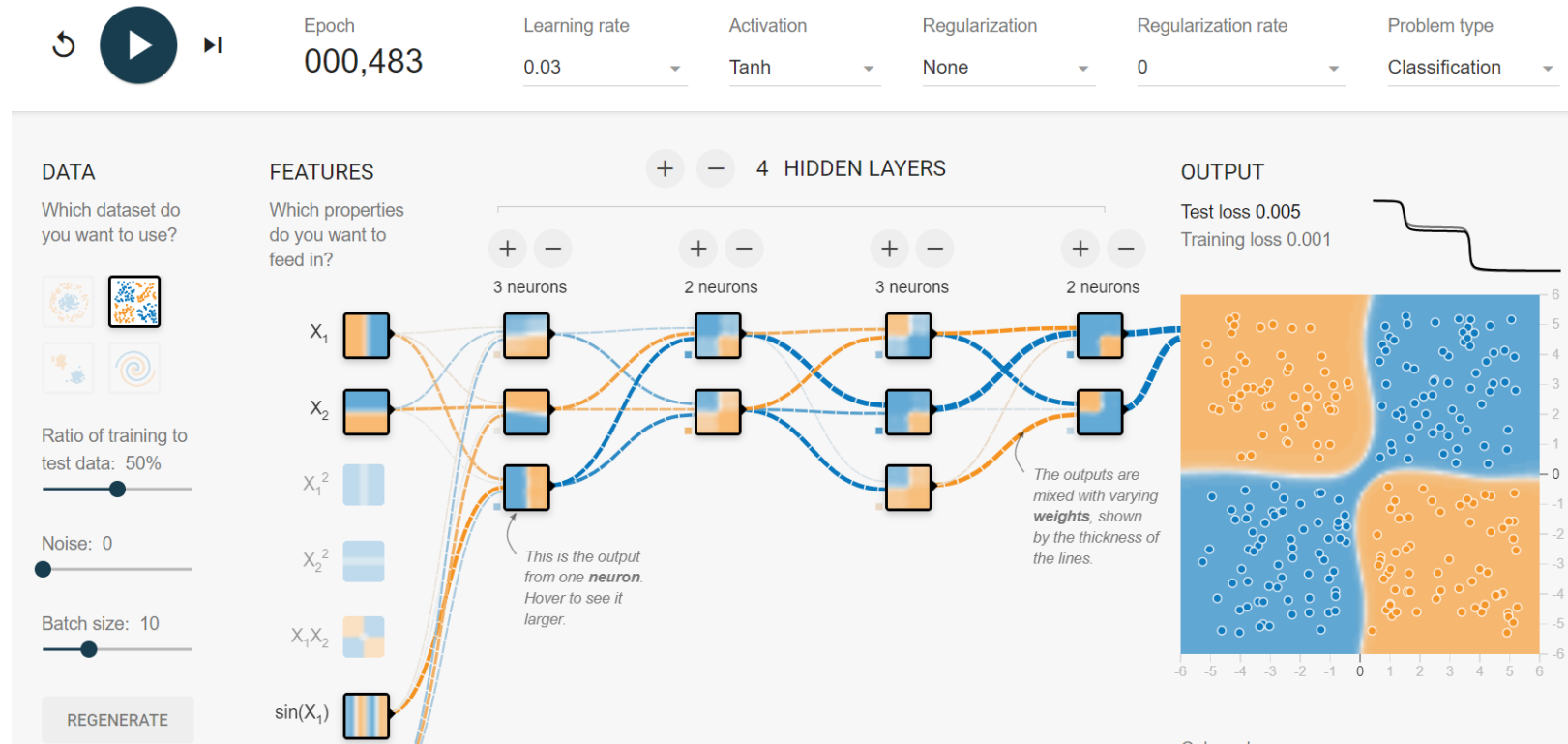
Diffusion



Concepts fondamentaux d'un réseau de neurones

Le réseau de neurones : architecture, fonctions d'activation et de pondération

Tensorflow playground

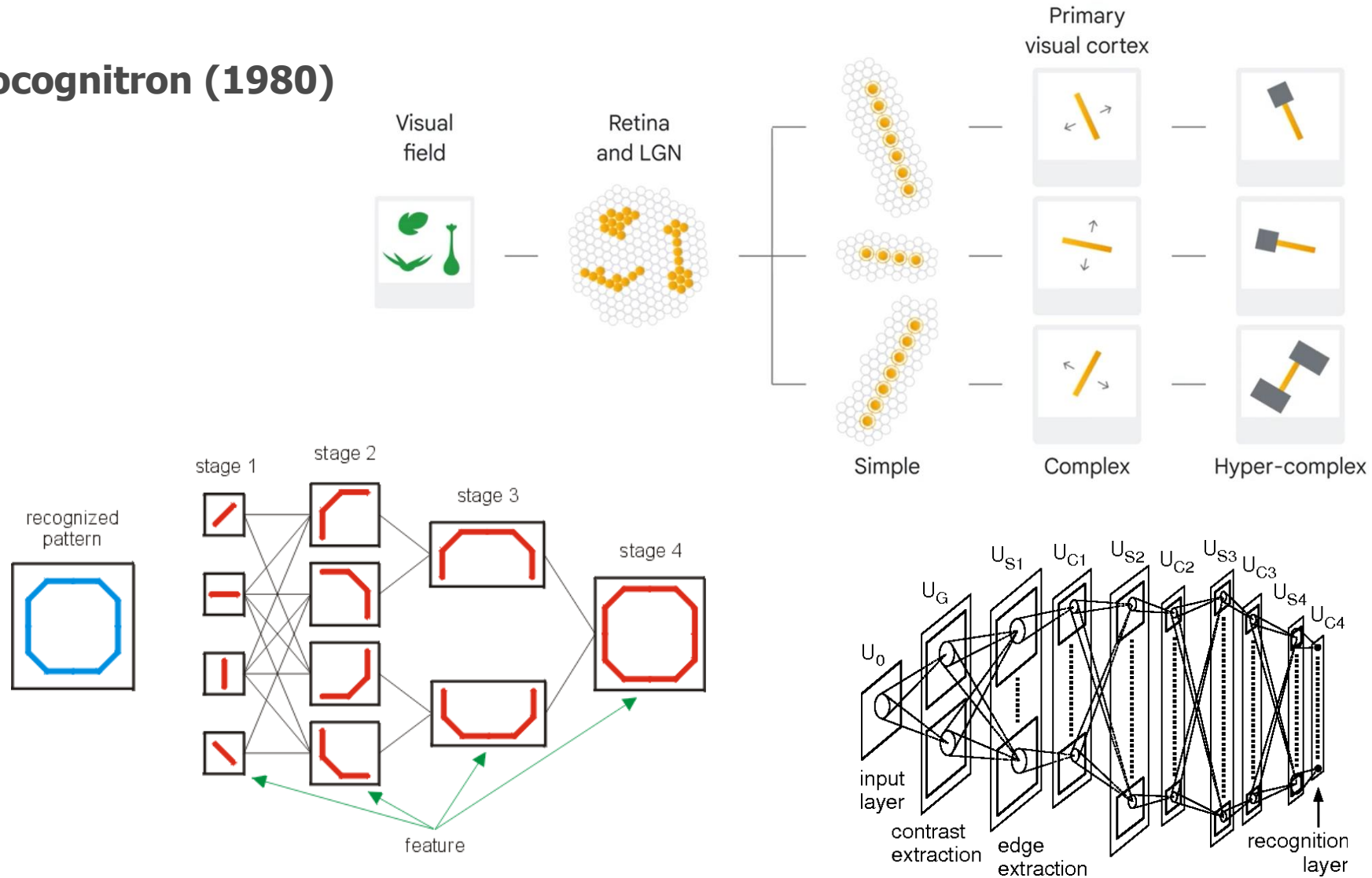


<https://playground.tensorflow.org>

Convolutional Neural Networks (CNN)

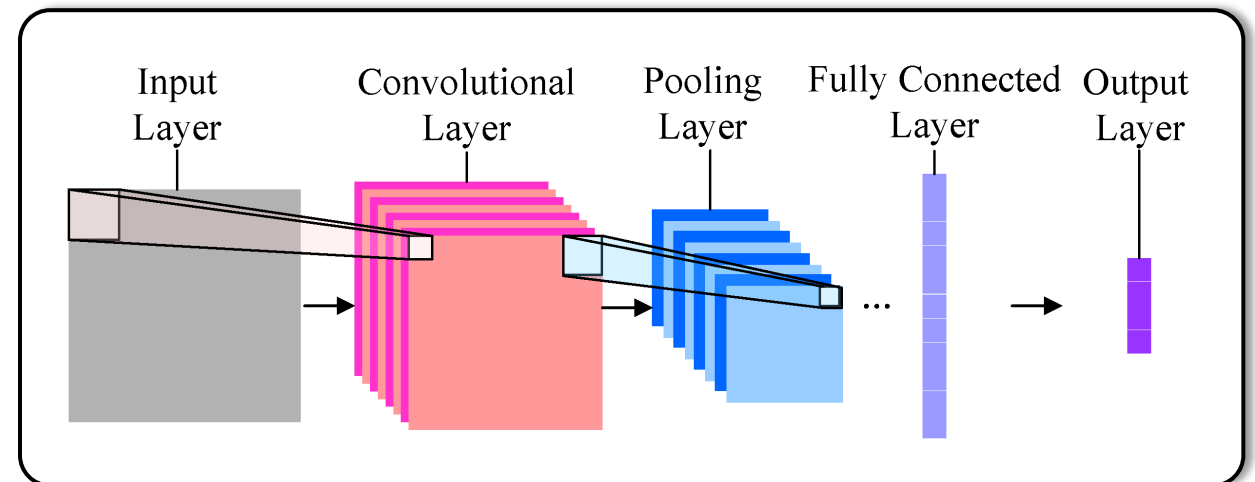
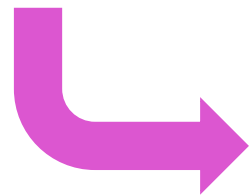
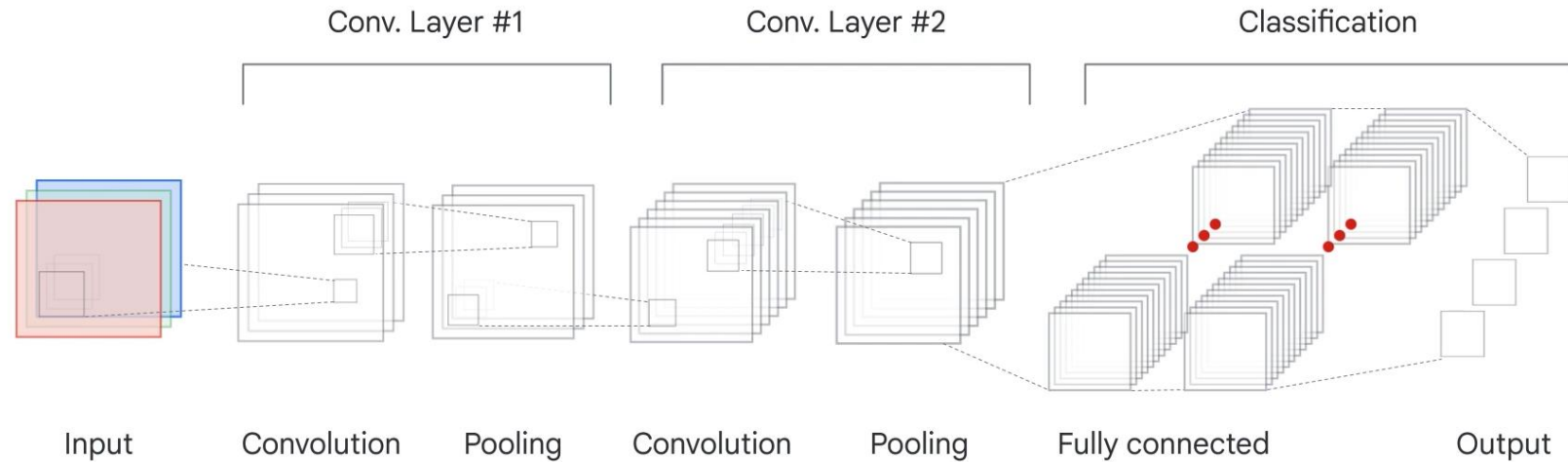
Présentation des CNN : principes fondamentaux et applications

Neocognitron (1980)



Convolutional Neural Networks (CNN)

Présentation des CNN : principes fondamentaux et applications



Convolutional Neural Networks (CNN)

Présentation des CNN : principes fondamentaux et applications

Deep learning with COTS HPC systems

Adam Coates
Brody Huval
Tao Wang
David J. Wu
Andrew Y. Ng
Stanford University Computer Science Dept., 353 Serra Mall, Stanford, CA 94305 USA

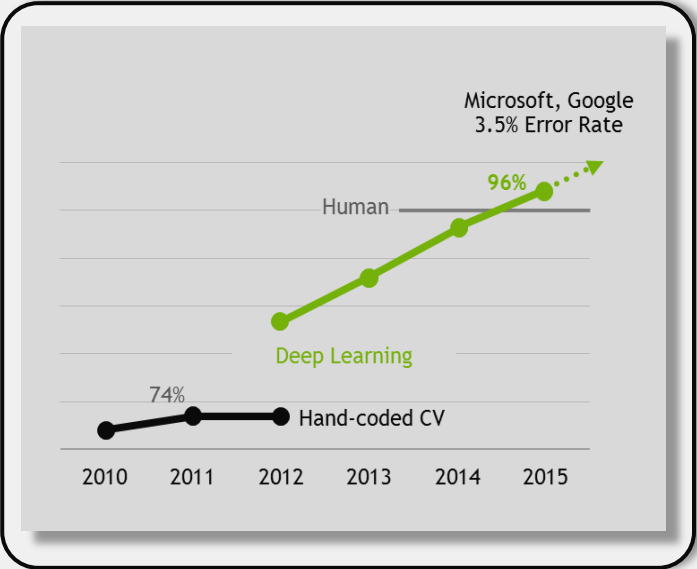
Bryan Catanzaro
NVIDIA Corporation, 2701 San Tomas Expressway, Santa Clara, CA 95050

Abstract

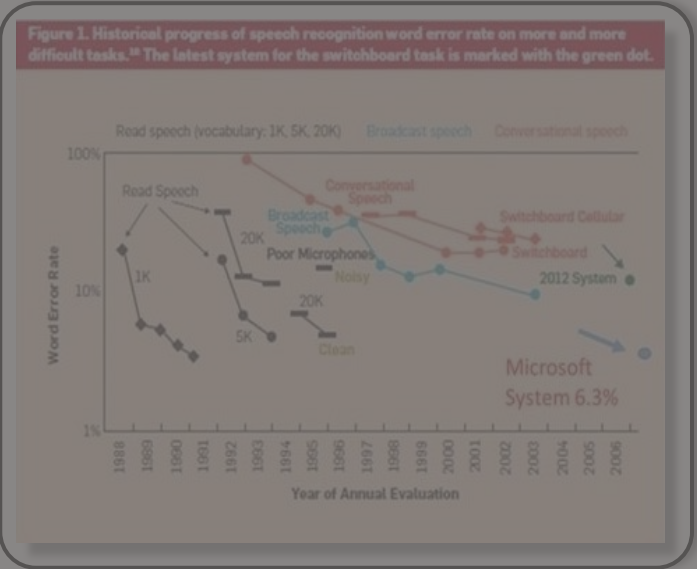
Scaling up deep learning algorithms has been shown to lead to increased performance in benchmark tasks and to enable discovery of complex high-level features. Recent efforts to train extremely large networks (with over 1 billion parameters) have relied on cloud-like computing infrastructure and thousands of CPU cores. In this paper, we present technical details and results from our own system based on Commodity Off-The-Shelf High Performance Computing (COTS HPC) technology: a cluster of GPU servers with InfiniBand interconnects and MPI. Our system is able to train 1 billion parameter networks on

to detect objects when trained from unlabeled images alone (Cortes et al., 2012; Le et al., 2012). The very largest of these systems has been constructed by Le et al. (Le et al., 2012) and Dean et al. (Dean et al., 2012), which is able to train neural networks with over 1 billion trainable parameters. While such extremely large networks are potentially valuable objects of AI research, the expense to train them is overwhelming: the distributed computing infrastructure (known as "DistBelief") used for the experiments in (Le et al., 2012) manages to train a neural network using 10000 CPU cores (in 1000 machines) in just a few days, yet this level of resource is likely beyond those available to most deep learning researchers. Less clear still is how to continue scaling significantly beyond this size

2012 : Deep Learning et GPUs



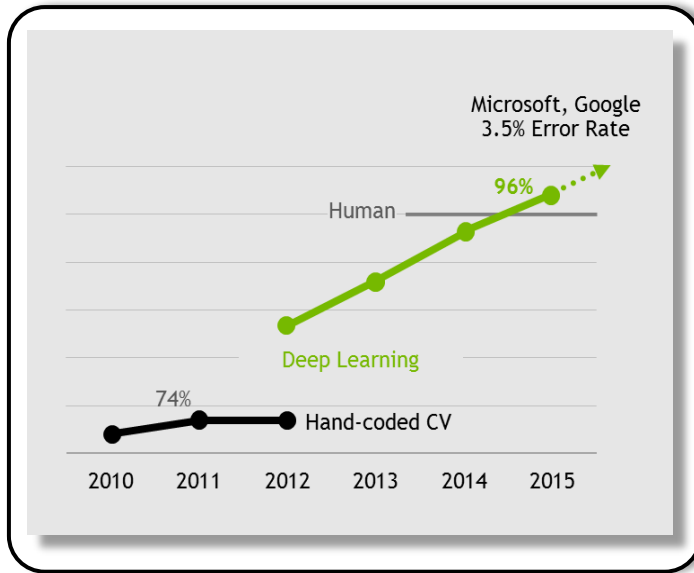
2015 : ImageNet challenge - de meilleurs résultats que l'homme dans la reconnaissance d'image grâce au Deep Learning



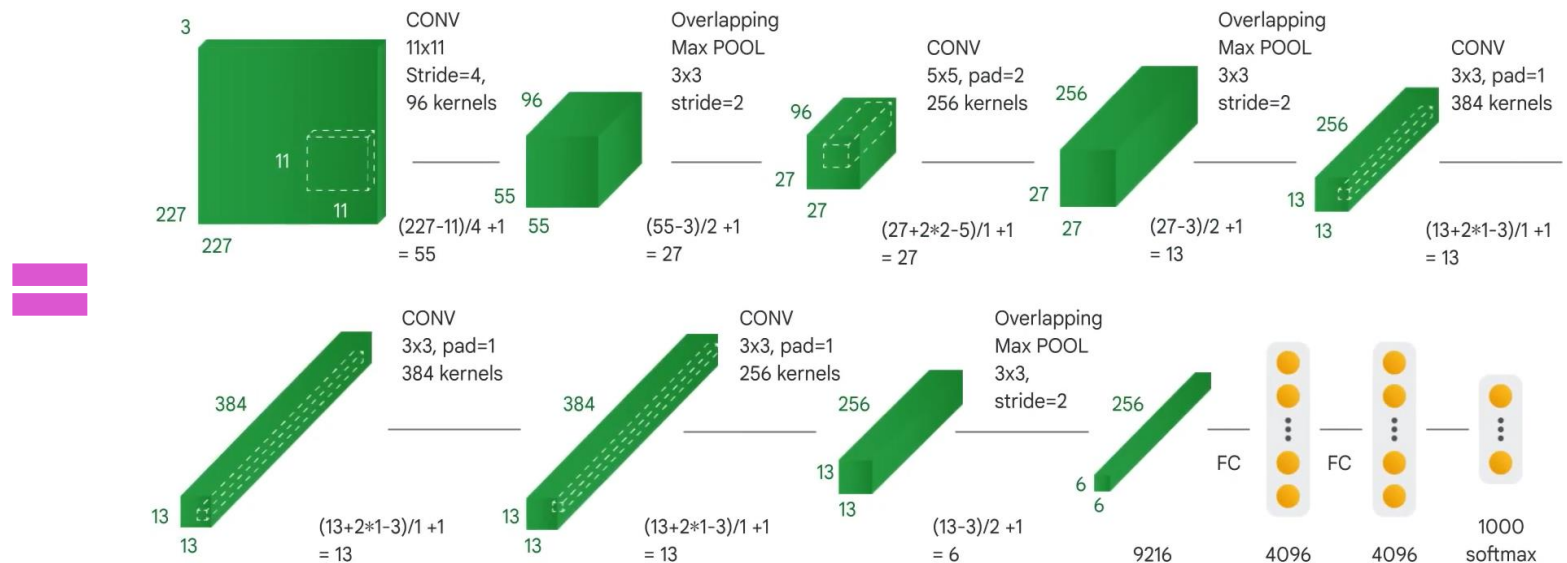
2016 : Microsoft annonce que son IA comprend la parole aussi bien qu'un humain

Convolutional Neural Networks (CNN)

Présentation des CNN : principes fondamentaux et applications



2015 : ImageNet challenge - de meilleurs résultats que l'homme dans la reconnaissance d'image grâce au Deep Learning



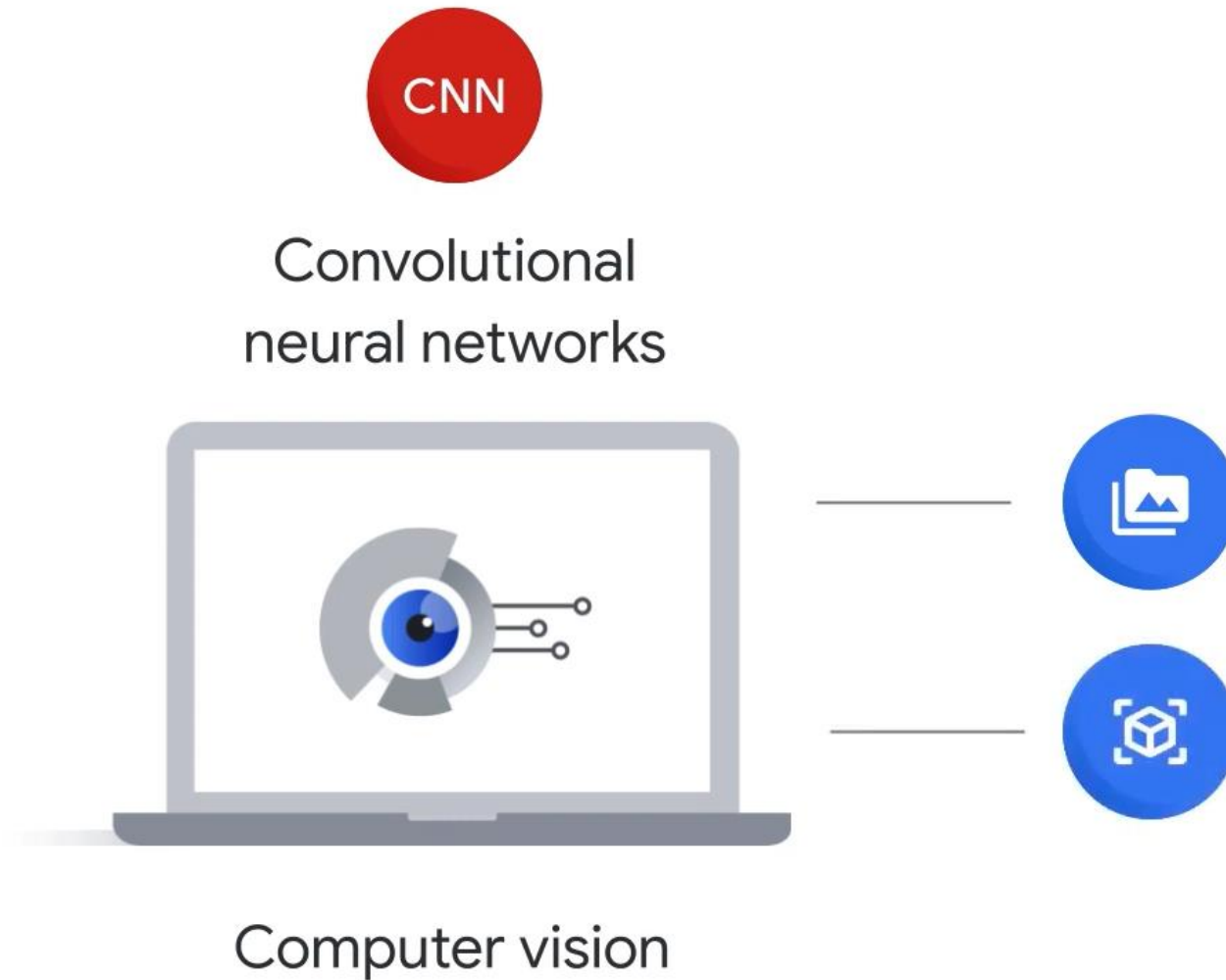
ImageNet classification with deep convolutional neural networks

Authors: Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton

<https://dl.acm.org/doi/10.5555/2999134.2999257>

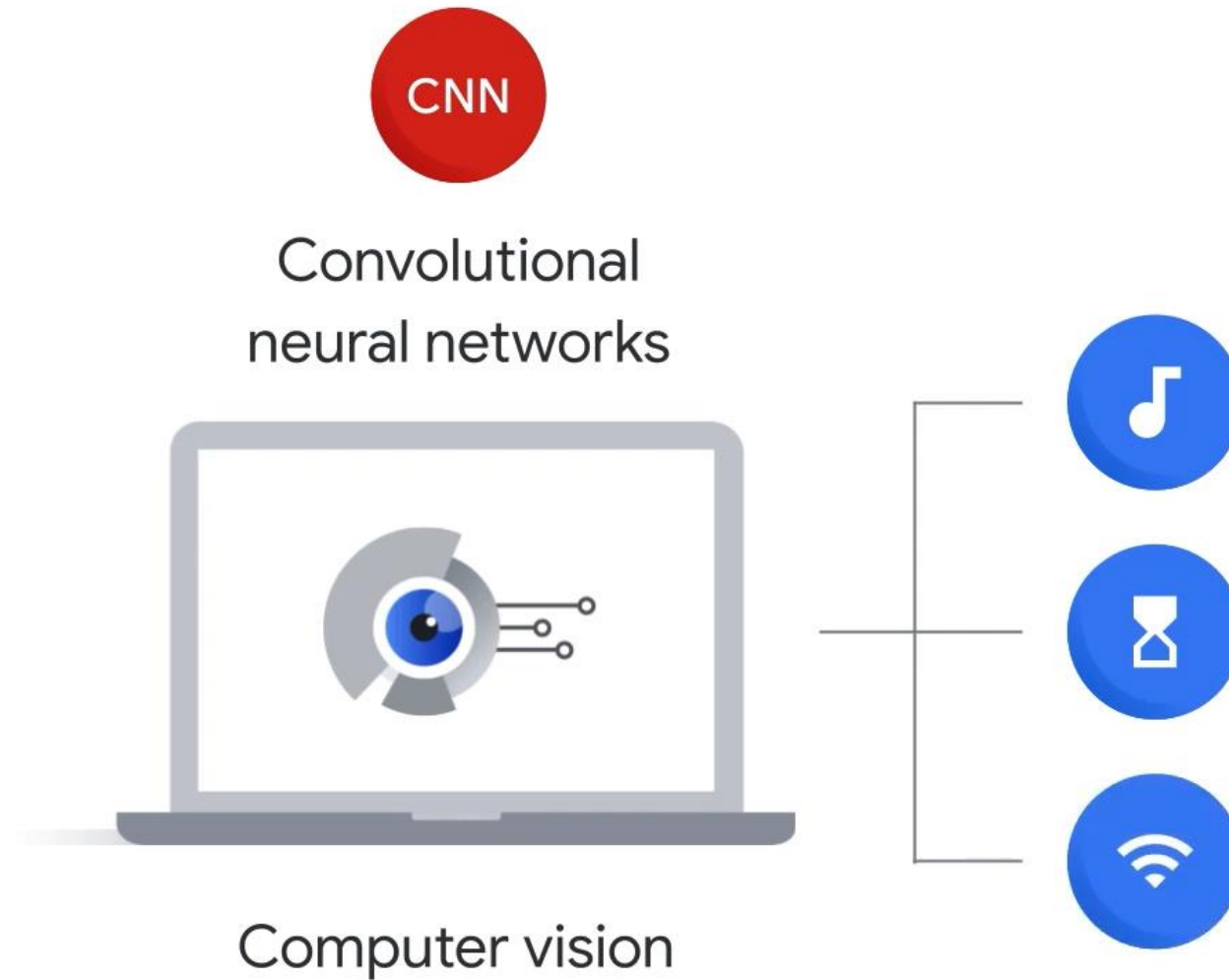
Convolutional Neural Networks (CNN)

Présentation des CNN : principes fondamentaux et applications



Convolutional Neural Networks (CNN)

Présentation des CNN : principes fondamentaux et applications



Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN



12 Mp

Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN



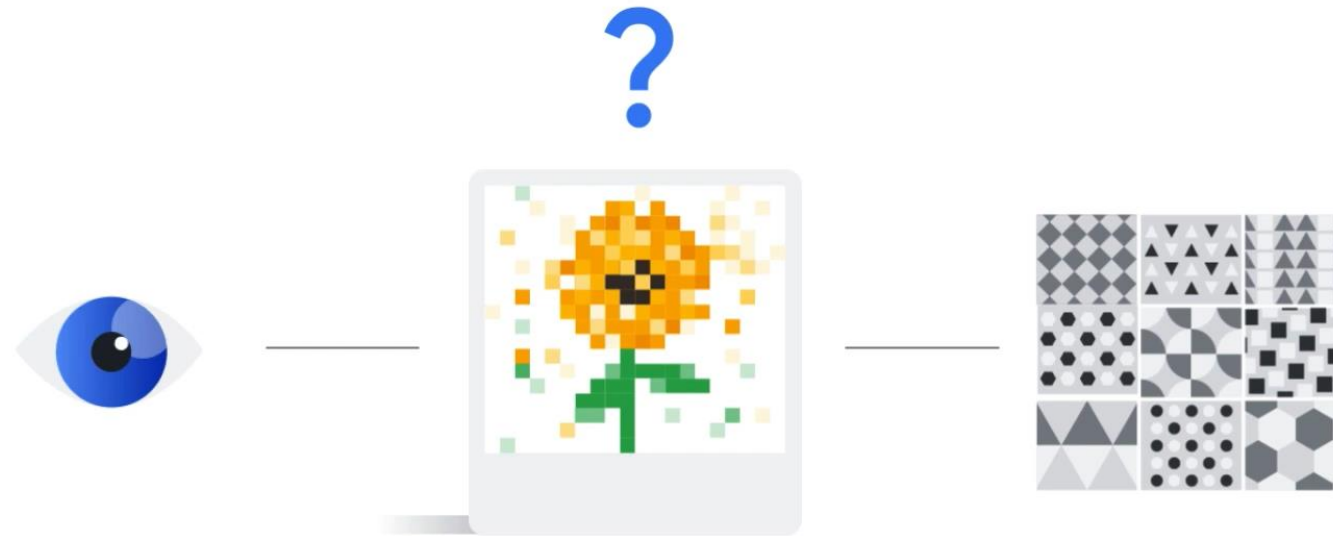
12 Mpx

Imaginons cette entrée connectée à une première couche Dense de 100 neurones.

Combien cela fait-il de paramètres ?

Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN



Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN

Avant 2012



Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN

Avant 2012



Convolutional Neural Networks (CNN)

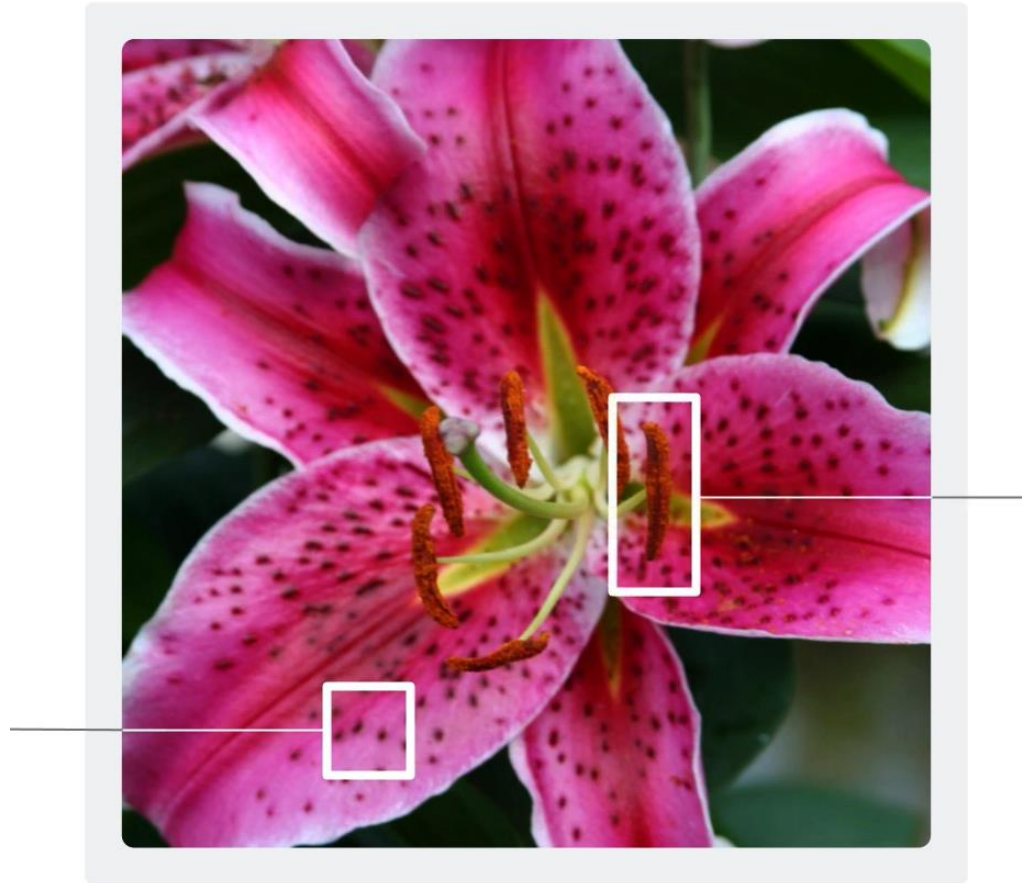
Fonctionnement fondamental d'un CNN

Avant 2012



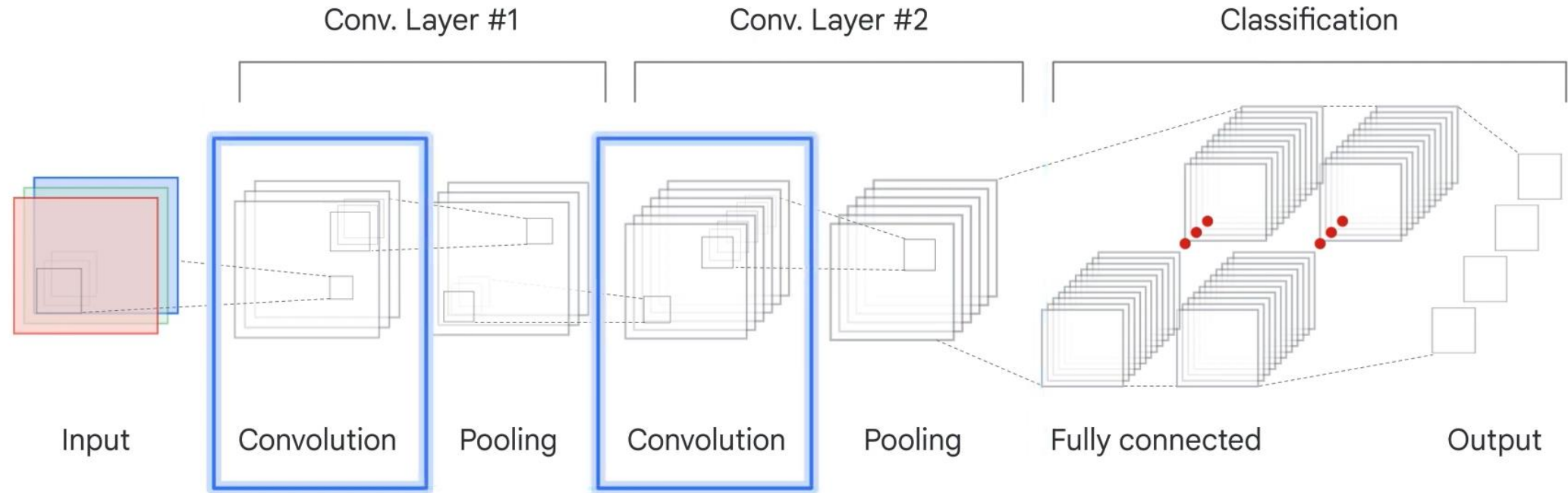
Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN



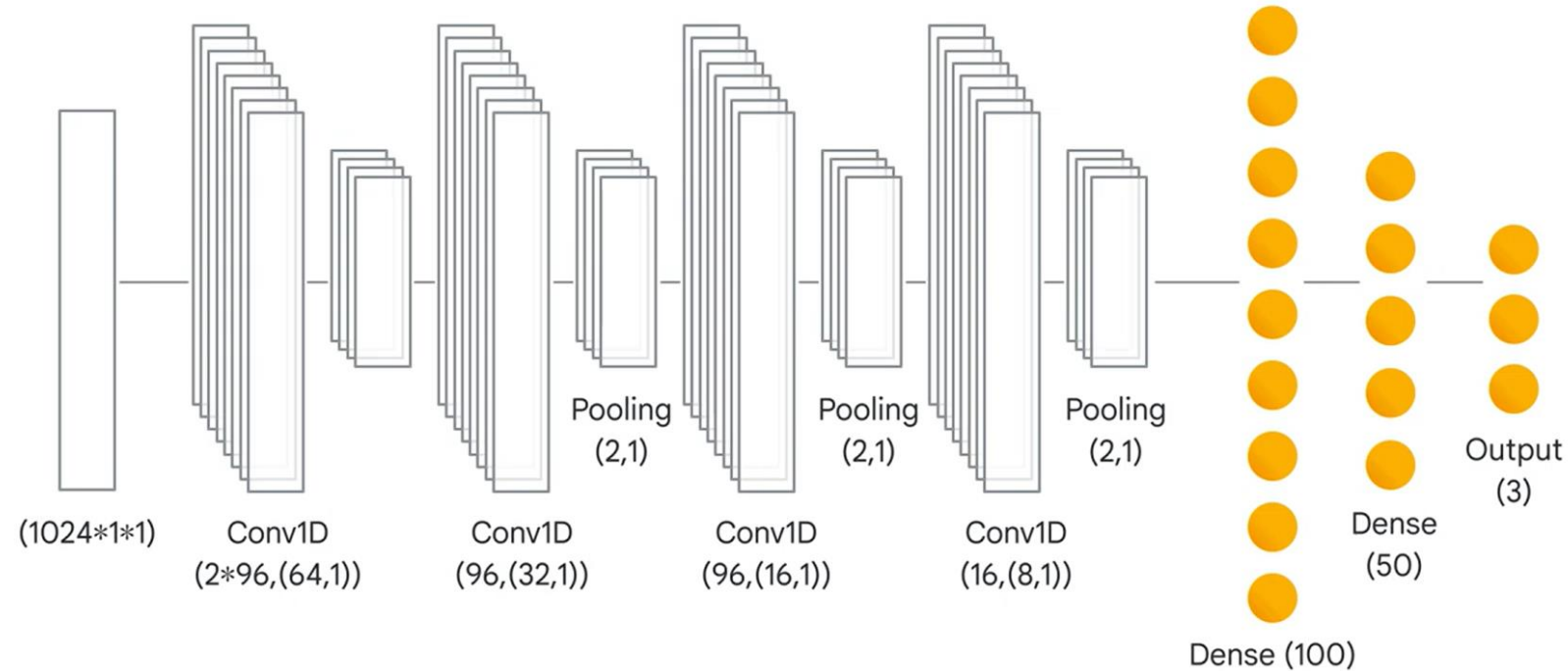
Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN



Convolutional Neural Networks (CNN)

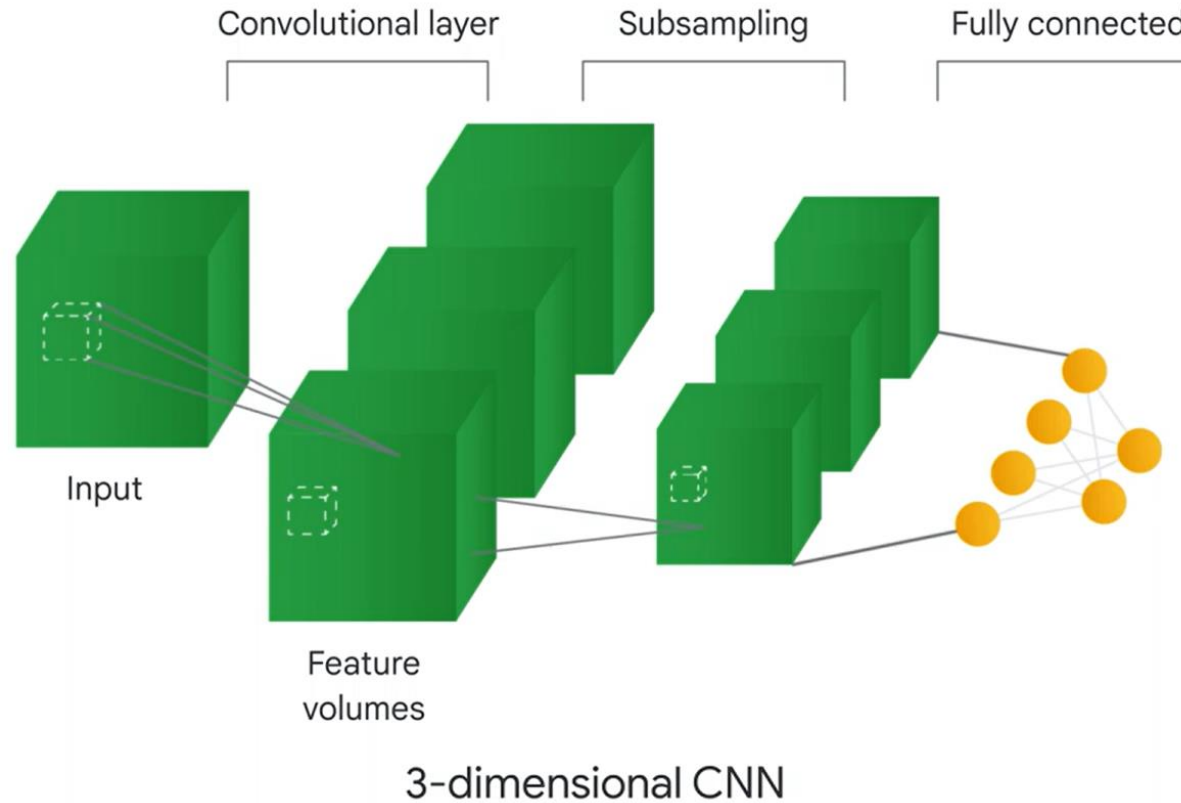
Fonctionnement fondamental d'un CNN



1-dimensional CNN

Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN



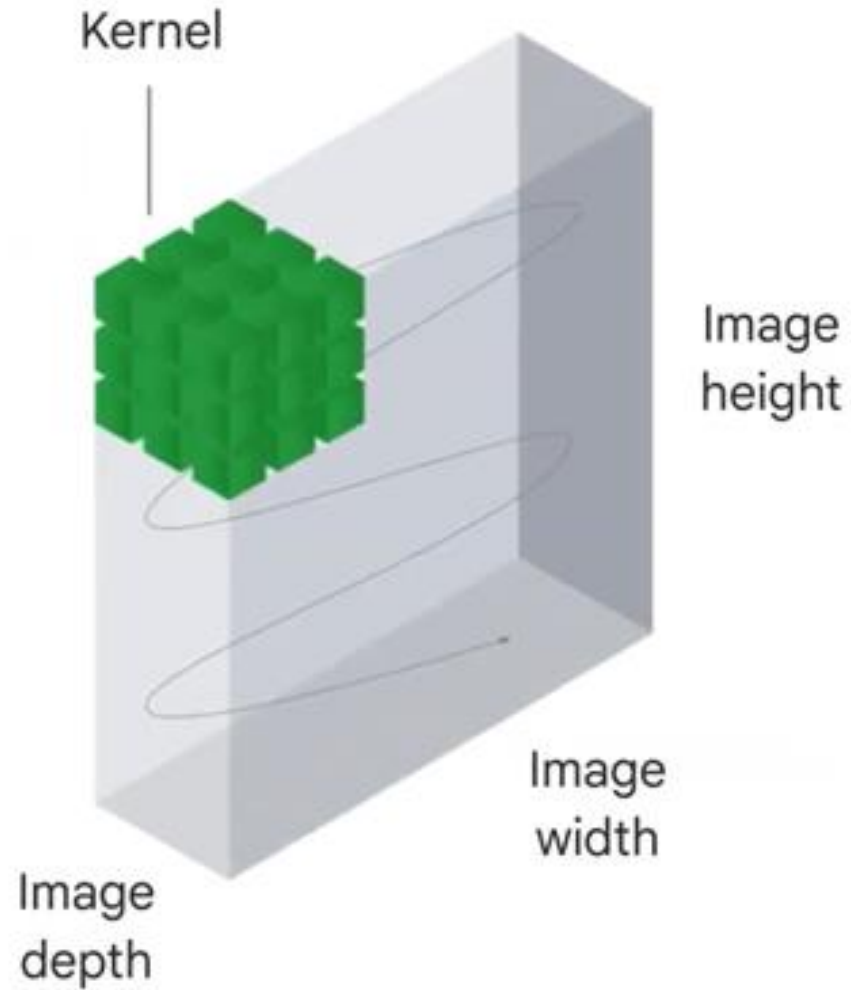
Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN

Type	Données/applications	Popularité
1D	Séquences (séries temporelles, textes)	Moyenne
2D	Images (hauteur x largeur)	Très élevée
3D	Vidéos, volumes (hauteur x largeur x temps/profondeur)	Spécifique

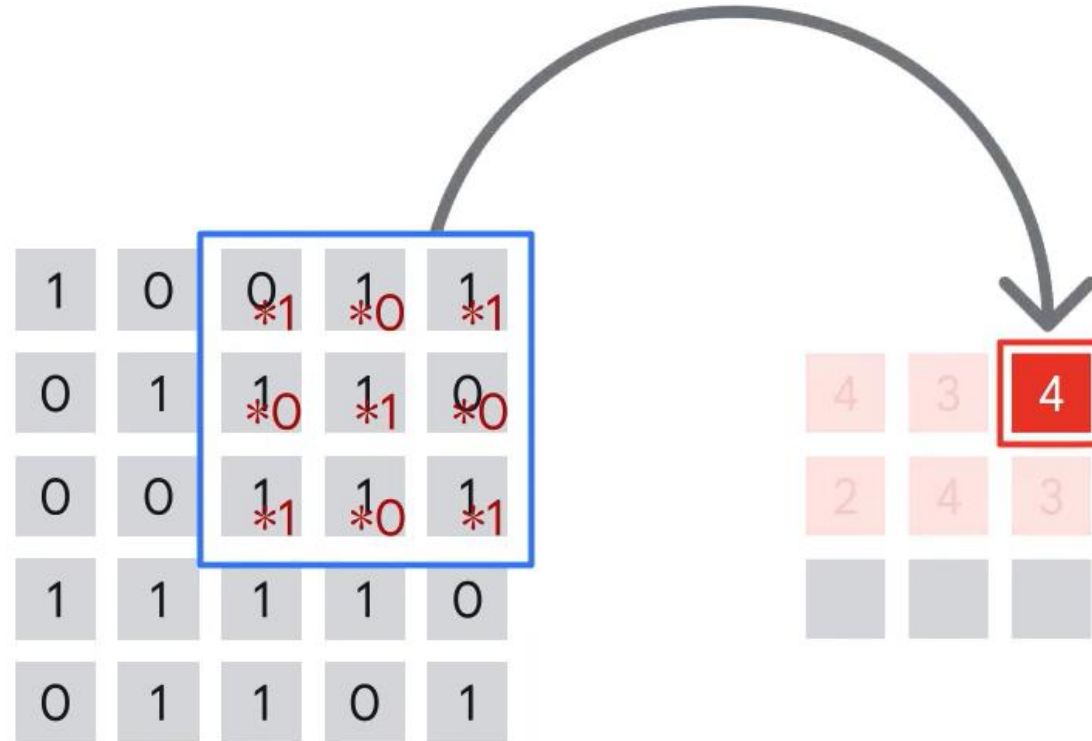
Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN



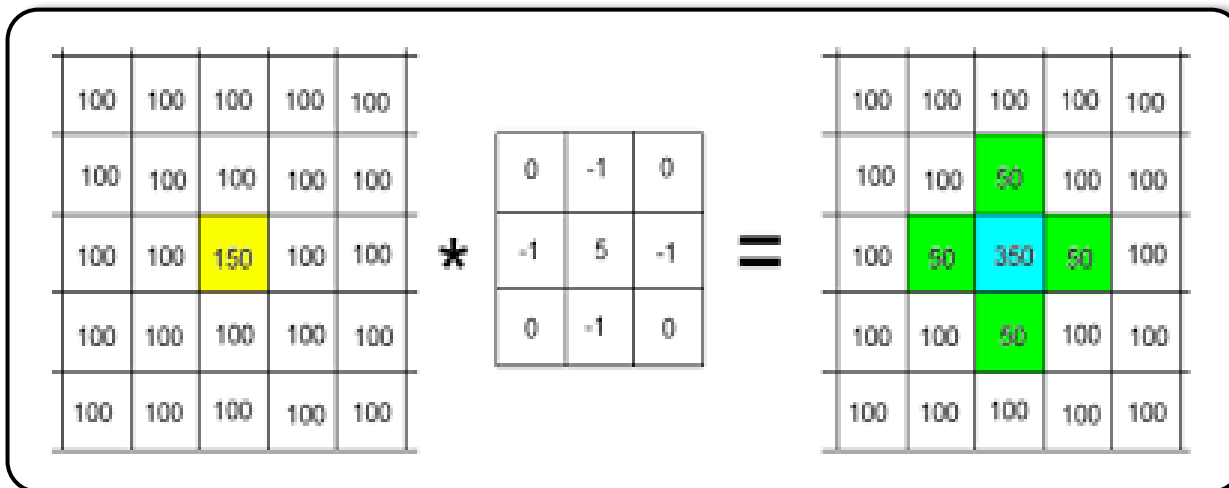
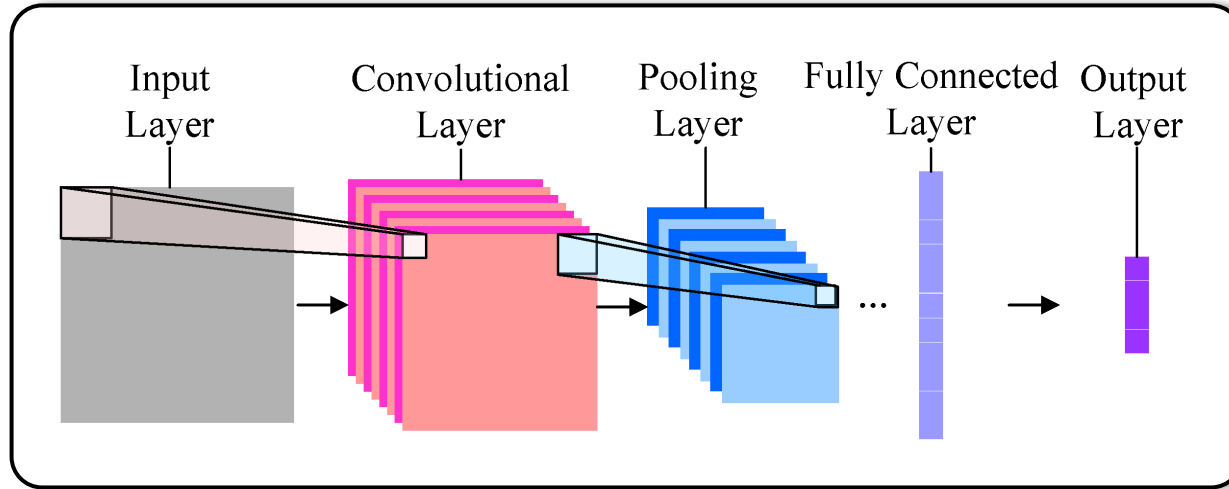
Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN



Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN



Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN



A vous de jouer !

A votre avis, à quoi servent ces kernels ?

	-1	
-1	4	-1
	-1	

Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN



A vous de jouer !

A votre avis, à quoi servent ces kernels ?

1	2	1
-1	-2	-1

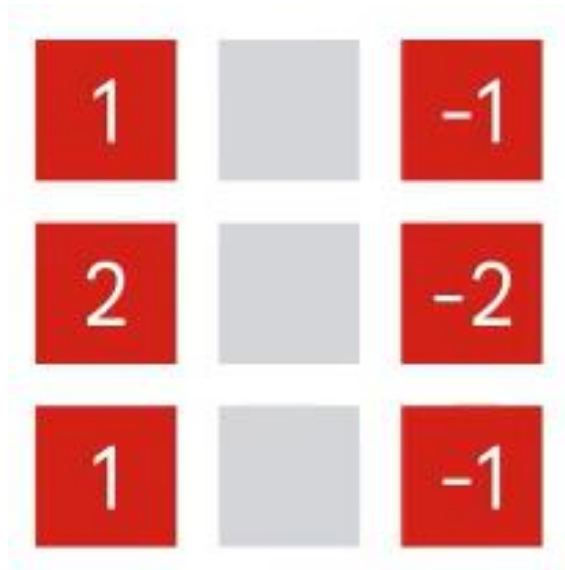
Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN



A vous de jouer !

A votre avis, à quoi servent ces kernels ?



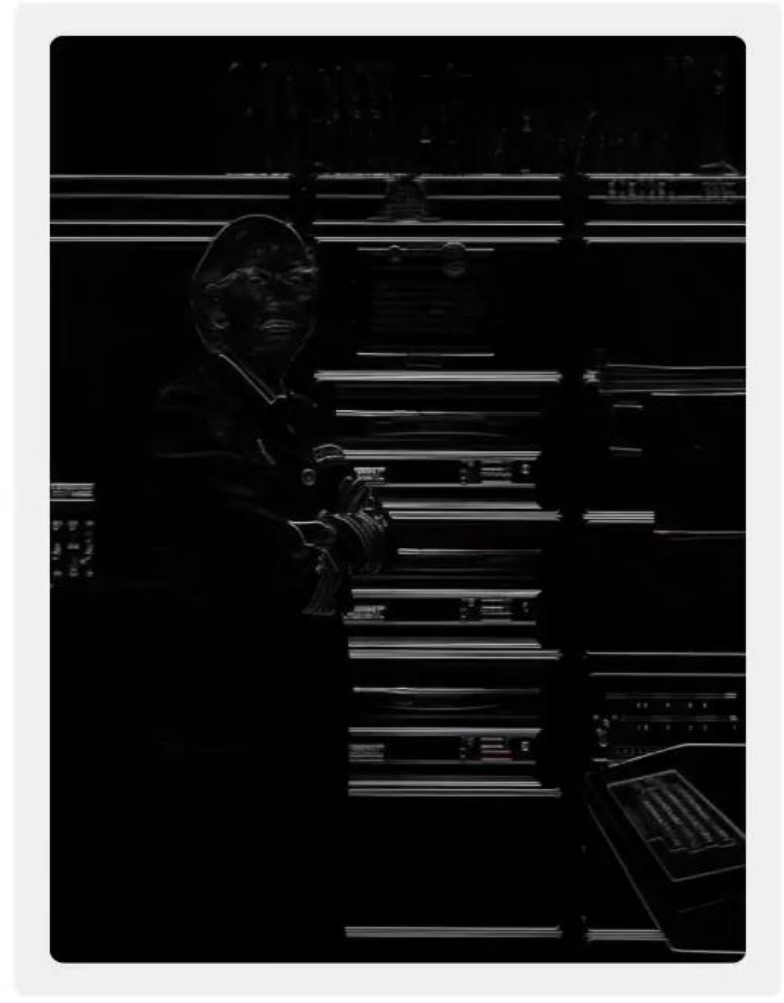
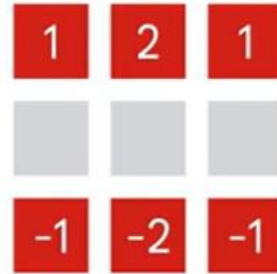
Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN



Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN

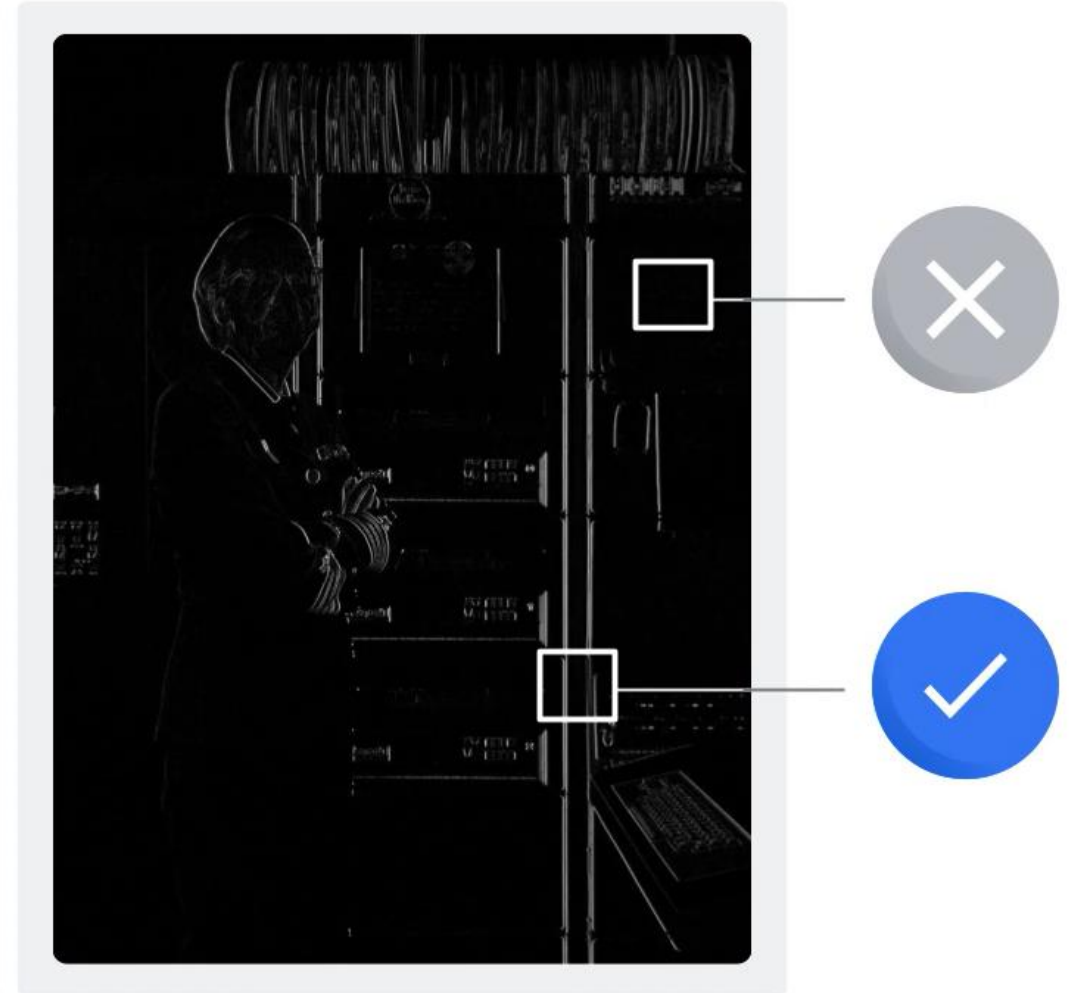


Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN



1		-1
2		-2
1		-1



Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN



Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN

Les couches convolutionnelles correspondent à des filtres.



Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN

Paramètres des couches convolutionnelles

```
tf.keras.layers.Conv2D(filters,  
                        kernel_size,  
                        strides=(1, 1),  
                        padding='valid',  
                        activation='relu')
```

Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN



A vous de jouer !

A votre avis, de combien de dimensions est composée l'entrée d'une couche convolutionnelle (d'un réseau 2D) ?

Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN

[batch, height, width, channels]



Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN

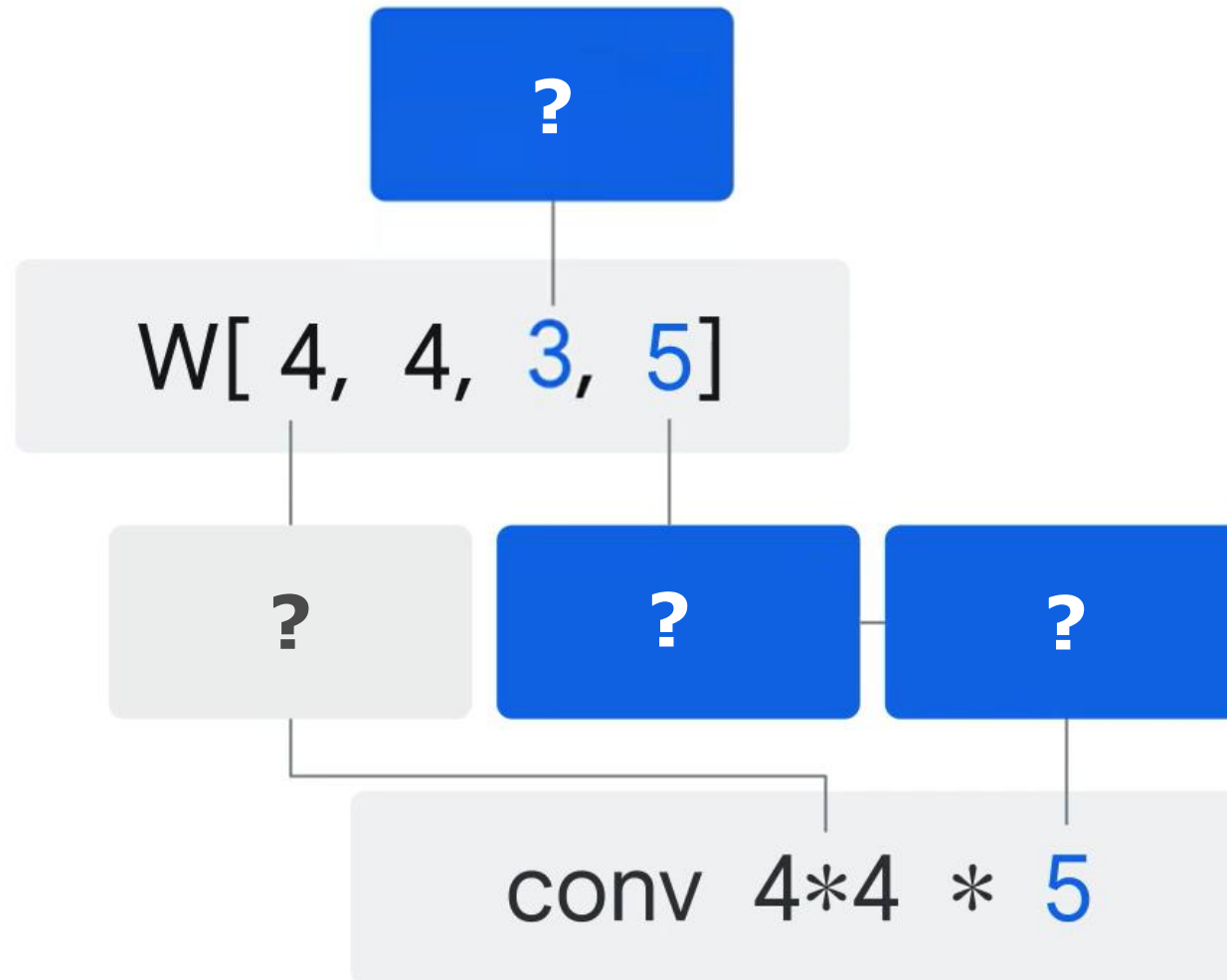


A vous de jouer !

A votre avis, de combien de dimensions est composée la sortie d'une couche convolutionnelle (d'un réseau 2D) ?

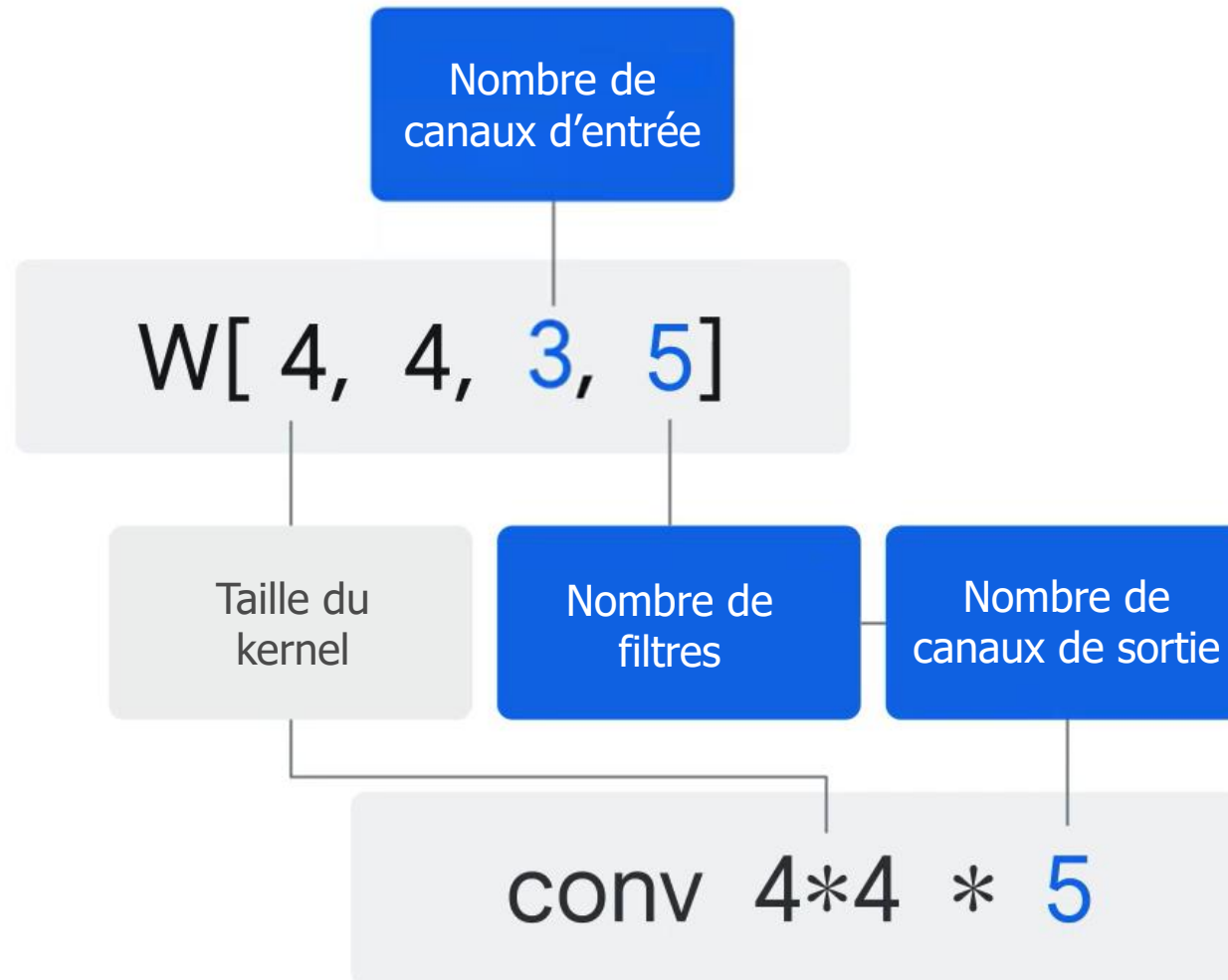
Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN



Convolutional Neural Networks (CNN)

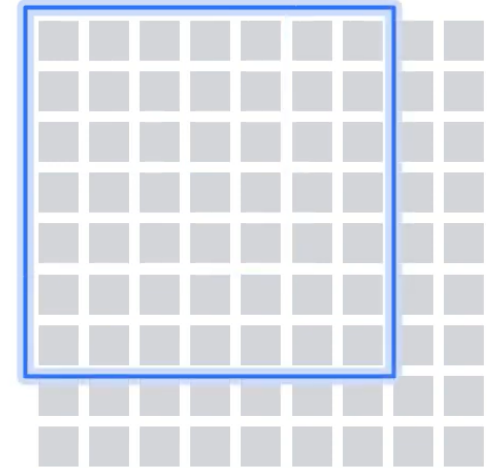
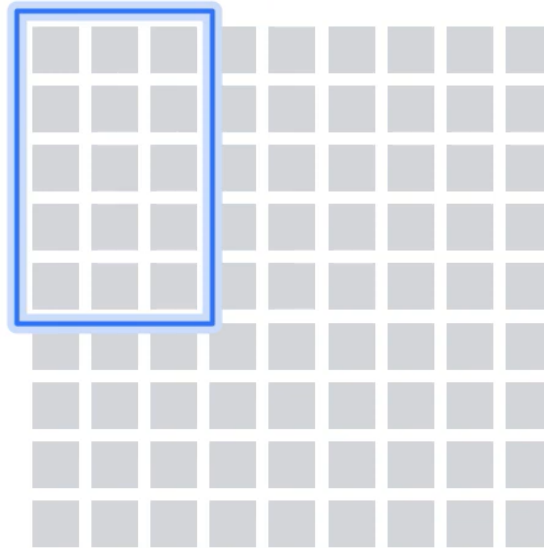
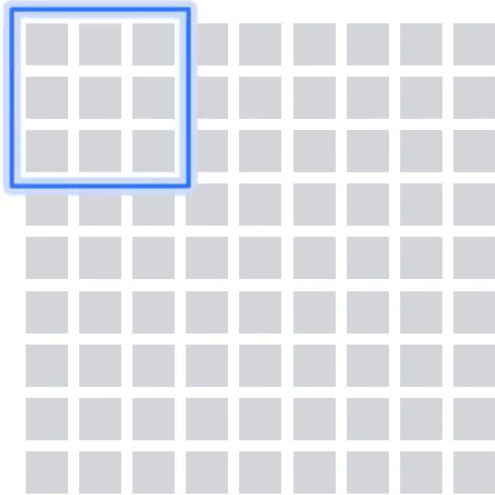
Fonctionnement fondamental d'un CNN



Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN

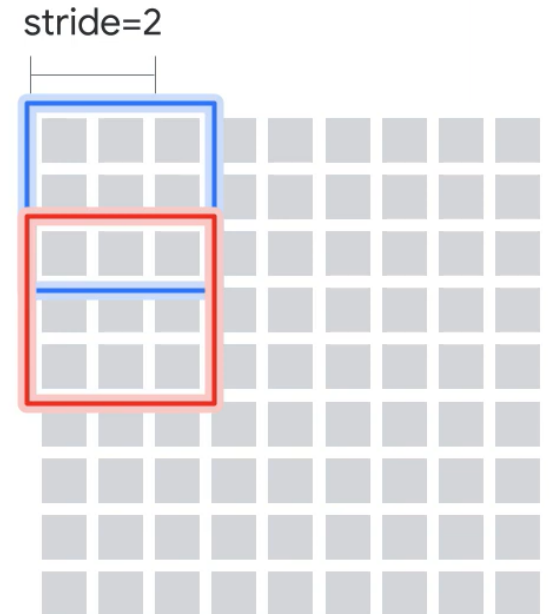
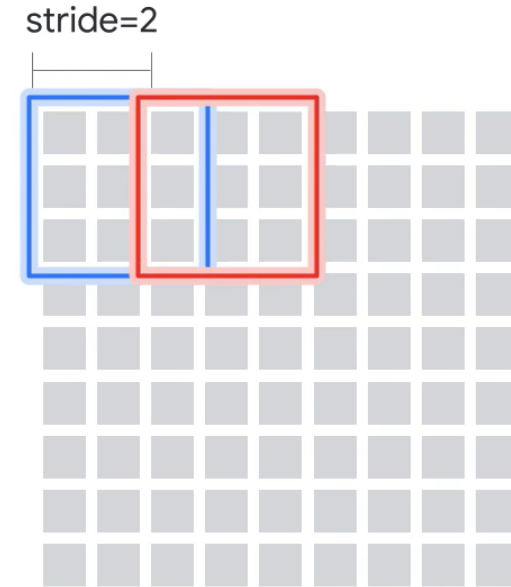
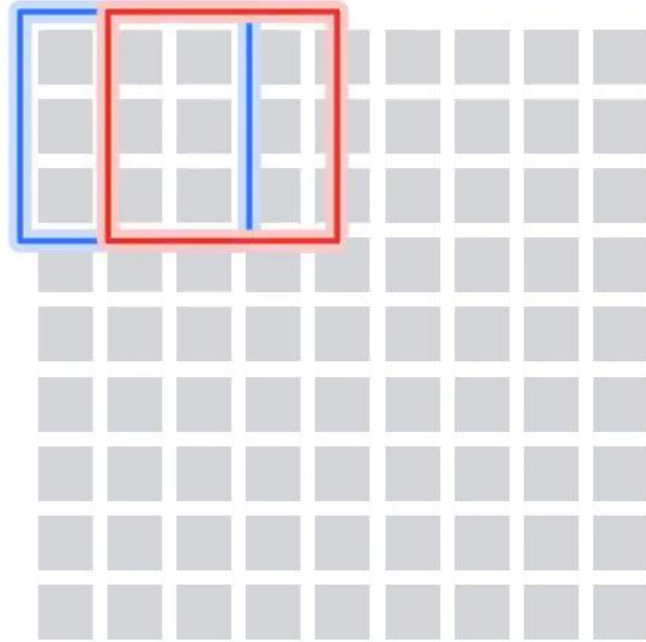
Taille du kernel



Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN

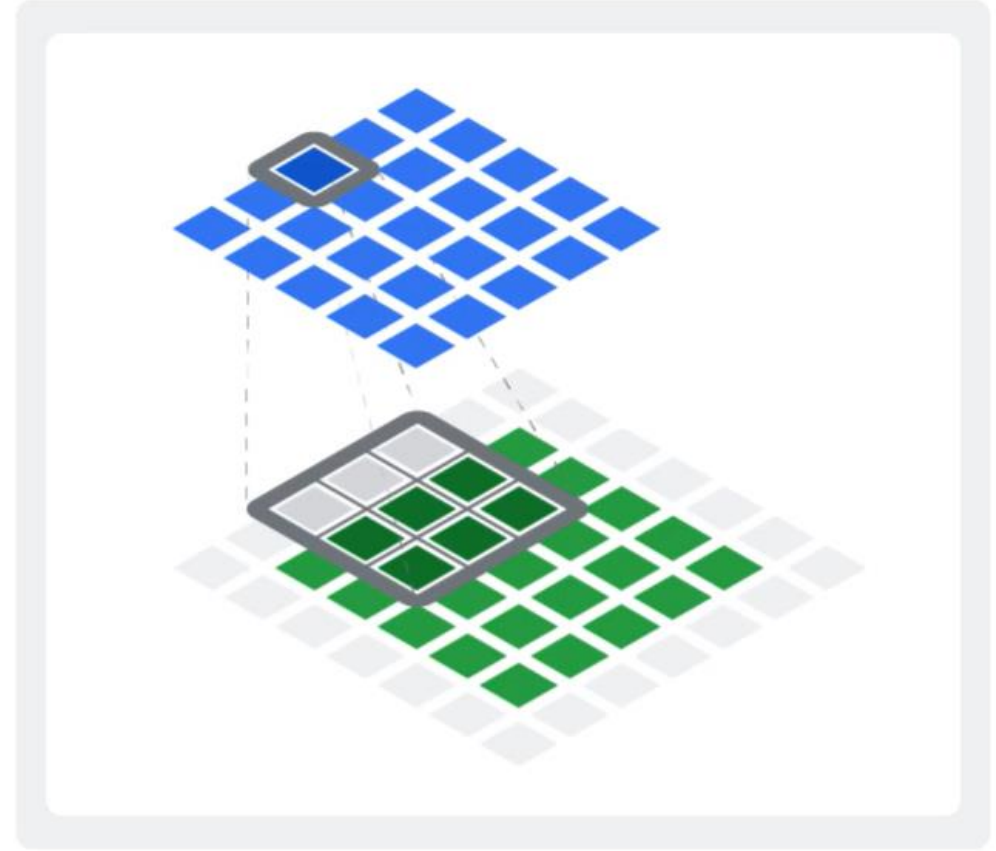
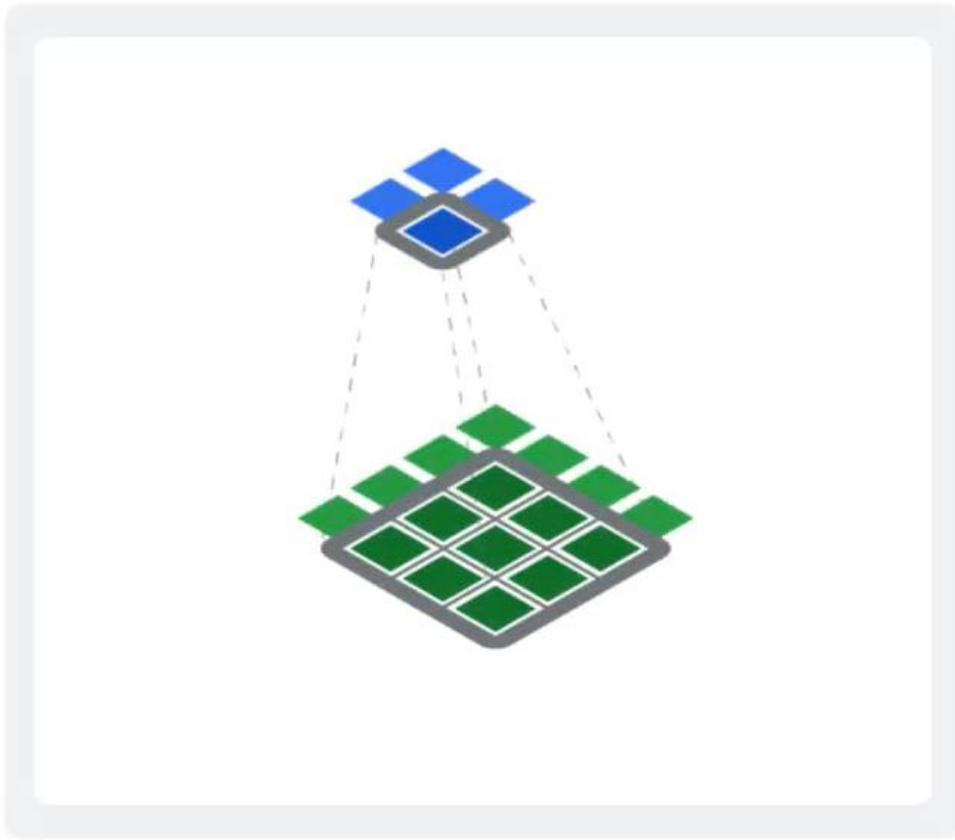
Stride



Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN

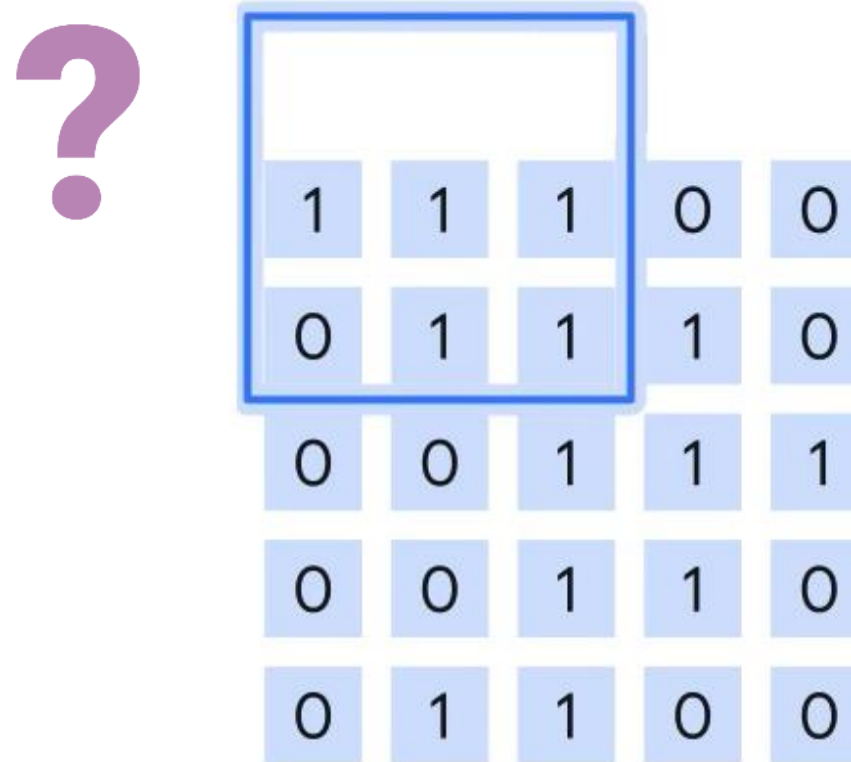
Padding



Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN

Padding



Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN

Padding

0	0	0	0	0	0	0
0	1	1	1	0	0	0
0	0	1	1	1	0	0
0	0	0	1	1	1	0
0	0	0	1	1	0	0
0	0	1	1	0	0	0
0	0	0	0	0	0	0

Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN



Quelle est la taille de la sortie en fonction de celle d'entrée et des paramètres ?

$$O = \frac{P + \frac{K - 1}{S} + 1}{2}$$

O = Size (width) of output image

I = Size (width) of input image

K = Size (width) of kernels used in the conv layer

P = Padding

S = Stride

Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN

Quelle est la taille de la sortie en fonction de celle d'entrée et des paramètres ?

$$O = \left\lfloor \frac{I - K + 2P}{S} \right\rfloor + 1$$

O = Size (width) of output image

I = Size (width) of input image

K = Size (width) of kernels used in the conv layer

P = Padding

S = Stride

Convolutional Neural Networks (CNN)

Démo



Transfer learning d'un modèle de la famille des CNN pour une problématique de classification d'images, avec TensorFlow/Keras

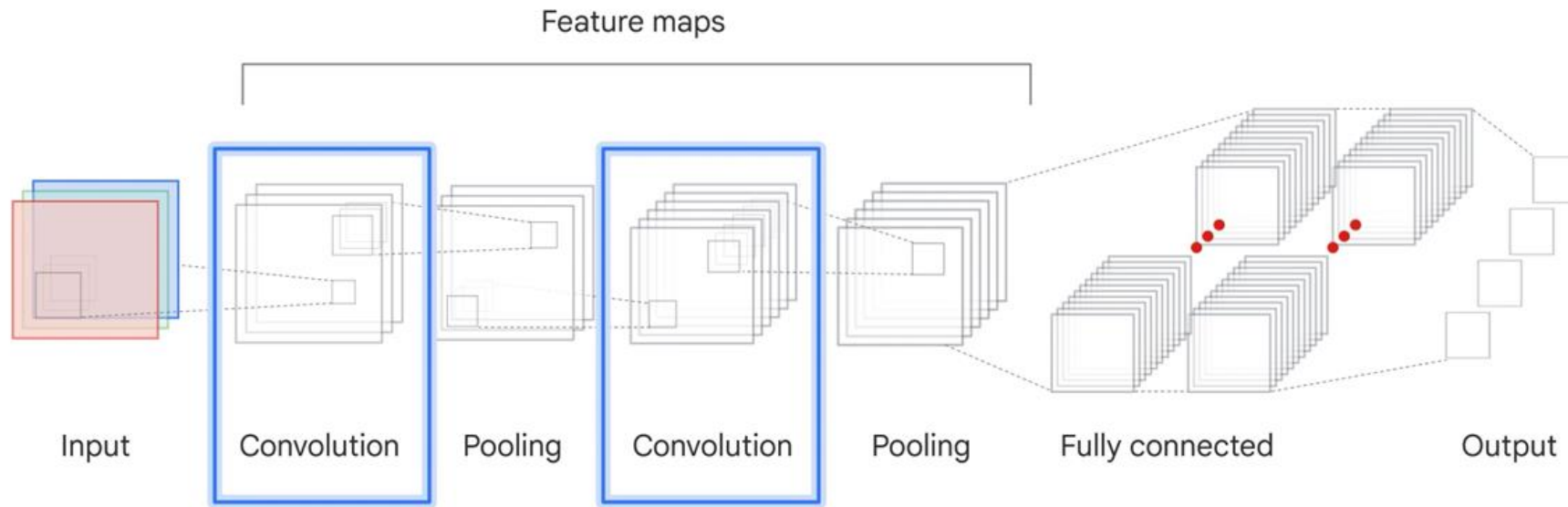
Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN

Couches de convolution



Nombre de neurones ?



?

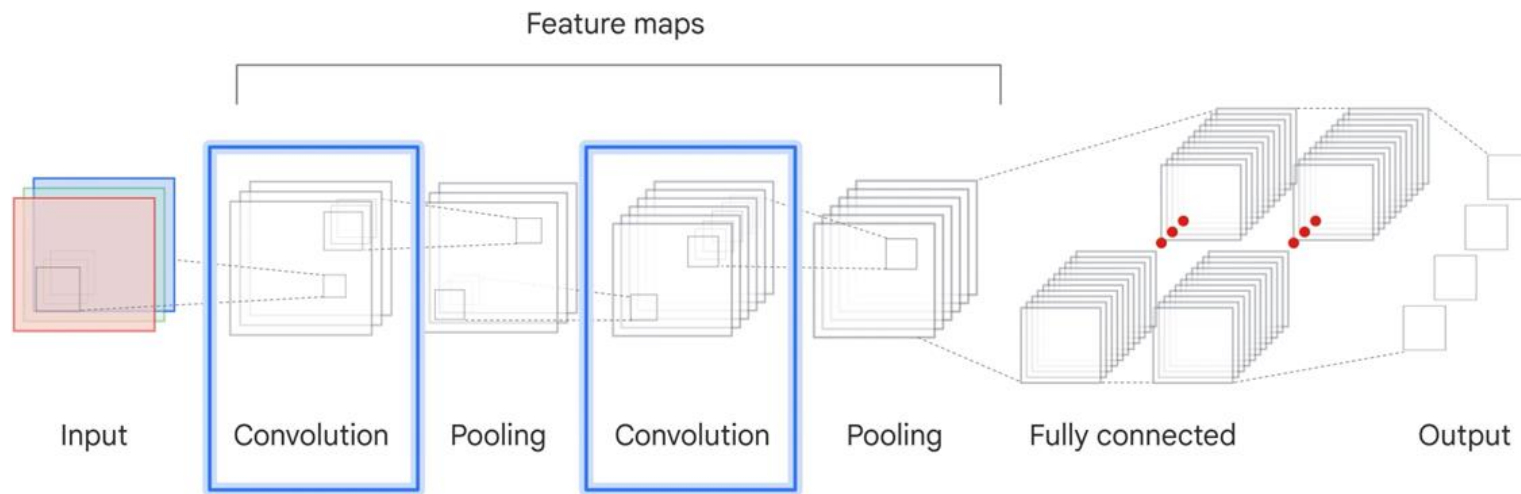
Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN

Couches de convolution



Nombre de paramètres ?



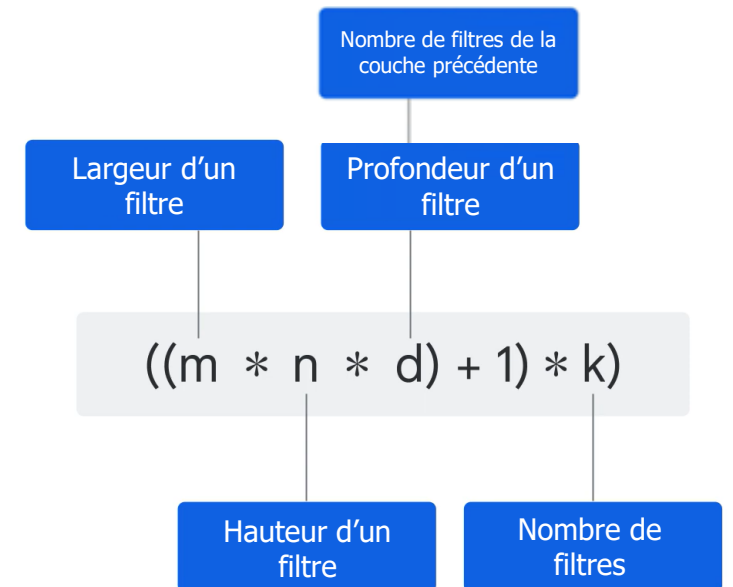
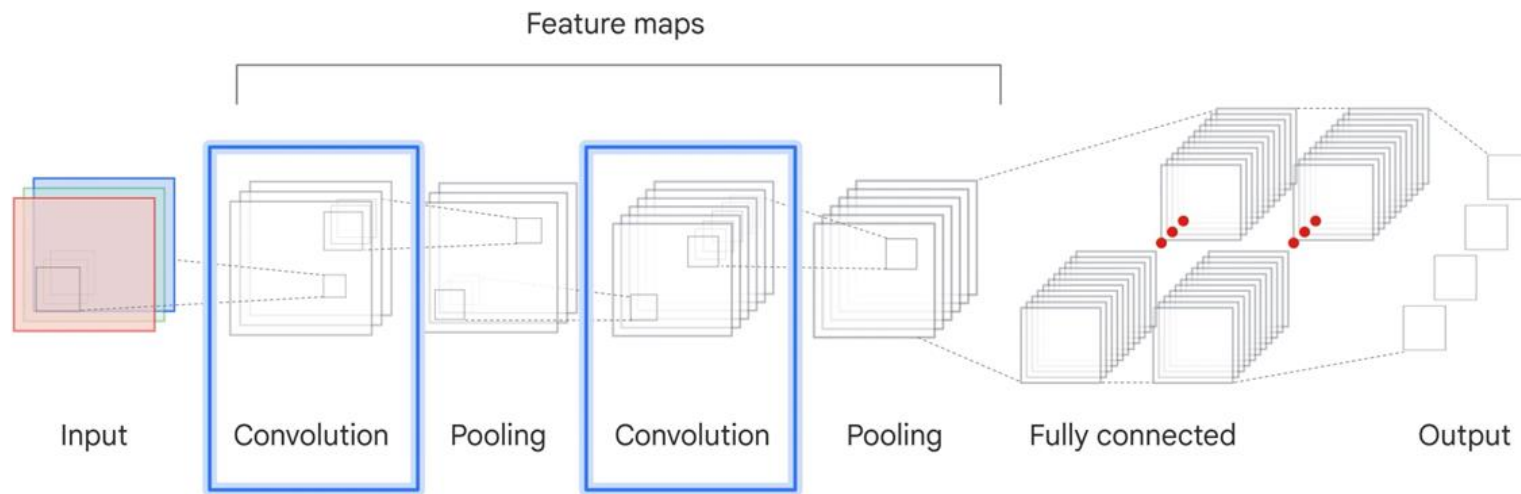
The diagram shows the formula for calculating the number of parameters in a convolutional layer. It consists of four blue boxes with question marks, connected by lines to a central gray box containing the formula $((m * n * d) + 1) * k$. The top-left and top-right boxes are connected to the m and n terms respectively. The bottom-left and bottom-right boxes are connected to the d and k terms respectively.

Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN

Couches de convolution

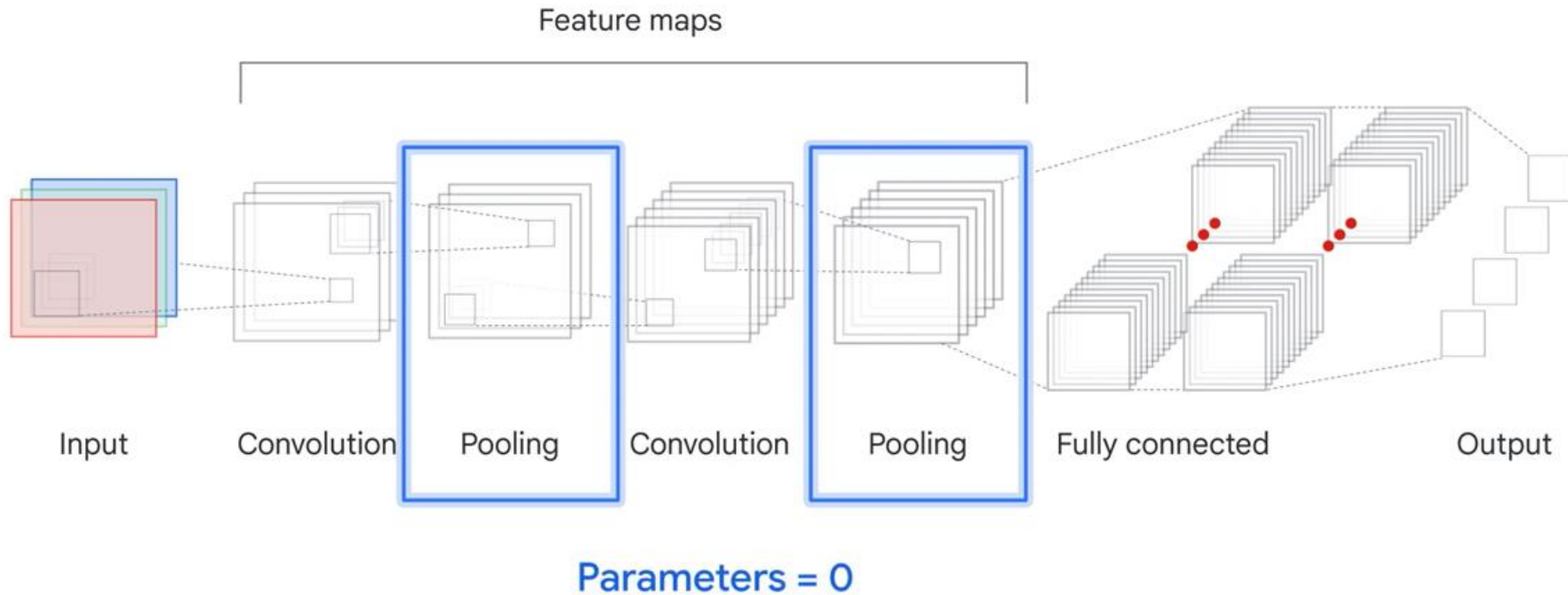
Nombre de paramètres ?



Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN

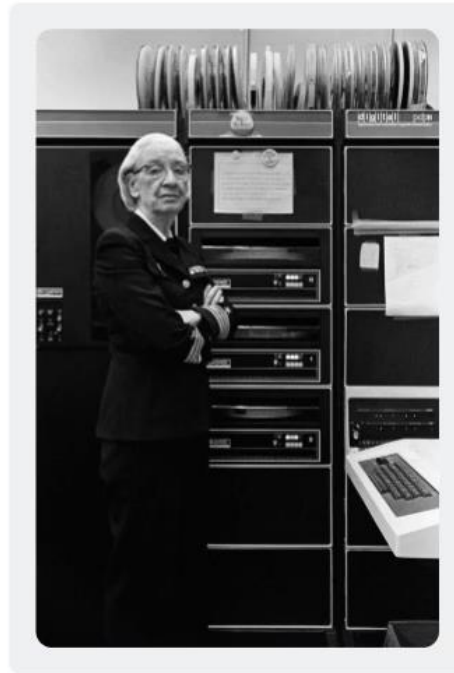
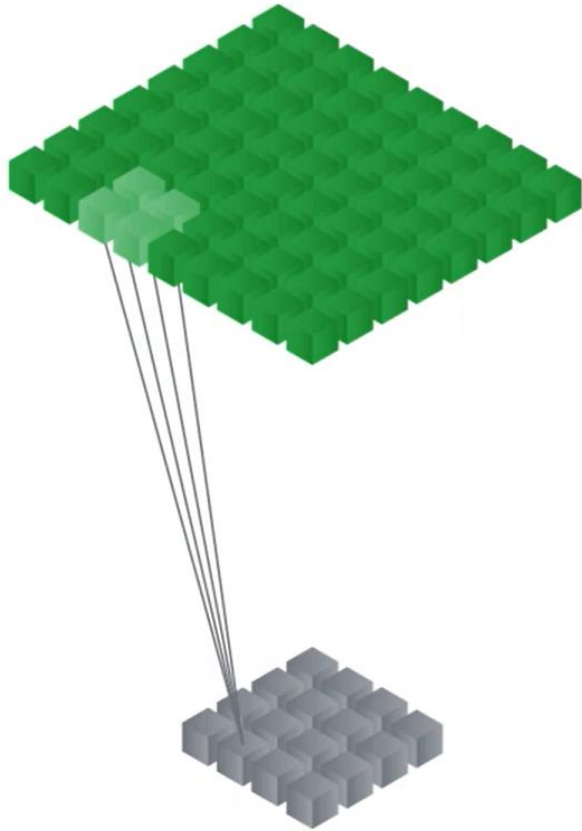
Couches de pooling



Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN

Couches de pooling



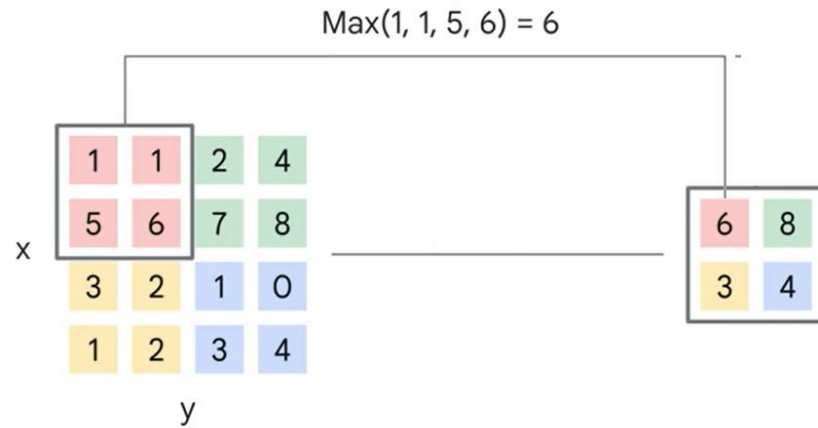
Max pool
Kernel: 2×2
Stride: 2



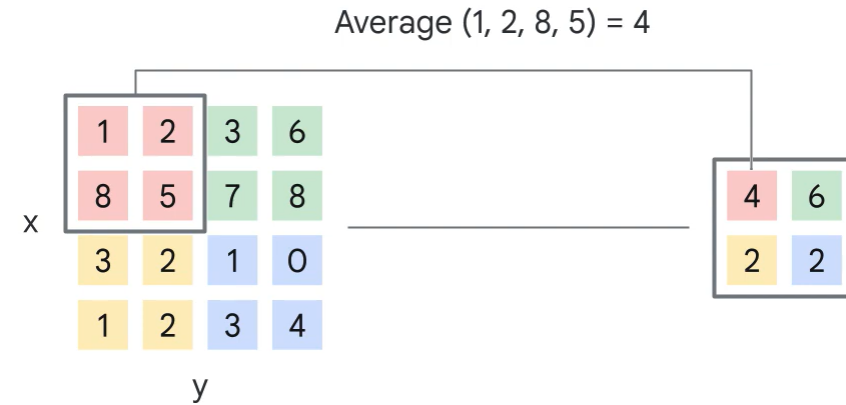
Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN

Couches de pooling



Max pooling



Average pooling

Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN

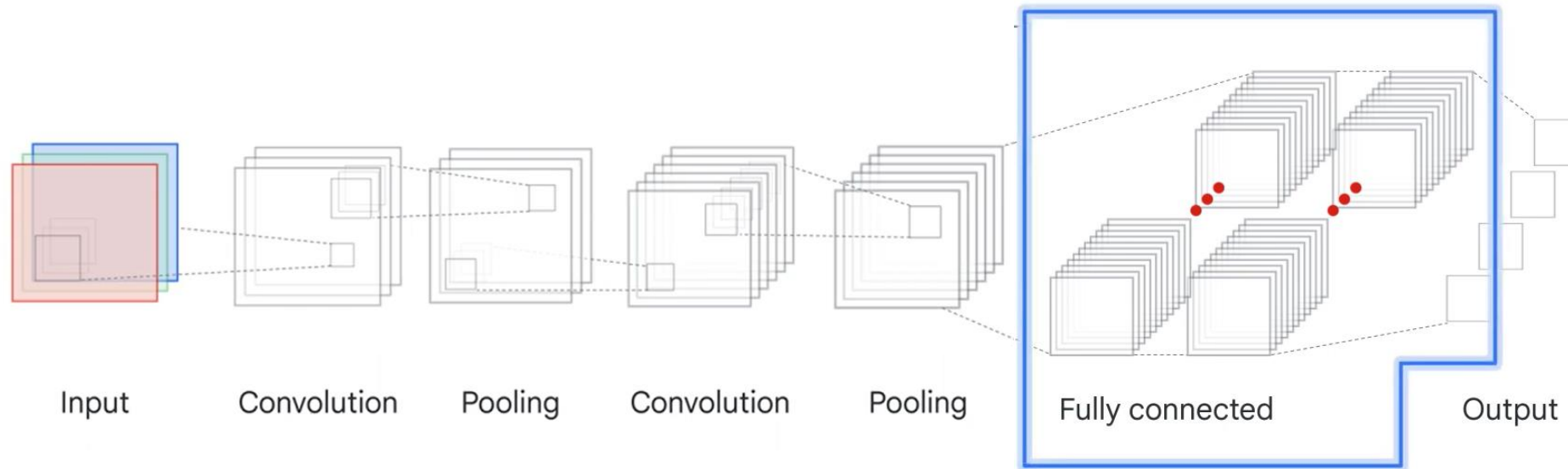
Pourquoi utiliser du pooling plutôt qu'un stride directement plus élevé dans la couche de convolution ?

- ➡ Découple les rôles d'extraction de caractéristiques et de réduction de taille.
- ➡ Améliore la robustesse aux petites variations dans les données.

Convolutional Neural Networks (CNN)

Fonctionnement fondamental d'un CNN

Couches denses (fully-connected)



Nombre de paramètres :

nombre de neurones de la couche précédente

x

nombre de neurones

Convolutional Neural Networks (CNN)

Quizz



Laquelle de ces affirmations relatives au nombre de paramètres dans les CNN est vraie ?

- ☐ La majorité des paramètres sont situés dans les couches de pooling.
- ☐ Un grand nombre de paramètres provient des couches denses à la fin, les couches convolutionnelles en contiennent moins.
- ☐ Entre 75% et 80% des paramètres environ proviennent des couches convolutionnelles.
- ☐ Les couches convolutionnelles, de pooling et denses contiennent environ autant de paramètres les unes que les autres.

Diffusion

Génération d'images à partir d'un texte

Algorithmes de diffusion



Image 1



Image 2

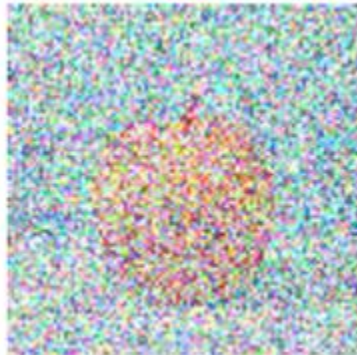


Image 3



Image 4

Entrée

Image bruitée



Sortie

**Image légèrement
moins bruitée**

Entrée	Sortie
Image 2	Image 1
Image 3	Image 2
Image 4	Image 3

Diffusion

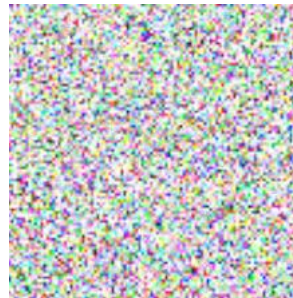
Génération d'images à partir d'un texte

Algorithmes de diffusion

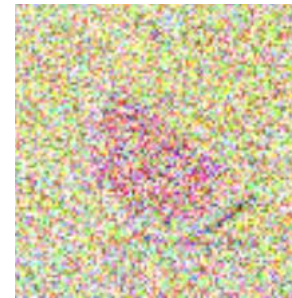


Image 1

Entrée



Sortie



Diffusion

Génération d'images à partir d'un texte

Algorithmes de diffusion



Image 1



Image 2

Entrée



Sortie



Diffusion

Génération d'images à partir d'un texte

Algorithmes de diffusion



Image 1



Image 2

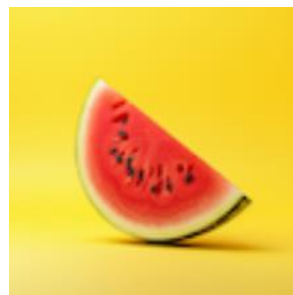


Image 3

Entrée



Sortie



Diffusion

Génération d'images à partir d'un texte

Algorithmes de diffusion



Image 1



Image 2



Image 3

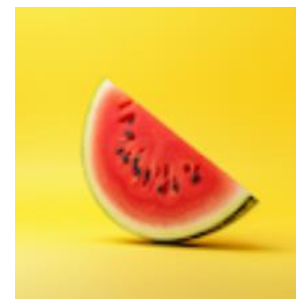


Image 4

~100 étapes

Diffusion

Génération d'images à partir d'un texte

Algorithmes de diffusion : ajout d'un texte



Image 1,
« pomme rouge »



Image 2

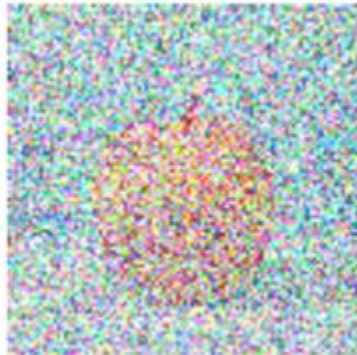


Image 3



Image 4

Entrée

Image bruitée



Sortie

**Image légèrement
moins bruitée**

Entrée	Sortie
Image 2, « pomme rouge »	Image 1
Image 3, « pomme rouge »	Image 2
Image 4, « pomme rouge »	Image 3

Diffusion

Génération d'images à partir d'un texte

Algorithmes de diffusion : ajout d'un texte



Image 1

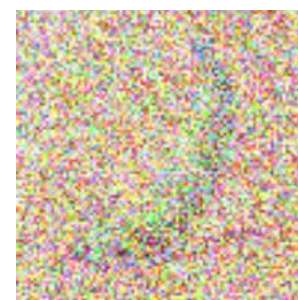
Entrée



+ « **banane verte** »



Sortie



Diffusion

Génération d'images à partir d'un texte

Algorithmes de diffusion : ajout d'un texte



Image 1



Image 2

Entrée



+ « **banane verte** »



Sortie



Diffusion

Génération d'images à partir d'un texte

Algorithmes de diffusion : ajout d'un texte



Image 1



Image 2



Image 3

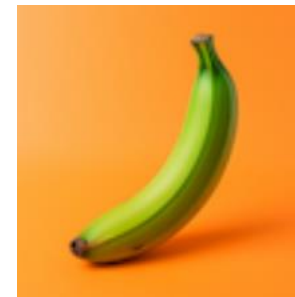
Entrée



+ « **banane verte** »



Sortie



Diffusion

Génération d'images à partir d'un texte

Algorithmes de diffusion : ajout d'un texte



Image 1



Image 2



Image 3

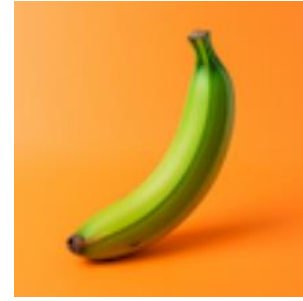


Image 4