

# Réseaux de convolution

- ✓ Limitations du perceptron multi-couches
- ✓ Quelques notions de traitement d'images
- ✓ Couche de convolution
- ✓ Réseau de convolution
- ✓ Exemples de réseaux de convolution
- ✓ Quand on manque de données ...

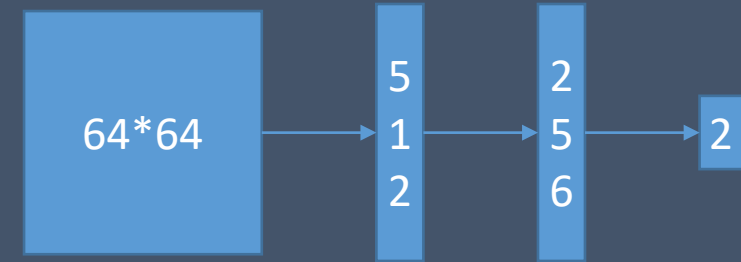
Françoise Bouvet  
francoise.bouvet@ijclab.in2p3.fr

# Limitations des MLP

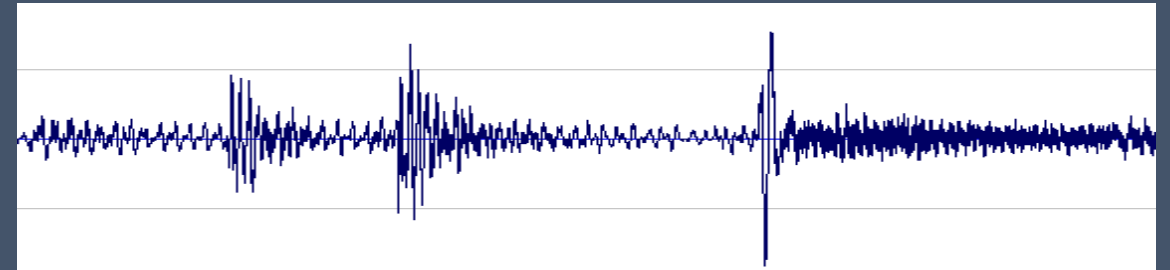
- ✓ Explosion du nombre de paramètres

- Ex :  $64*64 - 512 - 256 - 2$

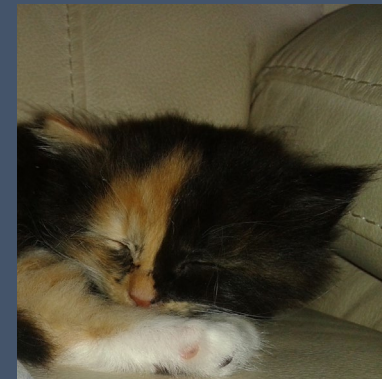
- >  $4096 * 512 + 512 * 256 + 256 * 2 = 2\,228\,736$  poids



- ✓ Image, signal : structure locale non prise en compte



- ✓ Invariance à la translation / rotation difficile à prendre en compte



# Quelques notions de traitement d'images

# Filtre de convolution

	9	5	3	
	11	7	5	
	12	4	4	

1	0	-1
1	0	-1
1	0	-1

$$\begin{aligned}
 &1*9 + 0*5 - 1*3 \\
 &+ 1*11 + 0*7 - 1*5 \\
 &+ 1*12 + 0*4 - 1*3 \\
 &= 20
 \end{aligned}$$

		20		

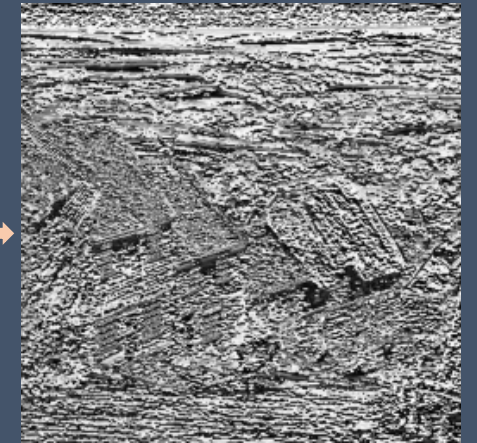


Masque 5\*5 uniforme (1/25)



1	1	1
0	0	0
-1	-1	-1

1	0	-1
1	0	-1
1	0	-1



Remplissage des bords : padding

- ✓ Pixels à 0
- ✓ Dupliquer les lignes/colonnes
- ✓ ...

# Pooling / stride

**Pooling** : regroupement de pixels voisins

6	9	5	3
6	11	7	9
5	12	4	4
8	3	6	6

Pooling 2\*2  
Stride 2

Max pooling

11	9
12	6

Average pooling

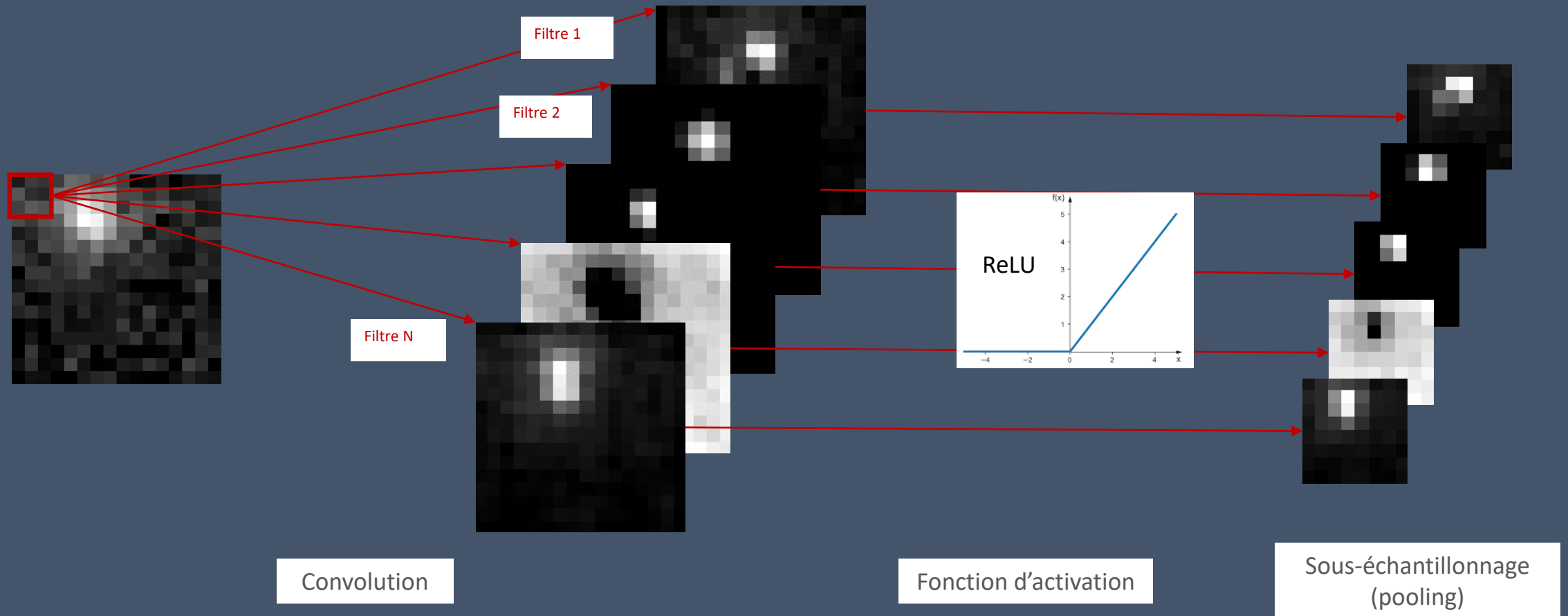
8	6
7	5

**Stride** : pas de déplacement (step)

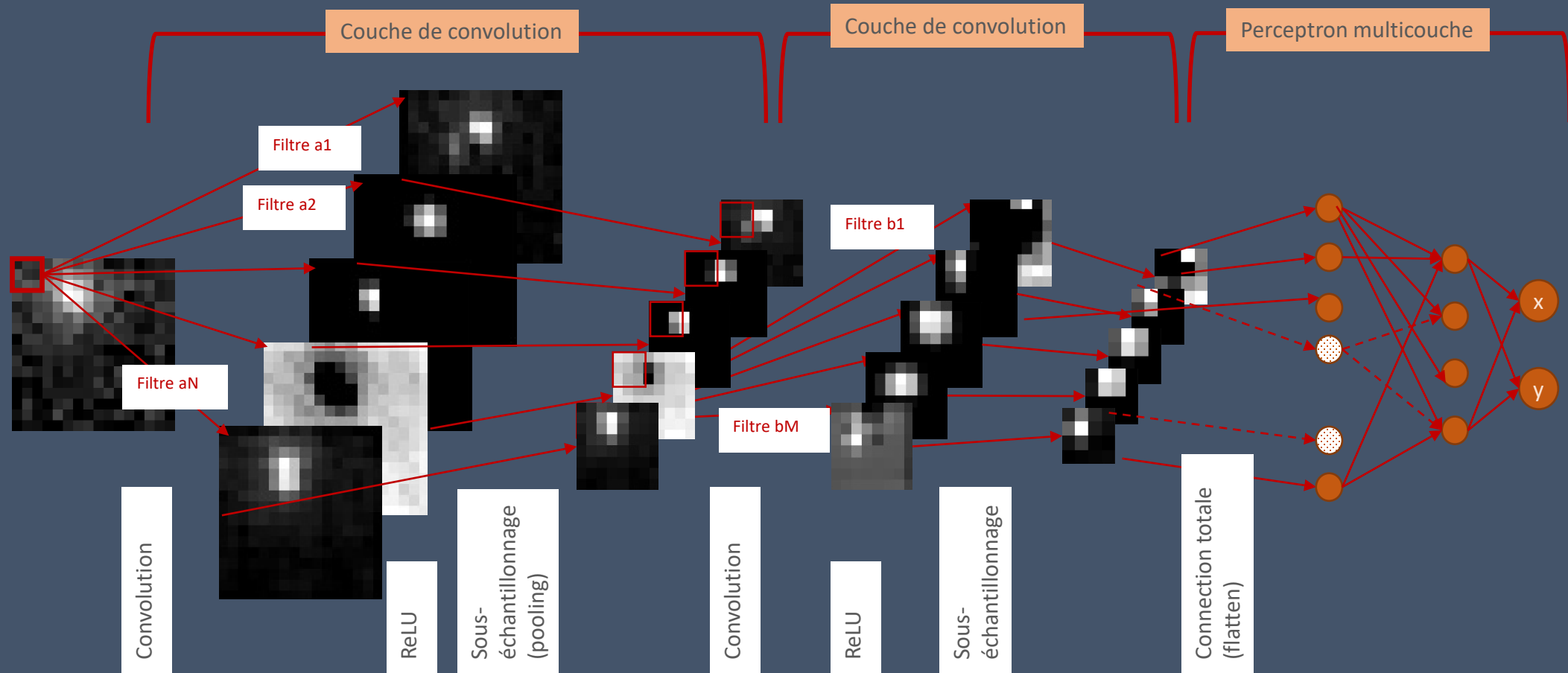

Pooling 3\*3  
Stride 2


# Structure du réseau de convolution

# Couche de convolution



# Réseau de convolution





# Réseau de convolution : paramètres

## Données du problème

- ✓ Nombre de neurones dans la couche d'entrée
- ✓ Nombre de neurones dans la couche de sortie

## Paramètres calculés

- ✓ Poids
- ✓ Masques des filtres de convolution

## Hyper-paramètres

### Structure du réseau

- ✓ Nombre de couches de convolution
- ✓ Nombre de filtres /couche
- ✓ Fonction d'activation
- ✓ Pooling (max/moy)
- ✓ Hyper-paramètres du MLP

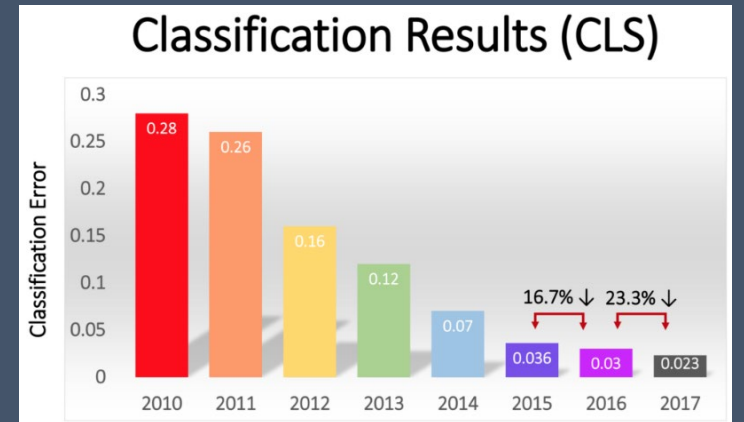
### Apprentissage

- ✓ Fonction de coût
- ✓ Nombre d'époques
- ✓ Paramètres de régularisation ( $\alpha$ ,  $\lambda$  ( $\lambda \|w\|^2$ , batch, dropout)
- ✓ Padding

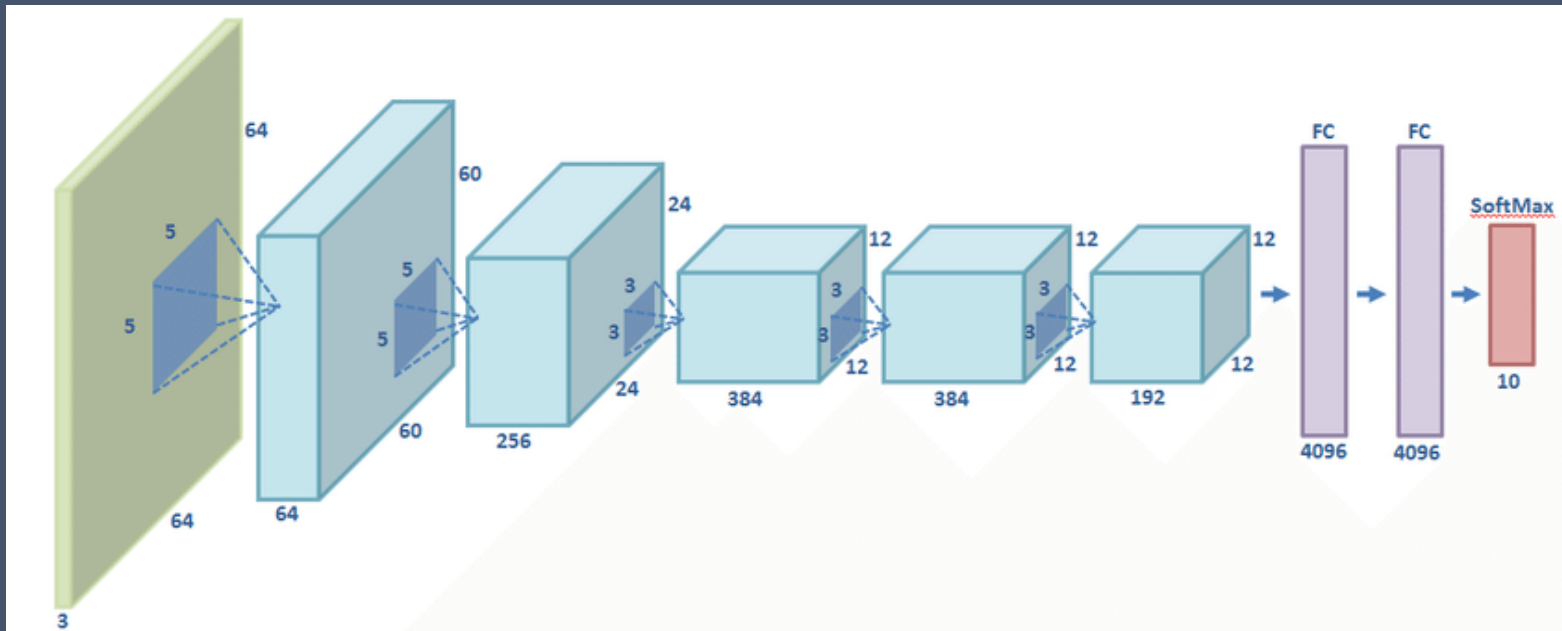
## Quelques exemples

# AlexNet, 2012

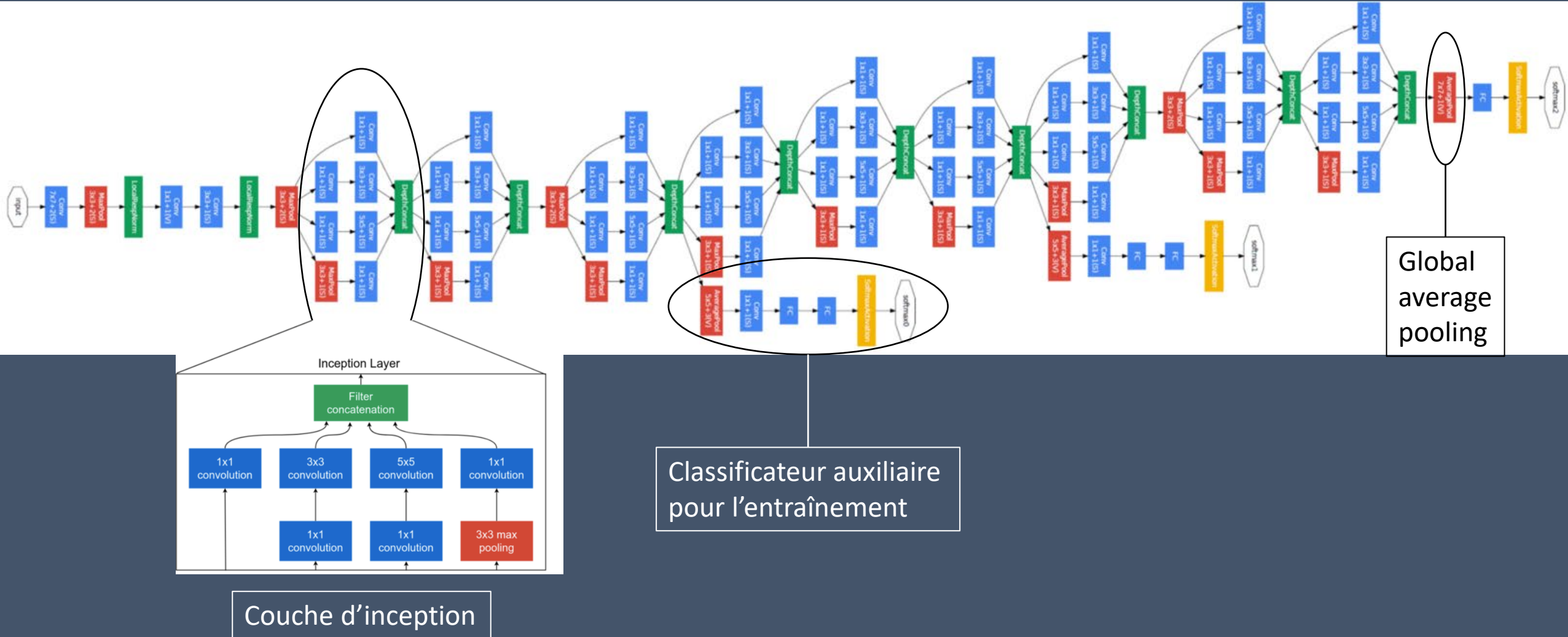
- A partir de 2006, constitution d'une grande base d'images labellisées
- ImageNet Large Scale Visual Recognition Challenge (ILSVRC)
  - $\sim 1.5 \cdot 10^6$  images (90% entraînement - 3% validation - 7% de test)
  - $\sim 1000$  catégories



*ILSVRC : évolution du taux d'erreur*



# GoogLeNet 2014



Quand on manque de données  
ou de temps de calcul

# Augmentation des données

Idée : **accroître artificiellement** la base de données d'apprentissage

Principe : générer des versions modifiées et réalistes des images

- ✓ Décalage horizontal/vertical
- ✓ Miroir
- ✓ Rotation
- ✓ Zoom / recadrage
- ✓ Changement de contraste / luminosité / couleurs



N'est appliqué qu'aux données d'apprentissage, pas à celles de validation et de test

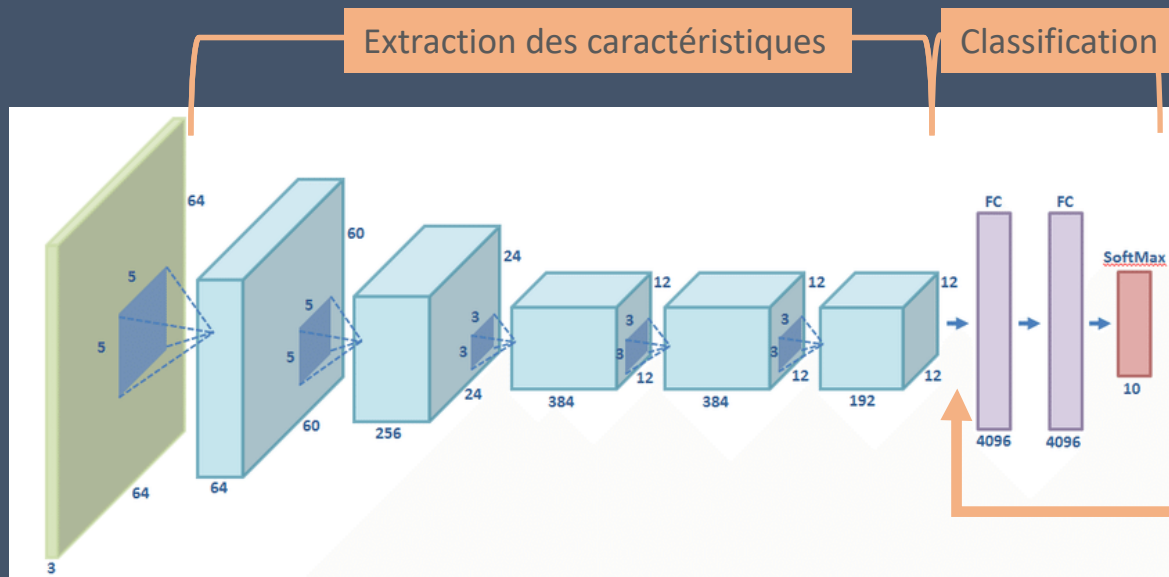
# Augmentation des données



# Apprentissage par transfert (Transfer Learning)

**Idée : réutiliser des réseaux déjà entraînés sur des bases de données énormes**

- ✓ Le modèle pré-entraîné est utilisé comme point de départ du second.
- ✓ Une part plus ou moins importante du modèle pré-entraîné est réutilisée en fonction de la similarité entre les données.



Branchement pour des images sémantiquement proches



# Quand les réseaux font n'importe quoi ...

