

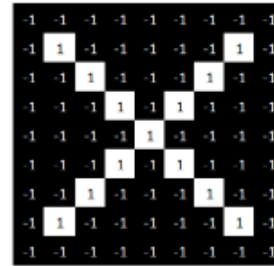
Les couches d'un CNN expliquées...

Extraction des features

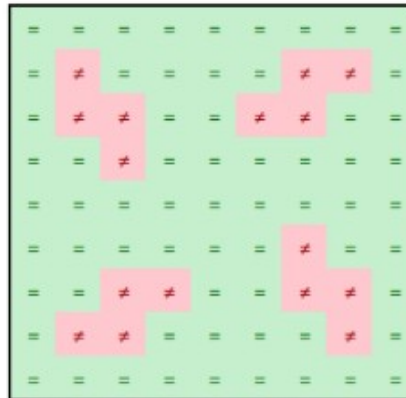
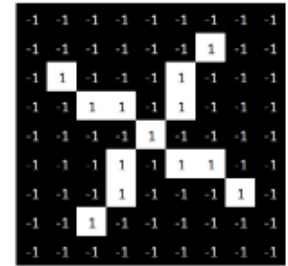
Comparaison les deux images suivantes :



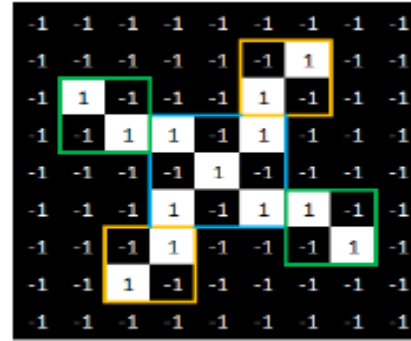
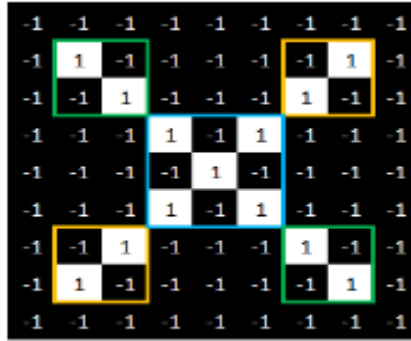
$\stackrel{?}{=}$



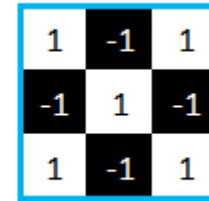
$\stackrel{?}{=}$

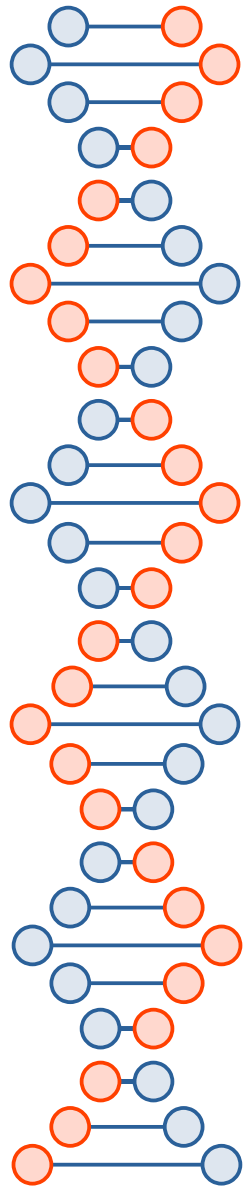


Extraction des features



Les features retenus :

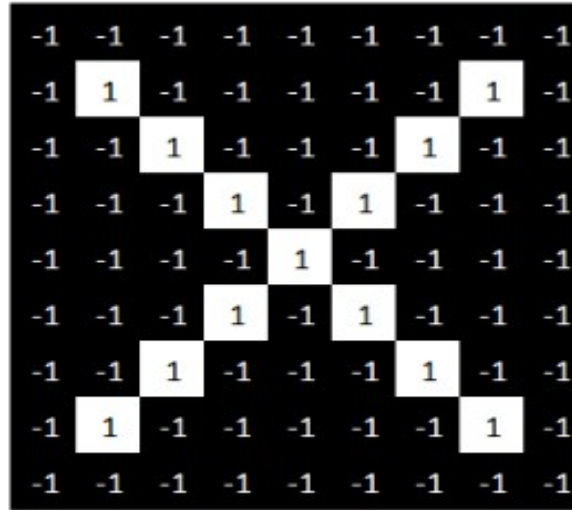




3 x 3

???

La convolution

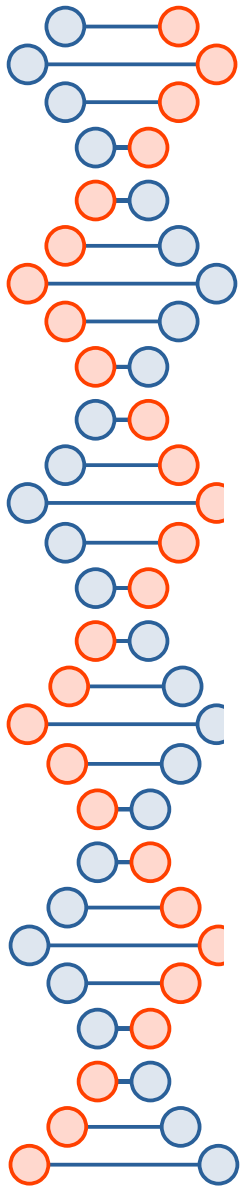


9 x 9
N x N

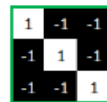
Caractéristiques d'une couche de convolution

- Un ensemble de features, filtres ou noyaux de **taille** $n \times n$
- Un **stride** s (pas du filtre)
- Un **padding** p ou marge (0 pour une convolution simple)

La convolution



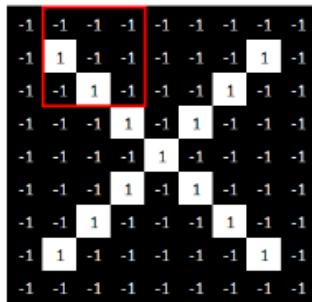
?



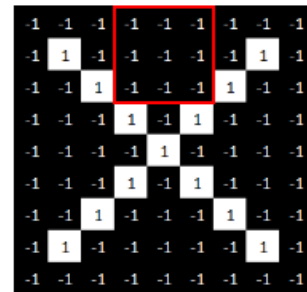
?

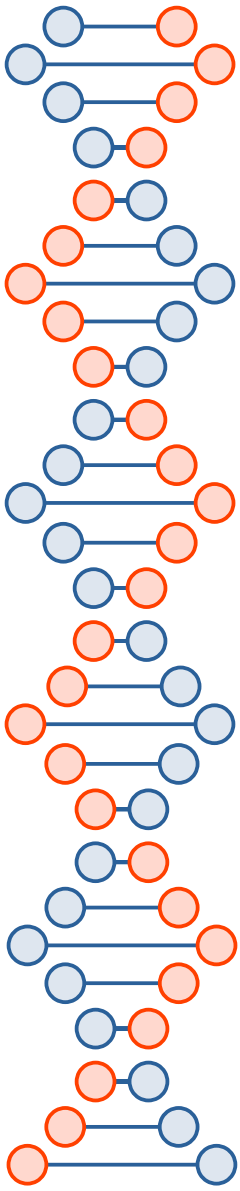


?



?

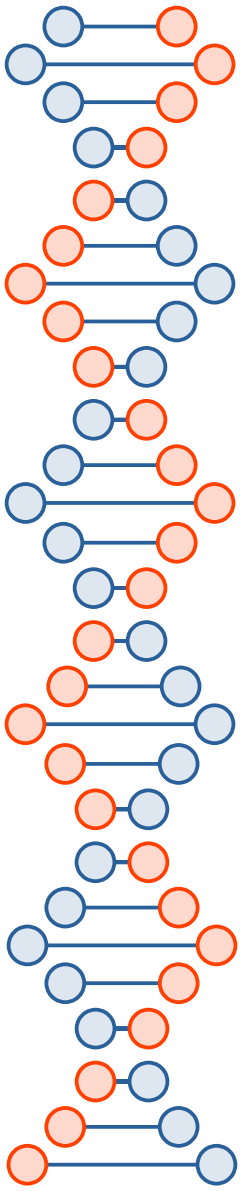




La convolution

Concrètement, pour comparer cette feature avec une partie de l'image, nous allons :

- multiplier les 9 valeurs des pixels de la caractéristique avec les 9 valeurs des pixels du morceau de l'image à trouver,
- additionner ces 9 résultats
- diviser par le nombre de pixels (ici 9)
- conclure (si le résultat est égal à 1, alors la feature a été identifiée dans l'image)



La convolution

Multiplier les 9 valeurs des pixels de la feature avec les 9 valeurs des pixels du morceau de l'image à trouver.

1	-1	-1
-1	1	-1
-1	-1	1

?

-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	-1	-1	1	-1	-1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1

-1	1	1
1	1	1
1	1	1

La convolution

Additionner ces 9 résultats



-1	1	1
1	1	1
1	1	1

$$\gg -1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 = 7$$

Diviser par le nombre de pixels

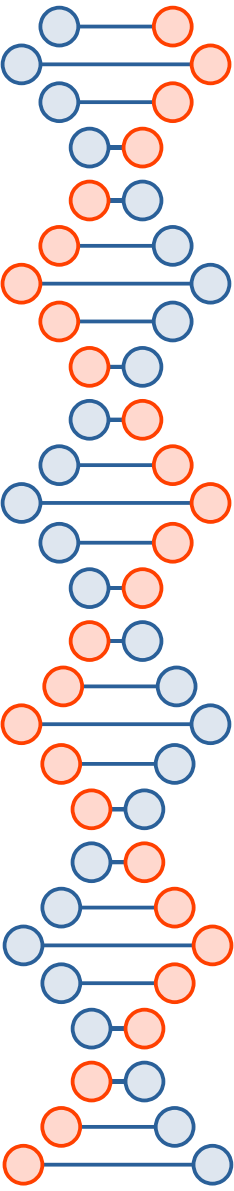
$$7/9=0,78$$

Conclure

0,78 est différent de 1, alors la caractéristique n'a pas été trouvée dans cette partie de l'image.

La convolution

En réalisant la même opération pour le reste de l'image, nous obtenons le tableau de valeurs suivant :

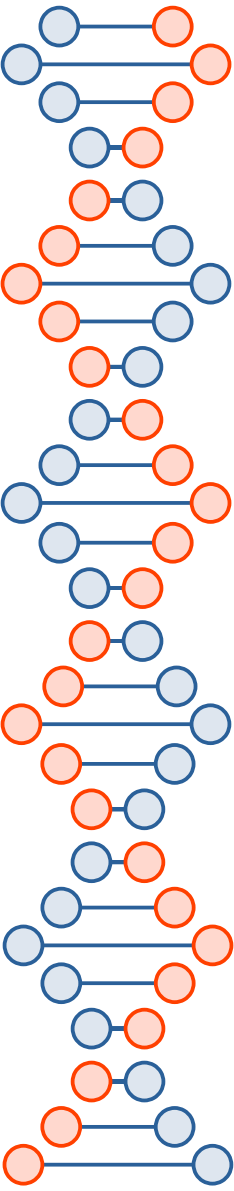

$$7 = [(9 + 2 * 0 - 3) / 1] + 1$$

0,8	-0,1	0,1	0,3	0,6	-0,1	0,3
-0,1	1,0	-0,1	0,3	-0,1	0,1	-0,1
0,1	-0,1	1,0	-0,3	0,1	-0,1	0,6
0,3	0,3	-0,3	0,6	-0,3	0,3	0,3
0,6	-0,1	0,1	-0,3	1,0	-0,1	0,1
-0,1	0,1	-0,1	0,3	-0,1	1,0	-0,1
0,3	-0,1	0,6	0,3	0,1	-0,1	0,8

7 x 7

$$N_{out} = [(N_{in} + 2p - n) / s] + 1$$

La correction (ReLU)



0,8	-0,1	0,1	0,3	0,6	-0,1	0,3
-0,1	1,0	-0,1	0,3	-0,1	0,1	-0,1
0,1	-0,1	1,0	-0,3	0,1	-0,1	0,6
0,3	0,3	-0,3	0,6	-0,3	0,3	0,3
0,6	-0,1	0,1	-0,3	1,0	-0,1	0,1
-0,1	0,1	-0,1	0,3	-0,1	1,0	-0,1
0,3	-0,1	0,6	0,3	0,1	-0,1	0,8

>>>

0,8	0,0	0,1	0,3	0,6	0,0	0,3
0,0	1,0	0,0	0,3	0,0	0,1	0,0
0,1	0,0	1,0	0,0	0,1	0,0	0,6
0,3	0,3	0,0	0,6	0,0	0,3	0,3
0,6	0,0	0,1	0,0	1,0	0,0	0,1
0,0	0,1	0,0	0,3	0,0	1,0	0,0
0,3	0,0	0,6	0,3	0,1	0,0	0,8



Le pooling

Overlapping max POOL de taille 2x2 effectué avec un stride de 2 et pad = 0

0,8	0,0	0,1	0,3	0,6	0,0	0,3
0,0	1,0	0,0	0,3	0,0	0,1	0,0
0,1	0,0	1,0	0,0	0,1	0,0	0,6
0,3	0,3	0,0	0,6	0,0	0,3	0,3
0,6	0,0	0,1	0,0	1,0	0,0	0,1
0,0	0,1	0,0	0,3	0,0	1,0	0,0
0,3	0,0	0,6	0,3	0,1	0,0	0,8

7 x 7

>>>

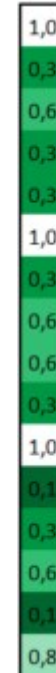
1,0	0,3	0,6	0,3
0,3	1,0	0,3	0,6
0,6	0,3	1,0	0,1
0,3	0,6	0,1	0,8

4 x 4

Cette couche possède les même caractéristiques que la convolution
(n, s, p)

Le flattening (mise à plat)

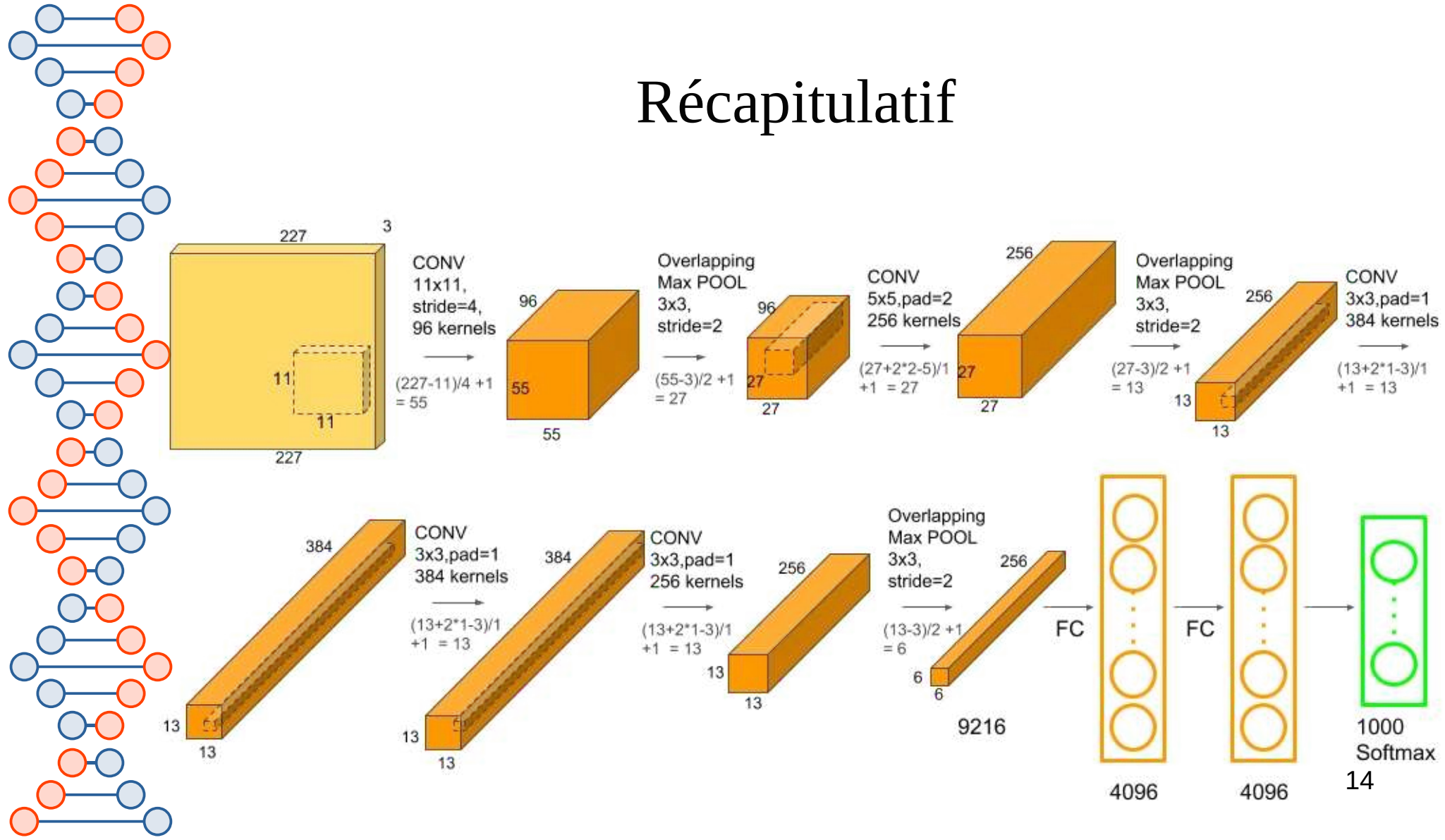
Cette opération consiste tout simplement à prendre la totalité des valeurs de notre matrices précédemment calculées, et à les empiler, en vue de les exploiter dans la couche d'entrée d'un réseau de neurones.



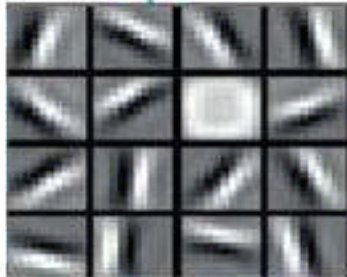
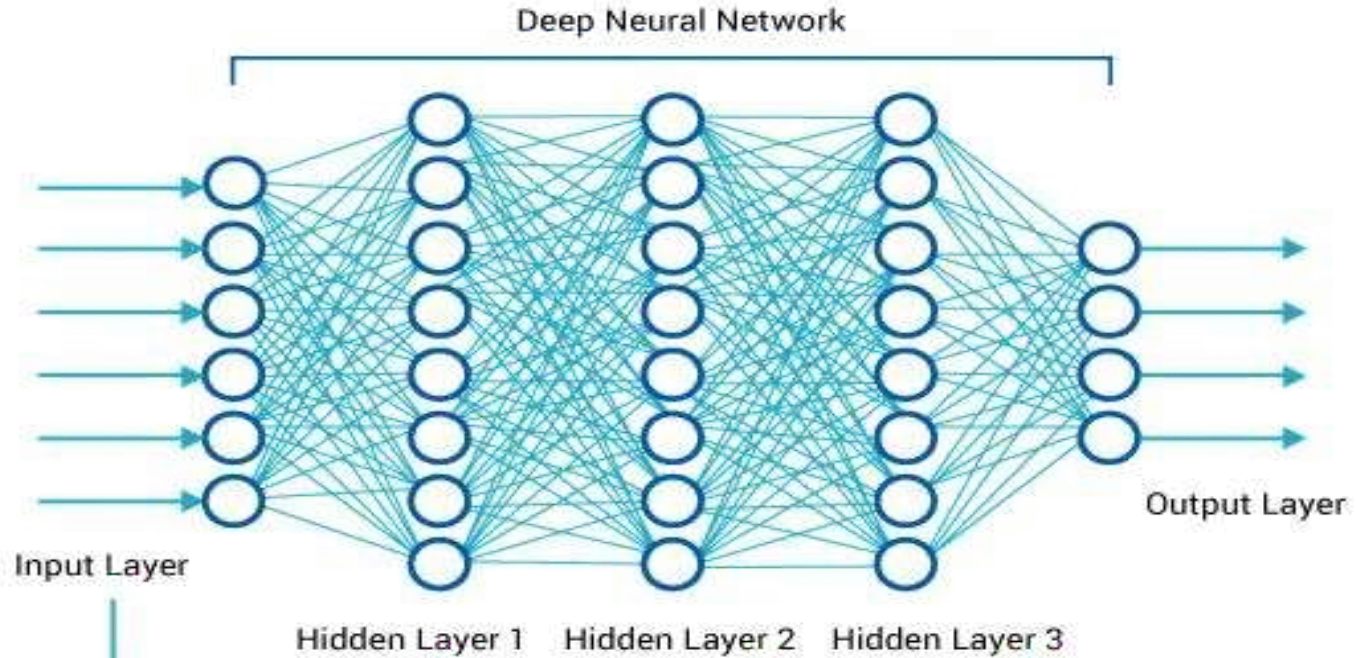
Fully-connected

- La couche fully-connected constitue toujours la dernière couche d'un réseau de neurones, convolutif ou non
- La dernière couche fully-connected permet de classifier l'image en entrée du réseau
- Elle renvoie un vecteur de taille N (nombre de classes dans notre problème de classification d'images).
- Chaque élément du vecteur indique la probabilité pour l'image en entrée d'appartenir à une classe.

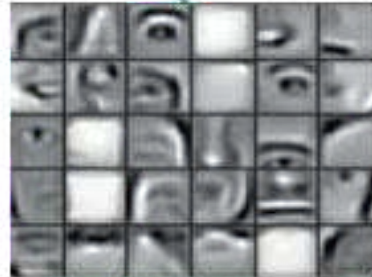
Récapitulatif



Récapitulatif



edges



combinations of edges



object models



Merci pour votre attention...