

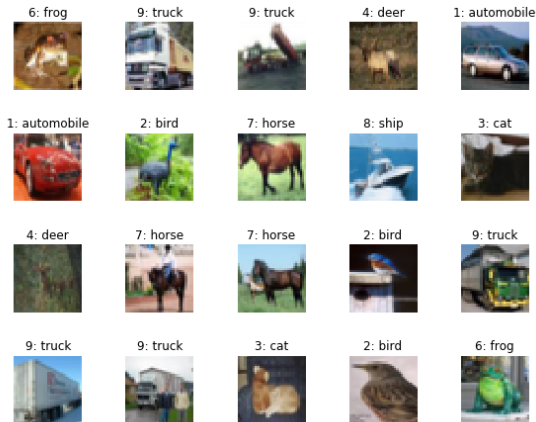
Convolutional neural networks

Ens'IA

Ensimag 2021-2022

2 janvier 2023

Reminder

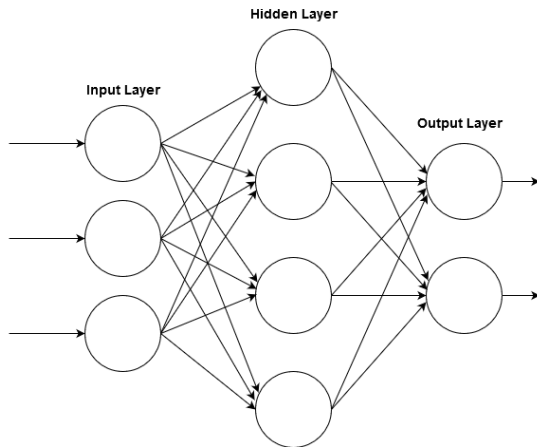


Images classification

→ We need a function f
such that $f(x) = y$

How to approximate this function? → Neural networks

Neural network



→ Succession of neuron layers

Convolution

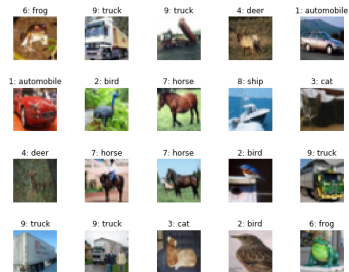
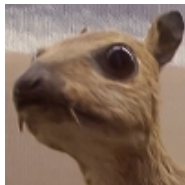


FIGURE 1 – CIFAR-10 vs. ImageNet

Convolution

Idea :

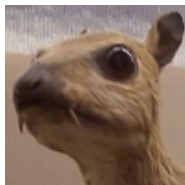
→ Image processing technique : find features in the image : color changes, lines, brightness variations...



Convolution

Idea :

→ Image processing technique : find features in the image : color changes, lines, brightness variations...



How to do that ?

→ Apply filters : a convolution kernel

$$H = \begin{pmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{pmatrix}$$

Convolution - Example

Detecting edges on an image :



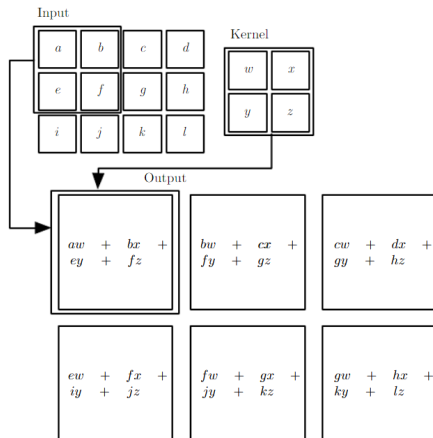
Vertical edges



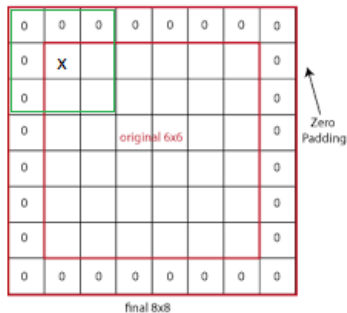
Horizontal edges

Convolution

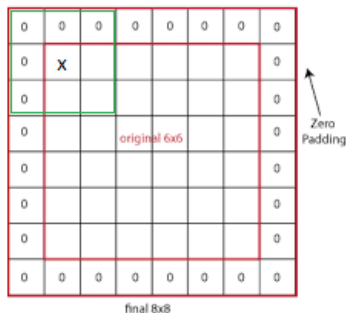
→ The mathematical operation



Convolution



Convolution



4 parameters to tune (+ the activation function) :

- The size of the filter : *kernel size*
- The number of filter : *depth*
- The number of pixels that the filter moves on each iteration : *stride*
- *zero padding*

Pooling

→ Problem : dimensions can increase rapidly !

Pooling

→ Problem : dimensions can increase rapidly !

→ Idea : reduce the dimensions of data by combining the outputs of the previous layer

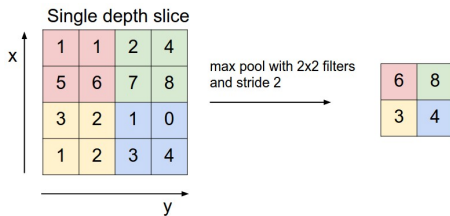
Pooling

→ Problem : dimensions can increase rapidly !

→ Idea : reduce the dimensions of data by combining the outputs of the previous layer

→ different types :

- *Max Pooling*
- *Average Pooling*

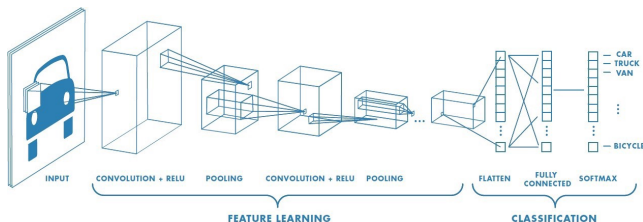


→ After the convolutional and pooling layers, we add fully connected layers.

→ After the convolutional and pooling layers, we add fully connected layers.

→ Summary :

- Convolution then Pooling Layers
- Fully connected layer



State of the art

Yolov3 for object detection



<https://www.youtube.com/watch?v=1LCb1PVqzeY>

State of the art

Yolo =

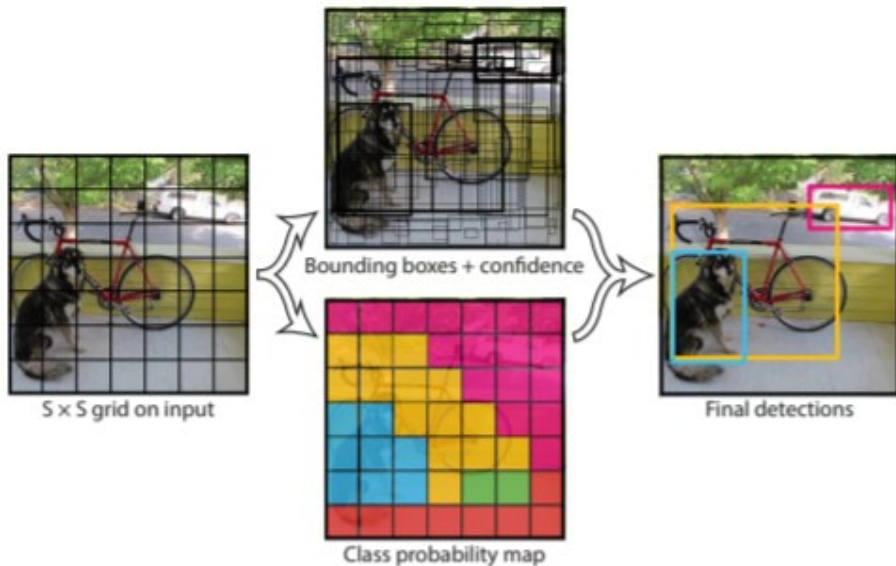
Yolo = You Only Look Once

Yolo = You Only Look Once

- ➊ Residual blocks : the image is divided into various grids. Each grid has a dimension of $S \times S$.
- ➋ Bounding box regression : outline that highlights an object in an image.
- ➌ Intersection over union (IOU) : a phenomenon in object detection that describes how boxes overlap.

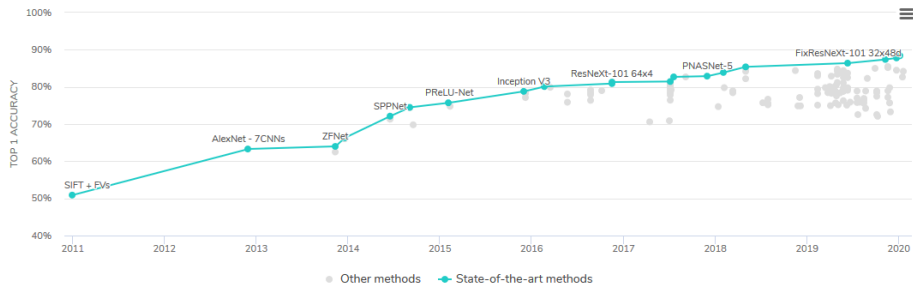
→ Only a single forward propagation through a neural network to detect objects (fast!)

State of the art



State of the art

Image Classification using ImageNet



Famous networks : ResNet, DenseNet, VGG

→ Training for many weeks

→ very effective

Idea : Some models were trained for weeks to learn a job

Idea : Some models were trained for weeks to learn a job
Can we "transfer" what it learned ?

Idea : Some models were trained for weeks to learn a job

Can we "transfer" what it learned ?

→ Yes.

- ① Choose a model that was used for a similar job
- ② Replace the fully connected layers
- ③ Train the fully connected layers
- ④ Use your new model

→ Faster

→ More effective

Sometimes, we don't have enough data...

We can use techniques to generate more data from an already existing one :

- Crop the image
- flip the image
- zoom/unzoom the image

Sometimes, we don't have enough data...

We can use techniques to generate more data from an already existing one :

- Crop the image
- flip the image
- zoom/unzoom the image

→ Also helps preventing overfitting

- CS231N (<http://cs231n.github.io/convolutional-networks/>)
- <http://neuralnetworksanddeeplearning.com/>
- <http://www.deeplearningbook.org/>
- <http://towardsdatascience.com/>
- Wikipedia
- <https://machinelearningmastery.com/transfer-learning-for-deep-learning/>
- <https://paperswithcode.com/sota/>
- <https://www.section.io/engineering-education/introduction-to-yolo-algorithm-for-object-detection/>