Machine Learning

NN and Deep Learning

Fabio Vandin December 10^{th}, 2021

Neural Networks

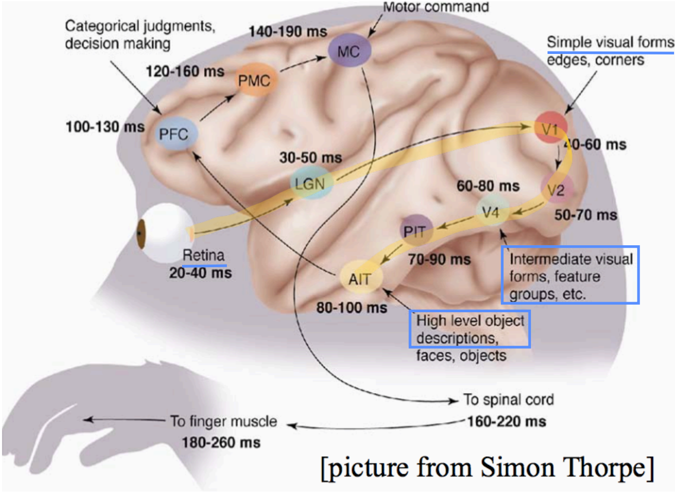
Informal definition: simplified models of the brain

- large number of basic computing units: *neurons*
- connected in a complex network

Deep Learning

Deep Learning = learning hierarchical representations

Biological inspiration *e.g. our visual cortex is hierarchical*

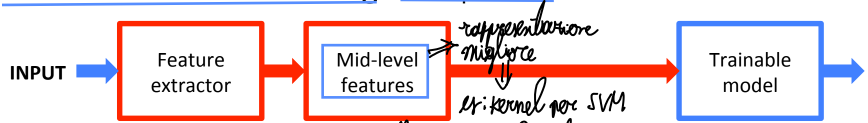


[picture from Simon Thorpe]

Hierarchical Representations and ML

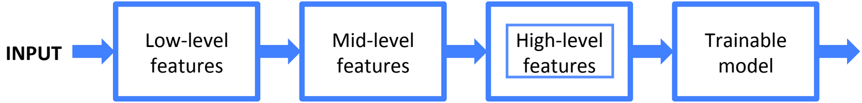
Traditional Machine Learning: Fixed/Handcrafted Feature Extractor

Advanced Machine Learning: Unsupervised mid-level features



non si imparano vettori manualmente

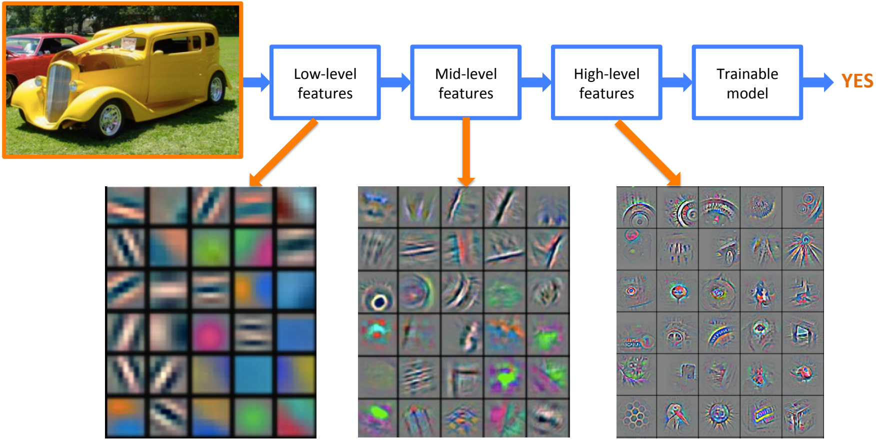
Deep Learning: Representations are hierarchical and trained



Deep Learning $\Rightarrow$ **End-to-end Learning**

Hierarchical Representations and ML (continue)

Example: object recognition (object = car)



Neural Networks: Historical Remarks

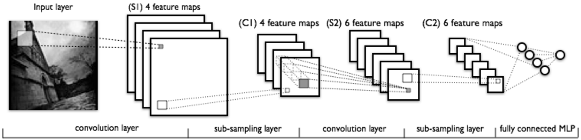
- 1940s-70s:
  - Inspired by learning/modeling the brain (Pitts, Hebb, and others)
  - Perceptron (Rosenblatt), Multilayer perceptron (Minksy and Papert)
  - Backpropagation (Werbos, 1975)
- 1980s-early 1990s:
  - Practical Backpropagation (Rumelhart, Hinton et al., 1986) and SGD (Bottou)
  - Initial empirical success
- 1990s-2000s:
  - Lost favor to implicit linear methods: SVM, Boosting
- 2006-:
  - Regain popularity because of unsupervised pre-training (Hinton, Bengio, LeCun, Ng, and others)
  - Computational advances and several new tricks allow training *huge* networks. Empirical success leads to renewed interest
  - 2012: Krizhevsky, Sustkever, Hinton: significant improvement of state-of-the-art on imagenet dataset (object recognition of 1000 classes), without unsupervised pre-training

Issues with “Traditional NNs”

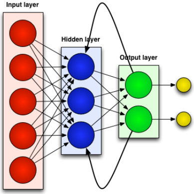
- Huge number of edges $\Rightarrow$ huge number of weights (due to *full connectivity*
- Domain structure is not taken into account
  - in some domains (e.g., images, text) there is a lot of *structure* $\Rightarrow$ parts of the input have specific relations!
  - what about having *similar/related* neurons?

Modern Neural Networks

- Convolutional Neural Networks (CNN): used mostly for computer vision, imaging, etc.



- Recurrent Neural Networks (RNN): used mostly for time series analysis, speech recognition, machine translation, etc.



- ...

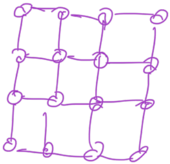
We will see an overview of CNNs

Convolutional Neural Networks (CNNs)

Convolutional NNs: employ the operation of *convolution*

Definition. Convolutional networks are neural networks that use convolution in place of general matrix multiplication in at least one of their layers

CNNs are a specialized kind of neural network for processing data that has a known grid-like topology



Convolutional Neural Networks (CNNs)

Convolutional NNs: employ the operation of *convolution*

Definition. Convolutional networks are neural networks that use convolution in place of general matrix multiplication in at least one of their layers

CNNs are a specialized kind of neural network for processing data that has a known grid-like topology

Example

Convolution in CNNs

Assume we have a two-dimensional input I and a two-dimensional function K, the (discrete) convolution S of I and K is:

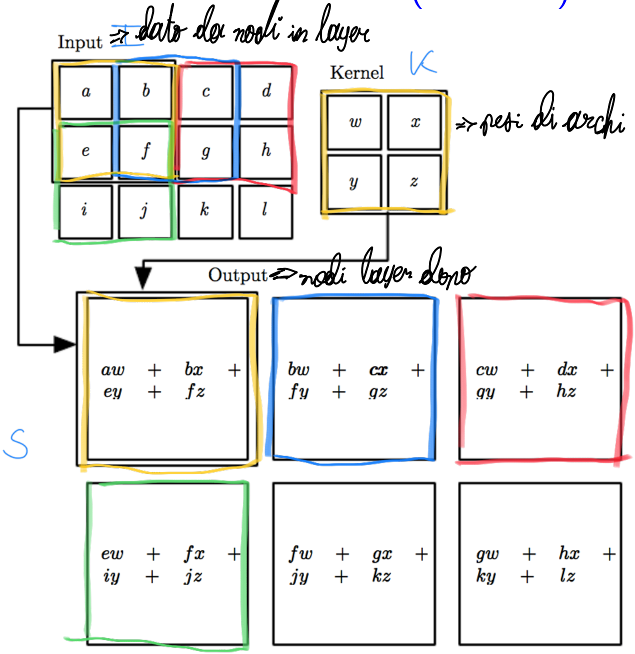
$$\boxed{S(i,j) = (I * K)(i,j) = \sum_m \sum_n I(i+m, j+n)K(m,n)}$$

Discrete convolution can be viewed as multiplication by a matrix, but the matrix has several entries constrained to be equal to other entries

Note: K is usually called a *kernel*

In CNNs $K(m,n)$ is zero for all but small values of m, n

Convolution in CNNs (continue)



Convolution: Example

INPUT

OUTPUT

KERNEL

	0	0	0	
	-1	1	0	
	0	0	0	

Filter: *edge enhance*