Convolution Neural Networks

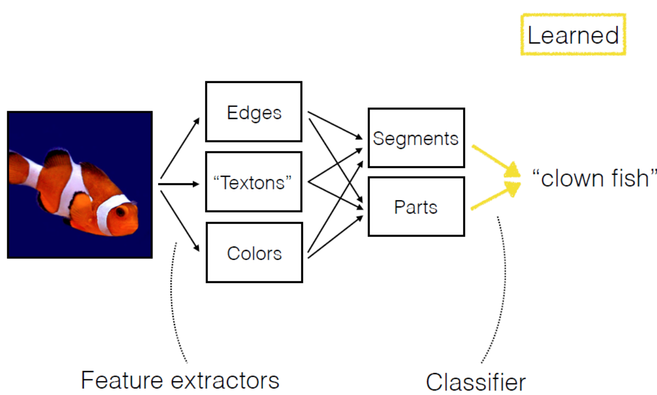
- Slides (./files/deep_learning_tutorial_2015.pdf) by Phillip Isola
- Slides (./files/cnn1.pdf) by Prof. Ali Ghodsi

Prof. Seungchul Lee
iSystems
UNIST
http://isystems.unist.ac.kr/

Table of Contents

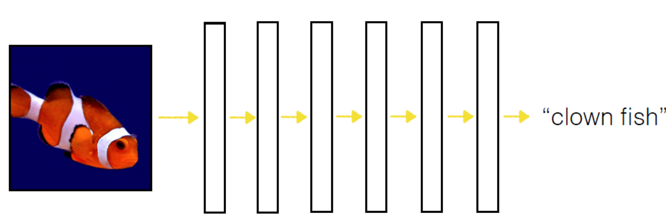
1. Tranditional Machine Learning vs. Neural Networks

Object recognition using machine learning



Neural Network

***Deep* Neural Network**



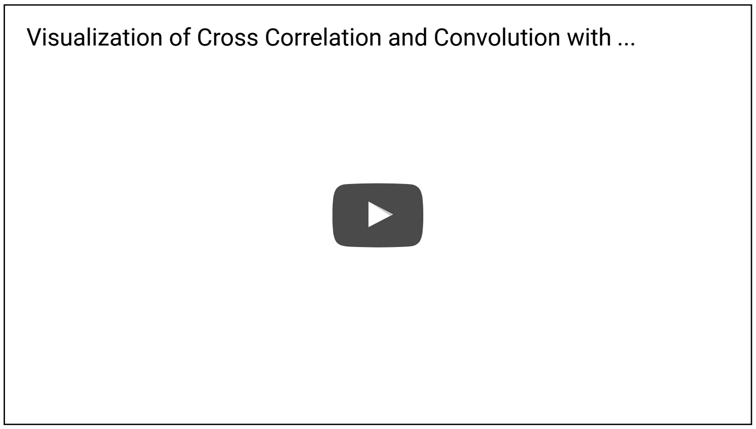
2. Convolution Neural Networks

CNNs are simply neural networks that use *convolution* in place of general matrix multiplication in at least one of their layers

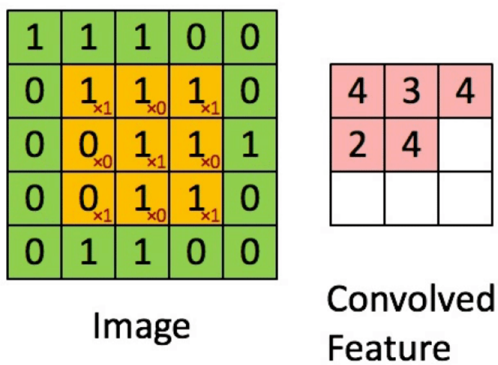
2.1. Convolution and cross-correlation

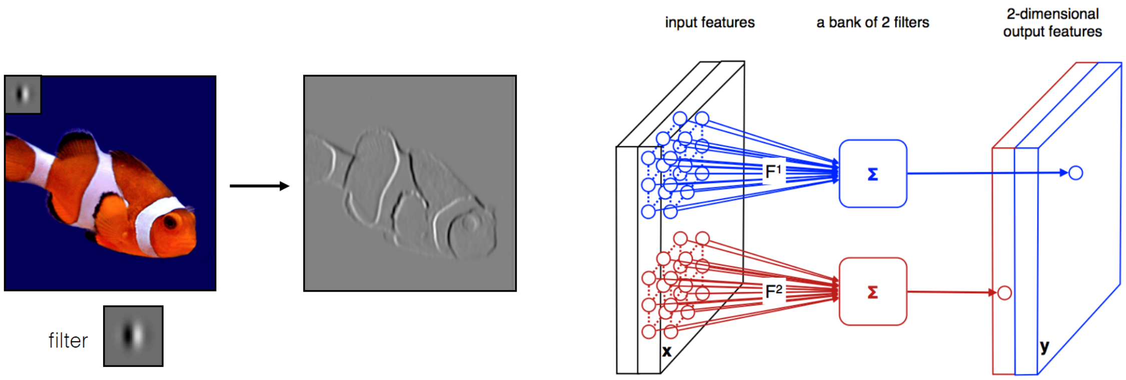
- Many machine learning libraries implement cross-correlation, but call it convolution

In [1]:
```
%%html
<iframe src="https://www.youtube.com/embed/Ma0YONjMZLI"
width="560" height="315" frameborder="0" allowfullscreen></iframe>
```



- Discrete convolution can be viewed as multiplication by a matrix





Sparse interations

- CNNs, typically have sparse connectivity (sparse weights)
- This is accomplished by making the kernel (convolution mask) smaller than the input

Parameter sharing

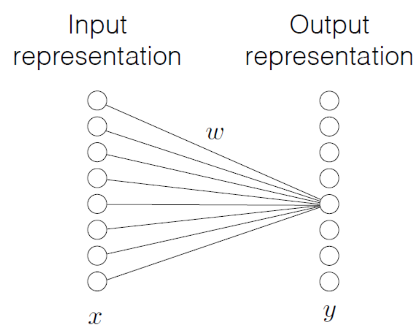
- In CNNs each number of the kernel is used at every position of the input
- Instead of learning a separate set of parameters for every location, we learn only one set

Equivariance

- A function $f(x)$ is equivariant to a function g if $f(g(x)) = g(f(x))$
- A convolution layer has equivariance to translation
- If we apply this translation to x, then apply convolution, the result will be the same as if we applied convolution to x, then applied the transformation to the output
- Note that convolution is not equivariant to some other transformation, such as changes in the scale or rotation of an image

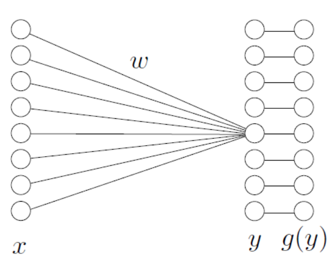
2.2. Computation in a neural net

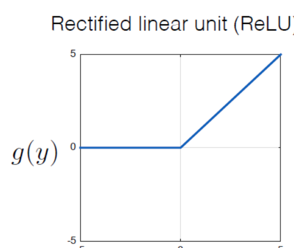
1) Linear combination



$$y_j = \sum_i w_{ij} x_i$$

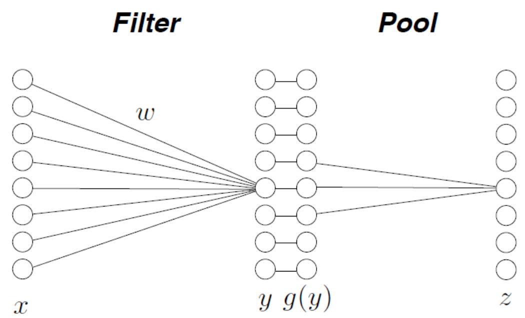
2) Nonlinear activation function





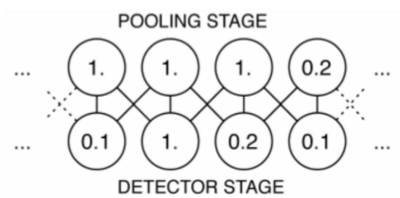
$$g(y) = \max(0, y)$$

3) Pooling functions

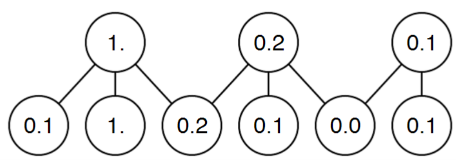


$$y_j = \sum_i w_{ij} x_i \qquad z_k = \max_{j \in \mathcal{N}(j)} g(y_j)$$

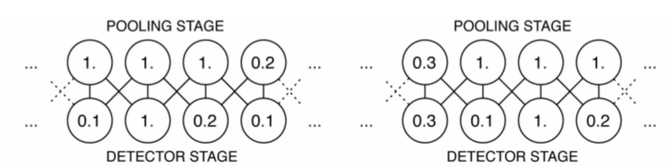
- The maximum of a rectangular neighborhood (max pooling operation)



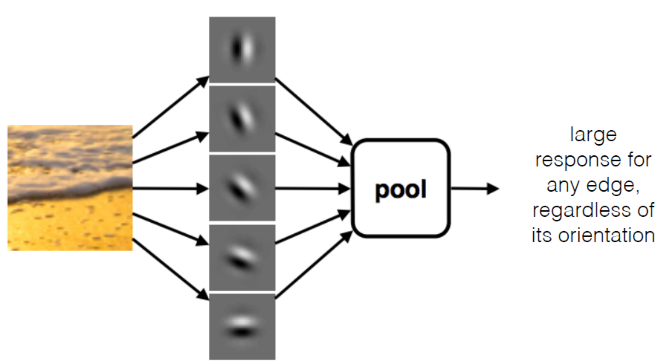
- Other candiates
  - the average of a rectangular neighborhood
  - the L_2 norm of a rectangular neighborhood
- Pooling with downsampling
  - reduce the representation size by a factor of 2, which reduces the computational and statistical burden on the next layer



- Pooling and translations
  - Pooling helps to make the representation become invariant to small translations of the input

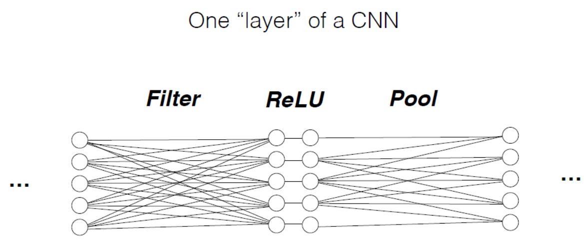


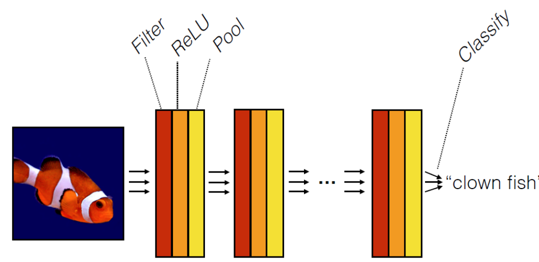
  - Invariance to local translation can be a very useful property if we care more about whether some feature is present than exactly where it is.
  - For example, we need not know the exact location of the eyes in a face



4) Classification

CNN summary

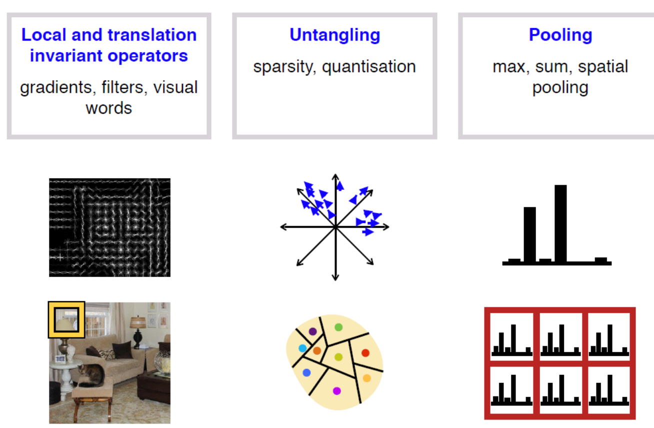




$$f(\mathbf{x}) = f_L(\dots f_2(f_1(\mathbf{x})))$$

2.3. Ingredients

- Select important features of the data
  - linear filters
  - pointwise nonlinearity
- Group features that all indicate the same thing
  - pooling
- Repeat to achieve greater abstraction

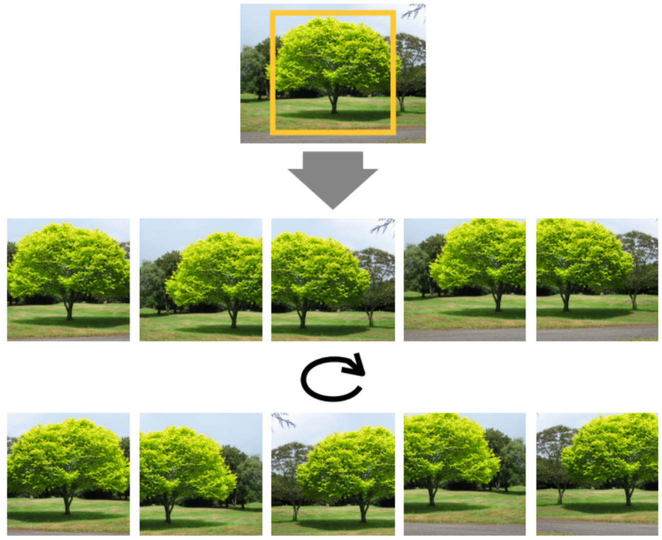


3. Learning

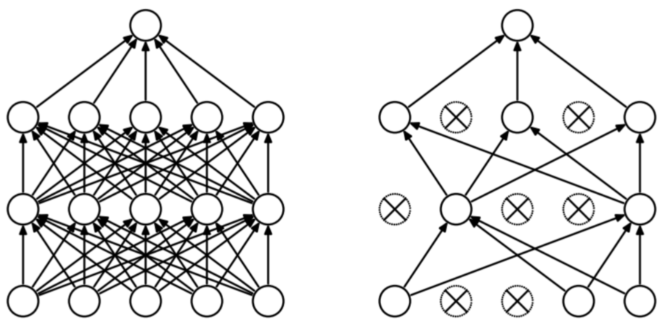
- Backpropagation and
- stochastic gradient descent

4. How to avoid overfitting?

- Convolutional nets use a prior that stuff in the world does not change identity as it translates
- Data Augmentation
  - Augment the training data by adding jittered versions of each image



- Dropout
  - Randomly choose edges not to update
  - Insensitive to local changes
  - acting as regularization



(a) Standard Neural Net (b) After applying dropout.

5. How do deep neural nets work?

- Hierarchy of simple, repeated computations
- Sift through data by filtering it
- Build up invariance by pooling alike features
- Can be learned with vanilla SGD

6. Software Tools

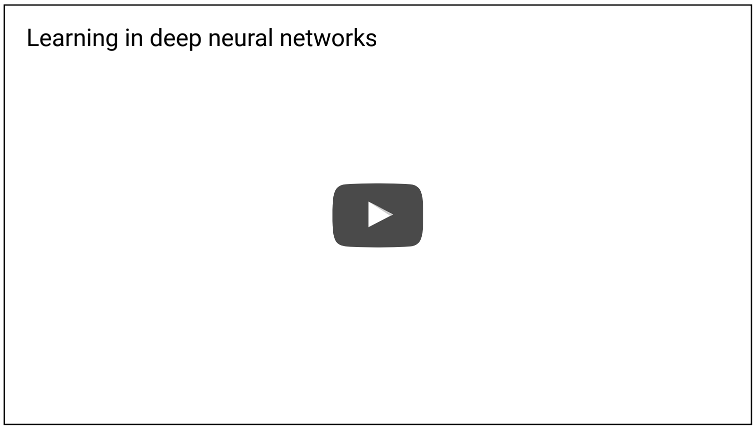
- Caffe
  - fast and popular
  - hard to use
  - C++ with linited Matlab and Python interfaces
- Theano
  - Symbolic computation and automatic differentiation python
- Torch
  - Lua

Online Video Lectures

- Slides (./files/deep_learning_tutorial_2015.pdf) by Phillip Isola

In [4]:

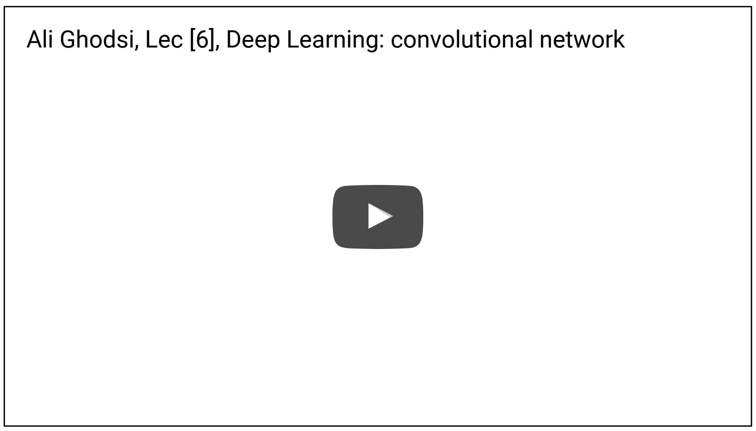
```
%%html
<iframe src="https://www.youtube.com/embed/bL1Zymz1b7g"
width="560" height="315" frameborder="0" allowfullscreen></iframe>
```



- Slides (./files/cnn1.pdf) by Prof. Ali Ghodsi

In [5]:

```
%%html
<iframe src="https://www.youtube.com/embed/ZMBp7_qqtLE"
width="560" height="315" frameborder="0" allowfullscreen></iframe>
```



In [3]:

```
%%javascript
$.getScript('https://kmahelona.github.io/ipython_notebook_goodies/ipython_notebook_toc.js')
```