Convolutional neural network (CNN)

History

Yann LeCun, Professor of Computer Science
The Courant Institute of Mathematical Sciences
New York University
Room 1220, 715 Broadway, New York, NY 10003, USA.
(212)998-3283 yann@cs.nyu.edu

- In 1995, Yann LeCun and Yoshua Bengio introduced the concept of convolutional neural networks.

About CNN's

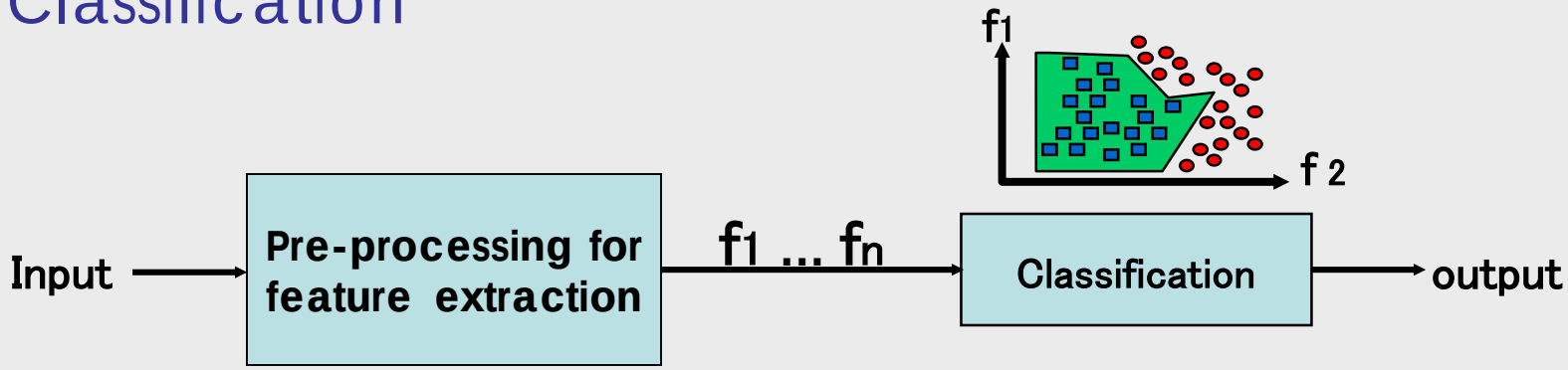
- CNN's Were neurobiologically motivated by the findings of locally sensitive and orientation-selective nerve cells in the visual cortex.
- They designed a network structure that implicitly extracts relevant features.
- Convolutional Neural Networks are a special kind of multi-layer neural networks.

About CNN's

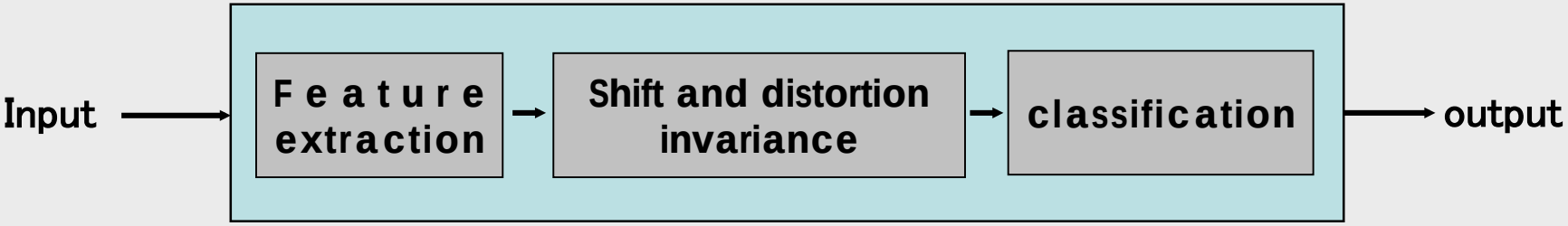
- Like almost every other neural networks they are trained with a version of the back-propagation algorithm.
- Convolutional Neural Networks are designed to recognize visual patterns directly from pixel images with minimal preprocessing.
- They can recognize patterns with extreme variability (such as handwritten characters).

Classification

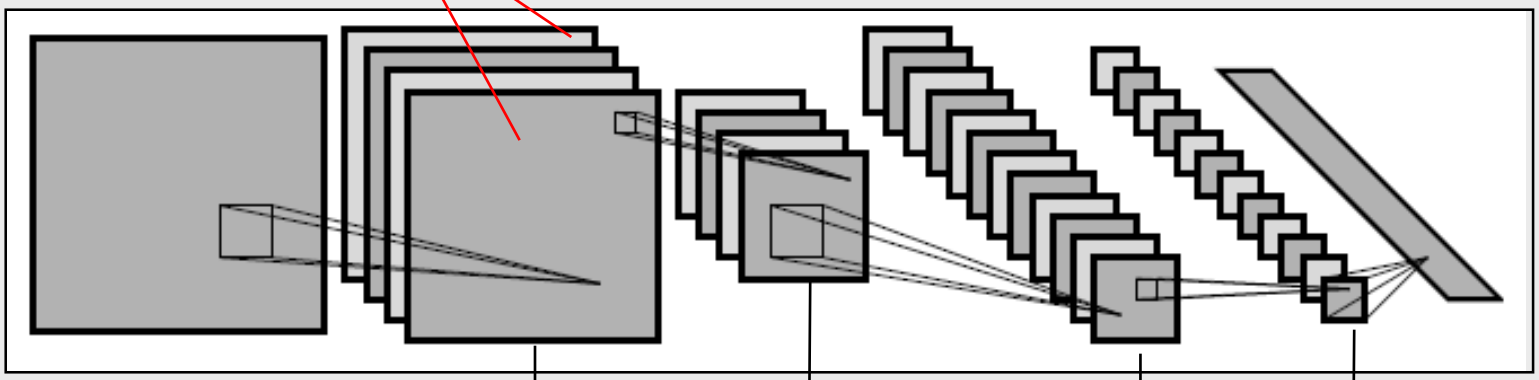
Classification



Convolutional neural network



CNN's Topology



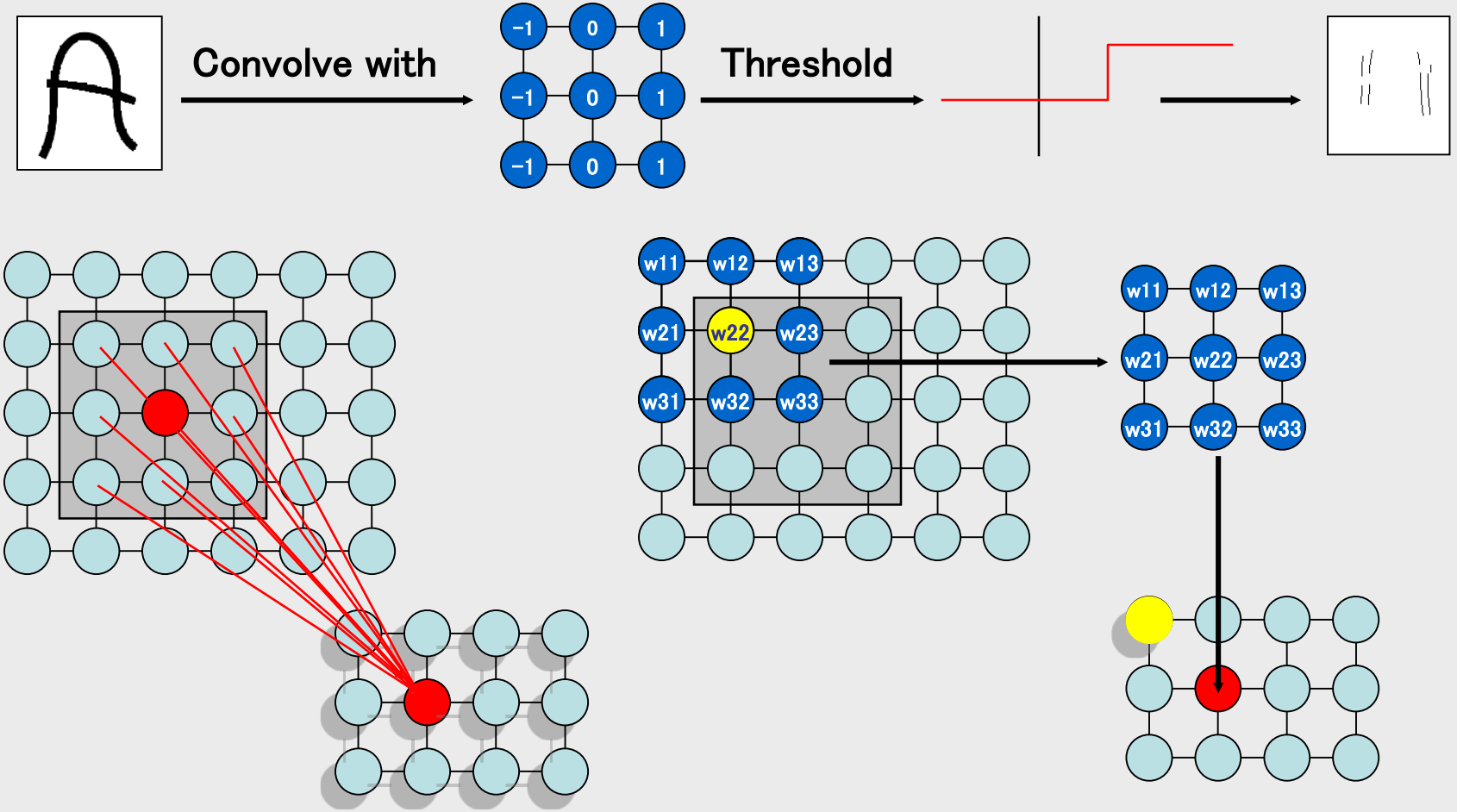
Feature extraction layer
Convolution layer

C

Shift and distortion invariance or
Subsampling layer

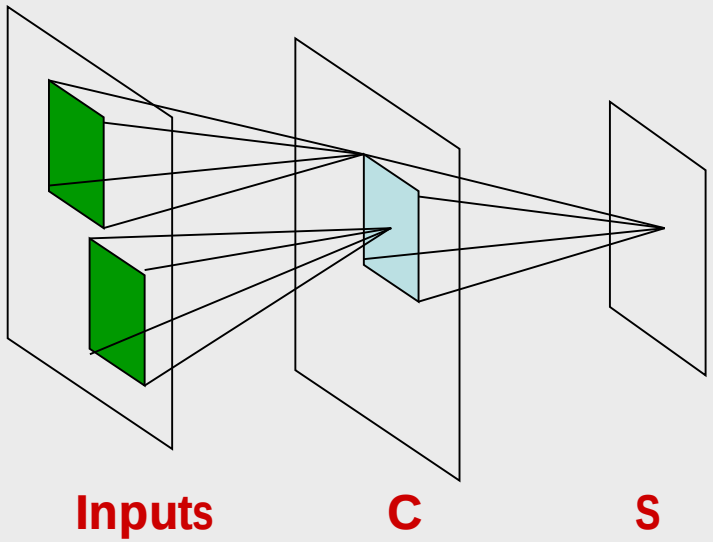
S

Feature extraction



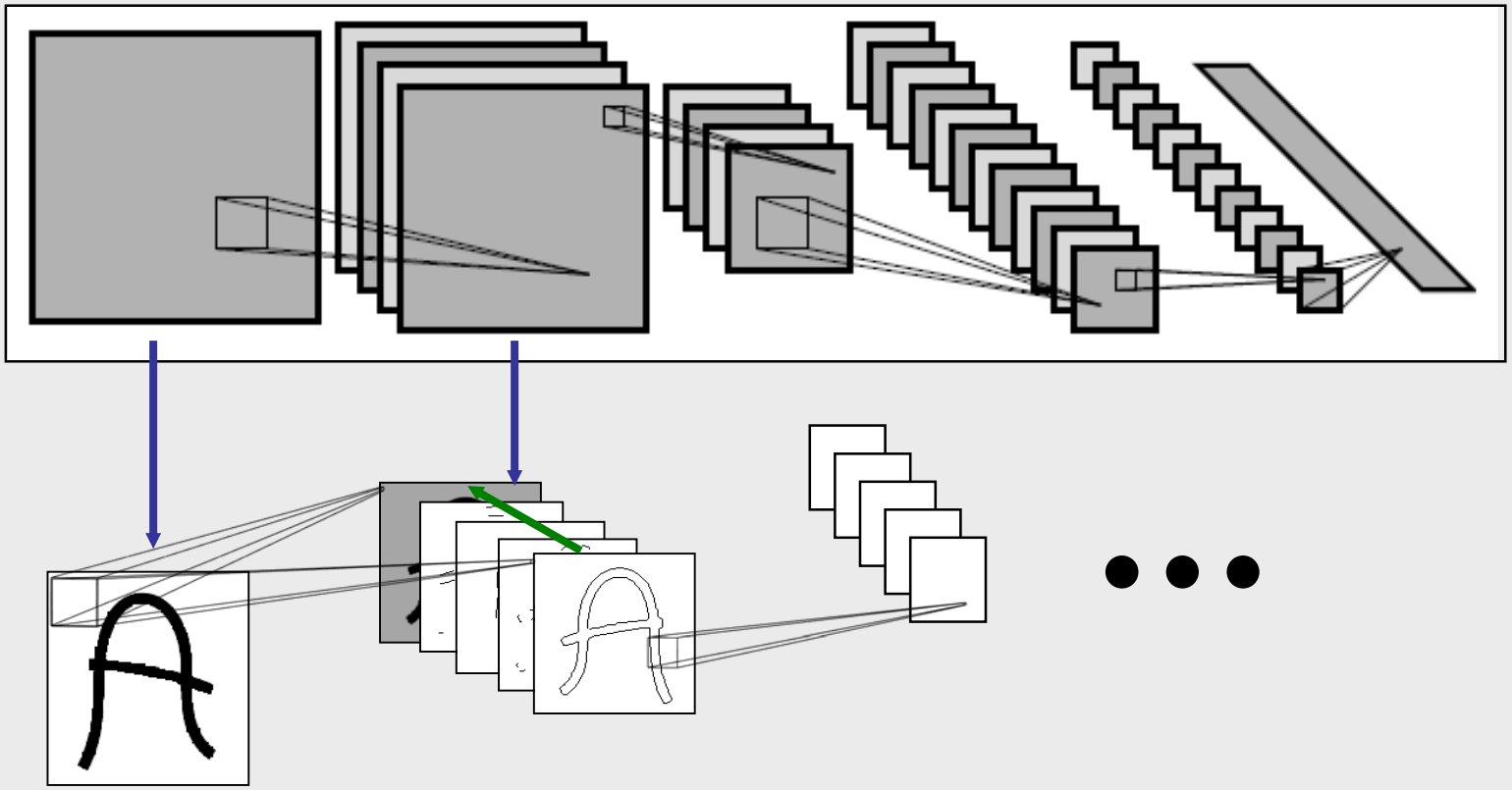
Feature extraction

- Shared weights: all neurons in a feature share the same weights .
- In this way all neurons detect the same feature at different positions in the input image.
- Reduce the number of free parameters.



Feature extraction

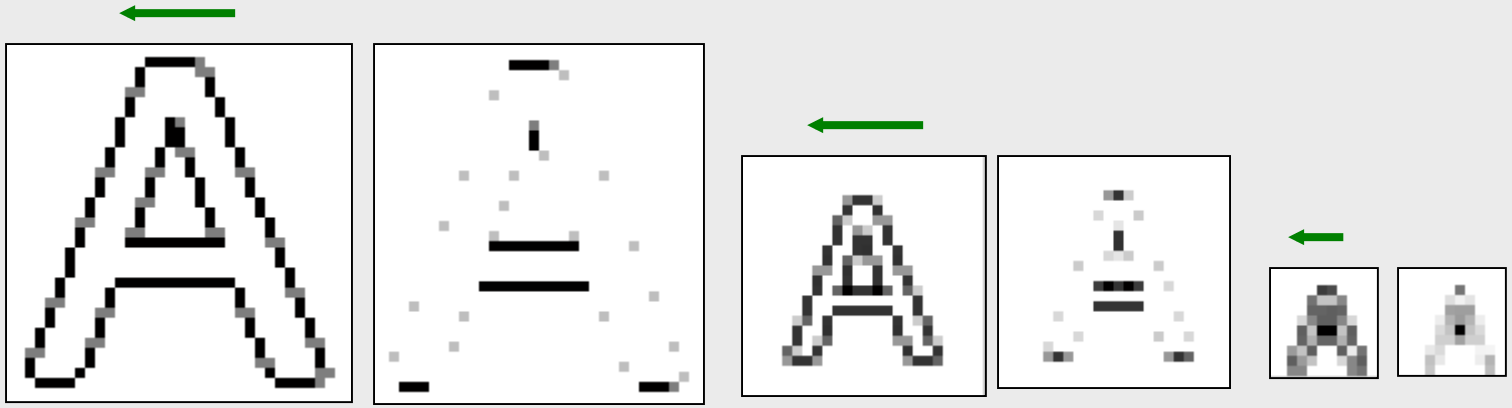
- If a neuron in the feature map fires, this corresponds to a match with the template.



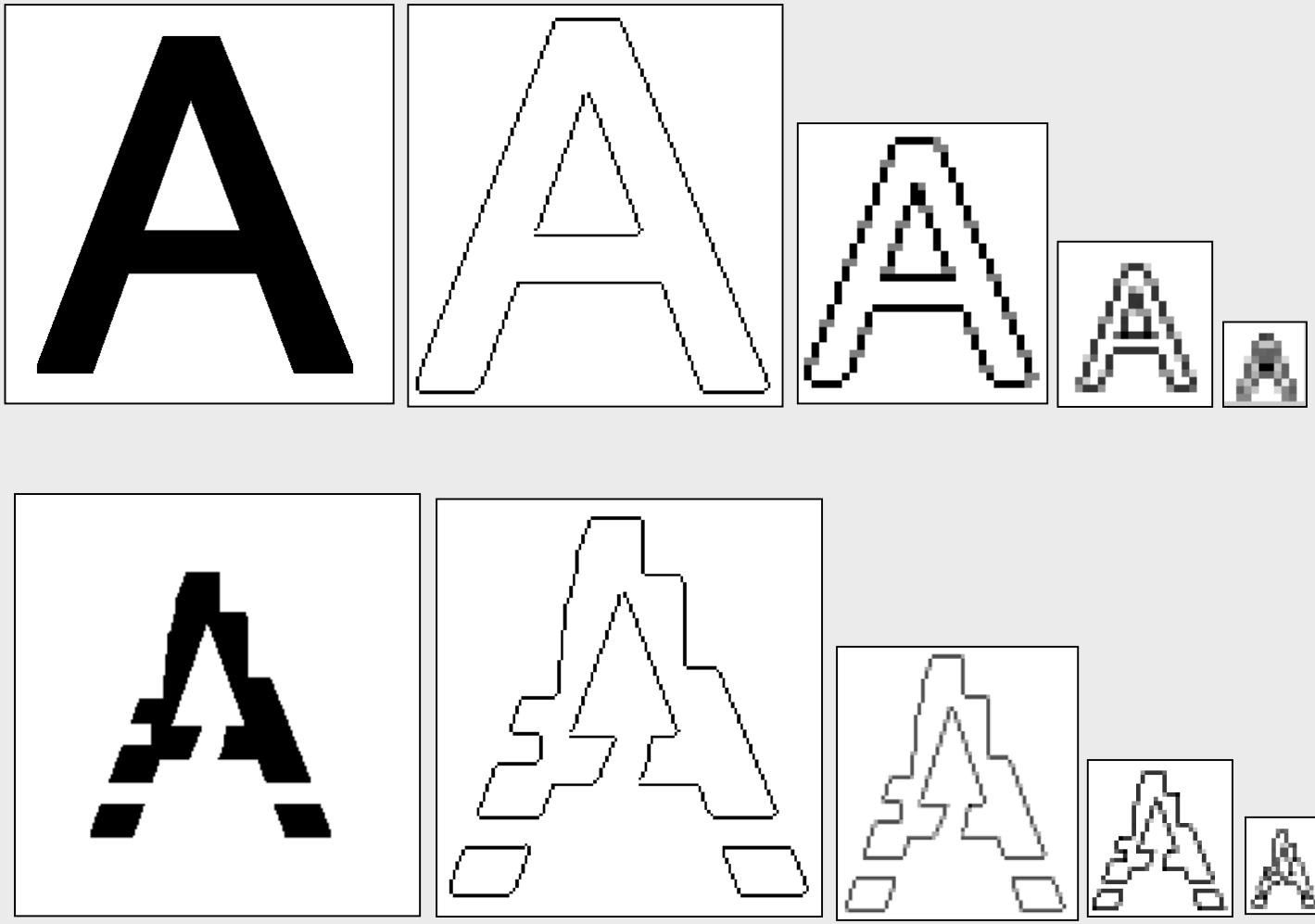
Subsampling layer

- the subsampling layers reduce the spatial resolution of each feature map

- By reducing the spatial resolution of the feature map, a certain degree of shift and distortion invariance is achieved.

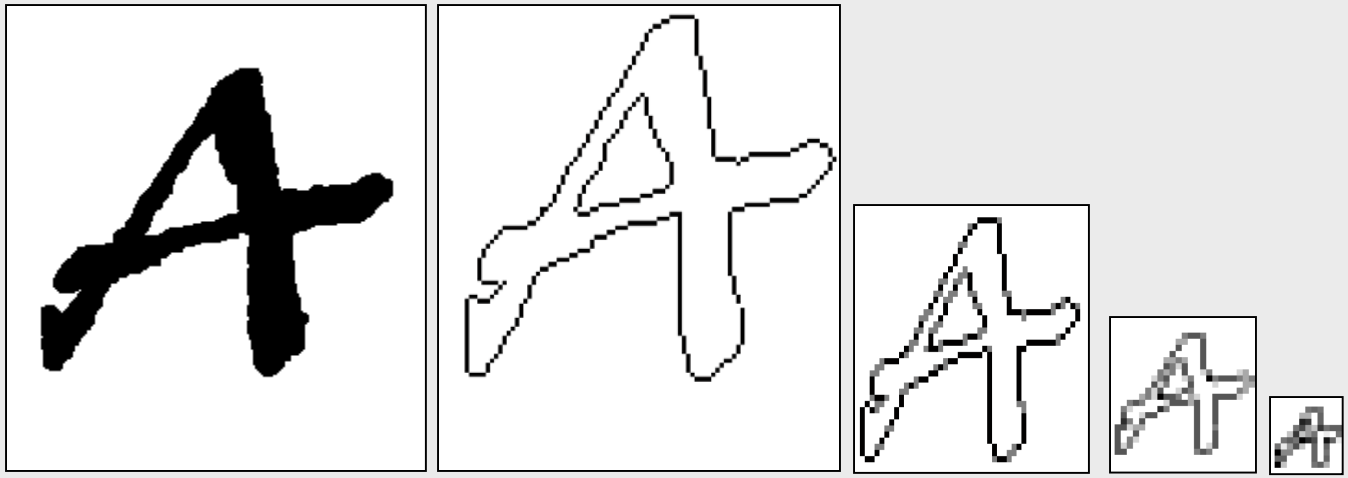


Subsampling layer



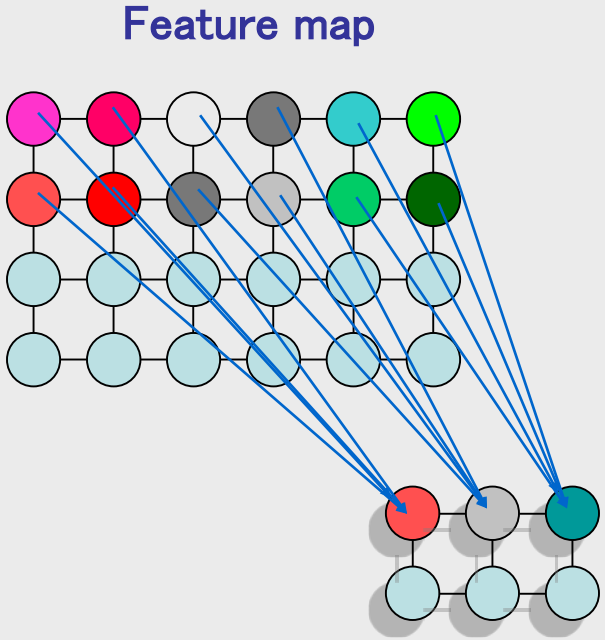
Subsampling layer

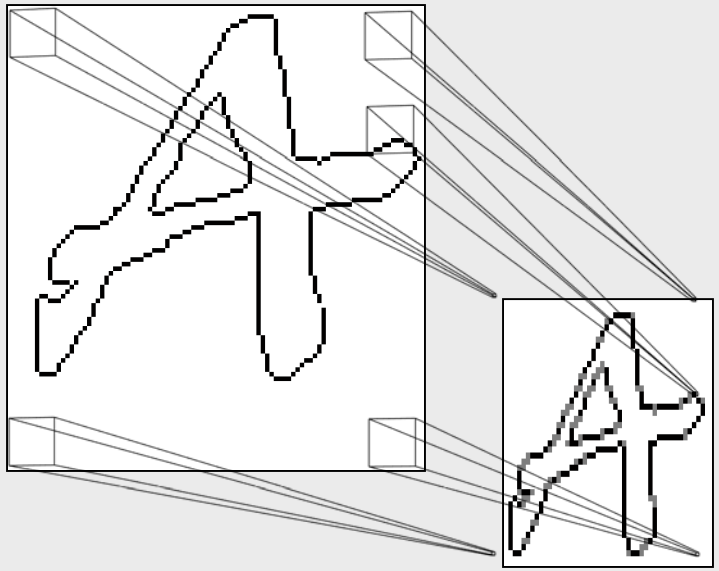
- The weight sharing is also applied in subsampling layers.



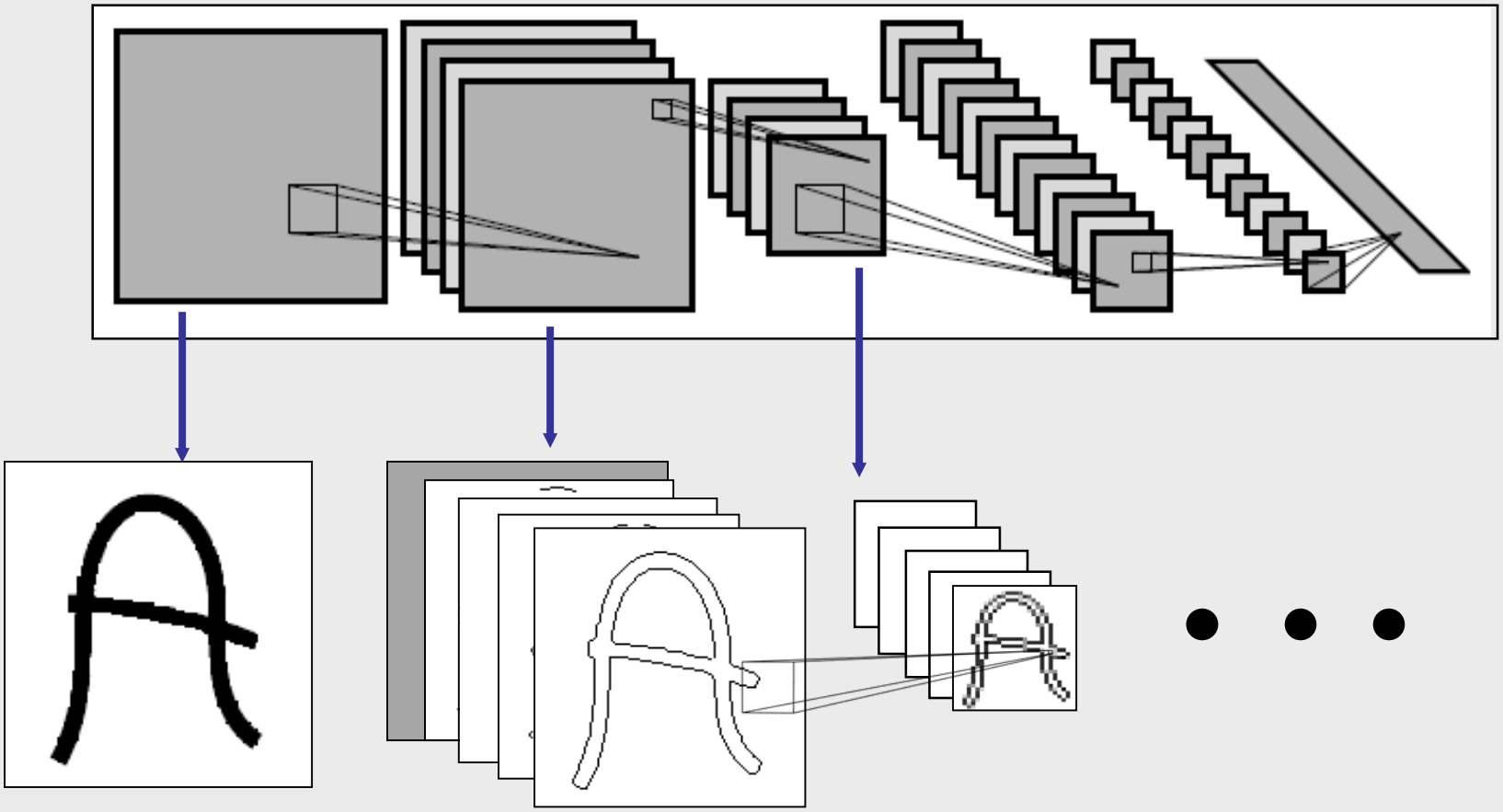
Subsampling layer

- the weight sharing is also applied in subsampling layers
- reduce the effect of noises and shift or distortion





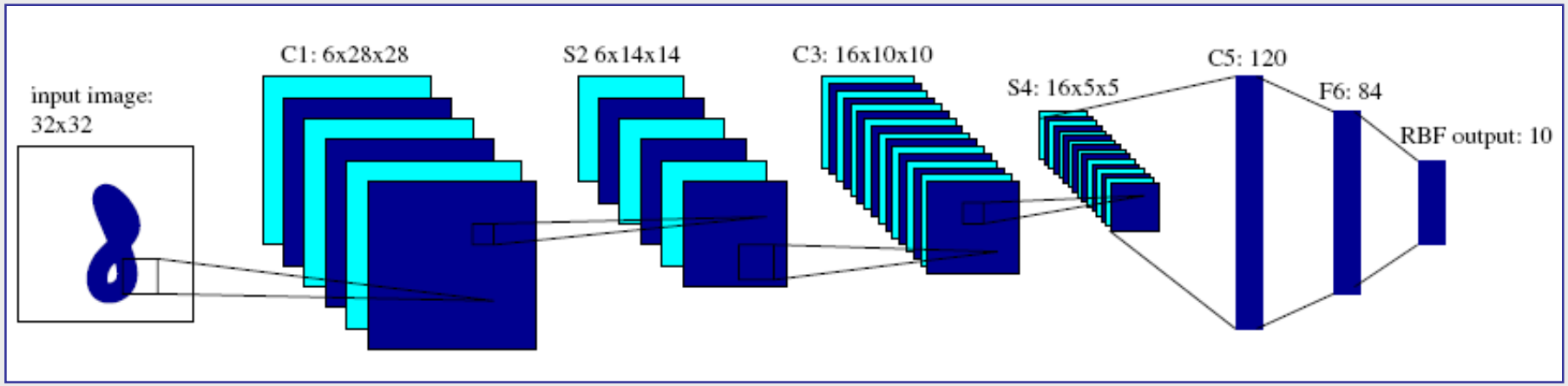
Up to now ...



LeNet 5

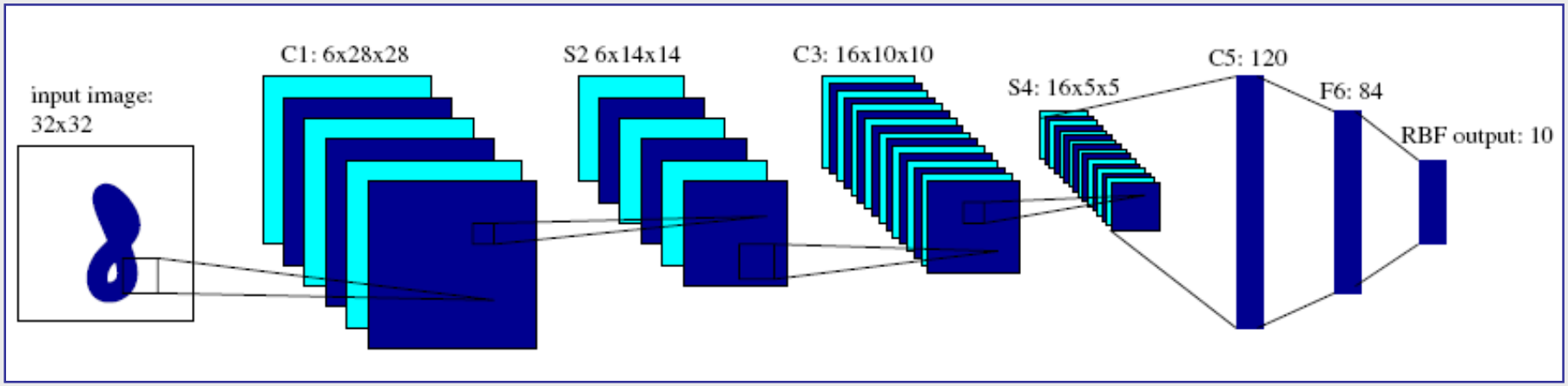
LeNet5

- Introduced by LeCun.
- raw image of 32 × 32 pixels as input.



LeNet5

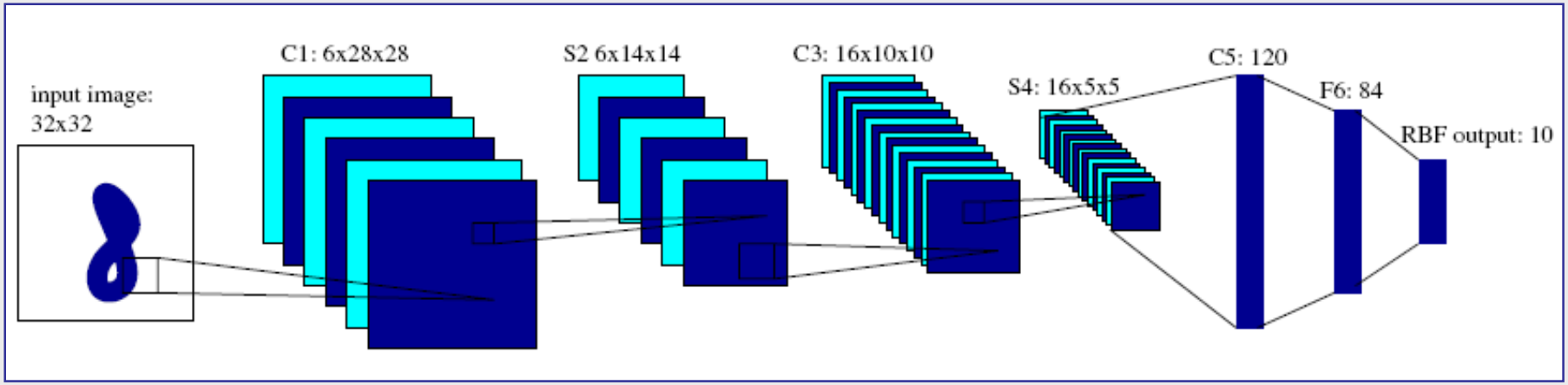
- C1,C3,C5 : Convolutional layer.
- 5 × 5 Convolution matrix.
- S2 , S4 : Subsampling layer.
- Subsampling by factor 2.
- F6 : Fully connected layer.



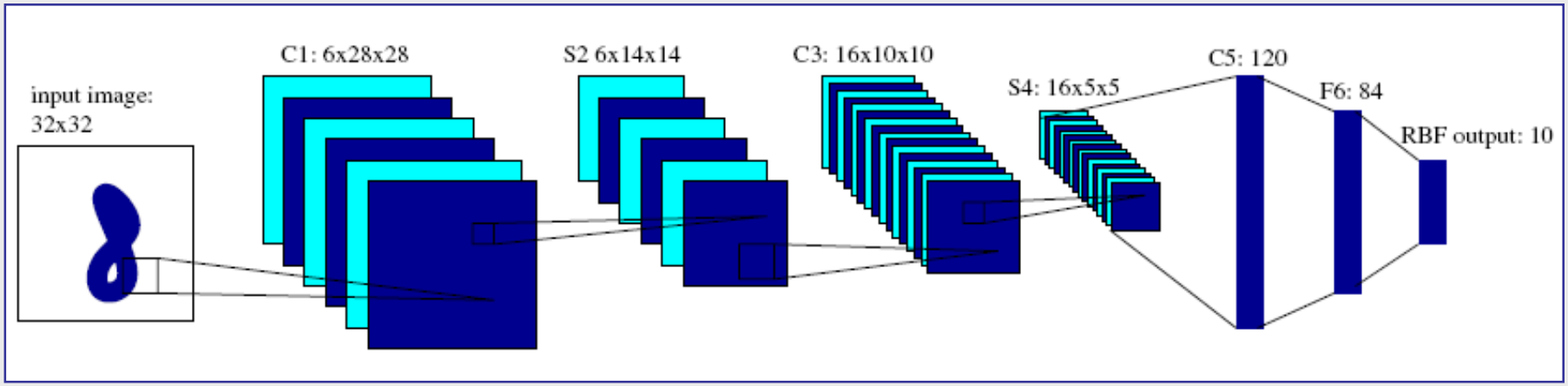
LeNet5

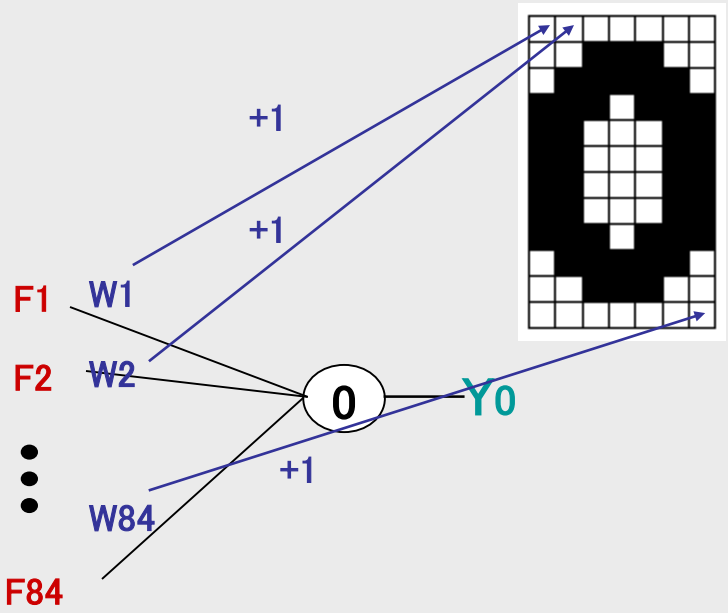
- All the units of the layers up to F6 have a sigmoidal activation function of the type:

$$y_j = \varphi(v_j) = A\tanh(Sv_j)$$



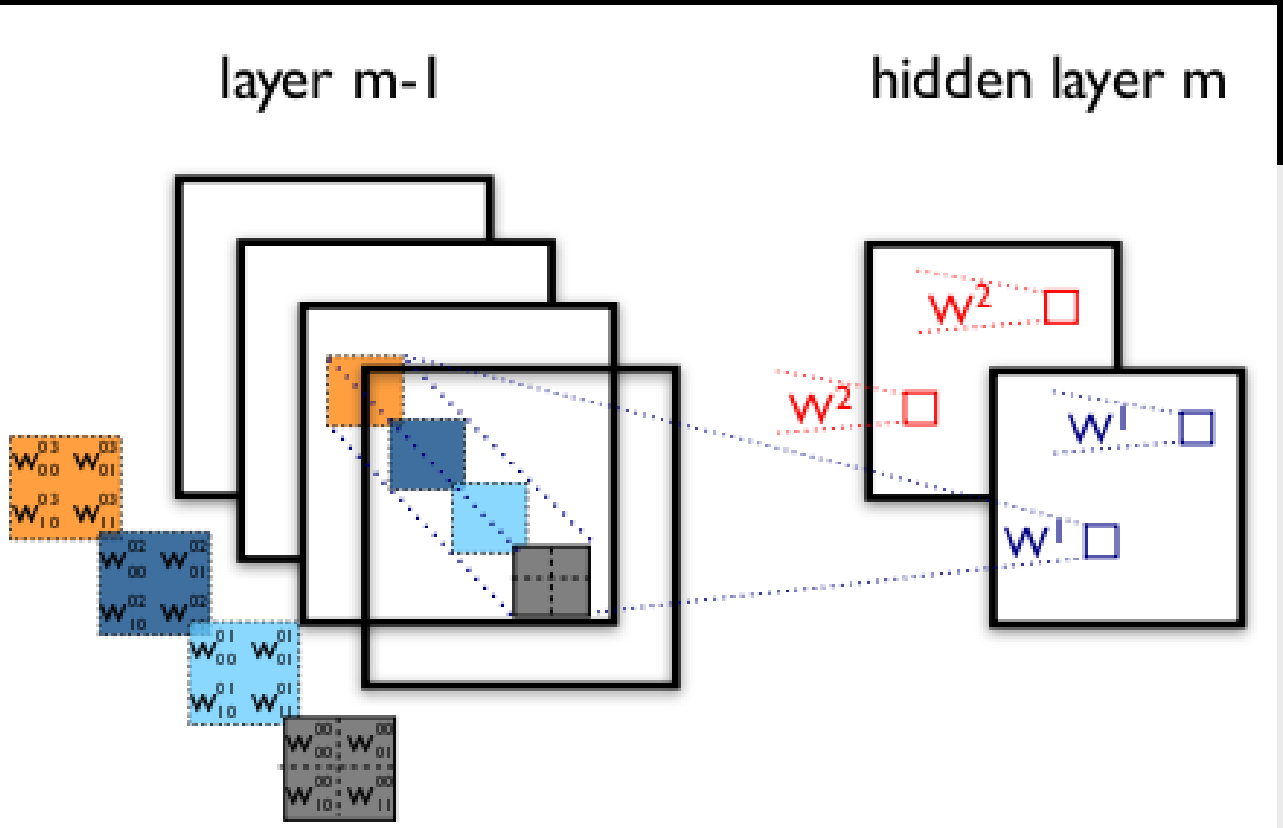
LeNet5

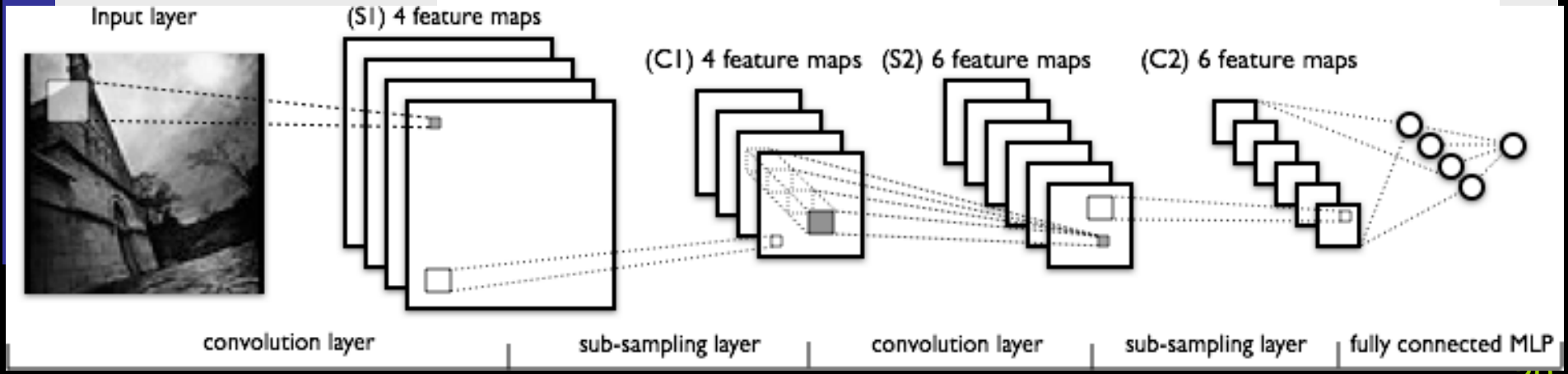




$$Y_j = \sum_{i=1}^{84} (F_i - W_{ij})^2, \; j = 0,\ldots,9$$

Summary


layer m-1
hidden layer m
w^2
w^2
w^1
w^1


Input layer
(S1) 4 feature maps
(C1) 4 feature maps
(S2) 6 feature maps
(C2) 6 feature maps
convolution layer
sub-sampling layer
convolution layer
sub-sampling layer
fully connected MLP