Object recognition 4

Machine learning

Denis Pelli

Psychology and Neural Science, NYU

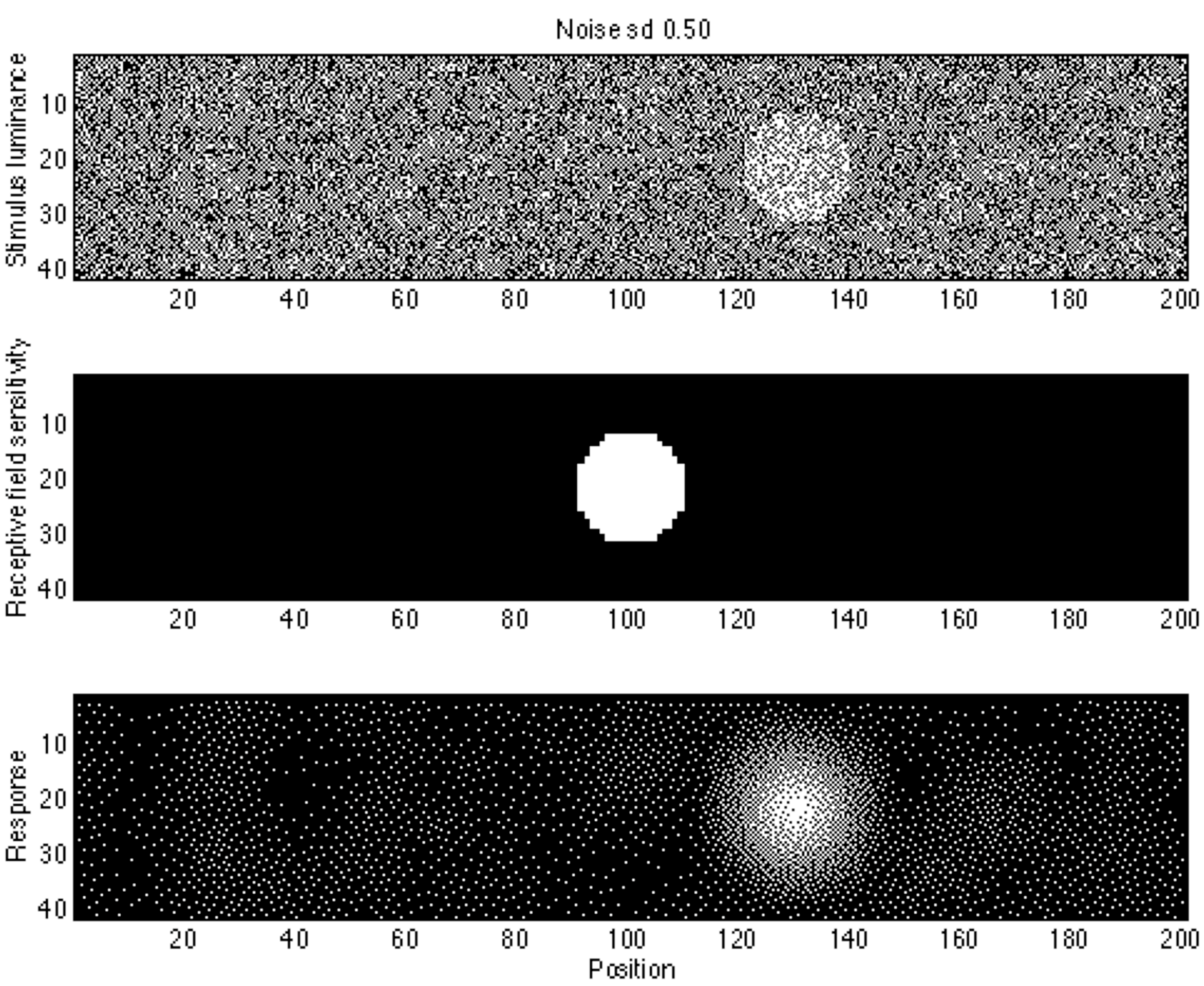
December 12, 2007

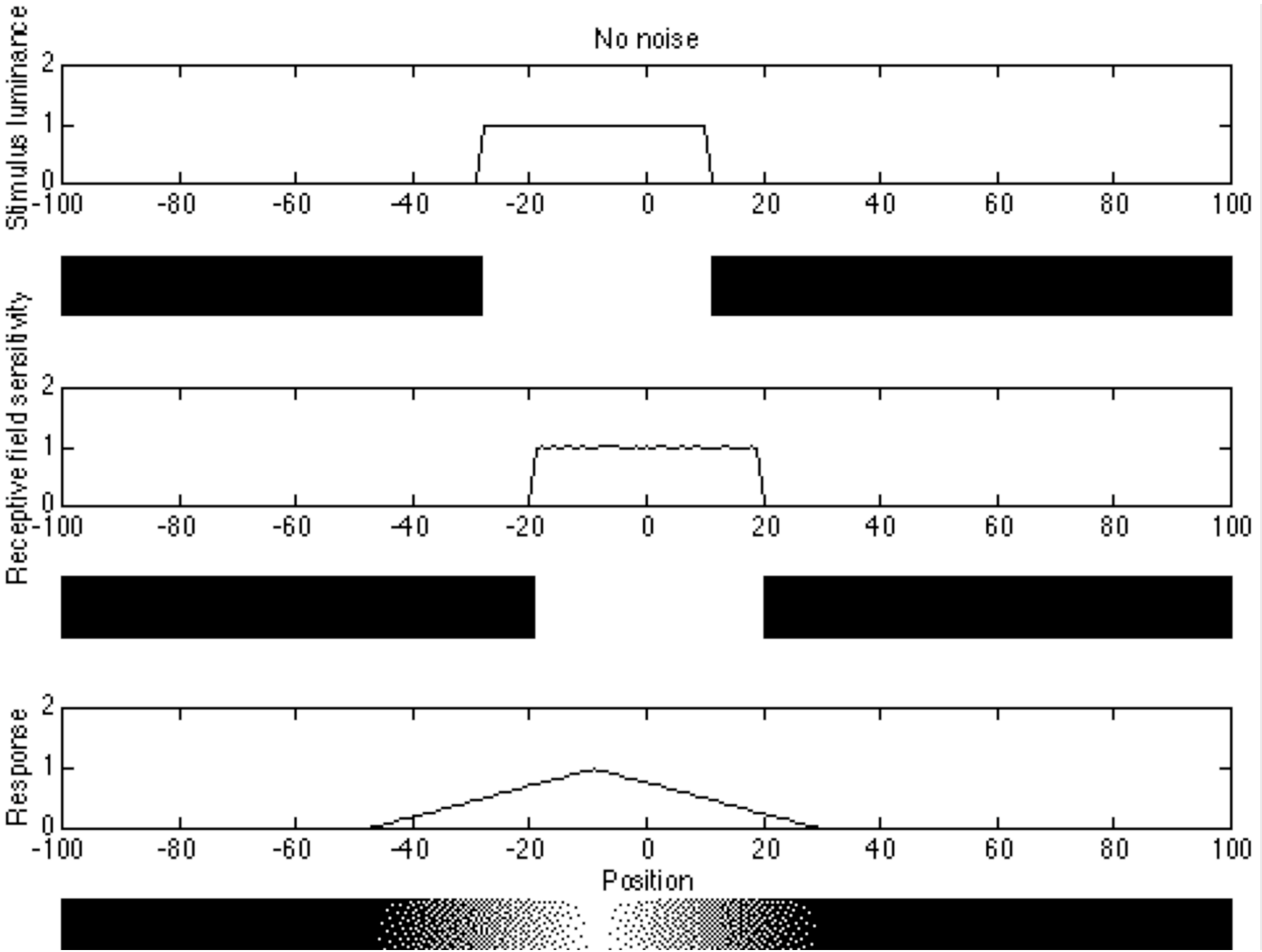
Outline

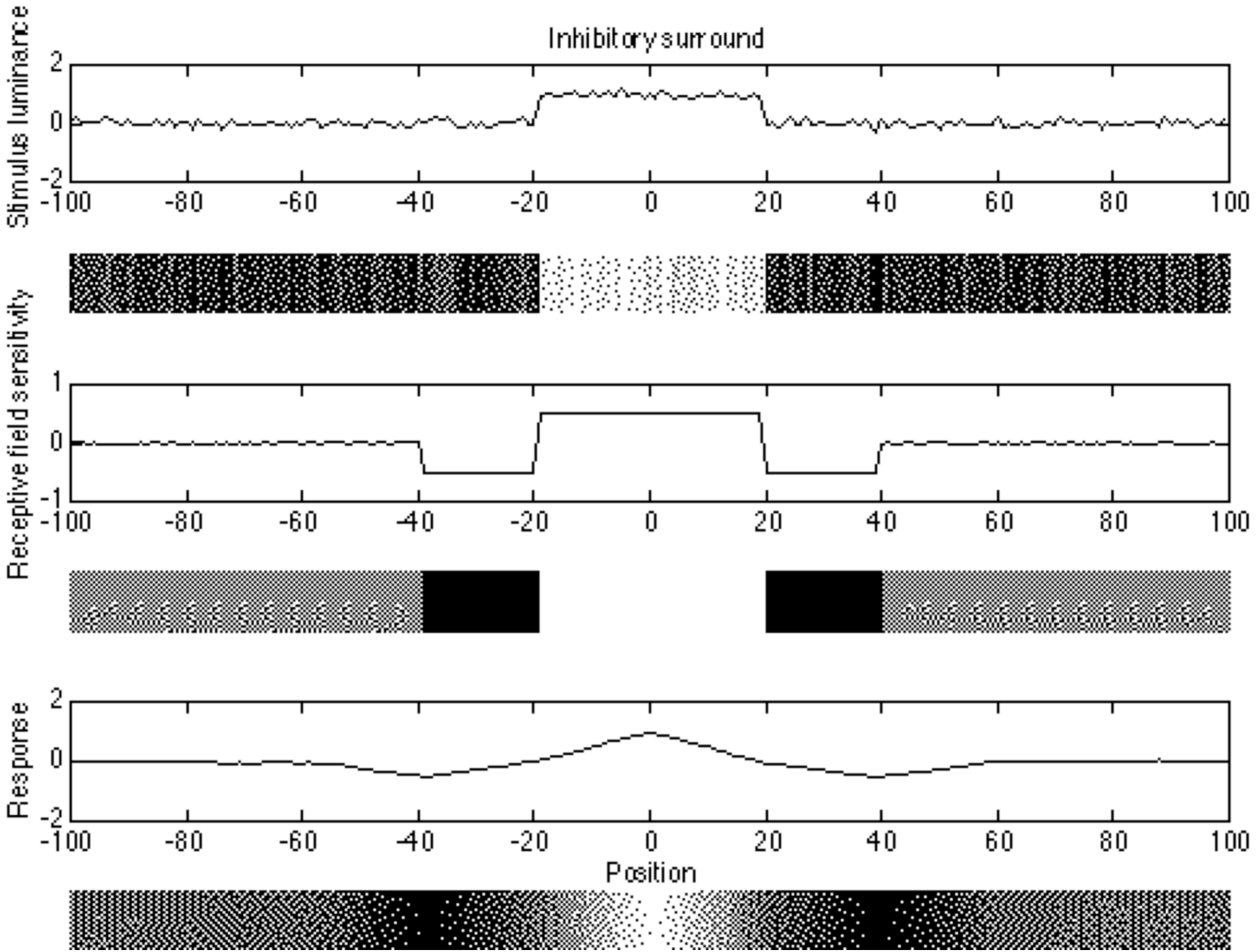
1. Receptive field
   Summing over space.
2. Machine learning
   Learning receptive fields (i.e. features) from the raw images.

Receptive field as a computational model.

$$r = w_1 s_1 + w_2 s_2 + \ldots + w_{10000} s_{10000} \qquad (1)$$


Noise sd 0.50
Stimulus luminance
Receptive field sensitivity
Response
Position
10
20
30
40
20
40
60
80
100
120
140
160
180
200


No noise
Stimulus luminance
Receptive field sensitivity
Response
Position
2
1
0
-100
-80
-60
-40
-20
0
20
40
60
80
100


Inhibitory surround
Stimulus luminance
2
0
-2
-100
-80
-60
-40
-20
0
20
40
60
80
100
Receptive field sensitivity
1
0
-1
Response
Position

Ranzato, M., Huang, F.-J., Boureau, Y.-L., & LeCun, Y. (2007). Unsupervised learning of invariant feature hierarchies with applications to object recognition. *Proc. Computer Vision and Pattern Recognition Conference (CVPR 2007)*
http://www.cs.nyu.edu/~ranzato/publications/ranzato-cvpr07.pdf

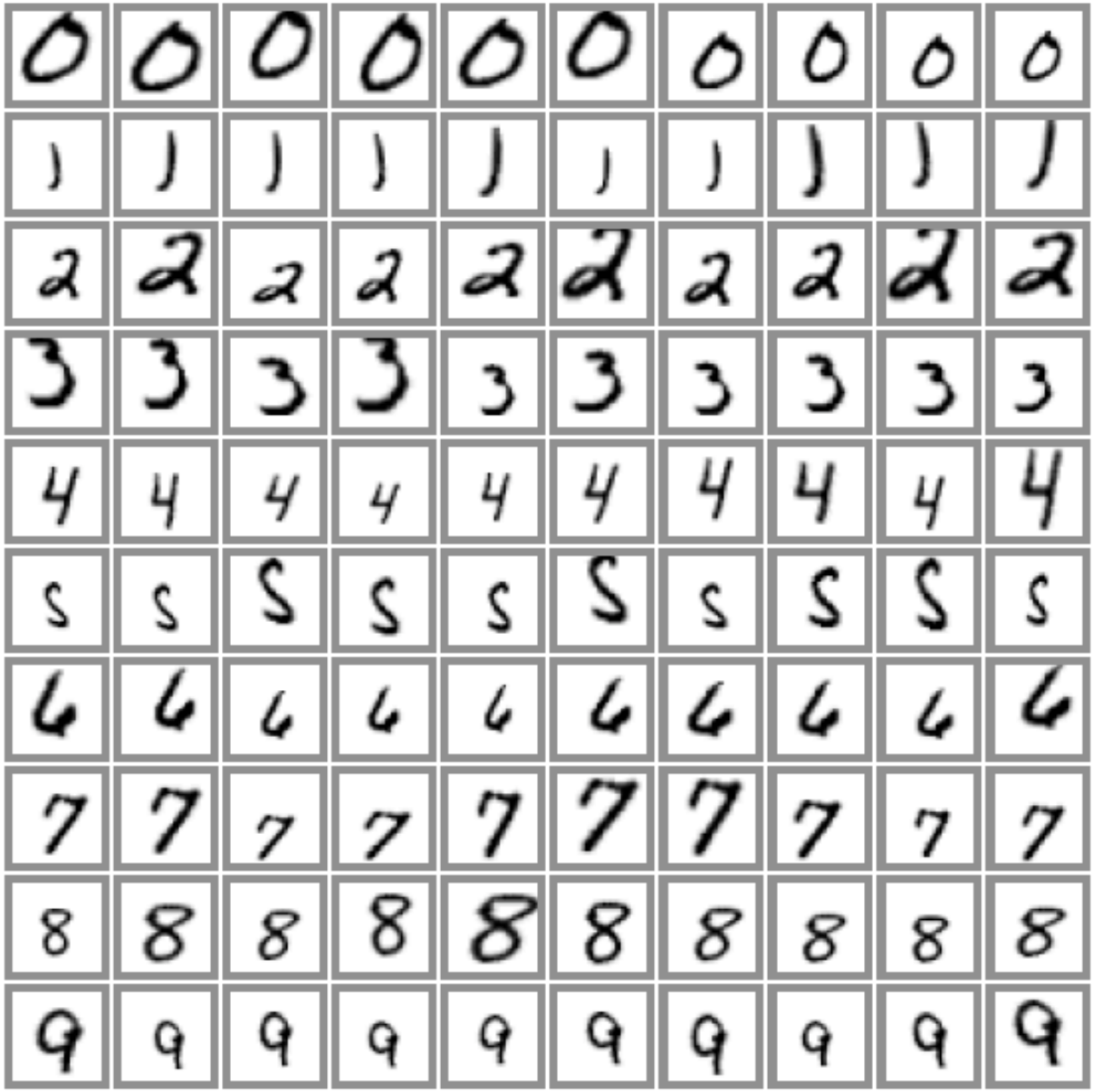


Fig. 7. Examples of distortions of ten training patterns.

LeCun, Bottou, Bengio, & Haffner (1998) *IEEE*

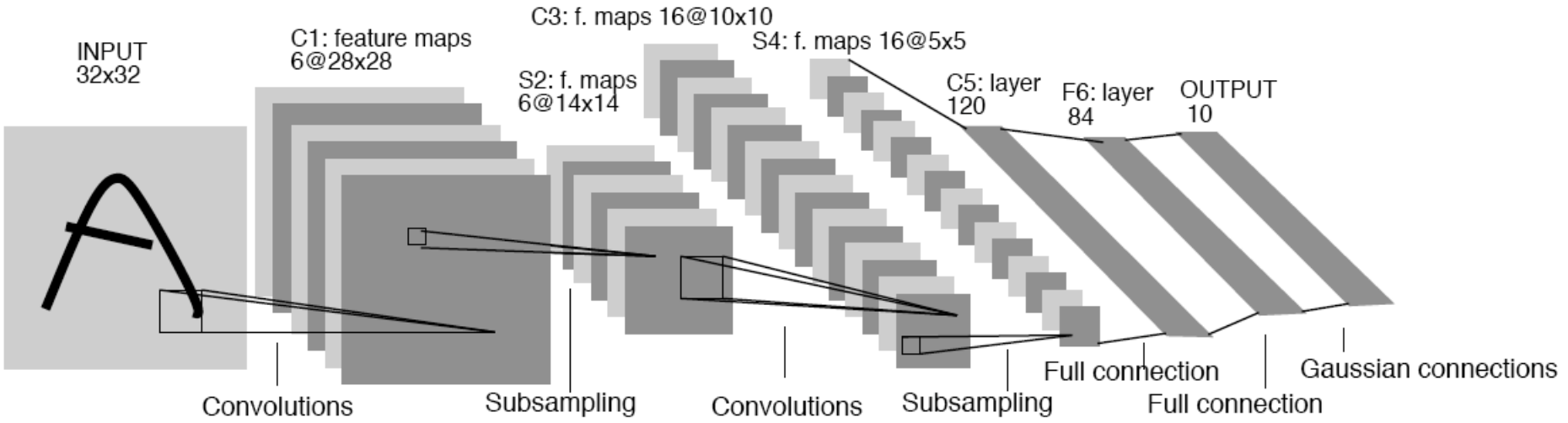
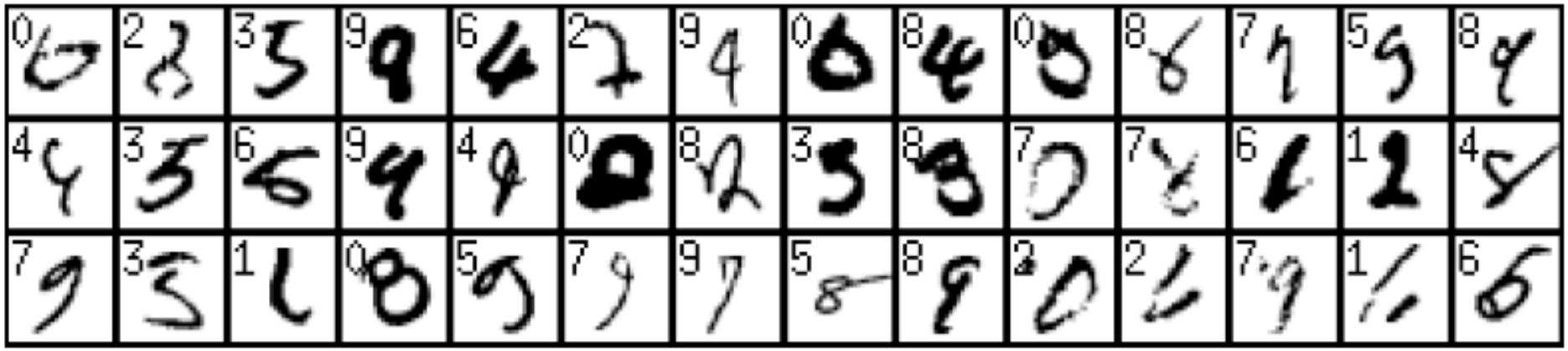


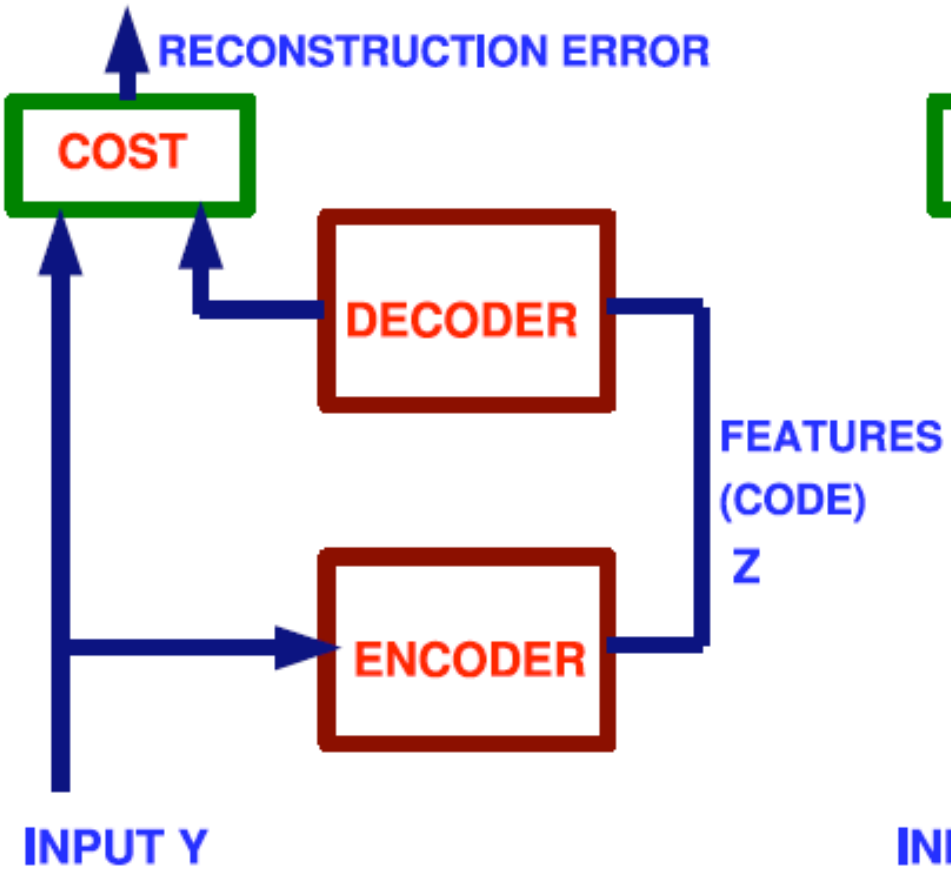
Fig. 2. Architecture of LeNet-5, a Convolutional Neural Network, here for digits recognition. Each plane is a feature map, i.e. a set of units whose weights are constrained to be identical.

LeCun, Bottou, Bengio, & Haffner (1998) *IEEE*

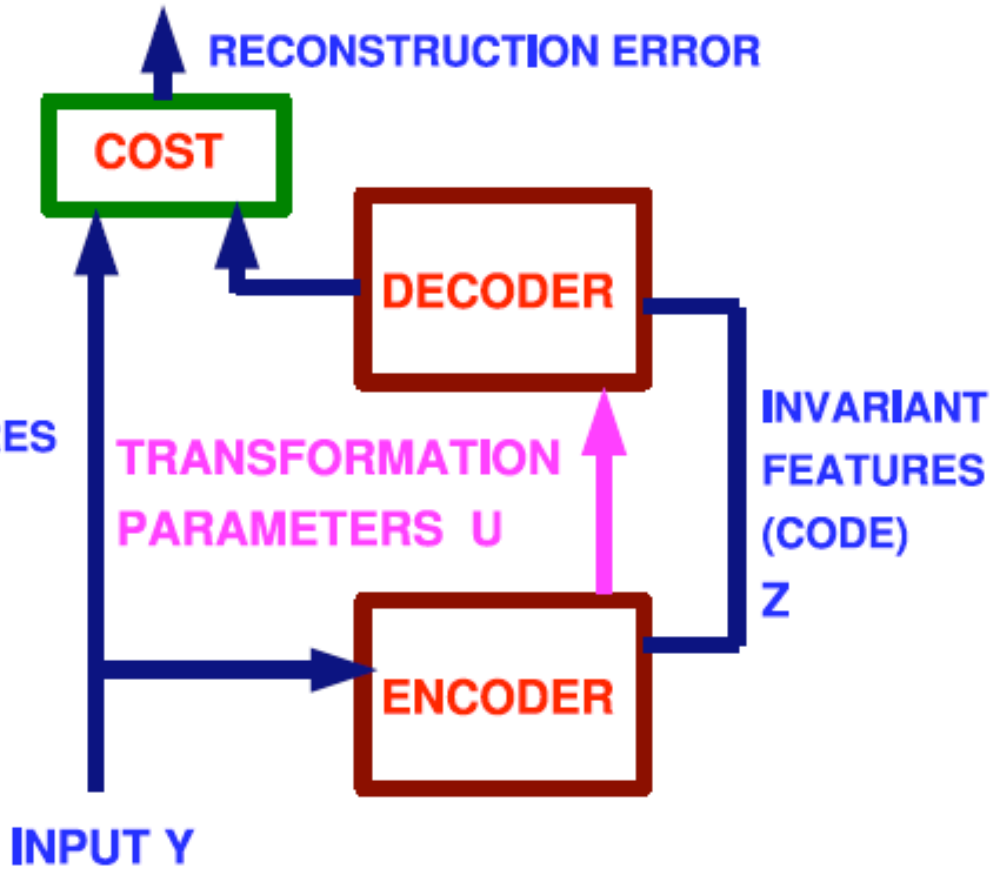


The 42 misclassified digits in the testing MNIST dataset. On the upper left corner there is the true label.

Ranzato & LeCun (2007) *ICDAR*

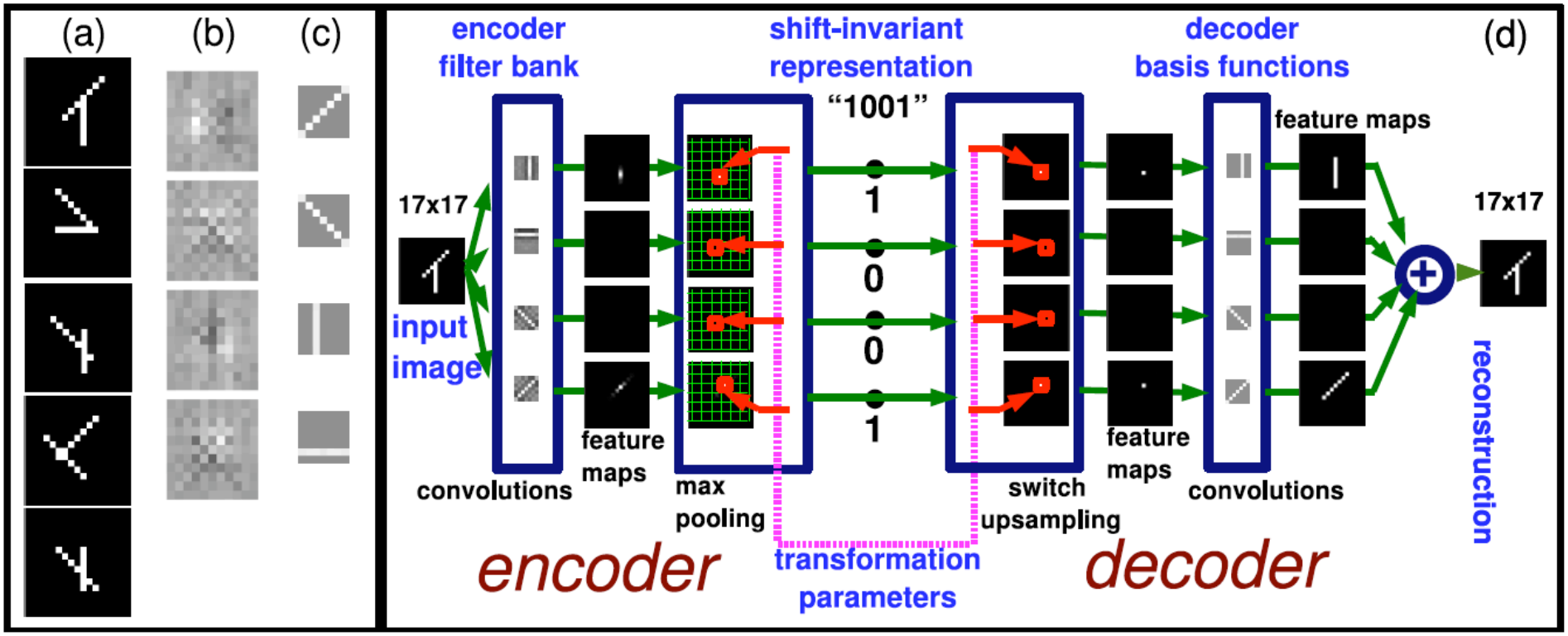


Standard Feature Extractor



Invariant Feature Extractor

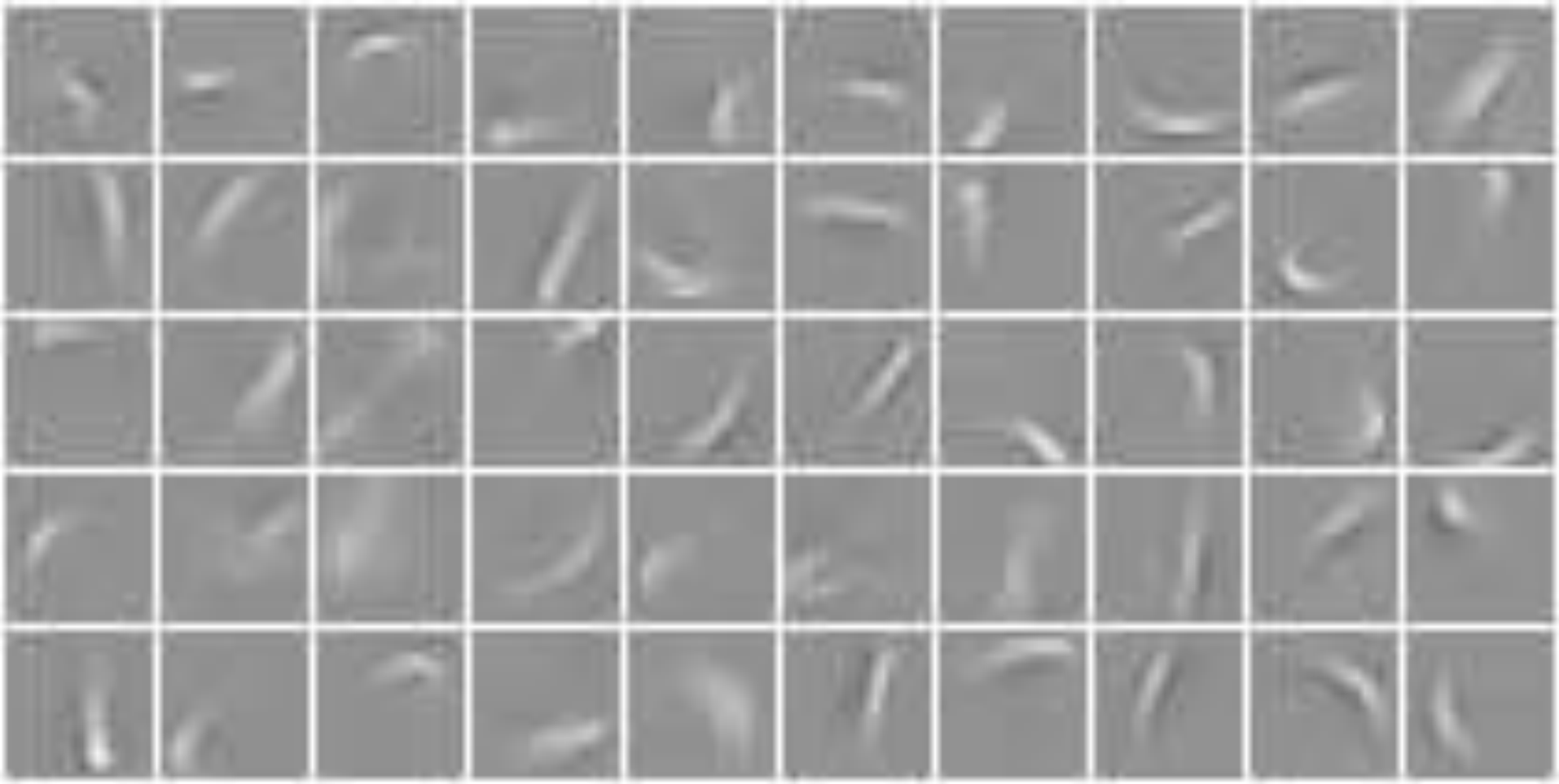
Ranzato, Huang, Boureau, & LeCun (2007) *CVPR*



Ranzato, Huang, Boureau, & LeCun (2007) *CVPR*

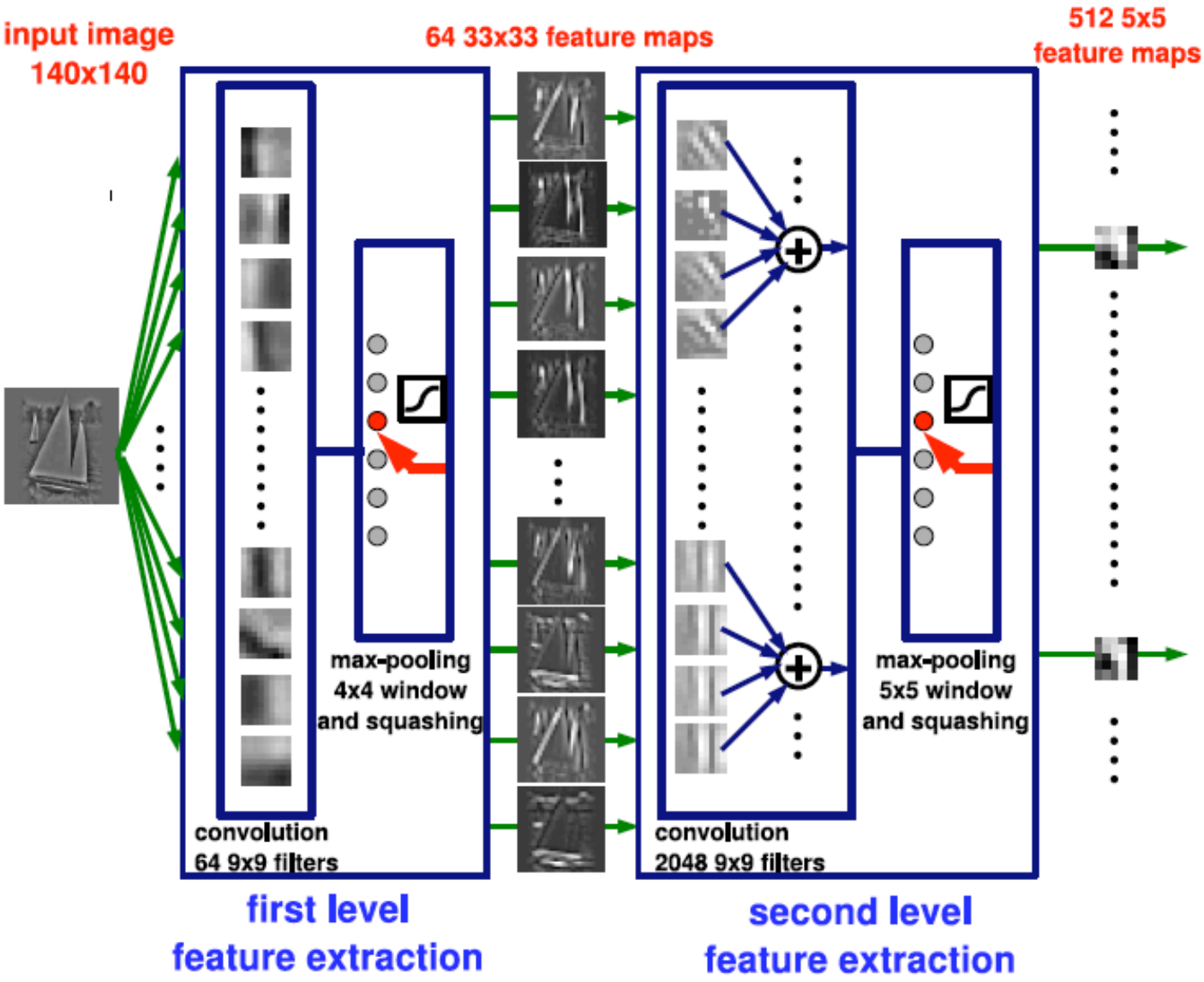
Fifty 7×7 sparse shift-invariant features learned by the unsupervised learning algorithm on the MNIST dataset. These filters are used in the first convolutional layer of the feature extractor.

Ranzato, Huang, Boureau, & LeCun (2007) *CVPR*



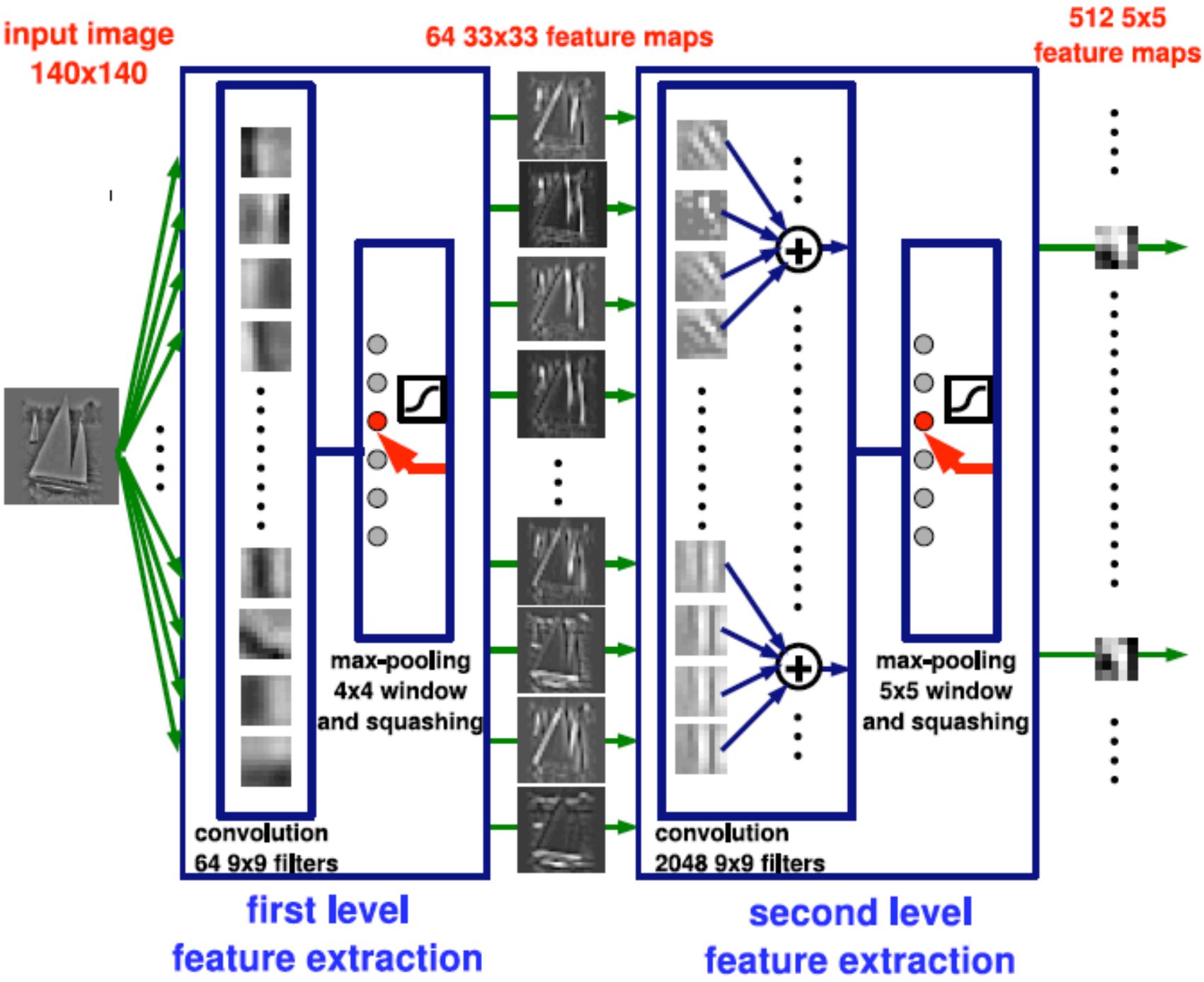
Fifty 20x20 filters learned in the decoder by the sparse and shift invariant learning algorithm after training on the MNIST dataset of 28x28 digits. A digit is reconstructed as linear combination of a small subset of these features positioned at one of 81 possible locations (9x9).

Ranzato, Huang, Boureau, & LeCun (2007) *CVPR*

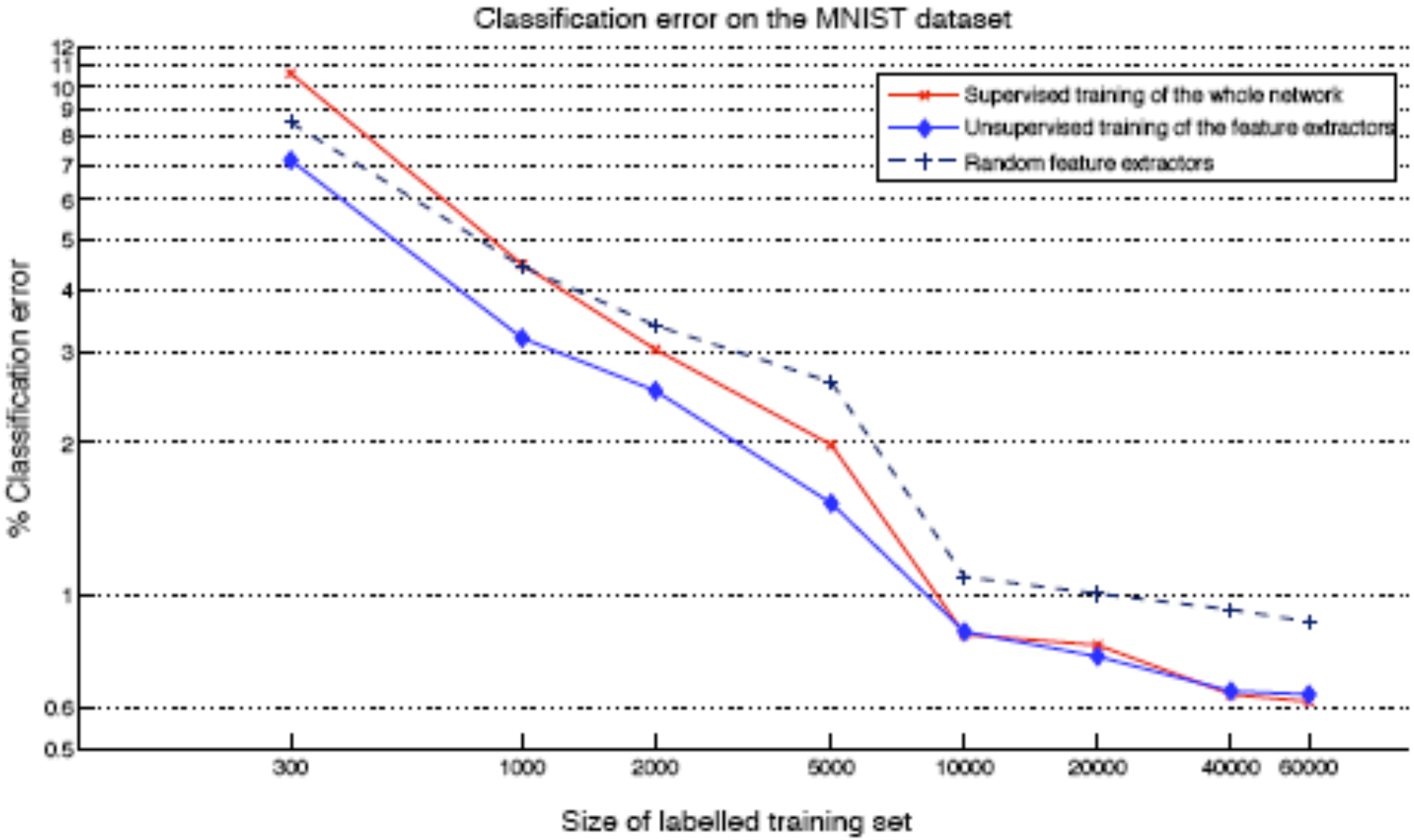


Example of the computational steps involved in the generation of two 5×5 shift-invariant feature maps from a preprocessed image in the Caltech101 dataset.

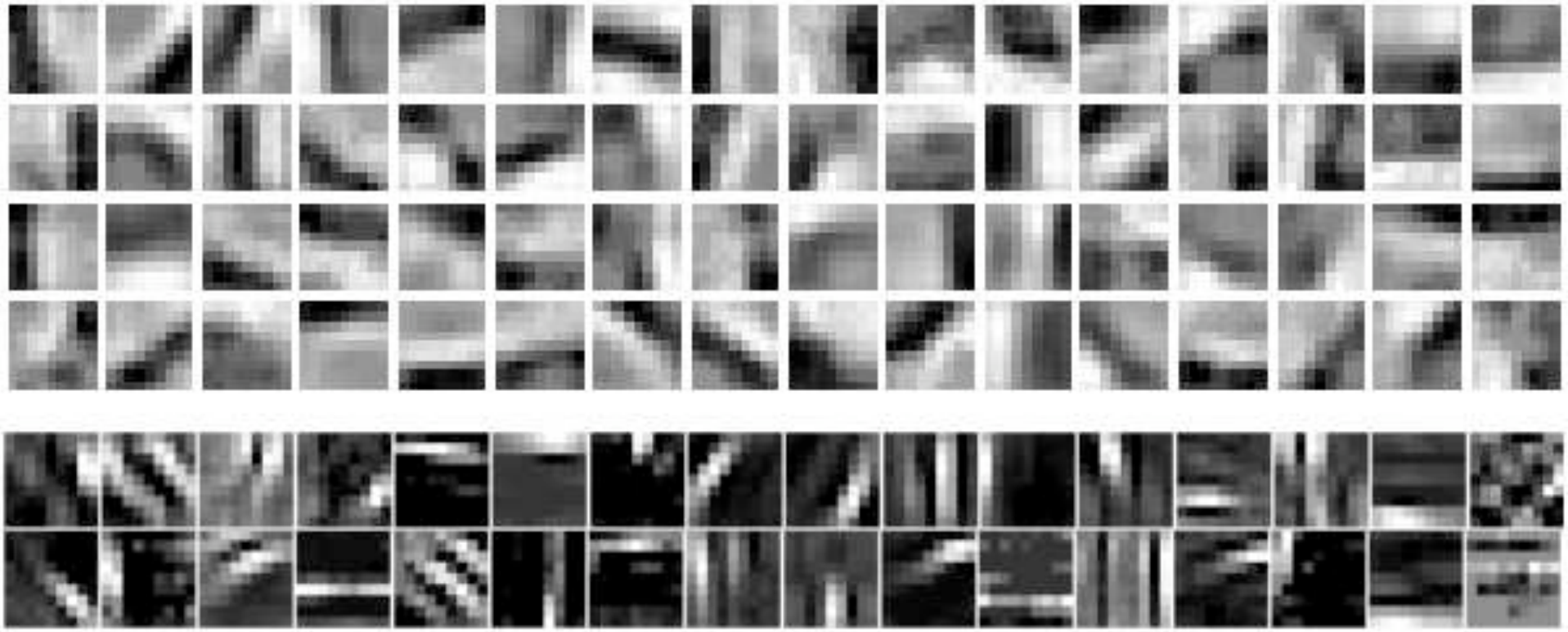
Ranzato, Huang, Boureau, & LeCun (2007) *CVPR*



Ranzato, Huang, Boureau, & LeCun (2007) *CVPR*



Ranzato, Huang, Boureau, & LeCun (2007) *CVPR*



Caltech 101 feature extraction. **Top Panel**: the 64 convolutional filters of size 9×9 learned by the first stage of the invariant feature extraction. **Bottom Panel**: a selection of 32 (out of 2048) randomly chosen filters learned in the second stage of invariant feature extraction.

Ranzato, Huang, Boureau, & LeCun (2007) *CVPR*

Outline

1. Receptive field
   Summing over space.
2. Machine learning
   Learning receptive fields (i.e. features) from the raw images.