MLP
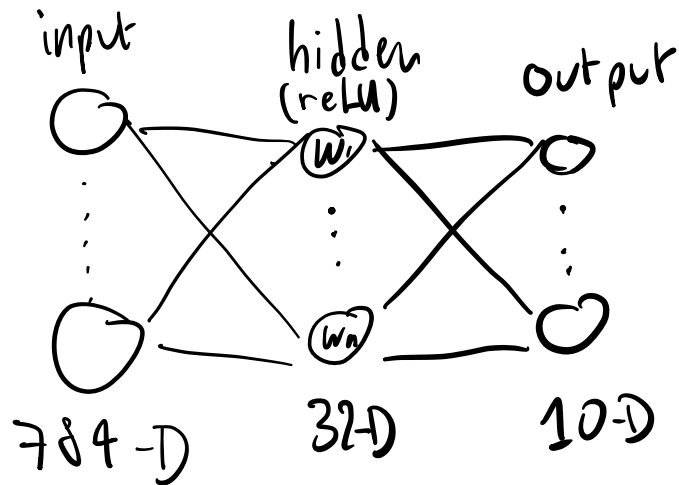

Number of parameters:

Input - hidden : $784 \cdot 32 = 25088$

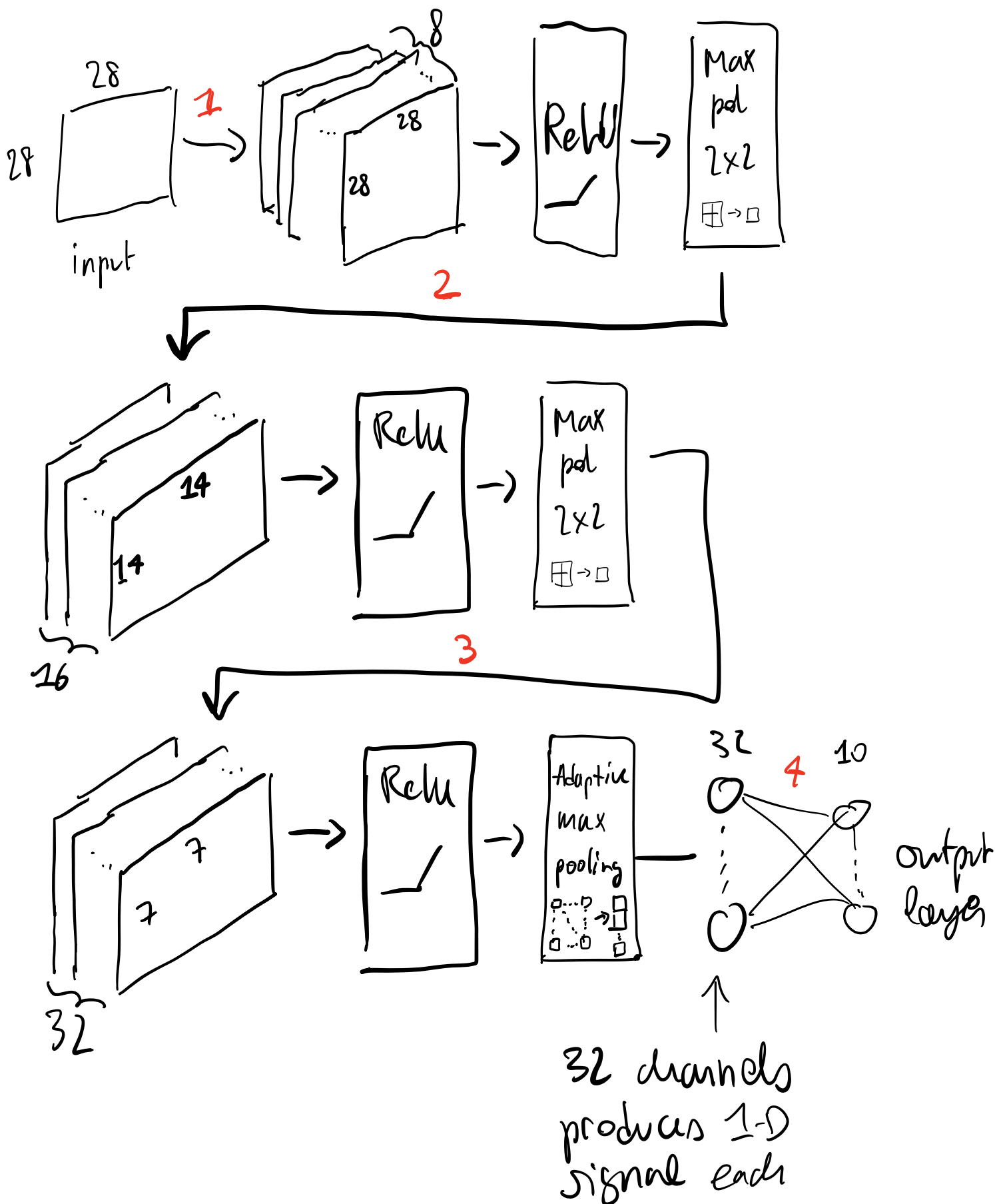1st layer Bias : $32 = 32$

hidden - output : $32 \cdot 10 = 320$

output Bias : 10

total : 25450

CONVOLUTIONAL NETWORK

28
28
input

1

8
28
28

ReLU → Max pool 2x2 ⊞→□

2

14
14
16

ReLU → Max pool 2x2 ⊞→□

3

7
7
32

ReLU → Adaptive max pooling

4

32   10

output layer

32 channels produces 1-D signal each

#parameters

for a simple convolutional layer

$$\boxed{\# \text{Weight} = C_{out} \times (K \times K \times C_{in} + 1)}$$

$\uparrow$ Bias

1: $\quad 8 \cdot (3 \cdot 3 \cdot 1 + 1) = 80$

2: $\quad 16 \cdot (3 \cdot 3 \cdot 8 + 1) = 1168$

3: $\quad 32 \cdot (3 \cdot 3 \cdot 16 + 1) = 4640$

4: $\quad 32 \cdot 10 \text{ (fully connected) } + 10 \text{ (Bias)}$

$\qquad = 330$

$$\boxed{\text{total } \# \text{ parameters} = 6218}$$

the number of parameters depends on how the layers are linked. each next layer channel will be connected to all precedent (input) channels. Each channels is represented by a kernel and so each element of each channel 's kernel must be connected to each elements of each elements of each input channel's level result plus 1 bias for each current channel

as each   curnt filter (one for each
channel) must   agnostic of patern
tran slations.

$\Rightarrow$ $C_{out} \times (K \cdot K \cdot C_{in} + 1)$