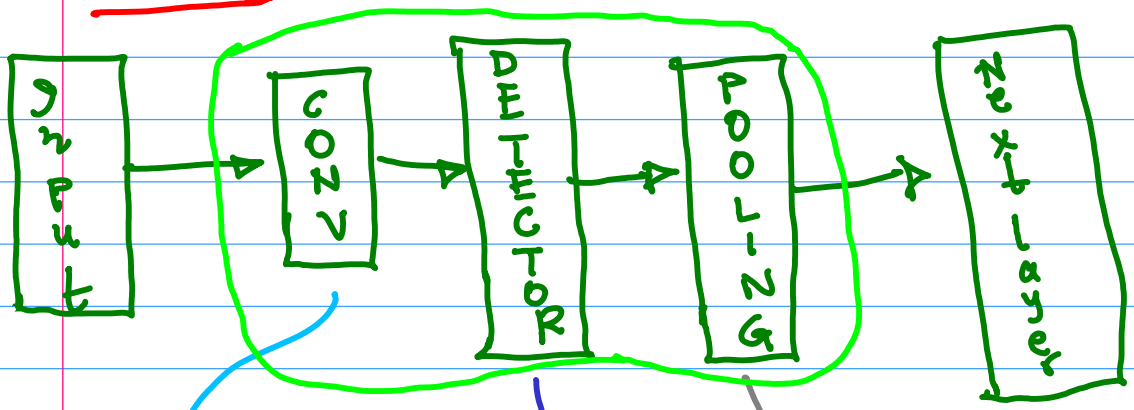
(4) Pooling

"one convolutional layer"



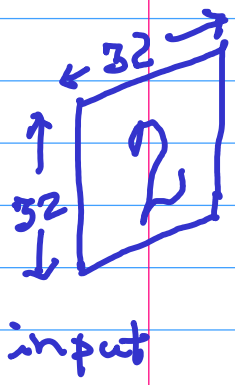
convolutional stage with linear activations

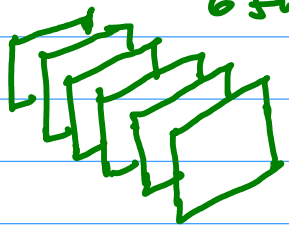
non-linearity: activation f'ns e.g., ReLU

to modify the output of this layer, further (some form of subsampling — MaxPooling / Average pooling (historical)

LeNet-5 (1989) LeCun, Bottou, Bengio, Haffner

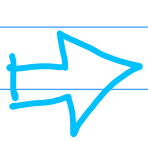
(handwritten & machine printed digit recognition)

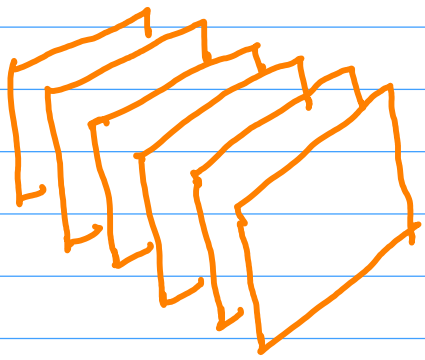




6 filters/ kernels/ masks, each of size 5x5

Use the 'perfect fit' model of convolution





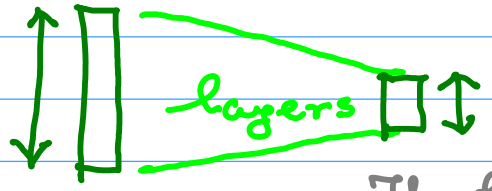
each is a 28x28 image

Important philosophy: this is one part of a 'contractive structure' (the other being pooling)

Why contractive?

large-sized input

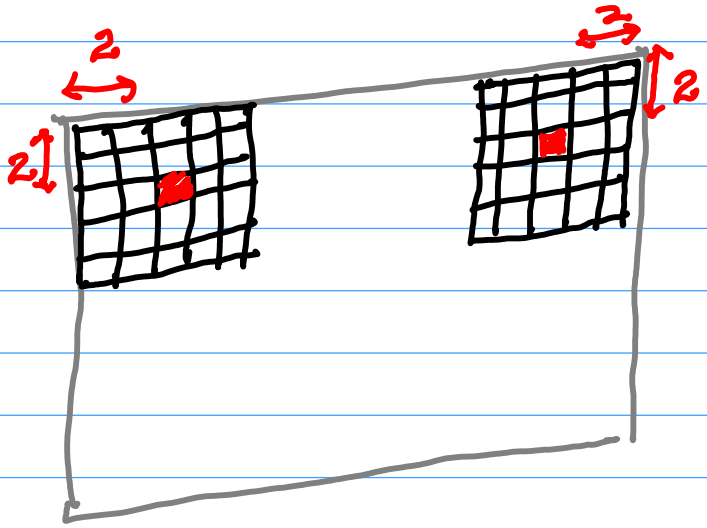
e.g., 32×32 image = 1024×1 input

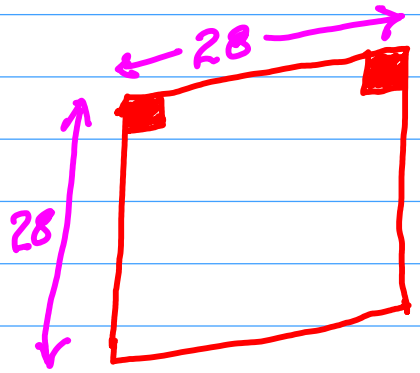


The final classifier is a 10-class classifier

The 'perfect fit' convolution, as opposed to the 'full' convolution: → A−B+1

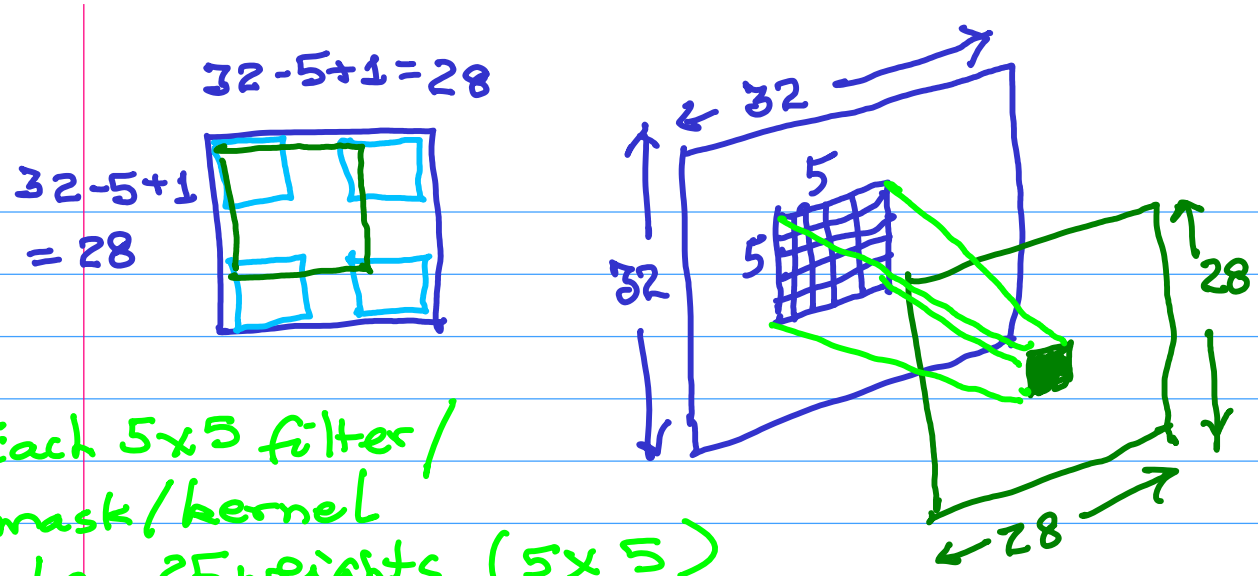
→ A+B−1



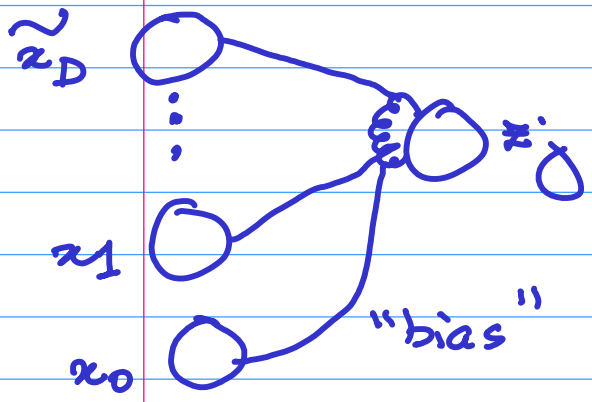


why 28? 32−2−2 = 28 (32−5+1)

The resultant of the 'perfect fit' convolution is of a smaller size.



Each 5x5 filter/ mask/kernel has 25 weights (5x5) + 1 bias term (also counted as a connection)



$$a_j = \sum_{i=1}^{\tilde{x}_D} w_{ji} x_i + w_{j0} \text{ or } b_j$$

25 weights + bias

$\tilde{x}_D$ is a small fraction of the total x_D

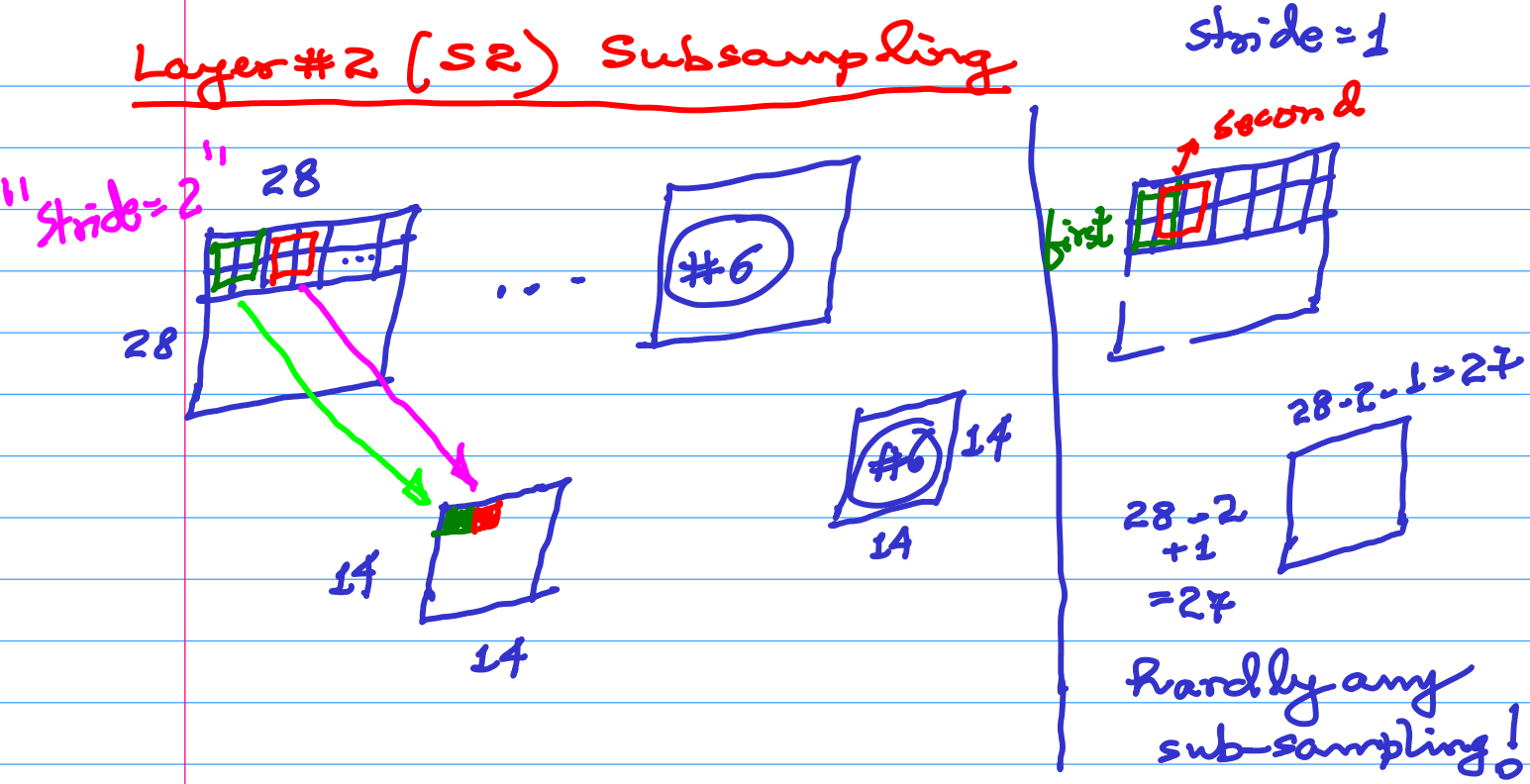
parameter: in each of the repeating structures

One for each of the 6 filters

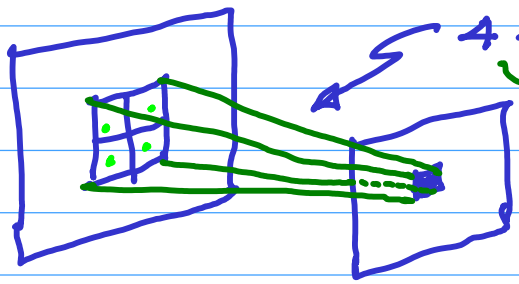
parameters = $6 \times (5 \times 5 + 1) = 6 \times 26 = 156$

connections: $6 \times (28 \times 28) \times (5 \times 5 + 1) = 122,304$

Layer #2 (S2) Subsampling



parameters & connections?



Parameters $6 \times (1 + 1\text{(bias)}) = 12$ parameters

Connections $6 \times (14 \times 14) \times (2 \times 2 + 1)$

$= 6 \times 14 \times 14 \times 5 = 5,880$ connections