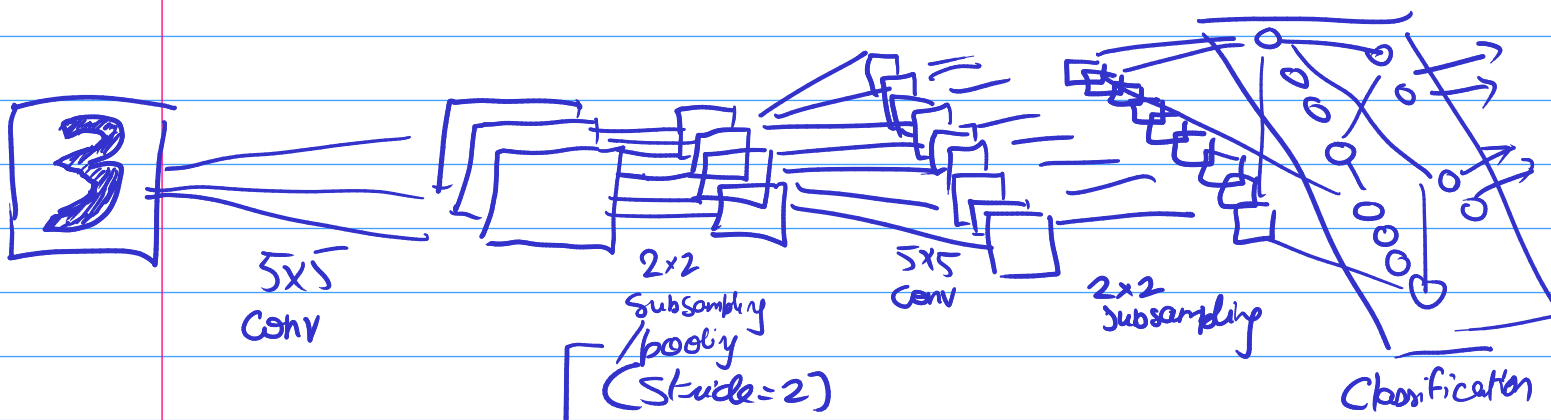

pooling (makes model less sensitive to changes such as shifting)

since you are taking max/avg of the values in the window, it doesn't matter where the window is.

Convolution → Strong prior? Low entropy

weights of one hidden unit → close to the neighbor, identical but shifted

Convolutional net is similar to a fully connected net but with an infinitely strong prior over its weights

- It says that the weights for one hidden unit must be identical to the weights of its neighbor, but shifted in space
- Prior also says that the weights must be zero, except for in the small spatially contiguous receptive field assigned to that hidden unit

Convolution with a kernel of width 3 s_3 is a hidden unit. It has 3 weights which are the same as for s_4

- Convolution introduces an infinitely strong prior probability distribution over the parameters of a layer

low variance

⇒ local interactions

$$y = w_0 + w_1 x + w_2 x^2 + w_3 x^3$$

(w_0, w_1, w_2, w_3)

Pooling also has strong prior

N convolutions → 1 pooling

(needn't be after 1 conv 1 pooling)

wed did till 9-4 : 9-10 → Neuroscientific basis (read when free)