

Convolutional Neural Network

Model

- In addition to deep layers, a CNN has convolutional layers where the weights in a kernel are learnt by the model
- This is justified by the way in which (1) human eyes look for specific patterns, and (2) similar features occur across an image
- A CNN layer has many stages
 - Input stage

$$x_{i,j}$$

- Convolutional stage (rejecting edge cases means we limit the range of i and j)

$$a_{i,j} = \sum_{k,l} w_{k,l} x_{i-k,j-l}$$

- Non-linear stage (the non-linear function f could be ReLu $f(a) = [a]_+$ or sigmoid $f(a) = \sigma(a)$)

$$y_{i,j} = f(a_{i,j})$$

- Pooling stage (mean also used)

$$y_{i,j} = \max_{|k| < \tau, |l| < \tau} (y_{i-k,j-l})$$

- From our knowledge of convolutions, certain kernels can perform numerical differentiation in different directions
- Numerical differentiation is edge detection
- The convolution function is the same as cross-correlation (only with the sign reversed), so we are measuring the similarity between the kernel and a region of pixels in the image – this allows us to do blob detection as well