Prof. Dr. Leif Kobbelt
Tim Elsner, Gregor Kobsik

Theoretical Exercise 8

Date Published: June 21st 2023

This is the sample solution.

Exercise 1 Featuremap Sizes

Assume that the input image to a CNN has width and height dimensions of 224×224. Compute the feature map sizes (width and height) after the image passes through a first convolution layer (with kernel size 7, stride 1, padding 0), a second convolution layer (with kernel size 3, stride 2, padding 1), and finally a max pooling layer (with kernel size 2).

Solution

We can use the following formular to compute the feature map size after the convolution layers

$$\left\lfloor \frac{W - F + 2P}{S} \right\rfloor + 1$$

, where W is the size of the input, F the kernel size, P the padding, and S the stride. Here we assume that $W \geq F$. For max pooling we can use the same formular (where we typically assume $F = S$)

- first convolution layer: $\left\lfloor \frac{224-7}{2} \right\rfloor + 1 = 218$
- second convolution layer: $\left\lfloor \frac{218-3+2}{2} \right\rfloor + 1 = 109$
- max pooling layer: $\left\lfloor \frac{109-2}{2} \right\rfloor + 1 = 54$

Exercise 2 Number of Parameters of a Neural Network

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 **LRN**	conv3-64 **conv3-64**	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 **conv3-128**	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 **conv1-256**	conv3-256 conv3-256 **conv3-256**	conv3-256 conv3-256 conv3-256 **conv3-256**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Compute the number of parameters in configuration A. Assume that the convolution layers use stride 1 and do not reduce the width and height of the input featuremaps.

Solution

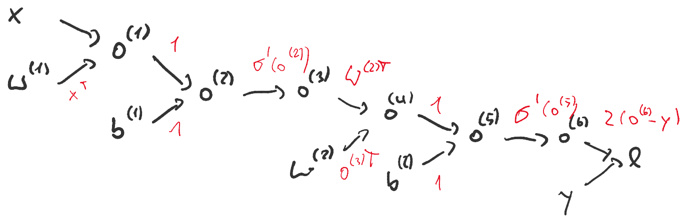
- $3 \cdot 64 \cdot 3 \cdot 3 + 64 = 1792$
- $64 \cdot 128 \cdot 3 \cdot 3 + 128 = 73856$
- $128 \cdot 256 \cdot 3 \cdot 3 + 256 = 295168$
- $256 \cdot 256 \cdot 3 \cdot 3 + 256 = 590080$
- $256 \cdot 512 \cdot 3 \cdot 3 + 512 = 1180160$
- $512 \cdot 512 \cdot 3 \cdot 3 + 512 = 2359808$
- $512 \cdot 512 \cdot 3 \cdot 3 + 512 = 2359808$
- $512 \cdot 512 \cdot 3 \cdot 3 + 512 = 2359808$
- $512 \cdot 7 \cdot 4096 + 4096 = 14684160$
- $4096 \cdot 4096 + 4096 = 16781312$
- $4096 \cdot 1000 + 1000 = 4097000$
- total $= 44782952$

Exercise 3 Backpropagation

Compute the gradient of the parameters of a 2-layer MLP f (with sigmoid activation functions) w.r.t. the loss function $||f(x; W^{(1)}, W^{(2)}, b^{(1)}, b^{(2)}) - y||_2^2$. Given the case that $y = f(x)$ and $z = g(y)$, you can make use of the multivariate chain rule $\frac{\partial z}{\partial x_i} = \sum_j \frac{\partial z}{\partial y_j} \frac{\partial y_j}{\partial x_i}$.

Solution

$$
\begin{aligned}
o^{(1)} &= W^{(1)} x \\
o^{(2)} &= o^{(1)} + b^{(1)} \\
o^{(3)} &= \sigma(o^{(2)}) \\
o^{(4)} &= W^{(2)} o^{(3)} \\
o^{(5)} &= o^{(4)} + b^{(2)} \\
o^{(6)} &= \sigma(o^{(5)}) \\
\ell &= ||o^{(6)} - y||_2^2
\end{aligned}
$$



$$\sigma(x) = \frac{1}{1+e^{-x}}$$

$$\sigma'(x) = \sigma(x)(1 - \sigma(x))$$

$$\frac{\partial \ell}{\partial b^{(2)}} = \sigma'(o^{(5)}) \odot 2(o^{(6)} - y)$$

$$\frac{\partial \ell}{\partial W^{(2)}} = \sigma'(o^{(5)}) \odot 2(o^{(6)} - y) \cdot o^{(3)T}$$

$$\frac{\partial \ell}{\partial o^{(3)}} = W^{(2)T} \cdot \sigma'(o^{(5)}) \odot 2(o^{(6)} - y)$$

$$\frac{\partial \ell}{\partial b^{(1)}} = \sigma'(o^{(2)}) \odot \frac{\partial \ell}{\partial o^{(3)}}$$

$$\frac{\partial \ell}{\partial W^{(1)}} = \sigma'(o^{(2)}) \odot \frac{\partial \ell}{\partial o^{(3)}} \cdot x^T$$