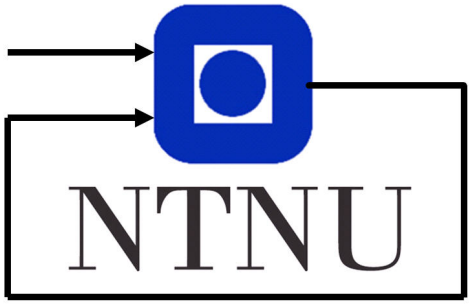
Image Processing - Assignment 2

Group 3
Martin Eek Gerhardsen

October 25



Department of Engineering Cybernetics

Contents

1 Convolutional Neural Networks

1.1 Task 1: Theory

a)

Listing the equations for calculating the resulting height and width:

$$W_{i+1} = (W_i - F_W + 2P_W)/S_W + 1 \tag{1}$$
$$H_{i+1} = (H_i - F_H + 2P_H)/S_H + 1 \tag{2}$$

Now defining:

$$\begin{aligned} S_W &= S_H = 1 \\ F_W &= F_H = 5 \\ H_2 &= H_1 = H \\ W_2 &= W_1 = W \end{aligned}$$

Then we get for the width:

$$\begin{aligned} W_2 &= (W_1 - F_W + 2P_W)/S_W + 1 \\ W - 1 &= W - 5 + 2P_W \\ 2P_W &= 4 \\ P_W &= 2 \end{aligned}$$

And for height:

$$\begin{aligned} H_2 &= (H_1 - F_H + 2P_H)/S_H + 1 \\ H - 1 &= H - 5 + 2P_H \\ 2P_H &= 4 \\ P_H &= 2 \end{aligned}$$

So we should therefore be padding with $P_H = 2$ and $P_W = 2$.

c)

Now defining:

$$\begin{aligned} S_H &= S_W = 1 \\ P_H &= P_W = 0 \\ H_1 &= W_1 = 512 \\ H_2 &= W_2 = 504 \\ F_H &= ? \\ F_W &= ? \end{aligned}$$

Then using the same equations as above, we calculate:

$$
\begin{aligned}
H_2 &= (H_1 - F_H + 2 * P_H)/S_H + 1 \\
504 &= (512 - F_H + 2 * 0)/1 + 1 \\
F_H &= 512 + 1 - 504 = 9
\end{aligned}
$$

$$
\begin{aligned}
W_2 &= (W_1 - F_W + 2 * P_W)/S_W + 1 \\
504 &= (512 - F_W + 2 * 0)/1 + 1 \\
F_W &= 512 + 1 - 504 = 9
\end{aligned}
$$

We also knew from the task that the kernel was supposed to be square, and its dimensions odd, which we can confirm from the result. The kernel must therefore be of the size $9x9$.

d)

If we start with subsampling with neighbourhoods of size $2x2$ and stride 2, then we'd decrease the width and height by half. Then, using the same information utilized above, we define:

$$
\begin{aligned}
H_1 &= W_1 = 512/2 = 256 \\
S_H &= S_W = 1 \\
P_H &= P_W = 0 \\
F_H &= F_W = 9 \\
H_2 &= ? \\
W_2 &= ?
\end{aligned}
$$

Then we calculate:

$$
\begin{aligned}
H_2 &= (H_1 - F_H + 2 * P_H)/S_H + 1 \\
&= (256 - 9 + 2 * 0)/1 + 1 = 248 \\
W_2 &= (W_1 - F_W + 2 * P_W)/S_W + 1 \\
&= (256 - 9 + 2 * 0)/1 + 1 = 248
\end{aligned}
$$

Then the spatial dimensions of the pooled feature maps in the first layer would be $248x248$.

e)

1.2 Task 2: Programming

2 Filtering in the Frequency Domain

2.1 Task 3: Theory

a)

From the convolution theorem, we can see that the Fourier transform of a convolution of two signals is multiplication of the Fourier transforms of those individual signals. So a stepwise description would be:

- Find the Fourier transform of the individual signals (i.e. using FFT)
- Pointwise multiply the transformed signals
- Use inverse Fourier transform to find the convolution of the original signals

b)

2.2 Task 4: Programming